Dynamics of competitive learning: The role of updates and memory

Ajaz Ahmad Bhat^{*} and Anita Mehta[†]

Theoretical Science Department, Satyendra Nath Bose National Centre, Block JD Sector 3, Salt Lake, Kolkata 700098, India

(Received 2 July 2011; revised manuscript received 29 December 2011; published 19 January 2012)

We examine the effects of memory and different updating paradigms in a game-theoretic model of competitive learning, where agents are influenced in their choice of strategy by both the choices made by, and the consequent success rates of, their immediate neighbors. We apply parallel and sequential updates in all possible combinations to the two competing rules and find, typically, that the phase diagram of the model consists of a disordered phase separating two ordered phases at coexistence. A major result is that the corresponding critical exponents belong to the generalized universality class of the voter model. When the two strategies are distinct but not too different, we find the expected linear-response behavior as a function of their difference. Finally, we look at the extreme situation when a superior strategy, accompanied by a short memory of earlier outcomes, is pitted against its inverse; interestingly, we find that a long memory of earlier outcomes can occasionally compensate for the choice of a globally inferior strategy.

DOI: 10.1103/PhysRevE.85.011134

PACS number(s): 05.70.Jk, 87.23.Ge, 87.19.lv, 02.50.Le

I. INTRODUCTION

The modeling of social behavior is of increasing concern to statistical physicists [1]. Studies of social and biological systems often reveal that, even when the interactions of a given individual are very localized in time and space, collective, regular behavior can emerge: this is analogous to the cooperative behavior manifested by emergent systems in the natural world. Such social regularities may well take the form of *learning*, when individuals adopt the behavior of other individuals. From the perspective of game theory [2], this can be seen as an adoption of a particular strategy, whose result may or may not be associated with a favorable outcome. It is then quite reasonable to expect that the effectiveness of a strategy in yielding favorable outcomes should influence how likely it is to persist and spread through the population; the resulting ideas of strategic learning [3] have found wide application, from economics [4] to cognitive science [5].

Against the backdrop of the above ideas, a model of strategic learning was introduced in [6], with one of two possible strategies (denoted as + and - in the remainder of this paper) being available to each agent on a lattice: the agents were referred to as "myopic" (aware only of their immediate neighbors) and "memoryless" (unaware of their own and others' past outcomes) in the paper on technology diffusion [4] that inspired the above model [6,7]. The question on which this body of work has centered is: despite these handicaps, can agents overall learn to use the superior one of two available technologies? Briefly, each agent changes (or does not change) strategy based on two elementary rules at every time step: a majority-based rule, reflecting its tendency to align with its local neighborhood, followed by a performance-based rule, where the agent adopts the strategy that "wins" in its neighborhood. This (relative) success is measured in terms of outcomes, where the probability of a successful outcome for strategy + (-) is $p_+(p_-)$. Also, the model of [6] added to the description of [4] by endowing the agents with memory: those

agents who make their choices on the basis of the last payoff alone are adjudged to be memoryless (with a corresponding parameter ε near one), while those who allow for memories of earlier outcomes may make decisions that run counter to immediate evidence (ε small).

Some related ideas have been examined in recent work. For example, the issue of consensus formation in a model of threshold learning [8] shows close analogies: in this model, the competition between the "noisy" signals from the immediate neighbors of an agent (cf. the majority rule in [6]) and the acceptance threshold that agents require to change their state (cf. the memory threshold in the performance-based rule of [6]), determine the phase diagrams obtained. Recent studies of coevolving Glauber dynamics on networks [9] are also relevant, since the model of [6] can be viewed as a competition between the Glauber dynamics of two sets of Ising spins, corresponding to strategy and outcome, respectively.

In the current paper, we take all these ideas further. First, we explore the effect of different updates. If new information propagates sequentially through the network, and the arrow of time is discernible in the decisions of individual agents, are the global phase diagrams any different from what they would be if information was transmitted and all decisions were taken simultaneously? Common sense tells us that sequential or parallel updates should make a difference to the nature of the phase diagram, and the results of the present paper confirm this. Also (unlike the work of [6,7], which examined the situation at coexistence) we look in this paper at the effects of disparate strategies $(p_+ \neq p_-)$. The final and possibly most important issue is that of memory, which acts as a threshold governing change [8]: what is the effect of the threshold ε , which tells the agent that longer-term inputs are significant and need to be considered when making a decision? We will find that, indeed, a longer memory of earlier outcomes can sometimes make up for the choice of a globally inferior strategy.

The plan of this paper is as follows. In Sec. II, we review the model of [6]. In Sec. III, we discuss the behavior of the model for a range of updating schemes, in the presence of memory. In Sec. IV, we examine the behavior of the model away from coexistence, as a function of *distinct* parameter values for the two strategies; in particular we discuss here the role of memory.

^{*}ajaz@bose.res.in

[†]anita@bose.res.in

In the concluding section, we discuss our results and put them in the context of other recent work in the field.

II. DEFINITION OF THE MODEL

The model of [6] involves two types of strategies, - and +, where the + strategy is globally superior [4] to the - strategy. As mentioned above, agents tend to follow the strategy adopted by the majority of their neighbors, modifying this choice in a second step (if necessary) according to which of these have proved to be the most successful.

Assuming that the agents sit at the nodes of a *d*-dimensional regular lattice with coordination number z = 2d, the efficiency of an agent at site *i* is represented by an Ising spin variable:

$$\eta_i(t) = \begin{cases} +1 & \text{if } i \text{ is } + \text{ at time } t, \\ -1 & \text{if } i \text{ is } - \text{ at time } t. \end{cases}$$
(1)

The evolution dynamics of the lattice is governed by two rules. The first is a *majority* rule, which consists of the alignment of an agent with the local field (created by its nearest neighbors) acting upon it, according to

$$\eta_i(t+\tau_1) = \begin{cases} +1 & \text{if } h_i(t) > 0, \\ \pm 1 & \text{w.p.} & \frac{1}{2} \text{ if } h_i(t) = 0, \\ -1 & \text{if } h_i(t) < 0. \end{cases}$$
(2)

w.p. stands for with probability. Here, the local field

$$h_i(t) = \sum_{j(i)} \eta_j(t) \tag{3}$$

is the sum of the efficiencies of the *z* neighboring agents *j* of site *i* and τ_1 is the associated time step. Next, a *performance* rule is applied. This starts with the assignment of an outcome σ_i (another Ising-like variable, with values of ± 1 corresponding to success and failure, respectively) to each site *i*, according to the following rules:

if
$$\eta_i(t) = +1$$
,
then $\sigma_i(t + \tau_2) = \begin{cases} +1 & \text{w.p.} & p_+ \\ -1 & \text{w.p.} & 1 - p_+, \end{cases}$
if $\eta_i(t) = -1$,
then $\sigma_i(t + \tau_2) = \begin{cases} +1 & \text{w.p.} & p_- \\ -1 & \text{w.p.} & 1 - p_-, \end{cases}$ (4)

where τ_2 is the associated time step and p_{\pm} are the probabilities of having a successful outcome for the corresponding strategy. With N_i^+ and N_i^- denoting the total number of neighbors of a site *i* who have adopted strategies + and -, respectively, and I_i^+ (I_i^-) denoting the number of successful outcomes within the set N_i^+ (N_i^-), the dynamical rules for site *i* are

if
$$\eta_i(t) = +1$$
 and $\frac{I_i^+(t)}{N_i^+(t)} < \frac{I_i^-(t)}{N_i^-(t)}$,
then $\eta_i(t + \tau_3) = \begin{cases} -1 & \text{w.p.} & \varepsilon_+ \\ +1 & \text{w.p.} & 1 - \varepsilon_+, \end{cases}$
if $\eta_i(t) = -1$ and $\frac{I_i^-(t)}{N_i^-(t)} < \frac{I_i^+(t)}{N_i^+(t)}$,
then $\eta_i(t + \tau_3) = \begin{cases} +1 & \text{w.p.} & \varepsilon_- \\ -1 & \text{w.p.} & 1 - \varepsilon_-. \end{cases}$ (5)

Here, the ratios $\frac{I_i(t)}{N_i(t)}$ are nothing but the average payoffs assigned by an agent to each of the two strategies in its neighborhood at time *t* (assuming that success yields a payoff of unity and failure, zero). Also, τ_3 is the associated time step and the parameters ε_{\pm} are indicators of the memory associated with each strategy. In their full generality, ε and *p* are independent variables: the choice of a particular strategy can be associated with either a short or a long memory. However, we would like in this paper to answer a question which was posed but not answered in [6]: can the presence of a good memory compensate for the choice of an inferior strategy? We therefore examine the extreme situation when a globally superior strategy ($p_+ \gg p_-$), combined with a shorter memory ($\varepsilon_+ \gg \varepsilon_-$), is in competition with its inverse: this is the situation that will be studied in Sec. IV.

Setting the time scales

$$\tau_2 \to 0, \quad \tau_1 = \tau_3 = 1, \tag{6}$$

the above steps of the performance rule are recast as effective dynamical rules involving the efficiencies $\eta_i(t)$ and the associated local fields alone:

if
$$\eta_i(t) = +1$$
,
then $\eta_i(t+1) = \begin{cases} +1 & \text{w.p.} & w_+[h_i(t)] \\ -1 & \text{w.p.} & 1 - w_+[h_i(t)], \end{cases}$
if $\eta_i(t) = -1$,
then $\eta_i(t+1) = \begin{cases} +1 & \text{w.p.} & w_-[h_i(t)] \\ -1 & \text{w.p.} & 1 - w_-[h_i(t)]. \end{cases}$ (7)

The effective transition probabilities $w_{\pm}(h)$ are evaluated by enumerating the 2^z possible realizations of the outcomes σ_j of the neighboring sites *i* and weighing them appropriately. For a 2-*d* square lattice, the possible local-field values at the interfacial sites are 0 and ± 2 . The corresponding transition probabilities for these field values are [6]

$$w_{+}(+2) = 1 - \varepsilon_{+}p_{-}(1 - p_{+}^{3}),$$

$$w_{-}(+2) = \varepsilon_{-}(1 - p_{-})[1 - (1 - p_{+})^{3}],$$

$$w_{+}(0) = 1 - \varepsilon_{+}p_{-}(1 - p_{+})(2 - p_{-} - 2p_{+} + 3p_{-}p_{+}),$$

$$w_{-}(0) = \varepsilon_{-}p_{+}(1 - p_{-})(2 - p_{+} - 2p_{-} + 3p_{-}p_{+}),$$

$$w_{+}(-2) = 1 - \varepsilon_{+}(1 - p_{+})[1 - (1 - p_{-})^{3}],$$

$$w_{-}(-2) = \varepsilon_{-}p_{+}(1 - p_{-}^{3}).$$
(8)

In [6], the model was explored at coexistence with an ordered sequential update applied to memoryless agents [4]:

$$p_+ = p_-, \quad \varepsilon_+ = \varepsilon_- = 1. \tag{9}$$

In the present paper, we go beyond this in two different ways. First, still at coexistence, we explore the effect of different updates on the $p - \varepsilon$ phase diagram of the model: next, we examine the model away from coexistence, for distinct values of p_{\pm} and ε_{\pm} . The basic quantities considered hereafter are the magnetization M, staggered magnetization M_{stag} , and the energy E. These quantities are defined for a finite sample of N agents (or sites) and Nz/2 bonds (or links) as

$$M = \frac{1}{N} \sum_{i} \eta_{i},$$

$$E = \frac{1}{Nz} \sum_{ij} (1 - \eta_{i} \eta_{j}),$$
(10)

$$M_{\text{stag}} = \frac{2}{N} \sum_{i} \eta_i$$
 if *i* is odd or even.

In the following we shall usually consider mean values $\langle M \rangle$, $\langle E \rangle$, and $\langle M_{\text{stag}} \rangle$.

III. EFFECT OF FINITE MEMORY AND OF DIFFERENT UPDATES

We begin this section with a review of the physical significance of updating schemes. Most generally, updates can be random or ordered as follows:

(1) Random: Here, sites are chosen at random for the consecutive application of rules.

(2) Ordered: Here, sites are chosen in an ordered fashion, i.e., after choosing every (i, j)th site, the (i, j + 1)th site is selected.

Since the sociological basis for this work was the propagation of innovation through connected societies [4], we choose to deal only with ordered updates here. However, even ordered updates have two subclasses: parallel and sequential. Assume a condition *A* such that, when an agent satisfies *A*, it changes strategy:

(1) Sequential update: In this type of update, we check the condition A on the (i, j)th site, then update the efficiency of the site and proceed to the (i, j + 1)th site using the updated value of the (i, j)th site.

(2) Parallel update: In this type of update, we check the condition A on the (i, j)th site, do *not* update the site but instead save the update decision in memory, and proceed to the next site. Once the whole lattice is swept, all the saved update decisions are implemented "simultaneously."

The choice of different updates generally corresponds to different physical situations: it has been shown that it also leads to a disparity in the convergence time of the systems concerned [10,11]. We therefore examine all possible combinations for our two update rules: (1) parallel updates for both majority rule and performance rules (PP), (2) parallel update for majority rule and sequential update for performance rule (PS), (3) sequential updates for both majority rule and performance rules (SS), and (4) sequential update for majority rule and parallel update for performance rule (SP).

In the following subsections, we explore the phase dynamics at coexistence for each of these update rules in turn, for both parameters p and ε . We state at the outset that all the updates (except for the SP update) which we consider result in models which are in the general university class of the voter model [12]: the inverse energy 1/E(t) is thus always proportional to the logarithm of time, ln t. When, as in the case of the SS update, the value of the slope is exactly $2/\pi$ [6], the *exact* universality class of the voter model is retrieved.



FIG. 1. (Color online) (SS update) Phase diagram of the model with an SS update. Plot of the inverse energy 1/E(t) at time t = 512 for a square lattice of size $N = 64^2$ in the *p*- ε plane. The black region shows the disordered phase and the yellowish (light gray) region shows the frozen phase.

A. SS update

This is the update that was used throughout [6]; however the phase behavior of the model was there only explored for the parameter p, whereas here we extend it to the parameter ε . In Fig. 1, we plot the inverse energy 1/E(t) in the *p*- ε plane at time t = 512 for a square lattice of size $N = 64^2$. This phase diagram shows clearly the existence of a disordered paramagnetic phase embedded in a largely frozen phase elsewhere. The disordered phase exists for $p_{c1} (= 0.56 \pm 0.01)$ 0.70 ± 0.01) when $\varepsilon \ge 0.980$. Our results agree with those of [6] for $\varepsilon = 1$, and we extend them all across the rest of the p- ε plane. We mention here that the average time required to reach consensus increases exponentially as p decreases in the frozen phase, leading to the presence of striped states [13] at limiting values of p. Figure 1 also makes it clear that the effect of increasing memory wipes out the disordered phase: this is as it should be, since the disordered phase is generated by the competition between the majority and performance-based rules, which is dulled by increasing memory.

Figure 2 shows snapshots of the dynamics of the model using a lattice of size $N = 512^2$ at times t = 8, 64, and 512 with random initial configurations and parameter values p = 0.72 (very close to the critical point p_{c2}) and $\varepsilon = 1.0$. The plots reveal characteristically voterlike [12] coarsening behavior.

In Fig. 3, we have plotted the inverse energy 1/E(t) against the natural logarithm of time ln t for values of p around the critical point $p_{c2} = 0.70 \pm 0.01$. Each of the curves is obtained by averaging over 200 independent samples of size 256². At the critical point, we obtain a straight line with a slope close to $2/\pi$ [6], a behavior characteristic of the exact voter model [12] that corresponds to

$$E(t) \approx \frac{\pi/2}{\ln t}.$$
 (11)

Similar behavior is obtained at the other critical point p_{c1} , in agreement with [6].

B. PP update

In this case, both environmental majority and performancebased rules are applied using parallel updates. As we will see,

AJAZ AHMAD BHAT AND ANITA MEHTA



FIG. 2. (Color online) (SS update) Snapshots of the dynamics of the SS-updated model. Each plot is a portion (of size 100^2) of a square lattice of $N = 256^2$ for p = 0.72 and $\varepsilon = 1.0$ at times t = 8 (top left), t = 64 (top right), and t = 512 (bottom).

although the universality class of the model is qualitatively unchanged, this update results in the appearance of novel ordered phases compared to the SS update. As before, we first plot the phase diagram for all values of p and ε , then show snapshots of the dynamics, and finally get a more quantitative feel for the behavior of key quantities as a function of p.

Accordingly, Fig. 4 (top and bottom) shows plots of the absolute values of magnetization |M| and staggered magnetization $|M_{\text{stag}}|$, at time t = 512 for a lattice size $N = 100^2$, in the p- ε plane using PP updates. In these phase diagrams, we see clear evidence of the existence of two distinct frozen phases separated by a disordered phase. Looking along the line $\varepsilon = 1$, disorder prevails for $p_{c1} with <math>p_{c1} = 0.43 \pm 0.01$ and $p_{c2} = 0.57 \pm 0.01$. Notice the symmetry of the two critical points about p = 0.5: we shall have more to say about this later on.



FIG. 3. (Color online) (SS update) Plot of the inverse energy 1/E(t) vs the natural logarithm of time ln *t* for different values of *p* close to $p_{c2} = 0.70$, with ε set to 1.0. The lattice size $N = 256^2$ and the *p* values are indicated on the curves. The curve corresponding to $p = p_{c2} = 0.70$ (shown in red) has slope $2/\pi$ approximately [see Eq. (11)].



FIG. 4. (Color online) (PP update) Plot of the absolute value of the magnetization |M| (top) and the absolute value of the staggered magnetization $|M_{stag}|$ (bottom) at time t = 512 for a lattice of $N = 100^2$ in the p- ε plane. In the top figure, the yellowish (light gray) region refers to the parallel frozen phase (PFP), while the black region includes both the antiparallel frozen phase (AFP) and the disordered region. In the bottom figure, the black region represents the disordered phase characterized by very low $|M_{stag}|$.

For *p* below p_{c1} , there is a frozen phase characterized by overall alignment of spins: we call this the parallel frozen phase (PFP). For *p* above p_{c2} , the frozen phase that appears is characterized by an antiparallel ordering of spins: we call this the antiparallel frozen phase (AFP). We mention also that in the AFP the lattice may have more than one antiparallel domain, with thin frustrated chains running in between them. This frustration can be attributed to the inability of the different domains to align with each other under periodic boundary conditions. The disturbances caused by these chains (in quantities such as |M| or E) due to misalignment decrease as $1/\sqrt{N}$ and also appear to vanish for large times. Again, we notice that the phase transition disappears for low ε ; in fact, at very low values of ε the evolving lattice may get trapped into striped states [13] at long times.

This can be understood as follows: the effect of a long memory (ε small) strongly reduces the relative impact of the performance-based rule. Depending on the value of ε , the performance rule may not be effective for several time steps, whereas the majority rule is implemented at every time step. In the limit of vanishing ε , then only the (zero-temperature) majority rule will be effective, leading to stripe formation as predicted by [13] for this situation.

Figure 5 comprises snapshots of the dynamics of the model for a 2*d* square lattice of size $N = 512^2$ and at time t = 256, with random initial configurations. The plots show a portion DYNAMICS OF COMPETITIVE LEARNING: THE ROLE OF ...



FIG. 5. (Color online) (PP update) Snapshots of the dynamics of the PP-updated model on a square lattice, for p = 0.41 (left), p = 0.50 (center), and p = 0.59 (right) at time t = 512, with $\varepsilon = 1.0$. The yellow (light gray) and black colors represent the two strategies, while the grayish grid corresponds to antiparallel arrangements of yellow (light gray) and black. The left picture represents the PFP (see text), the center picture represents the disordered phase, and the right picture represents the AFP (see text).

of size 100^2 of the square lattice for three values of p: p = 0.41 (near the critical point p_{c1} between the PFP and the paramagnetic phase), p = 0.50 (within the paramagnetic phase), and p = 0.59 (near the critical point p_{c2} separating the paramagnetic phase from the AFP), with $\varepsilon = 1$. The snapshot at p = 0.41 shows the lattice evolving toward consensus (parallel alignment) with the formation of domains of one type only. The snapshot at p = 0.50 shows the lattice in its disordered phase, while the one at p = 0.59 shows that the nature of the lattice ordering is antiparallel.

To investigate this more quantitatively, we plot the absolute value of magnetization |M|, the absolute value of staggered magnetization $|M_{\text{stag}}|$, and energy E(t) against p, with $\varepsilon = 1.0$, in Fig. 6. These measurements were recorded using a square lattice of size $N = 80^2$ at time $t = 10^6$. All the curves are averaged over 100 independent samples for each value of p. In the region $p \leq p_{c1}$, the values of magnetization |M| and staggered magnetization $|M_{\text{stag}}|$ are both equal to unity at saturation, implying a parallel alignment of the sites, whereas, for p above p_{c2} , the magnetization |M| is zero and the staggered magnetization $|M_{\text{stag}}|$ equals unity at saturation, indicating an antiparallel alignment of the sites. The energy graph is consistent with this interpretation, given the definition of the energy in Eq. (10): zero in the PFP, middling in the paramagnetic phase, and unity in the AFP.



FIG. 6. (Color online) (PP update) Plot of the absolute value of magnetization |M|, the absolute value of staggered magnetization $|M_{\text{stag}}|$, and energy *E* against *p*, with $\varepsilon = 1$ and lattice size $N = 80^2$. Each curve is drawn using symbols (and color) as indicated in the legend.



FIG. 7. (Color online) (PP update) Plot of the inverse energy 1/E(t) vs ln t for different values of p close to $p_{c1} = 0.43$ and $\varepsilon = 1$, for a lattice of size $N = 100^2$. The p values are indicated on the curves. The straight line corresponding to $p_{c1} = 0.43$ (shown in red) has slope $1/2\pi$ approximately.

In order to confirm the voter-like nature of the critical points, we plot the inverse energy 1/E(t) against the natural logarithm of time ln *t*, choosing *p* values near both critical points (see Figs. 7 and 8). Each curve is an average over 200 independent samples. Exactly at the critical points $p_{c1} = 0.43$ and $p_{c2} = 0.57$, a linear behavior of inverse energy with respect to ln *t* is found, with slopes of $1/2\pi$ and $-1/5\pi$, respectively. While the critical exponents are those of the voter model [12], the values of the slope are different from $2/\pi$: we find therefore that the PP update of the model belongs to the universality class of the generalized, rather than the exact, voter model [12].

To conclude this subsection: the main effect of the PP update is to change the nature of the ordering in one of the two frozen phases, so that antiparallel ordering is found in the high-*p* frozen phase. As before, the effect of increasing memory (going to low ε) is to smear out the phase transitions to the disordered phase, by undermining the effect of the outcome-based rule whose competition with the majority rule causes the appearance of disorder. Such instances of mixed domains have been found in recent work on coevolving



FIG. 8. (Color online) (PP update) Plot of the inverse energy 1/E(t) vs ln t for different values of p close to $p_{c2} = 0.57$, with $\varepsilon = 1$, for a lattice of size $N = 100^2$. The p values are indicated on the curves. The straight line corresponding to $p_{c2} = 0.57$ (shown in red) has slope $-1/5\pi$ approximately.



FIG. 9. (Color online) (PS update) Phase diagram in the p- ε plane of the PS-updated model, with a plot of the absolute value of staggered magnetization $|M_{\text{stag}}|$ at time t = 512, for a lattice size of $N = 100^2$. The black region represents the disordered phase (very low $|M_{\text{stag}}|$), while the yellowish (light gray) region represents frozen phases with high $|M_{\text{stag}}|$.

(parallel) dynamics [9]; some features of these results also appear in studies of threshold dynamics of societal systems [8]. For a real-life example of the AFP in the case of technology diffusion, we cite the results of [14], where the authors conclude that "in technology clusters where direct competitors are right next door, leading firms generate innovations that are technologically very distant from their neighbors" [14].

C. PS update

The behavior of the PS-updated model is qualitatively similar to that of the PP-updated model above. Again, there are two frozen phases PFP and AFP, separated by a disordered phase: the values of the critical points p_{c1} and p_{c2} are however shifted, such that the disordered region extends between $p_{c1} = 0.31 \pm 0.01$ and $p_{c2} = 0.69 \pm 0.01$ at $\varepsilon = 1.0$. We find once again that the two critical points are symmetrically placed with respect to p = 0.5, as in the PP update: we will give an argument for why this is so, in the following subsection.

To avoid repetition, we present only the phase diagram for the staggered magnetization as a function of p and ε : Fig. 9 shows the absolute value of the staggered magnetization $|M_{\text{stag}}|$ of the system at time t = 512 for a square lattice of size $N = 100^2$. The paramagnetic region, with low values of $|M_{\rm stag}|$, is colored black in the figure, whereas the frozen regions (containing either parallel or antiparallel ordering) with high values of $|M_{\text{stag}}|$ are colored yellow (light gray). These phases are investigated more quantitatively in Fig. 10, where we plot the absolute value of magnetization |M|, energy E, and the absolute value of staggered magnetization $|M_{\text{stag}}|$ against p, with ε equal to 1.0; each curve is an average over 100 independent runs. The region where both the magnetization |M| and staggered magnetization $|M_{stag}|$ curves saturate to one corresponds to parallel alignment, whereas $|M| \approx 0$ with $|M_{\rm stag}| \approx 1$ implies an antiparallel alignment of the spin types. The energy graph is consistent with this interpretation, given the definition of the energy in Eq. (10): zero in the PFP, middling in the paramagnetic phase, and unity in the AFP.

Finally, we present the variation of inverse energy with the natural logarithm of time, $\ln t$, near the critical points p_{c1}



FIG. 10. (Color online) (PS update) Plot of the absolute value of magnetization |M|, energy E, and the absolute value of staggered magnetization $|M_{\text{stag}}|$ against p with ε set to 1.0 for a lattice size $N = 100^2$. Each curve is drawn using symbols (and color) as indicated in the legend.

and p_{c2} in Figs. 11 and 12, respectively. Each of the curves is an average over 200 independent runs. At criticality, both plots show a linear proportionality between 1/E(t) and $\ln t$, with slopes of $4/3\pi$ and $-4/15\pi$ at $p_{c1} = 0.31$ and $p_{c2} =$ 0.69, respectively. Again, this indicates that the PS update of the model belongs to the generalized, rather than the exact, universality class of the voter model [12].

To conclude, the PS update yields qualitatively similar results to the PP update, with the appearance of two frozen phases PFP and AFP. Again, small values of ε indicating longer memories of outcomes lead to a smearing out of the phase transition, because of the decreasing effectiveness of the outcome-based rule.

D. Explanation of the nature of phase diagrams for different updates

In this subsection, we give arguments for the three most important features of the phase diagrams presented above: (1) the appearance of antiparallel ordering in both PP and PS updates, (2) the symmetry of the PFP and the AFP phases in both PP and PS updates, and (3) the positioning of the disordered phase in SS, PP, and PS updates.



FIG. 11. (Color online) (PS update) Plot of the inverse energy 1/E(t) vs ln t for different values of p close to $p_{c1} = 0.31$, with ε set to 1.0 for a lattice size $N = 256^2$. The p values are indicated on the curves. The straight line corresponding to $p_{c1} = 0.31$ (shown in red) has a slope of approximately $4/3\pi$.



FIG. 12. (Color online) (PS update) Plot of the inverse energy 1/E(t) vs ln t for different values of p close to $p_{c2} = 0.69$, with ε set to 1.0, for lattice size $N = 256^2$. The p values are indicated on the curves. The straight line corresponding to $p_{c2} = 0.69$ (shown in red) has slope $-4/15\pi$ approximately.

The clue which explains all of the above is the formation of "active" or disparate bonds by the rules of the model under different updates: these are clearly the units of antiparallel ordering. Consider, thus, configurations where a site is surrounded by a majority of its own kind: this would correspond to a local field of +2 for a + and -2 for a -. Here the majority of the bonds are "like" or "inactive." The transition probability for the increase of active bonds from such configurations is $1 - w_+(+2)$ [or $1 - w_-(-2)$] [see Eq. (8)]. The transition probabilities for the *decrease* of active bonds are given by an opposite scenario, yielding $w_-(+2)$ [or $w_+(-2)$] [see Eq. (8)]. We plot two of these transition probabilities in Fig. 13, corresponding, respectively, to an increase and a decrease of active bonds: the former peaks at p = 0.63 while the latter peaks at p = 0.37.

The net probability of having active bonds is the difference between these two transition probabilities and is plotted in Fig. 14. We see from this that the probability of having active bonds is greatest at p = 0.79 and least at p = 0.21. The last ingredient that we need to explain the AFP phase in the PP and PS updates is the fact that, once clusters with many active bonds, i.e., antiparallel ordering, are formed, the majority rule applied via the *parallel* update *preserves* such ordering. With all this in place we see that, as expected, the AFP phase in



FIG. 13. (Color online) Transition probabilities $[1 - w_+(+2)]$ drawn as a solid line [in green (gray)] and $w_-(+2)$ drawn as a dashed line (in black) against *p* [from Eq. (8)].



FIG. 14. (Color online) The difference in the transition probabilities $[1 - w_+(+2)]$, with $w_-(+2)$ plotted against p [see Eq. (8)].

both PP and PS updates shows up in qualitatively the same regions as predicted by Fig. 14, with a peak, in both cases, at around p = 0.79. Correspondingly, the PFP in both PP and PS updates shows up in the region predicted in this figure, with a peak in both cases at around p = 0.21. Notice (Fig. 14) that the peak and the dip in the probability of active bonds are symmetric about p = 0.5, thus explaining the symmetry that we have observed in Figs. 4 and 9; p = 0.5 is thus the natural point for the appearance of the disordered phase in both PP and PS updates, as will be confirmed by an inspection of Figs. 4, 9, and 14.

The only remaining point to be explained is the appearance of the disordered phase in the SS update. In this case too, the analysis leading to Fig. 14 for the probabilities of having active bonds remains valid. However, the sequential update of the majority rule always favors strictly parallel ordering, so that typically clusters of active bonds are destroyed once formed. When the probability of their formation is strongest, i.e., at p = 0.63 (see Fig. 13), the competition between the majority and outcome-based rules is at its most intense, and a disordered phase may be expected to appear. Indeed, the midpoint of the disordered phase for the SS update is shown in Fig. 1 to be in exact agreement with this predicted peak, given as it is by $(p_{c1} + p_{c2})/2 = 0.63$.

E. SP update

In the case of this update, the phase diagram, Fig. 15, shows nothing but a frozen phase. As is evident from the plot of inverse energy 1/E(t) versus ln t (Fig. 16), there is a continuous increase in 1/E(t) for all values of p at $\varepsilon = 1.0$ (where the phase transition is expected to be the most visible). This suggests that the two rules, majority and performance based, do not compete with each other at all (this is what had led to the appearance of the disordered phase in all the other updates). We suggest that this might be because the sequential update (with its more immediate conversions) in the case of the majority rule completely dominates the slower parallel update for the outcome-based rule: this in turn leads to an increasing tendency for consensus, independent of the value of p, with which our results are consistent.



FIG. 15. (Color online) (SP update) Phase diagram of the SPupdated model. Plot of the absolute value of magnetization |M| at time t = 512 for lattice size $N = 64^2$ in the p- ε plane. No phase transition is visible.

IV. AWAY FROM COEXISTENCE: WHEN STRATEGIES ARE DISTINCT

Evidently, the real use of a competitive learning model such as this one is when the agents have a choice of distinct strategies. The full exploration of the behavior of the model at coexistence as carried out in this paper as well as in earlier work [6,7] was aimed at an understanding of its phase diagram. However, in the exploration of the behavior of the model away from coexistence, we hope to gain an understanding of the relative importance of parameters such as superiority of strategy (modeled by p) and memory (modeled by ε), when these are in competition. The behavior in asymmetric conditions (using $p_+ > p_-$ and $\varepsilon_+ > \varepsilon_-$) is formulated in terms of the application of two biasing "fields" [6]:

$$H = p_+ - p_-, \quad B = \varepsilon_+ - \varepsilon_-, \tag{12}$$

such that one strategy is favored over the other.

In the following subsection, we look at a linear-response formulation of our question in terms of unequal p's, viewed as a biasing field, keeping ε the same for both strategies. In the final subsection, we look at unequal strategies *as well*



FIG. 16. (Color online) (SP update) Plot of the inverse energy 1/E(t) vs ln (t) for different values of p and $\varepsilon = 1$, for a lattice size $N = 256^2$. The value of p for each curve is given by a different color as indicated. No phase transition is visible.



FIG. 17. (Color online) (SS update) Plot of magnetization M against biasing field H for different values of p. Each curve is drawn using different symbols (and color) as shown in the legend, at time t = 2000 with $N = 100^2$ and $\varepsilon = 1.0$.

as unequal memories, to find out whether inferior strategies applied with a good memory of past outcomes can win overall.

A. Linear-response theory: strategies with unequal *p*

Linear-response theory is premised on the basis that an order parameter such as the magnetization undergoes a sharp change in the neighborhood of a critical point. In both the SS and PP updates of this model, there are two critical points p_{c1} and p_{c2} separating a paramagnetic phase from two frozen phases. In this subsection, we look at the linear-response behavior of the model in the vicinity of both critical points, starting from the disordered phase: clearly the response will depend both on the value of p as well as on the value of the biasing field H [defined in terms of the difference of the p's in Eq. (12)]. In the following, we examine the response by choosing a given value of p and writing $p_{\pm} = p \pm H/2$, keeping ε fixed.

We first consider the SS-updated model. Figure 17 shows a plot for magnetization M against the biasing field H at various values of p, that are within the paramagnetic phase at $\varepsilon = 1.0$. Each curve is obtained after averaging over 100 initial configurations using a square lattice of size $N = 100^2$. For each p in the paramagnetic phase, we see a linear behavior of M against H around $H \sim 0$, with all subsequent increases in the field strength leading to saturation, as expected. For a given p value we observe a functional dependence of the form

$$M = \tanh(bH),$$

where

taking

$$p_{\text{central}} = (p_{c1} + p_{c2})/2.$$

 $b \propto (p_{\text{central}} + p)^2$,

The quality of the fit to tanh(bH) is seen Fig. 18: the black fitting curve almost completely coincides with a sample curve taken from Fig. 17.

These results also admit of an alternative representation, shown in Fig. 19, where it is clear that the relative values of the bias correspond to different regions of domination of each strategy in phase space.



FIG. 18. (Color online) (SS update) Plot of magnetization M against field H for p = 0.58 and $\varepsilon = 1.0$ at time t = 2000 with $N = 100^2$. The fit of a tanh(*bH*) curve almost completely overlaps with our numerical results.

We next examine the linear-response behavior of the PPupdated model, again in the vicinity of the two critical points. Figure 20 is a plot showing the variation in magnetization M along the field H for different p values at $\varepsilon = 1.0$. For the lower values of p, in the vicinity of p_{c1} , we see very similar behavior to that presented in Fig. 17, corresponding to an expected tanh(bH) behavior as shown in Fig. 18: the PFP phase lying to the left of p_{c1} is, after all, identical to the frozen phases in the SS update. As we approach the vicinity of p_{c2} , the curves are markedly different: the nature of the ordered phase is one that corresponds to magnetization values of 0 (see the orange curve drawn using plus symbols in Fig. 20), which is again consistent with the AFP phase that lies to the right of p_{c2} .

To establish this more firmly we look at plots of the absolute value of the staggered magnetization $|M_{\text{stag}}|$ as a function of bias H, in Fig. 21. The green (triangle), blue (square), and black (plus) lines denote increasing values of $p < p_{c2}$, where the staggered magnetization increasingly approaches zero, as expected in the disordered phase: however the red (star) line, corresponding to $p > p_{c2}$ shows an abrupt jump in the value of $|M_{\text{stag}}|$ to unity. Combined with the analysis of the previous paragraph, this shows convincingly that the phase we refer to as AFP indeed corresponds to antiparallel ordering.



FIG. 19. (Color online) (SS update) Plot of magnetization M in the p_+ - p_- plane at time t = 7000, for a lattice of size $N = 100^2$ with $\varepsilon = 1.0$. The + strategy dominates in the yellow (light gray) region, while the – strategy dominates in the brown (black) region.



FIG. 20. (Color online) (PP update) Plot of magnetization M against field H for different values of p, each curve indicated by a different symbol (and color) as shown in the legend, at time t = 2000 with $N = 100^2$ and $\varepsilon = 1.0$.

We present below an alternative representation of the above results for ease of visualization. In Fig. 22, the magnetization M is plotted in the p_+-p_- plane: as before, the regions of brown (black) [respectively, yellow (light gray)] correspond to domination by – strategies (respectively, + strategies). Notice, however, that the coexistence line has an island of very low magnetization: in actual fact, this corresponds to the regions of *both* the paramagnetic and AFP phase. This is clearer in the plot of the absolute value of the staggered magnetization $|M_{stag}|$, shown in Fig. 23, where the black portion of the island along the coexistence line corresponds to the disordered phase, while the faintly brown (gray) portion corresponds to the AFP.

These plots allow us to go beyond the previous analysis in defining the domain of stability of the AFP phase: we see clearly from Figs. 22 and 23 that the AFP phase exists for $p > p_{c2}$ only if the biasing field is within the bounds defined by $H = |p_+ - p_-| \le 0.19 \pm 0.02$. In qualitative terms, this implies that, at least in the absence of memory, when the two strategies have nearly equal success rates, neighboring agents may adopt different strategies [14] in equilibrium.

Having thoroughly investigated the linear-response regime for the SS- and PP-updated models, we will now examine the effect of the memory parameter ε in the next subsection.



FIG. 21. (Color online) (PP update) Plot of the absolute value of staggered magnetization $|M_{\text{stag}}|$ against field *H* for different values of *p*, each curve indicated by a different symbol (and color) as shown in the legend, at time t = 2000 with $N = 64^2$ and $\varepsilon = 1.0$.



FIG. 22. (Color online) (PP update) Plot of magnetization M in the p_+ - p_- plane for a lattice of size $N = 100^2$ at time t = 7000, with $\varepsilon = 1.0$. As before, the regions of + and - strategy domination are colored yellow (light gray) and brown (black); the orange (gray) region corresponds to *both* the paramagnetic and the AFP region (see text).

B. Role of memory parameters: the case of unequal ε

The principal competition in this model is between two strategies with different global success rates p, which determines the relative dominance of each one in phase space. The memory parameter ε plays a more subtle role in this competition: although it cannot be a determinant of phase behavior in the way that the success rates are [as a consequence of the rules elucidated in Eq. (8)], it *can*, as we will show, cause a surprising change in the dominance of an ostensibly superior strategies and good memories might indeed win against agents who had better strategies but worse memories. Here, we will make this prediction more quantitative.

The phase diagram of the model away from coexistence involves four parameters, $p_{\pm}, \varepsilon_{\pm}$, so that its representation is a nontrivial problem. In the following, we choose to fix p_+ to 0.5 and to vary the other three parameters: a sample threedimensional plot is shown in Fig. 24. We analyze the three visible faces in detail, before remarking on the phase behavior within the cube: the color coding is such that green (gray) represents dominance of + strategies, blue (black) represents



FIG. 23. (Color online) (PP update) Plot of staggered magnetization $|M_{\text{stag}}|$ in the p_+ - p_- plane, for a lattice of size $N = 100^2$ at time t = 7000, with $\varepsilon = 1.0$. Here, yellow (light gray) represents the region of parallel ordering, black represents the disordered phase, and the light brown (gray) represents AFP order.



FIG. 24. (Color online) (PP update) A three-dimensional plot of magnetization M with parameters ε_+ (along x), ε_- (along y) and p_- (along z), setting $p_+ = 0.50$ for a lattice of size $N = 64^2$ at time T = 2000. Green (gray) denotes the dominance of the +'s, while blue (black) denotes that of the -'s. The other colors represent cases of intermediate ordering.

dominance of – strategies, and other colors represent mixed states:

(1) The leftmost face of the cube corresponds to the plane $\varepsilon_{+} = 0$; this implies that the agents using + strategies will never convert, no matter what the outcome-based rule says. The minimum occupancy of + strategies for random configurations should thus be of the order of N/2, which can only increase depending on the conversions of agents using - strategies into the camp of the +'s. The bottom line corresponds to $p_{-} = 0$, which is when such conversions are maximal (so that all N sites are +): the green (light gray) color is at its most pronounced here, changing gradually over to other colors only as $\varepsilon_{-} \rightarrow 0$ to the right of the line, when agents using - strategies too begin to refuse to convert, irrespective of the outcome rules. As the values of p_{-} increase beyond 0.5 (the fixed value for p_+), we note that the dominance of the + strategy gradually gives way to states with a mixture of strategies; when $\varepsilon_{-} \rightarrow 0$, this tendency is at its most pronounced, while, when $\varepsilon_{-} \rightarrow 1$, this is at its least pronounced, since local conversions can sometimes go against global success rates.

(2) The front face of the cube corresponds to $\varepsilon_{-} = 0$; this implies that agents using – strategies will never convert, no matter what the outcome-based rule says. This is a reflection of the previous case, where the minimum number of – sites is once again N/2, which can only be increased as the conversions from the +'s add to it.

(3) The top face of the cube corresponds to $p_{-} = 1$, where globally a predominance of the – strategy is expected. This is found over almost all the range of ε_{-} except at low values of ε_{+} , where agents using the + strategy refuse to convert, despite their globally poorer performance.

The interior of the cube can show markedly different behavior, which we illustrate via a sample slice shown in Fig. 25. In this figure, red (lower black) and blue (upper black) regions correspond to the dominance of + and - strategies, respectively. Here we set the value of p_+ to 0.7 and look at a slice of its phase space cube, as before: choosing ε_+ to be



FIG. 25. (Color online) (PP update) A three-dimensional plot of magnetization M for the parameters ε_+ (lying between 0.7 and 0.8), ε_- and p_- , with $p_+ = 0.70$, for a lattice of size $N = 64^2$ at time T = 2000. The red (lower black) region represents the dominance of the + strategy and the blue (upper black) region represents the dominance of the - strategy. The green (light gray band) region represents AFP.

between 0.7 and 0.8, we look at the dominating strategy as a function of the variables ε_{-} and p_{-} . If the memory parameters had not existed, we would have expected the – strategy [blue (upper black) region in the figure] to predominate only for $p_{-} > 0.7$; the red (lower black) region would have been covering the entire slice below this, corresponding to the dominance of the + strategy. However, the reality is rather different. The + strategy does indeed predominate for $p_{-} < 0.7$, provided that agents using the – strategy have imperfect memory, but there is a striking predominance of the – strategy (even for very low values of p_{-}) provided that the memory of the agents employing this strategy is much better than those of the other kind ($\varepsilon_{-} \ll \varepsilon_{+}$).

A last feature to mention is the green (gray) region in Fig. 25: here, there is a region of alternating + and – ordering (AFP), corresponding to ongoing competition between the two strategies. This phenomenon is most pronounced when the two strategies are equally successful and are *both* accompanied by weak memories of earlier outcomes. In Fig. 26, the structure of the full AFP is shown (by selecting phase points with low values of absolute magnetization |M| and high absolute staggered magnetization $|M_{stag}|$) as a function of p_{-} and ε_{\pm} , fixing $p_{+} = 0.5$.

V. DISCUSSION

The work of this paper extends work done on a problem of strategic learning [6,7] which, although originally suggested by a problem on technology diffusion [4], has much wider ramifications (e.g., in relation to threshold learning dynamics [8]).

In any agent-based modeling scheme, it is important to know whether agents react sequentially or collectively to the spread of information. Our results show that these issues make



FIG. 26. (Color online) (PP update) A three-dimensional plot of absolute staggered magnetization $|M_{\text{stag}}|$ with parameters ε_+ (along *x*), ε_- (along *y*), and p_- (along *z*) for $p_+ = 0.50$ and lattice size $N = 64^2$, at time T = 2000. The graph delineates the region of AFP for these parameter values.

a quantitative as well as a qualitative difference to the results, changing not just exponents but also the entire nature of the phase diagram in most cases. Given that, typically, the propagation of technologies through well-connected societies is of interest [4], we choose ordered rather than random updates and examine the response of the model of [6] to all possible combinations of sequential and parallel updating. From the viewpoint of theoretical physics, a major result is that this model is robustly in the universality class of the voter model [12], for all but one of the updates. This strong relationship with the voter model results from the model of [6,7] being driven by interfacial noise alone, i.e., the absence of surface tension [12].

Another major result, still to do with updates, is the appearance of a phase of antiparallel ordering (AFP) in the high-performing limits of p, for both the PP and the PS updates. While the technicalities behind this are explained in the text, we give here a more intuitive reason for this, from the perspective of strategic learning. The parallel scheme can be viewed as a more "equilibrated" update than the sequential one, since it gives a chance for the entire lattice to be updated "simultaneously." It is then natural that in the regime that both agents are high-performing they should be equally preferred: this lies behind the "alternating" order inherent in the AFP regime. By contrast, since the sequential paradigm corresponds to a "nonequilibrium" update, where every agent responds to the updated value of its neighbors, the above logic leads to a disordered phase where every prescription of the outcomebased rule is countermanded by the following majority rule. Using once again the illustration of propagating technologies [4], when all the populace have equal and simultaneous access to information about two high-performing technologies, we will see the coexistence of both [14] (as predicted by the AFP phase), whereas when information about each one is passed on *sequentially*, the conflicting information so obtained can result in sheer disorder. Finally, we mention here that our investigation of different updates on this game-theoretic model has been applied to related game-theoretic models of cognitive learning and synaptic plasticity [15,16], where updates relate to the directionality of synapses in a network.

Moving away from the domain of critical behavior at coexistence, we have looked at the behavior of the competitive learning model when the two strategies have distinct attributes (this, after all, is truer to the title of competitive learning). To begin with, we have examined the response of the model to unequal strategies p_{\pm} , and have found in general that the smarter strategy wins (for equal values of the memory parameter ε), as might be expected. An interesting feature is that the region of antiparallel ordering (AFP) found earlier still persists in the presence of bias, provided that the difference in p is below a well-defined bound: in other words, when

two distinct strategies are *almost* equally successful, one will typically find that they can coexist in society. Finally, we have looked at the effect of memory: we have found that while memory has a secondary role in determining the phase behavior of the model, it has a particularly striking effect in turning around the results of any bias in p. A major result of our paper is thus that decisions based on a good memory of earlier outcomes can, within limits, compensate for the choice of inferior strategies.

ACKNOWLEDGMENTS

A.A.B. would like to thank Dr. G. Mahajan and Mr. B. Chakraborty for helpful discussions. A. M. acknowledges a grant from the Department of Science and Technology (Govt. of India) through the project "Generativity in Cognitive Networks," through which A.A.B. was supported.

- C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. 81, 591 (2009).
- [2] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games* (MIT Press, Cambridge, 1998).
- [3] H. P. Young, *Strategic Learning and Its Limits* (Oxford University Press, Oxford, 2004).
- [4] K. Chatterjee and S. H. Xu, Adv. Appl. Probab. 36, 355 (2004).
- [5] C. F. Camerer, Trends Cogn. Sci. 7, 225 (2003).
- [6] A. Mehta and J. M. Luck, Phys. Rev. E 60, 5218 (1999).
- [7] G. Mahajan and A. Mehta, Theory Biosci. 129, 271 (2010).
- [8] J. C. Gonzalez-Avella, V. M. Eguíluz, M. Marsili, F. Vega-Redondo, and M. San Miguel, PLoS ONE 6, e20207 (2011).
- [9] S. Mandra, S. Fortunato, and C. Castellano, Phys. Rev. E 80, 056105 (2009).

- [10] I. Kanter, Physica D 42, 273 (1990).
- [11] H. Kfir and I. Kanter, Physica A 330, 259 (2003).
- [12] I. Dornic, H. Chate, J. Chave, and H. Hinrichsen, Phys. Rev. Lett. 87, 045701 (2001).
- [13] K. Barros, P. L. Krapivsky, and S. Redner, Phys. Rev. E 80, 040101 (2009).
- [14] M. Zhao and J. Alcacer, Global Competitors as Next-Door Neighbors: Competition and Geographic Concentration in the Semiconductor Industry, Ross School of Business Paper No. 1091, 2007.
- [15] G. Mahajan and A. Mehta, Europhys. Lett. 95, 48008 (2011);
 95, 69901 (2011).
- [16] A. A. Bhat, G. Mahajan, and A. Mehta, PLoS ONE 6, e25048 (2011).