Forgetfulness can help you win games

James Burridge
Department of Mathematics, University of Portsmouth, Portsmouth, PO1 2UP, United Kingdom

Yu Gao and Yong Mao
School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
(Received 11 May 2015; revised manuscript received 26 August 2015; published 14 September 2015)

We present a simple game model where agents with different memory lengths compete for finite resources. We show by simulation and analytically that an instability exists at a critical memory length, and as a result, different memory lengths can compete and coexist in a dynamical equilibrium. Our analytical formulation makes a connection to statistical urn models, and we show that temperature is mirrored by the agent's memory. Our simple model of memory may be incorporated into other game models with implications that we briefly discuss.

DOI: 10.1103/PhysRevE.92.032119 PACS number(s): 02.50.Le, 05.65.+b, 89.75.Fb

I. INTRODUCTION

All successful forms of life must eventually engage in competition for resources. The equilibrium analysis of these competitions began with von Neumann [1] and Nash [2]. The theory of games has since found applications in genetics, ecology, economics, and sociology [3–6]. Computational implementation of games leads to agent-based models, which may be of particular importance in understanding the behavior of financial systems [6,7]. For example, the particularly successful minority game model [8–12] captures the competition between intelligent agents with a restricted form of memory. Recent work suggests such games may be generalized leading to clearly separated regimes of behavior [13]. In general, understanding the complex collective behavior arising from the nonlinear interactions between individuals is a major challenge for statistical physics [14,15].

In this paper we present a simple discrete time game model: individual agents reside in one of two urns, each offering a stochastic yield to be shared by the residents. Each agent has a memory of these payouts for the previous τ rounds, and may switch urns, using this memory to aid its decision. Such stochastic yield sharing is known as "exploitative competition" in ecological settings, where the urns may represent different prey species [16] or foraging patches [17,18]. Stock market investors may also be viewed as foragers [19], searching for under-exploited stocks, although price dynamics adds an extra layer of complexity to the game in this case. Some form of intelligence is essential in order to compete [20]. As with the minority game [8], agents' memory in our model is a tool for decision making, but our agents' memory is used to make direct estimates of the highest paying choice, rather than to second guess opponents' next moves. Whereas minority game agents possess both a detailed short term bit-string memory, and a secondary memory in the form of their strategy scores, our agents' only tool is a finite version of the latter memory type; they are more primitive. The behavior we uncover is related to the finite length of this memory: it is possible for payoffs to be remembered for too long. In the minority game the optimal length bit-string memory maximizes coordination between agents, but in our game the profitability of different memory lengths is due to a competition between obtaining accurate statistics and maintaining stability.

In common with the thermal minority game [10], dynamical urn models [21], and some evolutionary games [22], our agents have a "selection temperature," which captures the level of noise in the switches they make in search of yields. We find that the additional noise inherent in their finite memory samples, which is greater for shorter memories, leads the system to behave as if its agents have a higher temperature. Therefore, increasing memory "cools" the system. However, at a critical memory, a Hopf bifurcation [23] emerges producing stable cycles in the numbers of agents in each urn. Perhaps not surprisingly, a long memory is advantageous when the system possesses a stable fixed point, but the presence of these cycles allows short-memory agents to compete, and a mixed memory system will evolve toward the bifurcation point. Our theoretical formulation follows that of statistical urn models such as Ehrenfest's dog flea model [24], which played an important role in the early development of statistical mechanics, and more recently allowed analytical investigation of effects such as slow relaxation and condensation in nonequilibrium statistical mechanics [21].

While we have restricted our analysis to a yield-sharing game, our simple model of memory may be applied in general to social systems, where agents make decisions to switch between behaviors based on their memory of past payoffs or interactions. Simple systems in which memory could naturally be incorporated include the Hawk Dove [3] and Rock Scissors Paper [25] games, or the voter model [26]. The fact that systems of interacting units display a remarkable range of collective behavior controlled by temperature suggests that these behaviors, including condensation and the emergence of order [26], which have social interpretation, may also have a connection with memory [27], but that long memory can also introduce instability.

II. MODEL DEFINITION

Consider the case of two urns and a total of n agents. We let the urn yields, $U_1(t)$, $U_2(t)$, at round t be random variables uniformly distributed on $[0,\omega n]$ and $[0,n]$, respectively, where $\omega > 1$ so that urn 1 yields more on average than urn 2. This choice of yield distribution is motivated by simplicity—only one extra model parameter, ω, is required to describe it. The behavior we uncover is not tied to the particular choice of

yield distribution; it requires only that the yields have finite variance, allowing the central limit theorem [28] to be applied to sums of payoffs. We allow agents access to the arithmetic mean of the last τ payoffs, but we note that other forms of sampling could be used. Letting ϕ_t be the fraction of agents in urn 1, then the difference in the average payoffs between urn 1 and urn 2 is

$$\Delta_t := \frac{1}{\tau}\sum_{s=0}^{\tau-1}\left[\frac{U_2(t-s)}{n(1-\phi_{t-s})} - \frac{U_1(t-s)}{n\phi_{t-s}}\right]. \tag{1}$$

We refer to τ as the "memory length" of the agents. Agent dynamics is encoded in transition probabilities between urns, which are deterministic functions of Δ_t. At each round, each agent will switch urns using the probabilities

$$W_{1\to 2}(\Delta) = \frac{\epsilon}{2}[1+\tanh(\beta\Delta)], \tag{2}$$

$$W_{2\to 1}(\Delta) = \frac{\epsilon}{2}[1-\tanh(\beta\Delta)]. \tag{3}$$

These so called "Fermi Functions" are used in evolutionary game models [22,29,30], where changes in strategy are made based upon perceived increases in payoff, but with some degree of noise or irrationality in the switching process. The level of this stochasticity in decision making, which has been experimentally measured in humans [31] is captured by the parameter β, the "inverse temperature." For finite β, agents may decide to switch strategies even though their estimate of the payoff difference is unfavorable. In the limit $\beta\to\infty$ agents will only move if their estimate of the payoff difference indicates that the move is favorable. The parameter ϵ controls the rate at which strategy switching takes place compared to the rate at which yield information arrives, or equivalently the inertia in agent's decision making. It may also be seen as the frequency with which opportunities to switch strategy arise. In the limit $\epsilon\to 0$, at most one agent will move at each round. We note that the effects we uncover in the remainder of this paper do not require that the transition probabilities take the particular forms of Eqs. (2) and (3). In the Appendix we show that the same qualitative behavior is observed for transition probabilities that increase in proportion with perceived payoff difference ($\pm\Delta$), provided the difference is positive and $|\Delta| < \epsilon^{-1}$.

III. SIMULATION

A. Instability

We simulate the model for a series of values of τ when $n = 10^6$. Two different values of ϵ are used; in Fig. 1 we have $\epsilon^{-1} = 10^6 \gg \tau$ and in Fig. 2 we have $\epsilon = 10^{-3}$. For $\epsilon = 10^{-6}$, the expected number of moves at each step is <1, and ϕ appears very stable. For larger ϵ, ϕ experiences much larger fluctuations about the steady-state value, driven by the yield process. For shorter memory values these fluctuations are random, but as τ approaches ϵ^{-1}, periodic oscillations appear and dominate. The appearance of these stable oscillations at critical memory, τ_c, calculated analytically in Sec. IV, is known as a Hopf bifurcation [23]. From Figs. 1 and 2 we see that the number of rounds taken for the system to evolve to steady state, in the probabilistic sense, starting from an equal distribution between the two urns, is $\approx\epsilon^{-1}$. By allowing agents access only to the mean of their memory, we implicitly assume that changes in the expected payoff over the course of their memory, brought about by oscillations, are too subtle for them to infer from noise.

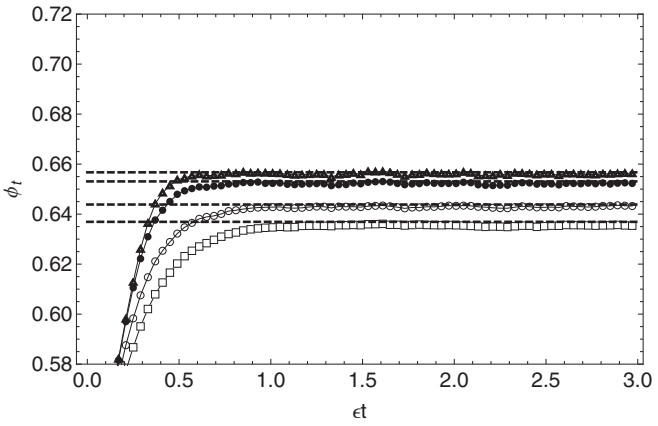

FIG. 1. Evolution of ϕ_t when $n = 10^6$, $\omega = 2$, $\beta = 5$, $\epsilon = 10^{-6}$, and $\phi_0 = 0.5$. Memory values are $\tau \in \{5, 10, 50, 500\}$ (squares, circles, dots, triangles, respectively). Dashed lines are analytical equilibrium values [see Eq. (12)]. Critical memory (see Sec. IV) is $\tau_c \approx 1.8\times 10^5$.

B. Coexistence

We now investigate how agents with two different memories compete against one another by interpreting the payoff as reproduction rate. We define δ and γ as rates of death and reproduction per unit payoff, respectively. Reproduction is assumed to occur before death in each round, but in practice the probability of any one agent reproducing and dying in the same round is extremely small for the γ, δ values we choose. Letting $p_i^\tau(t)$ be the number of agents with memory τ in urn i at time t we set the probability of birth for each agent in urn i to be

$$\mathbf{P}(\text{birth}) = \frac{\gamma U_i(t)}{\sum_\tau p_i^\tau(t)}. \tag{4}$$

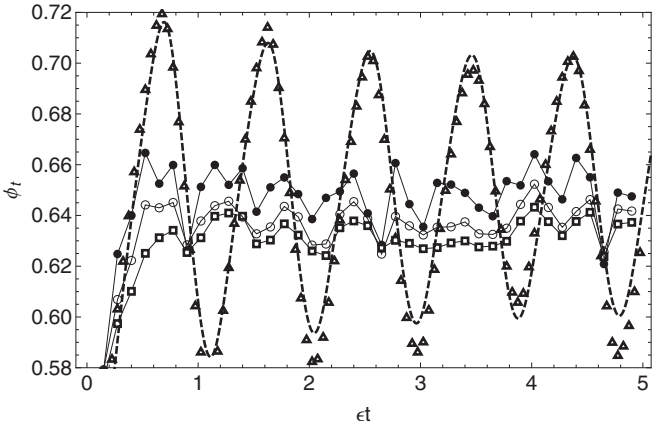

FIG. 2. Evolution of ϕ_t when $n = 10^6$, $\omega = 2$, $\beta = 5$, $\epsilon = 10^{-3}$, and $\phi_0 = 0.5$. Memory values are $\tau \in \{5, 10, 50, 500\}$ (squares, circles, dots, triangles, respectively). Dashed line is solution to Eq. (15) when $\tau = 500$ and ω, β, ϵ are as above. Critical memory (see Sec. IV) is $\tau_c \approx 390$.

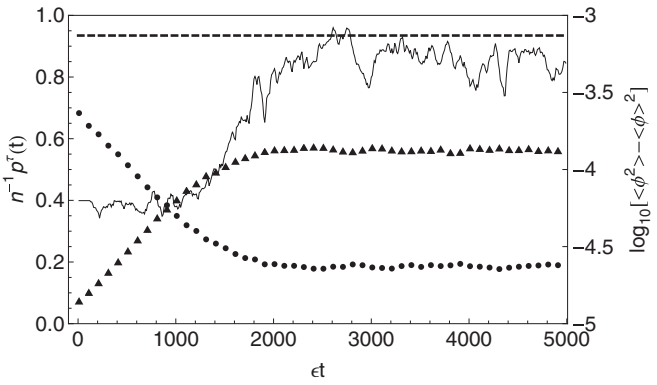

FIG. 3. Scaled populations $p^{\tau}(t) := p_1^{\tau}(t) + p_2^{\tau}(t)$ for $\tau = 10$ (circles) and $\tau = 1000$ (triangles) when total initial population is $n = 10^6$, $\epsilon = 10^{-3}$, $\beta = 5$ with $\gamma = 10^{-4}$ and $\delta = 2 \times 10^{-4}$. Also shown (thin black line) is evolution of variance of ϕ_t, over a moving time window of 10^5 steps, during population dynamics simulation. Straight dashed line shows variance of homogeneous population with the same ϵ, β, ω values at critical memory $\tau_c \approx 390$, where τ_c is calculated analytically using the theory of Hopf bifurcation (Sec. IV). Note: rapid initial equilibration of population values (bringing birth and death into balance) is not visible on time scale of plot.

The denominator in this term is the total population in urn i, reflecting the fact that the yield $U_i(t)$ is shared between all agents in the urn. When a given agent reproduces, he produces a copy of himself in his current urn, having the same memory length, τ. The death probability for each agent is set equal to δ. If populations are fixed in size and the system is not in an oscillatory state, then we expect that in equilibrium the longer memory agents will dominate the high yielding urn. Their long memory allows them to perceive smaller statistical advantages that are obscured by noise for the short-memory agents. Using the thermodynamic analogy, the higher temperature (shorter memory) agents are more likely to make moves that leave them in an urn with a lower expected payoff, corresponding to a higher "energy" state. Above zero temperature, and in the absence of oscillations, the high yield urn will be under-exploited, placing high memory agents at an advantage. This effect can be observed in Fig. 3, where we have simulated a mixed population of two memories $\tau \in \{10, 1000\}$ beginning with a ratio of 10:1 short-memory to long-memory agents. We see that initially the advantage afforded the long-memory agents causes their population to grow, whereas the short-memory agents reduce in number. Were this advantage to be sustained indefinitely then we would expect the short-memory agents to eventually disappear, but in fact the populations stabilize. This effect appears because the long-memory agents cause oscillations to develop once they are in sufficiently high concentration. In the presence of oscillations the short-memory agents have an advantage because they can quickly observe opportunities offered by the oscillating payoffs. We therefore expect the system to evolve to the point where oscillations are just beginning to form. We may observe this evolution by making use of the variance of ϕ_t as an order parameter that captures proximity to the Hopf bifurcation point. In Fig. 3 we see that at a critical ratio of short- to long-memory agents, the variance climbs rapidly, stabilizing just below the value seen in a system where all agents have memory τ_c but all other parameters are equal. In this way the Hopf bifurcation may be viewed as a self-organized state.

IV. ANALYSIS

A. Equilibrium

We consider the behavior of the model as $\epsilon \to 0$, allowing us to view it as an urn model in the Ehrenfest class [21], where agents independently make transitions using state (ϕ_t) dependent probabilities. Provided $\tau \ll \epsilon^{-1}$, the fraction ϕ_t may be approximated by a constant ϕ during the window over which payoff averaging takes place. In this case, by the central limit theorem, the marginal distributions of Δ_t for each t are approximately normal $N(\bar{\Delta}, \sigma^2/\tau)$ where, from Eq. (1),

$$\bar{\Delta}(\phi,\omega) := \frac{1}{2}\left(\frac{1}{1-\phi} - \frac{\omega}{\phi}\right), \tag{5}$$

$$\frac{\sigma^2(\phi,\omega)}{\tau} := \frac{1}{12\tau}\left[\frac{\omega^2}{\phi^2} + \frac{1}{(1-\phi)^2}\right]. \tag{6}$$

We now introduce a intermediate time scale T satisfying $\tau \ll T \ll \epsilon^{-1}$ and define the time average $\langle\cdot\rangle$, over a window of length T,

$$\langle W_{i\to j}(\Delta)\rangle(t) := \frac{1}{T}\sum_{s=t-T+1}^{t} W_{i\to j}(\Delta_s). \tag{7}$$

This average is a random variable which, for constant ϕ, has expected value $\mathbf{E}[W_{i\to j}(\Delta)]$, where the expectation is taken over the marginal distribution of Δ. The condition $\tau \ll T \ll \epsilon^{-1}$ ensures that ϕ is approximately constant over the window and that the variance of $\langle W_{i\to j}(\Delta)\rangle$ is proportional to T^{-1} (because Δ_{t_1} and Δ_{t_2} are dependent only when $|t_2 - t_1| < \tau \ll T$). As $\epsilon \to 0$, then assuming T is sufficiently large, the probability that an agent will make a transition $i \to j$ during interval T approaches $T\langle W_{i\to j}(\Delta)\rangle \approx T\mathbf{E}[W_{i\to j}(\Delta)]$, equivalent to a memoryless (Ehrenfest class) model, where transition probabilities, Eqs. (2) and (3), are replaced with their expectations $\mathbf{E}[W_{i\to j}(\Delta)]$. Averaging over the normally distributed difference Δ we find that

$$\langle W_{1\to 2}(\Delta)\rangle \approx \mathbf{E}[W_{1\to 2}(\Delta)] \approx \frac{\epsilon}{2}[1 + \tanh(\alpha\bar{\Delta})], \tag{8}$$

where

$$\alpha = \sqrt{\frac{2\tau\beta^2}{2\tau + \pi\beta^2\sigma^2}}. \tag{9}$$

To obtain this result, we have made the approximation $\tanh(\beta\Delta) \approx \mathrm{erf}(\sqrt{\pi}\beta\Delta/2)$, allowing us to make use of the exact relationship $\mathbf{E}[\mathrm{erf}(\sqrt{\pi}\beta\Delta/2)] = \mathrm{erf}(\sqrt{\pi}\alpha\bar{\Delta}/2)$. The constant α acts as an effective inverse temperature and we see that increasing τ "cools" the system closer to the inverse temperature β, and in the limit $\beta \to \infty$, $\alpha \propto \sqrt{\tau}$. To complete our analogy to a thermal urn model we now write the probability of finding the agents in a particular arrangement, or microstate, i, such that a fraction ϕ are in urn 1, as $p_i(\phi) \propto e^{-\alpha E}$ where E is an "energy" function. Considering two microstates separated by a single transition, and defining $\delta\phi = 1/n$, then detailed balance requires

that in equilibrium $2\alpha\bar{\Delta} = \partial_\phi(\alpha E)\delta\phi$. This condition allows $E(\phi)$ to be computed, in principle, by integration. A closed form approximation $E(\phi) \approx -n\ln[\phi^\omega(1-\phi)]$ is obtained by noting that α depends weakly on ϕ compared to E so that $\partial_\phi(\alpha E) \approx \alpha\partial_\phi E$. Summing over all microstates corresponding to macrostate ϕ we have a Boltzmann probability distribution for ϕ,

$$\mathbf{P}(\phi) = \frac{n!}{(n\phi)![n(1-\phi)]!}\frac{e^{-\alpha(\phi)E(\phi)}}{\mathcal{Z}}, \tag{10}$$

where $\mathcal{Z}$ is the partition function. Taking the thermodynamic limit $n \to \infty$, and making use of Stirling's approximation, we find that the most likely fraction, $\bar{\phi}$, satisfies

$$\frac{1}{2n}\frac{\partial}{\partial\phi}\ln\mathbf{P}(\phi) = \alpha\bar{\Delta} - 2\phi + 1 = 0. \tag{11}$$

As the memory increases and the system cools we expect the agents to arrange themselves so that yields are shared more fairly. We therefore linearize Eq. (11) about the perfectly fair state, $\phi = \omega/(1+\omega)$, where agents in both urns receive the same expected payoff, finding that

$$\bar{\phi} \approx \frac{f(\tau) + \frac{\beta(1+\omega)^2}{2}}{2f(\tau) + \frac{\beta(1+\omega)^3}{2\omega}}, \tag{12}$$

where $f(\tau) = \sqrt{1 + \pi\beta^2(1+\omega)^2/(12\tau)}$. The accuracy of this approximation is verified in Fig. 1. For larger values of ϵ (Fig. 2), agents move more quickly so the averaging effect Eq. (8) damps fluctuations in transition rates less strongly, creating larger fluctuations in ϕ_t. For finite β the system cannot reach perfect fairness for any memory length, but in the limit $\beta \to \infty$ where the transition probabilities, Eqs. (2) and (3), become step functions, we have that

$$\bar{\phi} \approx \frac{\omega}{\omega+1}\left[1 - \frac{\sqrt{\pi}(\omega-1)}{\sqrt{3\tau}(\omega+1)^2} + \mathcal{O}(\tau^{-1})\right]. \tag{13}$$

From this we see that the distance away from the fair state decreases as $\tau^{-1/2}$ as the memory of the agents becomes large.

We have shown that provided the timescales of switching and agent memory are sufficiently separated, then our model behaves as a memoryless dynamical urn model [21,24] with time-averaged transition rates. This averaging is equivalent to a rescaling of temperature, and we have given an approximate analytic expression for this new temperature, α, in terms of the underlying "selection temperature" of agents, β, and their memory length, τ. The study of the dynamics of particles in urns began with Ehrenfest's dog-flea model [24], where agents make random transitions between two urns at a fixed rate without reference to energy or temperature. Thermal models [21] include an energy, E, and have transition probabilities that respect detailed balance with respect to the Boltzmann distribution. These probabilities are decreasing functions of $\beta\Delta E$, where ΔE is the energy change associated with the transition, and β is the inverse thermodynamic temperature, so that reducing temperature makes energy increasing transitions less likely. In our case, $\bar{\Delta}$ plays the role of an energy change and α is the inverse temperature. Increasing memory reduces α, making payoff reducing transitions less likely. However, if memory is increased to the extent that it becomes comparable to the timescale of switching between urns ($\approx\epsilon^{-1}$) then the model begins to behave quite differently to classical thermal urn models [21,24]. It becomes important that the average payoff agents use to make decisions is calculated from the history of the system. This is because the populations in the urns are able to change significantly during the course of a single agent's memory, and so using delayed information for decision making can falsely identify the optimal urn. To understand mathematically why increasing τ too far, when ϵ is finite, destabilizes the system, we must make use of the theory of delay differential equations [23].

B. Instability

As τ increases, fluctuations in Δ_t due to the yield process are reduced but for finite ϵ we can no longer treat ϕ_t as a constant over the averaging window. It is instructive, therefore, to study the effect of variations in ϕ_t, neglecting the variations in yield. Promoting t to a continuous variable and replacing the urn yields with their mean values we have

$$\Delta_t \approx \frac{1}{2\tau}\int_{t-\tau}^{t}\left[\frac{1}{1-\phi_s} - \frac{\omega}{\phi_s}\right]ds. \tag{14}$$

We then approximate the evolution of ϕ_t using the following delay differential equation:

$$\dot{\phi}_t = (1-\phi_t)W_{2\to1}(\Delta_t) - \phi_t W_{1\to2}(\Delta_t). \tag{15}$$

A numerical solution to this equation is shown in Fig. 2, along with simulation results using the same parameter values. The oscillations in the simulation are accurately captured by Eq. (15), but the stochastic yield disrupts their perfect periodicity. To discover the parameter values at which stable oscillations develop we linearize Eq. (15) by writing $\phi_t = \bar{\phi} + \psi_t$, where ψ_t are small fluctuations and $\bar{\phi}$ is the constant fixed point, not necessarily stable, of Eq. (15). In terms of these new variables,

$$\Delta_t \approx \bar{\Delta}(\bar{\phi},\omega) + 6\frac{\sigma^2(\bar{\phi},\sqrt{\omega})}{\tau}\int_{t-\tau}^{t}\psi_s ds, \tag{16}$$

where the functions $\bar{\Delta}$ and σ^2 are defined in Eqs. (5) and (6). After expanding the tanh functions in the transition rates to first order about $\bar{\Delta}(\bar{\phi},\omega)$, we obtain the following linear delay equation:

$$\dot{\psi}_t = -\epsilon\left[\psi_t + \frac{A}{\tau}\int_{t-\tau}^{t}\psi_s ds\right], \tag{17}$$

where $A = 3\beta\,\mathrm{sech}^2[\beta\bar{\Delta}(\bar{\phi},\omega)]\sigma^2(\bar{\phi},\sqrt{\omega})$. To determine the stability of this equation we introduce an exponential trial solution $\psi_t = e^{\lambda t}$ where $\lambda = x + iy$. Substitution into Eq. (17) yields a characteristic equation with real and imaginary parts given by

$$x^2 - y^2 + \epsilon x + \frac{\epsilon A}{\tau}(1 - e^{-\tau x}\cos\tau y) = 0, \tag{18}$$

$$2xy + \epsilon y + \frac{\epsilon A}{\tau}e^{-\tau x}\sin\tau y = 0. \tag{19}$$

For sufficiently small memory, τ, the real part, x, of the solutions to Eqs. (18) and (19) is negative so the fixed point $\bar{\phi}$ is stable. As we increase τ, λ crosses through the imaginary

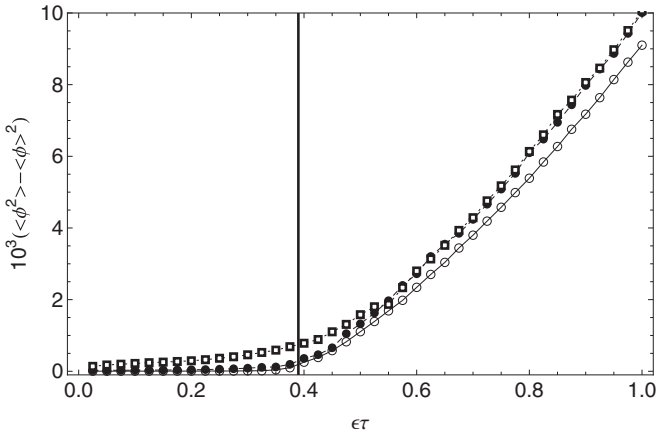

FIG. 4. Estimated variance of ϕ_t in steady state, as a function of memory length τ, from simulations with $n = 10^6$, $\omega = 2$, $\beta = 5$, and $\epsilon = 10^{-3}$ (squares), 10^{-4} (dots), 10^{-5} (circles). Variance estimates computed using time average over 10^6 time steps for $\epsilon \in \{10^{-3}, 10^{-4}\}$ and 10^7 time steps for $\epsilon = 10^{-5}$. Vertical black line marks theoretical Hopf bifurcation point $\epsilon\tau_c = 0.39$, computed from Eq. (21).

axis, destabilizing the fixed point and creating oscillations of exponentially increasing magnitude in the linearized version Eq. (17) of the full delay Eq. (15). Although the fixed point of the full Eq. (15) shares this transition to instability, we find that the resulting oscillations are bounded. The appearance of these stable oscillations as τ passes through a critical value, which we denote τ_c, constitutes the Hopf Bifurcation [23]. To compute τ_c we set $x = 0$ in Eq. (19) so that $\mathrm{sinc}(\tau y) = A^{-1}$. Expanding the sinc function to second order about its root at π/τ and solving the resulting quadratic, we find that

$$y \approx \frac{\pi}{2\tau}(3 - \sqrt{1 - 4A^{-1}}) := \frac{\kappa}{\tau}, \tag{20}$$

which defines a new constant κ. Substitution of this solution into Eq. (18) yields the following expression for the critical memory length:

$$\tau_c = \frac{\kappa^2}{\epsilon A(1 - \cos\kappa)}. \tag{21}$$

In order to test this analysis, in Fig. 4 we have plotted estimates of the variance of ϕ_t in equilibrium as a function of memory length for three values of $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}\}$. The transition from stable fixed point to limit cycles should, according to Eq. (21), occur when the product $\epsilon\tau$ reaches a critical value, beyond which we expect the oscillations to increase the variance of ϕ_t. This behavior is clearly observed in the Fig. 4. For the largest ϵ value, the variance is significantly greater than zero for $\tau < \tau_c$ because the shorter memory length has a reduced damping effect on stochastic fluctuations in the yield process.

V. CONCLUSION

We have introduced a simple thermal urn model of competition between agents with memory. Increasing memory allows agents to more accurately determine the most productive strategy and reduces the temperature of the model. However, if a sufficiently high concentration of long-memory agents is present a limit cycle appears that reduces their competitiveness, leading to a self-organized Hopf bifurcation driven by population dynamics. The simplicity of our memory model, its connection to classical urn models, together with the fact that limit cycles arise naturally, suggest it might be fruitfully generalized and employed to study a different games. For example, our approach may be applied to the Rock Scissors Paper game [3], where agents, interacting pairwise, recall their last τ interactions [33]. Other natural extensions include the introduction of multiple urns to represent different sources of yield or game strategies, or heterogeneity in switching rates and a more general distribution of memory lengths. By introducing multiple urns we might expect to observe more complex patterns of oscillation [32] and regimes of behavior [12]. Experimental research into the nature of human and animal memory [34–37] places emphasis on the "forgetting function," which describes how memories decay with time. Such a function, or greater powers of statistical inference, could be naturally incorporated into our analysis, and their effects on stability explored.

ACKNOWLEDGMENT

The authors thank the referees for their useful comments.

APPENDIX: ALTERNATIVE TRANSITION PROBABILITIES

In this Appendix we address the following question: Is the "Fermi Function" (smoothed step) transition probability essential to the effects we uncover? To motivate this question, we note that the step form,

$$W_{1\to2}(\Delta) = \frac{\epsilon}{2}[1 + \tanh(\beta\Delta)], \tag{A1}$$

produces transition rates that remain approximately constant for values of the payoff difference $\Delta \gg \beta^{-1}$. As β becomes large, agents become progressively less sensitive to the magnitude of the payoff difference, and in the limit $\beta \to \infty$ they react only to its sign.

As an alternative we consider the following form of transition probability:

$$\tilde{W}_{1\to2}(\Delta) = \begin{cases} \min(\epsilon\Delta, 1) & \text{if } \Delta \geqslant 0 \\ 0 & \text{if } \Delta < 0 \end{cases}, \tag{A2}$$

with $\tilde{W}_{2\to1} = \tilde{W}_{1\to2}(-\Delta)$. Here we have no concept of irrationality in agent behavior, and stochasticity in decision making is driven purely by noise inherent in the finite memory of agents. This corresponds to taking the limit $\beta \to \infty$ in the original rates. In contrast to the original form of Eq. (A1), here the transition probability is proportional to payoff difference provided that $\Delta < \epsilon^{-1}$. For the values of ϵ we consider, this condition is always met so, in contrast to the original rates, agents remain sensitive to the magnitude of Δ. With these new transition rates, the connection to thermal urn models is lost because they do not satisfy detailed balance with respect to a Boltzmann distribution for which Δ plays the role of an energy change.

We first explore the evolution of a system where all agents have identical memory. In Fig. 5 we consider the case $\epsilon = 10^{-3}$ where we see that provided agents' memory is sufficiently

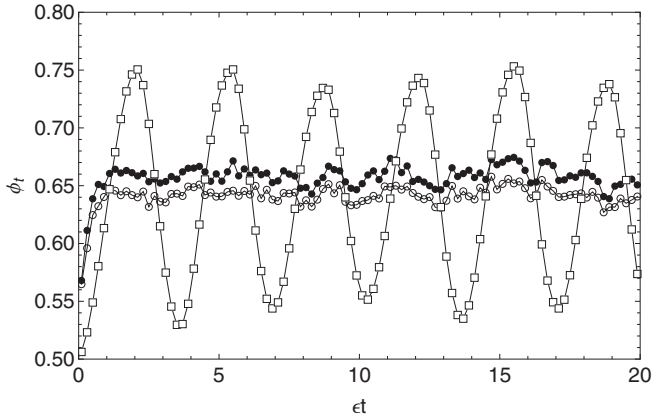

FIG. 5. Evolution of ϕ_t (fraction of agents in urn 1) using alternative transition probabilities when $n = 10^6$, $\epsilon = 10^{-3}$, $\omega = 2$, and $\phi_0 = 0.5$. Memory values are $\tau \in \{5, 100, 1750\}$ (open circles, dots, squares).

short, the system remains stable. As with the original rates, increasing memory brings the system closer to the fair state $\phi = \omega/(1+\omega)$, but for long enough memory, high-amplitude regular oscillations appear. Qualitatively, therefore, the system behaves in the same way for both choices of rate. However, the critical memory length at which oscillations appear differs between the two choices.

We now consider the case of mixed memory with population dynamics. In Fig. 6 we have simulated a population of two memory lengths $\tau \in \{100, 4000\}$ in the case $\epsilon = 10^{-3}$. From Fig. 5 we see that these memory values lie below and above the critical length, respectively. We see that the behavior of the system matches its behavior in the case of smoothed step rates Eq. (A1): initially the population of short-memory agents declines. While the system possesses a stable fixed point (no oscillations) agents with a longer memory make more accurate estimates of the true payoff difference between the urns and are less likely to make detrimental moves. Once oscillations

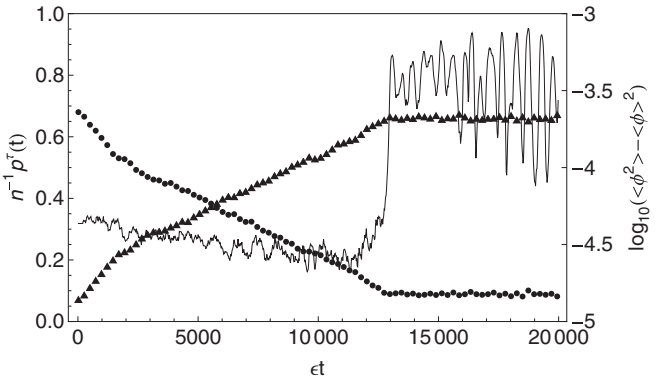

FIG. 6. Scaled populations $p^{\tau}(t) = p_1^{\tau}(t) + p_2^{\tau}(t)$ for $\tau = 100$ (circles) and $\tau = 4000$ (triangles) when initial population is $n = 10^6$ with memory types in ratio short:long = 10:1. Parameter values are $\epsilon = 10^{-3}$, $\gamma = 10^{-4}$, and $\delta = 2 \times 10^{-4}$. Also shown (thin black line) is evolution of variance of ϕ_t, over a moving time window of 10^5 steps, during population dynamics simulation. Note: rapid initial equilibration of population values (bringing birth and death into balance) is not visible on time scale of plot.

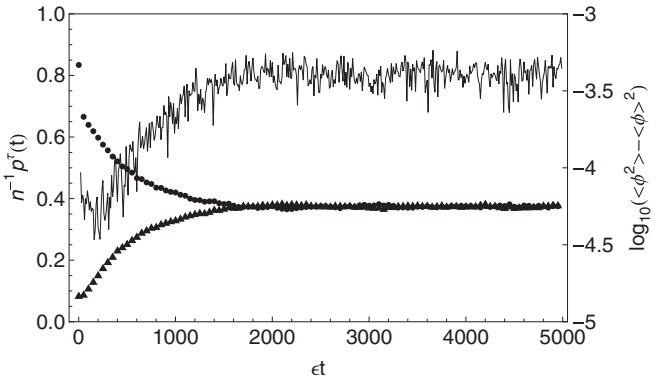

FIG. 7. Scaled populations $p^{\tau}(t) = p_1^{\tau}(t) + p_2^{\tau}(t)$ in the original model with step rates when inverse temperature $\beta \to \infty$. Memory lengths are $\tau = 10$ (circles) and $\tau = 1000$ (triangles) and initial population is $n = 10^6$ with memory types in ratio short:long = 10:1. Parameter values are $\epsilon = 10^{-3}$, $\gamma = 10^{-4}$, and $\delta = 2 \times 10^{-4}$. Also shown (thin black line) is evolution of variance of ϕ_t, over a moving time window of 10^5 steps, during population dynamics simulation. Note: rapid initial equilibration of population values (bringing birth and death into balance) is not visible on time scale of plot.

appear, short-memory agents have an advantage because they respond more quickly to opportunities created by oscillating payoffs. The onset of oscillations takes place once long-memory agents are in sufficient concentration and is marked by a dramatic jump in the window averaged variance ϕ_t. This jump coincides with a stabilization of the population dynamics, indicating that the advantage of long-memory players has disappeared. Figure 6 demonstrates that the phenomenon of a dynamical equilibrium between competing memory lengths appears both with "smoothed step" probabilities (W) and "proportional" probabilities ($\tilde{W}$).

To complete our analysis we perform a population dynamics simulation using the original transition probabilities (W) in the limit $\beta \to \infty$, obtaining a pure step functional form

$$W_{1\to 2}(\Delta) = \begin{cases} \epsilon & \text{if } \Delta \geqslant 0 \\ 0 & \text{if } \Delta < 0 \end{cases}, \tag{A3}$$

so that individual agents react only to the sign of the perceived payoff difference. In this way we are able to compare the two types of transition probability when agents have no intrinsic irrationality of behavior. The simulation results are shown in Fig. 7. We use the same ϵ value in both simulations, but agents make fewer moves with proportional rates because the payoff difference is typically small: $\Delta \ll 1 \Rightarrow \tilde{W}_{1\to 2}(\Delta) \ll \epsilon$. This delays the appearance of payoff differences between memory lengths that drive the population changes, so these changes take place on a longer time scale for proportional transition probabilities. This effect is evident in Figs. 6 and 7.

In both Figs. 6 and 7 we see that the population of short-memory agents initially declines, but eventually stabilizes. This stabilization occurs coincidentally with a jump in the variance of ϕ_t, indicating that Hopf Bifurcation is responsible for the dynamical equilibrium between memory lengths. It is interesting to note that the jump in variance is more dramatic in the case of proportional probabilities (Fig. 6). We suggest that this occurs because the proportional form

of transition probability damps fluctuations more effectively than the step form, so that when oscillations do appear in ϕ_t they have a more significant effect on the behavior of the system, and therefore on the relative competitiveness of agents with different memory lengths. Variance builds more gradually with step probabilities because fluctuations are damped less effectively, and prior to the fixed point losing stability, we see decaying oscillations (under damping) that increase variance when coupled with stochastic fluctuations. The relative equilibrium frequency of short- to long-memory agents depends, in a nontrivial way, on the functional form of switching probabilities, memory length, and population dynamics, and can vary quite considerably (Figs. 6 and 7). A theoretical calculation of these equilibrium population sizes remains an open challenge.

We conclude by noting that the proportional transition probability we have considered in this appendix is a representative of a wider class of functional forms that are zero when $\Delta \leqslant 0$ and increase continuously with Δ for $\Delta > 0$. When Δ is small, the behavior of such systems will depend on the first derivative of the transition probability at $\Delta = 0^+$, which in the case we have considered is equal to ϵ. We therefore expect to see similar behavior for all such forms when fluctuations about the fixed point are small.

[1] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour* (Princeton University Press, Princeton, NJ, 1953).

[2] J. Nash, Non-cooperative games, Ann. Math. **54**, 286 (1951).

[3] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[4] R. B. Myerson, *Game Theory: Analysis of Conflict* (Harvard University Press, Boston, 1991).

[5] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).

[6] J. D. Farmer and D. Foley, The economy needs agent-based modelling, Nature **460**, 685 (2009).

[7] A. Chakraborti, D. Challet, A. Chatterjee, M. Marsili, Y.-C. Zhang, and B. K. Chakrabarti, Statistical mechanics of competitive resource allocation using agent-based models, Phys. Rep. **552**, 1 (2015).

[8] D. Challet and Y.-C. Zhang, Emergence of cooperation and organization in an evolutionary game, Physica A **246**, 407 (1997).

[9] D. Challet, M. Marsili, and Y. C. Zhang, *Minority Games* (Oxford University Press, Oxford, 2005).

[10] A. Cavagna, J. P. Garrahan, I. Giardina, and D. Sherrington, A Thermal Model for Adaptive Competition in a Market, Phys. Rev. Lett. **83**, 4429 (1999).

[11] J. P. Garrahan, E. Moro, and D. Sherrington, Continuous time dynamics of the thermal minority game, Phys. Rev. E **62**, R9 (2000).

[12] A. Cavagna, Irrelevance of memory in the minority game, Phys. Rev. E **59**, R3783 (1999).

[13] T. Galla and J. D. Farmer, Complex dynamics in learning complicated games, Proc. Natl. Acad. Sci. USA **110**, 1232 (2013).

[14] J.-P. Bouchaud, An introduction to statistical finance, Physica A **313**, 238 (2002).

[15] R. Mantegna and H. E. Stanley, *An Introduction to Econophysics* (Cambridge University Press, Cambridge, 1999).

[16] P. Henschel, K. A. Abernethy, and L. T. J. White, Leopard food habits in the Lope National Park, Gabon, Central Africa, Afr. J. Ecol. **43**, 21 (2005).

[17] J. S. Brown and M. L. Rosenzweig, Habit selection in slowly regenerating environments, J. Theor. Biol. **123**, 151 (1986).

[18] D. G. C. Harper, Competitive foraging in mallards: 'Ideal free' Ducks, Anim. Behav. **30**, 575 (1982).

[19] S. Saavedra, R. D. Mailmgren, N. Switanek, and B. Uzzi, Foraging under conditions of short-term exploitative competition: The case of stock traders, P. R. Soc. B **280**, 20122901 (2013).

[20] C. Belisle and J. Cresswell, The effects of a limited memory capacity on foraging behavior, Theor. Popul. Biol. **52**, 78 (1997).

[21] C. Godreche and J. M. Luck, Nonequilibrium dynamics of urn models, J. Phys.: Condens. Matter **14**, 1601 (2002).

[22] A. Traulsen, J. M. Pacheco, and M. A. Nowak, Pairwise comparison and selection temperature in evolutionary game dynamics, J. Theor. Biol. **246**, 522 (2007).

[23] T. Erneux, *Applied Delay Differential Equations* (Springer, Berlin, 2009).

[24] P. Ehrenfest, Über zwei bekannte Einwände gegen das Boltzmannsche H-theorem, Phys. Z **8**, 311 (1907).

[25] Z. Wang, B. Xu, and H. J. Zhou, Social cycling and conditional responses in the Rock-Paper-Scissors game, Sci. Rep. **4**, 5830 (2014).

[26] C. Castellano, S. Fortunato, and V. Loreto, Statistical physics of social dynamics, Rev. Mod. Phys. **81** 591 (2009).

[27] F. Caccioli, S. Franz, and M. Marsili, Ising model with memory: Coarsening and persistence properties, J. Stat. Mech: Theory Exp. (2008) P07006.

[28] G. Grimmett and D. Stirzaker, *Probability and Random Processes* (Oxford University Press, Oxford, 2001).

[29] G. Szabo and C. Toke, Evolutionary prisoner's dilemma game on a square lattice, Phys. Rev. E **58**, 69 (1998).

[30] A. Traulsen, M. A. Nowak, and J. M. Pacheco, Stochastic dynamics of invasion and fixation, Phys. Rev. E **74**, 011909 (2006).

[31] A. Traulsen *et al.*, Human strategy updating in evolutionary games, Proc. Natl. Acad. Sci. USA **107**, 2962 (2010).

[32] P. Manfredi and L. Fanti, Cycles in dynamics economic modelling, Econ. Model. **21**, 573 (2004).

[33] J. Burridge (submitted to Phys. Rev. E).

[34] W. K. Estes, Research and theory on the learning of probabilities, J. Am. Stat. Assoc. **67**, 81 (1972).

[35] P. Killeen, Writing and overwriting short-term memory, Psychon. B Rev. **8**, 18 (2001).

[36] R. G. Cook and A. P. Blaisdell, Short-term item memory in successive same-different discriminations, Behav. Proc. **72**, 255 (2006).

[37] L. Averell and A. Heathcote, The form of the forgetting curve and the fate of memories, J. Math. Psychol. **55**, 25 (2011).