Extended estimator approach for 2×2 games and its mapping to the Ising Hamiltonian

D. Ariosa[1] and H. Fort[2]
[1]*IPMC, Institute of Physics of Complex Matter, IPMC/FSB/EPFL, Lausanne, Switzerland*
[2]*Instituto de Física, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay*
(Received 21 September 2004; published 25 January 2005)

We consider a system of adaptive self-interested agents interacting by playing an iterated pairwise prisoner's dilemma (PD) game. Each player has two options: either cooperate (C) or defect (D). Agents have no (long term) memory to reciprocate nor identifying tags to distinguish C from D. We show how their 16 possible elementary Markovian (one-step memory) strategies can be cast in a simple general formalism in terms of an estimator of expected utilities Δ^*. This formalism is helpful to map a subset of these strategies into an Ising Hamiltonian in a straightforward way. This connection in turn serves to shed light on the evolution of the iterated games played by agents, which can represent a broad variety of individuals from firms of a market to species coexisting in an ecosystem. Additionally, this magnetic description may be useful to introduce noise in a natural and simple way. The equilibrium states reached by the system depend strongly on whether the dynamics are synchronous or asynchronous and also on the system connectivity.

DOI: 10.1103/PhysRevE.71.016132 PACS number(s): 89.75.−k, 87.23.Ge, 89.65.Gh, 89.75.Fb

I. INTRODUCTION

The self-organization into cooperative equilibrium states is a ubiquitous phenomenon in nature: electrons in a superconducting material, local magnetic moments in a ferromagnet, molecules that cooperate to form cells, cells that cooperate to form living creatures that in turn cooperate to form societies, and so on. This problem of how populations of self-interested agents (or agents who pursue to satisfy some goal locally or individually) cooperate (or manage in order to satisfy this goal globally or collectively) can be regarded from different points of view. The way biologists, economists, and physicists approach this problem is often connected with paradigms or core concepts of their fields, respectively: Darwin's evolution, homo economicus and statistical thermodynamics. These three approaches can be summarized in turn by three extremal principles: fitness maximization in biology, profit maximization in economics, and minimization of the free energy in physics.

Game theory offers a common and flexible framework to perform the comparison of the different approaches. It coalesced in its normal form [1] during the Second World War with the work of von Neumann and Morgenstern [2] who first applied it in economics. Later, in the 1970s, it was the turn of biology mainly with the work of Maynard-Smith [3], who applied game theory to evolution and proposed the concept of evolutionarily stable strategy (ESS) for understanding biological phenomena. The problem of the evolution of cooperative behavior (how can cooperation emerge in a world of egoists without central authority?) was analyzed extensively by Axelrod [4] in the 1980s. The computer tournaments he organized demonstrated that cooperation based upon reciprocity can emerge and prove stable. Applications include politics, economics, international affairs, etc. Indeed, neither consciousness nor a brain are required to play games: the results of recent experiments with two variants of a RNA virus can be interpreted as both variants engaged in a two-player game [5].

Very recently, game theory entered into physics as an alternative approach to physical problems. For instance, energies could be represented as payoffs and phenomena like phase transitions understood as many-agents games. As a particular application we have seen a proliferation of papers devoted to quantum games and quantum strategies [6–9], issues connected with efficient quantum algorithms for quantum computing and quantum cryptography [10]. Conversely, physics can be useful to understand the behavior of adaptive agents playing games that are used to model several complex systems in nature. An example of this is the application of the techniques developed in nonequilibrium statistical physics to biological contexts [11,12].

The most popular exponent of game theory is the *prisoner's dilemma* (PD) game introduced in the early 1950s by Flood [13] to model the social behavior of "selfish" individuals—individuals which pursue exclusively their own self-benefit. The PD game is an example of a 2×2 game in normal form: (i) there are two players, each confronting two choices—to cooperate (C) or to defect (D)—(ii) each player makes his choice without knowing what the other will do, and (iii) there is a 2×2 matrix specifying the payoffs of each player for the four possible outcomes [C,C], [C,D], [D,C], and [D,D]. This payoff matrix is written as

$$M = \begin{pmatrix} R & S \\ T & P \end{pmatrix}. \tag{1}$$

A player who plays C gets the "reward" R or the "sucker's payoff" S depending if the other player plays C or D, respectively, while if he plays D he gets the "temptation to defect" T or the "punishment" P depending if the other player plays C or D, respectively. These four payoffs obey the relations

$$T > R > P > S, \tag{2a}$$

plus the condition

$$2R > S + T. \tag{2b}$$

Condition (2b) is required in order to avoid the possibility of collusion between the pair of players. The dilemma is that, independently of what the other player does, defection D yields a higher payoff than cooperation C ($T>R$ and $P>S$); i.e., D is the *dominant* strategy for every player. But by playing D in a sequence of encounters, both players do worse than if both had cooperated ($P<R$). Indeed, when the PD game is played repeatedly, there are many strategies that outperform the dominant D strategy of the one-shot game. This is the content of the so-called *Folk theorem* [14]. Different mechanisms have been proposed to escape from the noncooperative one-shot dominant strategy in the case of iterated prisoner's dilemma (IPD) game. We might call the preponderant approach in social sciences *direct reciprocity* because reciprocity between agents is considered as the basis for cooperation [4]. In order to reciprocate, the agents need first to discriminate between cooperators and defectors. Therefore, either memory of previous interactions or features ("tags") [15] permitting one to distinguish those agents who respond to the cooperation and those who do not are required. In other words, cooperation becomes an equilibrium because no one will gain from defecting due to the retaliation and losses they would suffer. This is the philosophy behind a popular strategy known as *tit for tat* (TFT): cooperate on the first move, and then cooperate or defect exactly as the other player did on the preceding move.

There are other approaches that do not require memory or tags. For instance, Nowak and May [16], working in a biological context, proposed a different approach to the problem of the evolution of cooperation which neglects all strategical complexities or memories of past encounters. Instead, from this perspective, which we will call the *spatial evolution* approach, spatial effects by themselves, in a classical Darwinian setting, are sufficient to the evolution of cooperation. Another alternative, not belonging to the evolutionary game theory tradition, was proposed very recently [17]. It involves self-interested agents without memory of past encounters, without tags which, in principle, do not need any spatial structure (pairs of players can be selected at random instead of being chosen from a fixed neighborhood). At each time step, a pair of agents are selected at random to play. Each player i uses a simple "measure of success" to evaluate if he did well or badly in the game, which consists in comparing his utilities Δ_i with his estimate of expected income Δ_i^*. He updates next his behavior or state (C or D) in consonance; i.e., he keeps his behavior if he did well and modifies (inverts) it if he did badly. We call this the *estimator* approach. It is a generalization of the philosophy underlying the strategy of "win-stay, lose-shift" known as PAVLOV [18], which is also very popular within adaptable agent modeling (for instance, in the PD game, PAVLOV corresponds to take the estimate $\Delta_i^*=\Delta^*$ with Δ^* somewhere in between the punishment P and the reward R).

In this article, (1) we provide a unifying framework for all elementary one-step memory strategies in terms of an extension of the estimator formulation and (2) we present a mapping between a subset of these strategies and an Ising Hamiltonian (that generates a Monte Carlo dynamics identical to the one produced by them). It turns out that these *Ising mappable strategies* (IMS's) include TFT. This connection, besides its conceptual interest, opens the possibility to apply tools of statistical mechanics to extract relevant properties such as the ground-state configuration (related to the asymptotic limit of the iterated game) and spatial correlation functions as much as to introduce stochasticity through temperature in a simple and uniform manner. We also analyze the subset of strategies covered by the estimator approach that are not mappable onto the Ising model but admits a fixed estimate (independent of the obtained utilities). The case of PAVLOV is the most relevant example. The equilibrium states into which the system self-organizes roughly fall into three types: "universal cooperation" or "all C," intermediate level of cooperation, and "universal defection" or "all D" depending on the fraction of C individuals at equilibrium c_{eq}—i.e., respectively, $c_{\mathrm{eq}}=1$, $0<c_{\mathrm{eq}}<1$, and $c_{\mathrm{eq}}=0$.

We would like to remark that the treatment we present here is nonevolutionary, at least from the traditional Darwinian point of view, since we do not consider competition of different strategies and subsequently the survival of the fittest.

II. EXTENDED ESTIMATOR APPROACH FOR ELEMENTARY MARKOVIAN STRATEGIES

A. Extended estimator formulation

Originally the estimate for each player Δ^* was taken fixed and the same for all the players [17]. Now we consider an extended estimate which, in general, depends on the two-player behavioral variables. Thus, for a given payoff matrix, like Eq. (1), a strategy can be defined by specifying its corresponding estimate. The state of a given player (i) is represented by a two-component vector $\boldsymbol{S}_i$. A general expression for $\boldsymbol{S}_i$ in terms of the behavioral variable c_i is given by

$$\boldsymbol{S}_i = \begin{pmatrix} c_i \\ 1-c_i \end{pmatrix}, \tag{3}$$

where the cooperation variable $c_i=1$ for the cooperative (C) state $\boldsymbol{S}_i=\boldsymbol{C}\equiv\binom{1}{0}$ and $c_i=0$ for the non-cooperative (D) state $\boldsymbol{S}_i=\boldsymbol{D}\equiv\binom{0}{1}$. In this basis, the payoff matrix $\boldsymbol{M}$ is given by Eq. (1). The utilities obtained by agent i playing with agent j can be expressed then as

$$\Delta_i = \boldsymbol{S}_i^{\boldsymbol{T}}\boldsymbol{M}\boldsymbol{S}_j = (R-S-T+P)c_ic_j + (S-P)c_i + (T-P)c_j + P, \tag{4a}$$

which generalizes as

$$\Delta_i = \sum_{nn} \boldsymbol{S}_i^T\boldsymbol{M}\boldsymbol{S}_{nn} = \boldsymbol{S}_i^T\boldsymbol{M}z\langle \boldsymbol{S}_{nn}\rangle, \tag{4b}$$

when agent i plays with all the z agents of his neighborhood (one at a time). In Eq. (4b) the subscript nn stands for nearest neighbors and $\langle \boldsymbol{S}_{nn}\rangle$ denotes the average over them.

In the extended version of the estimator formulation, a general strategy for a two players' game consists of flipping the state $\boldsymbol{S}_i$ *if and only if* the obtained utilties after playing against j are smaller than the estimate—i.e.,

$$\boldsymbol{S}_i' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\boldsymbol{S}_i \Leftrightarrow \Delta_i < \Delta^*(\boldsymbol{S}_i,\boldsymbol{S}_j), \tag{5a}$$

which, again, generalizes as

$$\mathbf{S}_i' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{S}_i \Leftrightarrow \Delta_i < z\langle \Delta^*(\mathbf{S}_i, \mathbf{S}_{nn})\rangle, \tag{5b}$$

when agent i plays with the z agents of his neighborhood. Here, $\Delta^*(\mathbf{S}_i, \mathbf{S}_j)$ is the estimate that defines the particular strategy, and $\mathbf{S}_i'$ is the new state of the player i after playing against j, according to the rule Eq. (5a). Note that Eq. (5a) is equivalent to the flipping condition at $T=0$ in a Metropolis [19] algorithm if $\Delta_i - \Delta^*$ is replaced by the energy variation when performing the updating of the local configuration: the state is flipped with probability 1, only if $\Delta_i - \Delta^* < 0$.

B. Mapping the iterated game into an Ising model

Let us now show how the estimator approach can be mapped into an Ising model. The equivalence between the two-valued behavioral variable of a given player (C or D) and a magnetic Ising spin $\sigma = 1/2$ is pretty obvious. Furthermore, the similitude between the update rule for c_i in the estimator approach when playing the IPD game and the Metropolis update rule used in Monte Carlo simulations will serve as a guide to establish this connection. Thus we begin by defining the Ising spins associated to the behavioral variables:

$$\sigma_i = c_i - 1/2 = \begin{cases} +1/2 & \text{for } (i) \text{ in the C state}, \\ -1/2 & \text{for } (i) \text{ in the D state}. \end{cases} \tag{6}$$

In the Ising language, the equation for the utilities (5a) and (5b) becomes

$$\Delta_i = (R - S - T + P)\sigma_i\sigma_j + \frac{1}{2}(R + S - T - P)\sigma_i + \frac{1}{2}(R - S + T - P)\sigma_j + \frac{1}{4}(R + S + T + P), \tag{7a}$$

which, in the case of z nearest neighbors, is generalized as

$$\Delta_i = \sum_{nn} \left((R - S - T + P)\sigma_i\sigma_{nn} + \frac{1}{2}(R + S - T - P)\sigma_i + \frac{1}{2}(R - S + T - P)\sigma_{nn} + \frac{1}{4}(R + S + T + P) \right). \tag{7b}$$

Hence, the flipping conditions become, respectively,

$$\sigma_i' = -\sigma_i \Leftrightarrow \Delta_i < \Delta^*(\sigma_i, \sigma_j), \tag{8a}$$

$$\sigma_i' = -\sigma_i \Leftrightarrow \Delta_i < z\langle \Delta^*(\sigma_i, \sigma_{nn})\rangle. \tag{8b}$$

Note that, since Δ_i is linear in σ_j, the interaction with multiple neighbors reduces to the interaction with the average spin $\langle \sigma_{nn}\rangle$.

Once we have made the translation between estimator and magnetic variables, we are ready to connect the Ising Hamiltonian with the iterated PD. We consider the Ising Hamiltonian in uniform magnetic field:

$$H = J \sum_{\langle i,j\rangle} \sigma_i\sigma_j + h \sum_i \sigma_i. \tag{9}$$

We will refer to J as the coupling constant and to h as the external field. The first sum is taken over all the interacting pairs of sites (i,j) and the second over all the sites.

The local energy change associated with the flipping of the spin at site i is given by

$$\Delta E_i = -2[Jz\langle \sigma_{nn}\rangle + h]\sigma_i. \tag{10}$$

The most general expression for the estimate $\Delta^*(\sigma_i, \sigma_j)$ is of the form of Eq. (7a) with arbitrary coefficients. In other words, the estimate can be defined as the fictitious payoff corresponding to an alternative matrix,

$$\mathbf{M}^* = \begin{pmatrix} R^* & S^* \\ T^* & P^* \end{pmatrix},$$

with obviously these four possibilities:

$$\Delta^*(1/2, 1/2) = R^*, \quad \Delta^*(1/2, -1/2) = S^*,$$

$$\Delta^*(-1/2, 1/2) = T^*, \quad \Delta^*(-1/2, -1/2) = P^*.$$

The way we choose to map the Ising system on IPD is by identifying ΔE_i with the difference $\Delta_i - \Delta^*(\sigma_i, \sigma_j)$ between the obtained payoff and the estimate. But only a reduced class of strategies can be mapped this way. Indeed, while expression (7a) contains a mixed term $(\sigma_i\sigma_j)$, two independent terms (σ_i and σ_j), and a constant term, expression (10) contains only mixed $(\sigma_i\sigma_j)$ and single (σ_i) terms. Thus, in order to perform the mapping $\Delta^*(\sigma_i, \sigma_j)$ must be chosen such that the terms proportional to (σ_j) and constant one cancel out. This restrictive condition in terms of the matrix M^* reads

$$\begin{aligned} &\tfrac{1}{2}(R^* - S^* + T^* - P^*)\sigma_j + \tfrac{1}{4}(R^* + S^* + T^* + P^*) \\ &\quad = \tfrac{1}{2}(R - S + T - P)\sigma_j + \tfrac{1}{4}(R + S + T + P) \\ &\quad \Rightarrow R^* + T^* = R + T \quad \text{and} \quad S^* + P^* = S + P. \end{aligned} \tag{11}$$

Acceptable strategies for the mapping are thus generated by specifying two independent parameters ϵ_1 and ϵ_2:

$$\epsilon_1 = S^* - S = P - P^* \quad \text{and} \quad \epsilon_2 = R^* - R = T - T^*. \tag{12}$$

With the above parametrization, the mapping takes a very simple form:

$$\Delta E_i = \Delta_i - z\Delta^*(\sigma_i, \langle \sigma_{nn}\rangle) = 2z\Big[(\epsilon_1 - \epsilon_2)\langle \sigma_{nn}\rangle - \tfrac{1}{2}(\epsilon_1 + \epsilon_2)\Big]\sigma_i$$

$$\Rightarrow J = \epsilon_2 - \epsilon_1 \quad \text{and} \quad h = \frac{z}{2}(\epsilon_1 + \epsilon_2). \tag{13}$$

The Metropolis algorithm, applied to single site updating, states that the probability for a given site to accept a given change (from 1 to 2) depends on the associated internal energy difference:

$$p_{1\to 2} = \frac{1}{1 + \exp\left(\frac{E_2 - E_1}{k_B T}\right)}.$$

Thus the temperature-dependent flipping probabilities are

TABLE I. Main features of nine elementary Markovian strategies: five with fixed estimate plus four with conditional estimate and mappable on the Ising model. ARD stands for asynchronous random dynamics; SFC stands for synchronous fully connected.

	Estimate				Steady-state cooperation	
Strategy	Fixed	Conditional		Character (p_R, p_T, p_S, p_P)	ARD	SFC
FROZEN	$\Delta^* < S$			(1,0,1,0)	Arbitrary, $c^* = c_0$	Depending on initial configuration (See Fig. 1)
RETALIATOR	$S < \Delta^* < P$			(1,0,0,0)	$c^* = 0$	Stable: $c^* = 0; 1; 1/2; c_0; 1 - c_0$ $c^* = (1/2)$
PAVLOV	$P < \Delta^* < R$	NO		(1,0,0,1)	$c^* = 1/2$	Oscillating: $1 <-> 0$; $c_0 <-> 1 - c_0$
AMBITIOUS	$R < \Delta^* < T$			(0,0,0,1)	$c^* = (3 - \sqrt{5})/2 = 0.38$	
ALTERNATOR	$T < \Delta^*$			(0,1,0,1)	$c^* = 1/2$	
		ε_1	ε_2			
ALWAYS D	NO	>0	>0	(0,0,0,0)	$c^* = 0$	(See Fig. 2) $c^* = 0$ or $c^* = 1$
TIT-FOR-TAT		>0	<0	(1,1,0,0)	Arbitrary, $c^* = c_0$	
CONTRADICTOR		<0	>0	(0,0,1,1)	$c^* = 1/2$	
ALWAYS C		<0	<0	(1,1,1,1)	$c^* = 1$	

$$p_f = \left[1 + \exp\left(\frac{2z\sigma_i \left[(\epsilon_1 - \epsilon_2)\langle \sigma_{nn} \rangle - \frac{1}{2}(\epsilon_1 + \epsilon_2) \right]}{k_B T} \right) \right]^{-1}. \tag{14}$$

At temperature $T = 0$, only the sign of the energy difference matters: the change is accepted with probability 1 if $\Delta E < 0$ and rejected otherwise.

C. Classifying Markovian strategies

A useful scheme for the classification of Markovian strategies is based on the four conditional probabilities p_R, p_T, p_S, and p_P that an individual plays C, in a two-players' game, after it scored R, T, S, or P, respectively, in the previous round [20,21]. The strategies can be represented as points in this four-dimensional strategy space (p_R, p_T, p_S, p_P). Here, we will restrict attention to *elementary Markovian strategies* in which p_R, p_T, p_S, and p_P are either equal to 0 or 1 (binary agents). Thus there are $2^4 = 16$ possible strategies. All these strategies can be formulated in terms of the extended or *conditional* estimate which in general depends on the state of both players. It turns out that the mapping procedure defined by Eqs. (9)–(14) generates a subset of four *Ising mappable strategies*, depending on the signs of ϵ_1 and ϵ_2. For instance, $\epsilon_1 > 0$ and $\epsilon_2 > 0$ is equivalent to $R < R^*$, $T > T^*$, $S < S^*$, and $P > P^*$ and hence this implies—according to Eq. (5a)—$p_R = p_T = p_S = p_P = 0$: i.e., (0,0,0,0) or ALWAYS D strategy. Another interesting situation is $\epsilon_1 > 0$ and $\epsilon_2 < 0$ that results in (1,1,0,0) or TFT: imitate in the next round what your opponent did in the present round. The other two strategies that complete the IMS subset are the opposite of these two: ALWAYS C and CONTRADICTOR. The remaining $16 - 4 = 12$ strategies are nonmappable onto the Ising model. This subset of strategies can in turn be divided into two types whether they admit or not a fixed estimate ($\Delta^* = \text{const}$). Thus, on the one hand, we have five *fixed estimate strategies* (FES's). Among them, we recognize the usual RETALIATOR (1,0,0,0), when $P > \Delta^* > S$, and PAVLOV (1,0,0,1) strategies, when $R > \Delta^* > P$, both especially relevant for evolutionary schemes. The three remaining less studied strategies are FROZEN (1,0,1,0), AMBITIOUS (0,0,0,1), and ALTERNATOR (0,1,0,1). On the other hand, we have the $16 - (5 + 4) = 7$ strategies, which cannot be formulated in terms of a fixed estimate and only accept a conditional estimate depending on the state of pairs of players. For instance, the ANTI-PAVLOV (0,1,1,0) implies estimator-matrix elements obeying $S^* < S$, $P^* < P$, $R^* > R$, and $T^* > T$.

The FES and IMS subsets, which have null intersection, comprise the most relevant strategies. In Table I we summarize the main features of these nine strategies. The five possible regions delimited by the four real numbers of M^* yield the five different FES's (upper part of Table I). The characters of the IMS's are the Metropolis flipping probabilities generated by Eq. (14) at zero temperature.

III. ANALYSIS OF ISING MAPPABLE STRATEGIES

Let us analyze the IMS strategies for the simplest case of $z = 1$. Notice that, from Eq. (10), for the TFT (CONTRADICTOR) strategy the sign of ΔE_i is equal to the sign of $\sigma_i \sigma_j$ $(-\sigma_i \sigma_j)$, $J < 0$ (>0) and h can take a positive, null or negative value; i.e., it corresponds to a ferromagnetic (antiferromagnetic) material in an arbitrary external magnetic field. On the other hand, in the case of ALWAYS-C and ALWAYS-D strategies the corresponding material can be a ferromagnet or an antiferromagnet and what determines the state is the direction of the external field, always stronger than the coupling constant J (cooperative state if h is negative and all the spins point in the upward direction, and noncooperative state if the field is positive and the spins point in the downward direction).

A measure for the attained degree of cooperation in the system is naturally provided by the average cooperation—i.e., the fraction c of C agents. The value c_{eq} of this fraction in the *ground state*, for IMS's, is related to the ground-state average magnetization $\langle\sigma_i\rangle_0$ of the Ising system:

$$c_{\rm eq} = \frac{1}{2} + \langle\sigma_i\rangle_0. \tag{15}$$

However, this thermodynamic ground state is not always attained during an IPD game, especially for highly connected systems. This is due to the fact that high connectivity implies the simultaneous updating of a given player and many of its opponents; thus, starting from a given configuration, the attainable configurations are just a small subset of the phase space. In the following, we have thus to distinguish between the steady-state (after a transient) average cooperation c^* and the thermodynamic average cooperation $c_{\rm eq}$. To bypass this nonergodicity, one has to use "generous" versions of the considered strategy, where the probabilities in the character (p_R, p_T, p_S, p_P) are different from strictly 0 or 1. From the thermodynamic point of view, this is equivalent to consider finite-temperature Monte Carlo dynamics.

IV. STABILITY OF COOPERATION

At this stage, it is necessary to make a distinction between *synchronous* and *asynchronous dynamics*. Indeed, while the characters of the different generated strategies have been defined on a "two-player" basis, the evolution will strongly depend on the type of dynamics and connectivity we assume for the system. When all the agents simultaneously update their states in each round, we talk about synchronous dynamics. Conversely, when the update is carried out for the subset of agents who just played, we talk about asynchronous dynamics. The asymptotic configuration also depends on the connectivity. Two limits will be addressed: the *asynchronous random dynamics* (ARD), in which a pair of players is randomly chosen for each round, and the *synchronous fully connected* (SFC) dynamics, where each player interacts simultaneously with all the $z=(N-1)$ remaining players in the system.

A. Stability in asynchronous random dynamics

In this case c^* for a large system $(N \gg 1)$ can be easily evaluated, both for FES's and IMS's, by equating the variation rate of the C population to zero:

$$\delta c = -c^2(1-p_R) - c(1-c)(1-p_S-p_T) + (1-c)^2 p_P = 0, \tag{16}$$

where the flipping probabilities are related to the system average cooperation c through the probabilities in the character (p_R, p_T, p_S, p_P). As an example, let's calculate the ARD equilibrium cooperation for the character (1, 0, 0, 1)—i.e., the PAVLOV strategy—in which a cooperator will flip if it scores S, while a defector will flip if it scores P:

$$c(1-c) = (1-c)^2 \Leftrightarrow c^* = \tfrac{1}{2} \quad \text{or} \quad c^* = 1.$$

Here, only $c^*=1/2$ is a stable solution since, for all but one cooperating agents, the system is rapidly driven away from $c=1$. This can be easily proven by noticing that, in the PAVLOV random-asynchronous strategy, minority defectors are most often satisfied since playing against a majority of cooperators, forcing them to defect at each round. In contrast, close to $c=1/2$, the probability for a cooperator to defect is about the same as the probability for a defector to cooperate. The same reasoning applies to RETALIATOR, for which the only stable cooperation is $c^*=0$. The steady state cooperation values for all considered strategies, within the ARD, have been calculated and included in the last but one column of Table I.

B. Stability in the synchronous fully connected system

In that case it is necessary to distinguish between the FES's and IMS's.

(I) FES. The equilibrium state strongly depends upon the initial configuration, producing a quite complex phase diagram in terms of initial configurations and values of Δ^*. From Eq. (4b), as $z=N-1$, the obtained utilities can be written as a function of c:

$$\Delta_i = \sum_{nn} \boldsymbol{S}_i^T \boldsymbol{M} \boldsymbol{S}_{nn} = \boldsymbol{S}_i^T \boldsymbol{M} \boldsymbol{z} \langle \boldsymbol{S}_{nn} \rangle = (\boldsymbol{N}-\boldsymbol{1}) \begin{cases} cR + (1-c)S, & \text{for } \boldsymbol{S}_i = \boldsymbol{C}, \\ cT + (1-c)P, & \text{for } \boldsymbol{S}_i = \boldsymbol{D}. \end{cases} \tag{17}$$

A stable configuration is reached when all players get a payoff greater than the estimate Δ^* or, in other words, when Δ^* is lower than the cooperator's utilities. Marginal stability is reached also when all players defect and Δ^* is lower than the defector's utilities. When Δ^* falls in between cooperator's and defector's utilities, defectors will remain in the D state, while former cooperators will all start to defect. Thus, the system is driven to $c=0$ in one round. From the latter configuration, two outcomes are possible: (i) if $\Delta^* < P$, all players (defectors) are satisfied with the payoff and the system is stable; (ii) if $\Delta^* > P$, all players are unhappy and the system rigidly flips to $c=1$ in the next round. But here, again, we are dealing with a bifurcation: the system will be stable only if $\Delta^* < R$ and will otherwise oscillate between $c=0$ and $c=1$. Similarly, for Δ^* greater than the defector's payoff, using the same kind of reasoning as above, we can distinguish four different steady-state configurations: stable $c^*=1-c_0$ (where c_0 is the initial average cooperation), oscillations between $c=0$ and $c=1$, stable $c^*=1$, and oscillations between $c=c_0$ and $c=1-c_0$. The $(c_0;\Delta^*)$ phase diagram in Fig. 1 summarizes the above discussion.

(II) IMS. In this case, the steady state corresponds to the asymptotic limit of the $T=0$ Metropolis dynamics. As we have anticipated, this steady state is not the thermodynamic ground state. Performing the same kind of analysis as we did for FES's, we obtain a phase diagram in which the steady state is either $c=0$ or $c=1$, depending both upon the initial average cooperation c_0 and upon the mapping parameters ϵ_1 and ϵ_2. In order to illustrate this situation, we displayed the phase diagram in Fig. 2: the distance of a given point to the origin, in this diagram, corresponds to the value of the initial average cooperation $(0 < c_0 < 1)$; its location in one of the four quadrants corresponds to the considered IMS's, as indicated for each quadrant. The phase space turns out to be

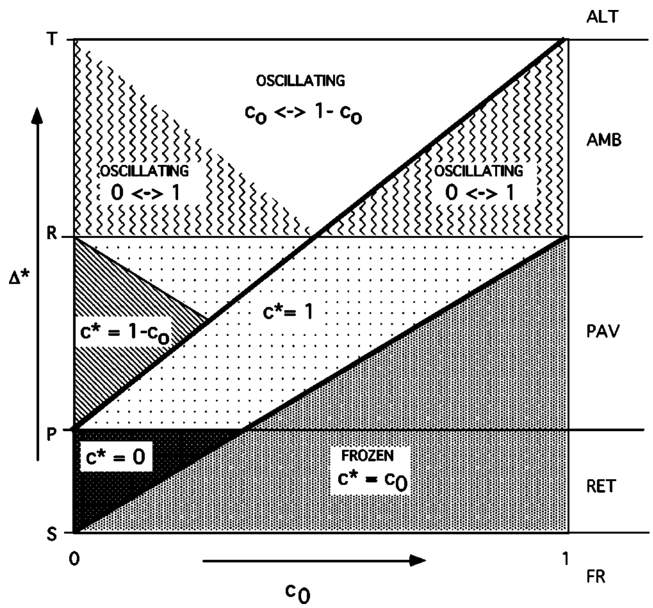

FIG. 1. Phase diagram indicating the FES steady-state average cooperation for SFC dynamics, as a function of the initial average cooperation c_0 and the fixed estimate Δ^*. Note the two bold lines across the diagram, representing the two payoff functions defined in Eq. (17).

divided into two regions for which the steady-state value of cooperation is 0 or 1. For ALWAYS D and ALWAYS C strategies, the steady state is trivially $c=0$ and $c=1$, respectively. The boundary in the TFT quadrant is given by the polar equation $c_0=1+2\theta/\pi$ and in the CONTRADICTOR quadrant by $c_0=-1+2\theta/\pi$. These two equations are the polar representation of the critical cooperation value $\tilde{c}=|\epsilon_1|/(|\epsilon_1|+|\epsilon_2|)$ dividing the c_0 axis in two regions in CONTRADICTOR and TFT strategies.

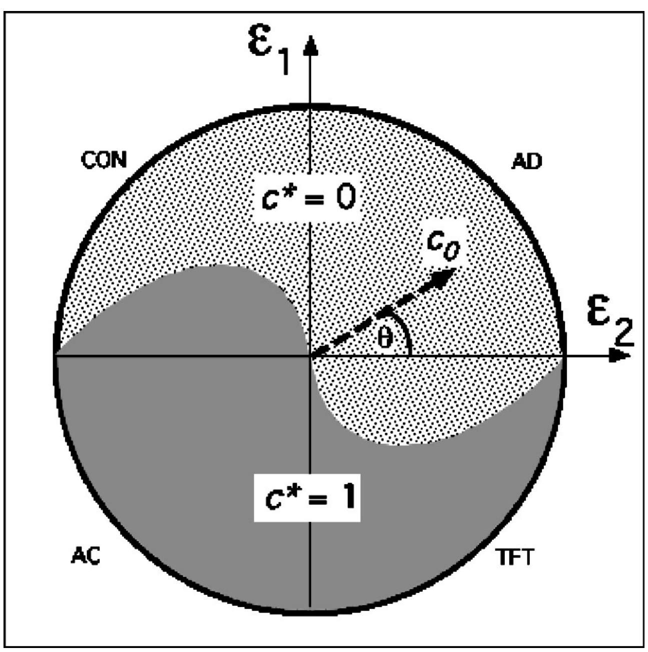

FIG. 2. Phase diagram indicating the IMS steady-state average cooperation for SFC dynamics, as a function of the initial average cooperation c_0 and the mapping parameters ϵ_1 and ϵ_2, as discussed in the text. CON, AD, and AC are abbreviations for CONTRADICTOR, ALWAYS C, and ALWAYS D, respectively.

V. THERMODYNAMICS PREDICTIONS FOR THE IMS CASE

An advantage of the formulation in terms of the magnetic variables is the possibility to introduce noise or fluctuations in a straightforward way by considering the case of nonzero temperature. The fluctuations can help this self-organizing system to explore and find more stable and, eventually, efficient equilibrium states in its "fitness" landscape. In other words, the injection of noise or random perturbations into the system will allow it to escape for an eventually "shallow" local minimum for the free energy and to reach a deeper valley. Let us analyze the IMS's in terms of the associated Ising model. The Hamiltonian (10), for fully connected systems $(z=N-1)$, reduces to the internal energy functional $U(\langle\sigma\rangle)$ and, at $T=0$, to the free energy functional $F(\langle\sigma\rangle)$:

$$H = J\sum_{\langle i,j\rangle}\sigma_i\sigma_j + h\sum_i \sigma_i = N\left[\frac{J}{2}N\langle\sigma\rangle^2 - \frac{J}{2}\langle\sigma^2\rangle + h\langle\sigma\rangle\right]$$

$$\equiv U(\langle\sigma\rangle) \Rightarrow F(\langle\sigma\rangle) = N\left[\frac{J}{2}N\langle\sigma\rangle^2 + h\langle\sigma\rangle - \frac{J}{8}\right]. \qquad (18)$$

According to Eq. (13), which defines the two parameters J and h, we can write

$$F(\langle\sigma\rangle) = \frac{N^2}{2}\left[(\epsilon_2-\epsilon_1)\langle\sigma\rangle^2 + \frac{N-1}{N}(\epsilon_1+\epsilon_2)\langle\sigma\rangle - \frac{\epsilon_2-\epsilon_1}{4N}\right]. \qquad (19)$$

The ground-state magnetization $\langle\sigma_i\rangle_0$ is the one that minimizes $F(\langle\sigma\rangle)$ within the interval $[-\frac{1}{2};+\frac{1}{2}]$. For large systems $(N \gg 1)$, the free energy functional reads

$$F(\langle\sigma\rangle) \approx \frac{N^2}{2}[(\epsilon_2-\epsilon_1)\langle\sigma\rangle^2 + (\epsilon_1+\epsilon_2)\langle\sigma\rangle].$$

When both ϵ_1 and ϵ_2 have the same sign, since $|\epsilon_2|-|\epsilon_1| < |\epsilon_2|+|\epsilon_1|$, the ground state is determined by the sign of the external field $(h \cong \epsilon_1+\epsilon_2)$, regardless of the sign of the coupling constant $(J=\epsilon_2-\epsilon_1)$. The minima will be then located on the edges of the allowed interval:

$$\text{for ALWAYS D} \quad (h>0): \quad \langle\sigma\rangle_0 = -\tfrac{1}{2}, \quad \text{i.e.,} \quad c_{\text{eq}}=0,$$

$$\text{for ALWAYS C} \quad (h<0): \quad \langle\sigma\rangle_0 = +\tfrac{1}{2}, \quad \text{i.e.,} \quad c_{\text{eq}}=1,$$

coinciding with the steady state obtained with the $T=0$ Metropolis dynamics (AD and AC quadrants in Fig. 2).

When $\epsilon_1>0$ and $\epsilon_2<0$ (TFT), we are mapped on a ferromagnetic system $(J<0)$ and the ground state coincides again with the ends of the interval, determined by the sign of the external field:

$$\text{for } \epsilon_1>\epsilon_2 \quad (h>0): \quad \langle\sigma\rangle_0 = -\tfrac{1}{2}, \quad \text{i.e.,} \quad c_{\text{eq}}=0,$$

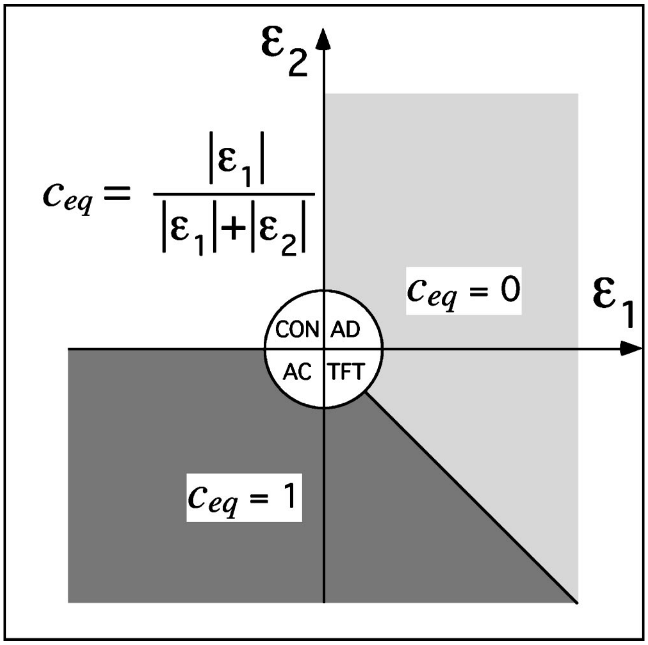

FIG. 3. Phase diagram indicating the IMS thermodynamic ground-state average cooperation for SFC dynamics, as a function of the two mapping parameters ϵ_1 and ϵ_2.

$$\text{for } \epsilon_1 < \epsilon_2 \quad (h<0): \quad \langle\sigma\rangle_0 = +\tfrac{1}{2}, \quad \text{i.e.,} \quad c_{\text{eq}} = 1.$$

This is in clear contrast to the phase diagram of Fig. 2. Here, the frontier between the ground states $c_{\text{eq}}=0$ and $c_{\text{eq}}=1$, in the TFT quadrant, is a straight diagonal line regardless of the initial configuration.

Finaly, when $\epsilon_1<0$ and $\epsilon_2>0$ (CONTRADICTOR), we are mapped on an antiferromagnetic system ($J>0$) and the ground state falls in the interior of the interval. The precise value of the ground-state magnetization $\langle\sigma\rangle_0$ is obtained for the derivative of $F(\langle\sigma\rangle)$ going to zero:

$$\frac{\partial F(\langle\sigma\rangle)}{\partial\langle\sigma\rangle} = 0 \Leftrightarrow \langle\sigma\rangle_0 = -\frac{1}{2}\frac{\epsilon_1+\epsilon_2}{\epsilon_2-\epsilon_1} = \frac{1}{2}\frac{|\epsilon_1|-|\epsilon_2|}{|\epsilon_1|+|\epsilon_2|} \Rightarrow c_{\text{eq}} = \frac{|\epsilon_1|}{|\epsilon_1|+|\epsilon_2|}.$$

Here again, the results differ drastically from the $T=0$ Metropolis dynamics. The ground state depends on the precise values of the model parameters, varying continuously between 0 and 1.

The above results—which are independent of the initial configuration—are summarized in the phase diagram of Fig. 3.

For fully connected systems, we expect the mean-field (MF) treatment to provide an exact solution. Indeed, the MF Hamiltonian can be written as follows:

$$H_{MF} = J\sum_{i\neq j}\sigma_i z\langle\sigma_j\rangle + h\sum_i \sigma_i = \sum_i \tilde{h}\sigma_i, \tag{20}$$

with $\tilde{h}=[zJ\langle\sigma\rangle+h]$. The corresponding partition function reads

$$Z_{MF} = \text{Tr}\{\exp(-\beta H_{MF})\} = \left[2\cosh\left(\frac{\beta\tilde{h}}{2}\right)\right]^N. \tag{21}$$

And the implicit equation for the average magnetization is

$$\langle\sigma\rangle = \frac{1}{N}\frac{\partial\ln(Z_{MF})}{\partial h} = -\frac{1}{2}\tanh\left\{\beta\frac{zJ\langle\sigma\rangle+h}{2}\right\}. \tag{22}$$

At the zero-temperature limit ($\beta\to\infty$), the right-hand term in Eq. (22) is a step function centered at $\langle\sigma\rangle=-h/zj$, jumping from $\pm\frac{1}{2}$ to $\mp\frac{1}{2}$, depending on the sign of J; its intersection with the linear left-hand term will give the MF solution.

When both ϵ_1 and ϵ_2 have the same sign, the intersection is unique, and the results are the same as indicated in Fig. 2.

When $\epsilon_1<0$ and $\epsilon_2>0$, the unique intersection is at

$$\langle\sigma\rangle_0 = \frac{-h}{zj} = \frac{1}{2}\frac{\epsilon_1+\epsilon_2}{\epsilon_1-\epsilon_2} \Rightarrow c_{\text{eq}} = \frac{1}{2} + \langle\sigma\rangle_0 = \frac{\epsilon_1}{\epsilon_1-\epsilon_2} = \frac{|\epsilon_1|}{|\epsilon_1|+|\epsilon_2|},$$

coinciding also with the results in Fig. 2.

When $\epsilon_1>0$ and $\epsilon_2<0$ (TFT), three intersections are possible, but only one corresponds to the free energy minimum (ferromagnetic system in external field). The two solutions are $c_{\text{eq}}=0$ (for $|\epsilon_1|>|\epsilon_2|$) and $c_{\text{eq}}=1$ (for $|\epsilon_1|<|\epsilon_2|$), again in perfect agreement with Fig. 2.

Thus, the ground state of the fully connected system is the same of the MF Hamiltonian. This ground state also applies, as a good approximation, to highly connected systems—i.e., when $(N-1)>z\gg 1$.

VI. DISCUSSION

In summary, it was shown that all the relevant elementary Markovian strategies can be formulated in terms of an extended (conditional) estimate. These strategies have been studied and their equilibrium states computed for two extreme situations: ARD and SFC. Two of the most popular agent strategies—PAVLOV and TFT—have been identified as particular examples of the estimator-based strategies belonging to different categories. While the second is mappable onto the Ising Hamiltonian the first is not.

The exploitation of the mapping between a subset of the space of simple Markovian strategies and the Ising model, presented in this work, will be analyzed in a future publication. Here we only explored some of its more straightforward consequences.

In addition, this correspondence can be extended beyond the Ising model. First, instead of considering uniform agents a more realistic assumption is the heterogeneity in the degree of selfishness. This can be accomplished by taking an estimate Δ^* that varies from agent to agent—i.e., a Δ_i^* dependent on the site i (or equivalently the fictitious payoffs varying from agent to agent). This would correspond to a spin glass instead of an Ising model. Second, the binary behavior assumption (C or D) is often criticized as unrealistic. In real life situations the agents exhibit different degrees of cooperation. This feature can be overcome by resorting to continuous behavioral variables c_i instead of binary ones as it was considered in [22]. In that case the mapping with magnetic systems would lead to more rich models like the XY model in which topological excitations govern its phase diagram. Work on these directions is in progress.

The connection between the generalized estimator approach and the spatial evolutionary games of [16] is also worth analyzing in the context of a general discussion of the cooperation between self-interested agents from the point of view of biologists, economists, and physicists. Moreover, a heterogeneous estimator formalism can be used to implement evolutionary games.

To conclude, it is worth remarking that the study we carried out for the PD game is valid for an arbitrary 2×2 game with a payoff matrix of the form of Eq. (1) but payoffs R, S, T, and P not obeying relations (2a) and (2b). For example, an alternative model of cooperation in human societies is the *Stag Hunt* game [23] also known as the *Assurance* game in which, $R>T>P>S$ and, thus, rational agents are pulled in one direction by considerations of risk and in another by considerations of mutual benefit [24]. Another relevant game to explore with the formalism presented here, which is useful in the context of population dynamics, is the *Hawk Dove* game introduced by Maynard-Smith [3]. In that case the punishment the players got when both play D is so strong that $R>T>S>P$.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation and by the EPFL.

[1] J. Hofbauer and K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, England, 1988).
[2] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, 1944).
[3] J. Maynard-Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, England, 1982).
[4] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).
[5] P. E. Turner and L. Chao, Nature (London) **398**, 441 (1999).
[6] D. A. Meyer, Phys. Rev. Lett. **82**, 1052 (1999).
[7] J. Eisert, M. Wilkens, and M. Lewenstein, Phys. Rev. Lett. **83**, 3077 (1999).
[8] C. F. Lee and N. Johnson, e-print quant-ph/0207012.
[9] C. F. Lee and N. Johnson, Phys. Lett. A **301**, 343 (2002).
[10] L. Goldenberg, L. Vaidman, and S. Wiesner, Phys. Rev. Lett. **82**, 3356 (1999).
[11] G. Szabó and C. Töke, Phys. Rev. E **58**, 69 (1998).
[12] G. Szabó, T. Antal, P. Szabó, and M. Droz, Phys. Rev. E **62**, 1095 (2000).
[13] M. Flood (unpublished).
[14] The iterated game has to be played an indefinite number of rounds; otherwise, if the players know when it ends, D remains as the dominant strategy. See R. B. Myerson, *Game Theory: Analysis of Conflict* (Harvard University Press, Cambridge, MA, 1991).
[15] J. Epstein, Complexity **4**(5), 41 (1999).
[16] M. Nowak and R. May, Int. J. Bifurcation Chaos Appl. Sci. Eng. **3**, 35 (1993); Nature (London) **359**, 826 (1992).
[17] H. Fort, J. Artificial Societies Social Simul. JASSS (UK) **6**(2), 4 (2003) (http://www.soc.surrey.ac.uk/JASSS/).
[18] D. Kraines and V. Kraines, Theory Decision **26**, 47 (1988).
[19] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. J. Chem. Phys. **21**, 1087 (1953).
[20] M. Nowak and K. Sigmund, Nature (London) **364**, 56 (1993).
[21] K. Brauchli, T. Killingback, and M. Doebeli, J. Theor. Biol. **200**, 405 (1999).
[22] H. Fort and S. Viola, Phys. Rev. E **69**, 036110 (2004).
[23] This game can be traced to Rousseau's "A Discourse on Inequality" in which he contrasts the payoff of hunting hare where the risk of noncooperation is small but the reward is equally small, against the payoff of hunting the stag where maximum cooperation is required but where the reward is so much greater.
[24] B. Skyrms, *The Stag Hunt and the Evolution of Social Structure* (Cambridge University Press, Cambridge, England, 2004).