


**Optimal dynamic incentive scheduling for Hawk-Dove evolutionary games**K. Stuckey \**Department of Aerospace & Mechanical Engineering, University of Southern California, Los Angeles, California 90089-1191, USA*R. Dua<sup>†</sup>*Department of Mathematics, University of Southern California, Los Angeles, California 90089-1191, USA*Y. Ma<sup>‡</sup>*Department of Physics & Astronomy, University of Southern California, Los Angeles, California 90089-1191, USA*J. Parker<sup>§</sup>*Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, USA*P. K. Newton <sup>||</sup>*Department of Aerospace & Mechanical Engineering, Mathematics, and The Ellison Institute, University of Southern California, Los Angeles, California 90089-1191, USA*

(Received 21 September 2021; accepted 5 January 2022; published 21 January 2022)

The Hawk-Dove evolutionary game offers a paradigm of the trade-offs associated with aggressive and passive behaviors. When two (or more) populations of players compete, their success or failure is measured by their frequency in the population, and the system is governed by the replicator dynamics. We develop a time-dependent optimal-adaptive control theory for this dynamical system in which the entries of the payoff matrix are dynamically altered to produce control schedules that minimize and maximize the aggressive population through a finite-time cycle. These schedules provide upper and lower bounds on the outcomes for all possible strategies since they represent two extremizers of the cost function. We then adaptively extend the optimal control schedules over multiple cycles to produce absolute maximizers and minimizers for the system.

DOI: [10.1103/PhysRevE.105.014412](https://doi.org/10.1103/PhysRevE.105.014412)**I. INTRODUCTION**

The Hawk-Dove game (also known as the Chicken or Snowdrift game) is a game-theoretic paradigm for studying the conflict between players (or populations of players) who use two opposing strategies: aggressive (Hawks) and passive (Doves). One way of framing the conflict is to consider competition in the animal world where two different species compete for a limited resource [1–4]. With no Hawks in the population, Doves will share the resources and avoid conflict. With no Doves, the Hawks will fight with each other for resources, taking the risk of injury or death. If Hawks are present in large enough numbers, the Doves will flee without fighting. A sufficient fraction of Doves, on the other hand, can cooperate and expel the Hawks from the population, thereby protecting the resource [5]. The challenge is to find conditions for the stable coexistence of the two opposing populations. In the context of military conflicts, the game is framed as the game of chicken, thought of as a situation in which two drivers

head towards each other in a single lane trying not to be the first to swerve away (Doves), each mindful of the fact that if neither swerves (Hawks), both will die. Key to this game is that the cost of losing is greater than the value of winning. Versions of this (static) game have been analyzed and used extensively in political science communities to study strategies associated with the problem of nuclear brinkmanship [6]. In this setup, the payoffs are fixed, and the interactions unfold based on the cost-benefit balance determined by these payoffs.

In a more complicated setting, one might want to measure repeated interactions in populations of competitors,  $\vec{x} = (x_1, x_2)^T \in \mathbb{R}^2$ , where winning and losing is reinforced by the relative frequencies of the two competing populations (frequency-dependent selection as in Darwinian evolution). For this, the replicator dynamical system is commonly used [7–9],

$$\dot{x}_i = x_i((A\vec{x})_i - \vec{x}^T(A\vec{x})) \quad (i = 1, 2), \quad (1)$$

with  $x_1 + x_2 = 1$ ,  $0 \leq x_1 \leq 1$ ,  $0 \leq x_2 \leq 1$ , where each variable has the interpretation of frequency in the population or the alternative interpretation as a probability of picking a member of one of the two subgroups randomly. It is sometimes useful to also think of the variables  $\vec{x} = (x_1, x_2)^T$  as strategies (heritable traits) that evolve, with the most successful strategy dominating, as in the context of Darwinian evolution [4] by natural selection. Here,  $A$  is the  $2 \times 2$  payoff

\*kstuckey@usc.edu

†rajvirdu@usc.edu

‡yongqiam@usc.edu

§joep@caltech.edu

||Corresponding author: newton@usc.edu

matrix,  $(A\bar{x})_i$  is the fitness of population  $i$ , and  $\bar{x}^T(A\bar{x})$  is the average fitness of both populations, so  $x_i$  in (1) drives growth (instantaneously) at times when the species population  $i$  is above the population average and decay at times when it falls below the system average. Cumulative growth or decay over a time  $T$  is governed by  $\int_0^T [(A\bar{x})_i - \bar{x}^T(A\bar{x})]dt$  which can either be positive or negative depending on the proportion of time spent above or below the system average. The fitness functions in (1) are said to be population dependent (selection pressure is imposed by the mix of population frequencies) and determine the growth or decay of each subpopulation. Because of this, these equations are also used extensively in the reinforcement learning community where success begets success and failure leads to a downward spiral of frequency in the population [10].

Using the standard Hawk-Dove payoff matrix [5],

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 0 \end{bmatrix}, \quad (2)$$

where the population  $x_1$  are the Hawks (aggressive), and  $x_2$  are the Doves (passive), and the strict Nash equilibrium  $\bar{x}^* \equiv (x_1^*, x_2^*)$  is the mixed state  $x_1^* = \frac{1}{3}$ ,  $x_2^* = \frac{2}{3}$  since  $\bar{x}^{*T}A\bar{x}^* > \bar{x}^T A\bar{x}$  for all  $\bar{x} \neq \bar{x}^*$ . This implies that the mixed state is also an evolutionary stable state (ESS) of the replicator system (1) as discussed in Ref. [11]. It is also useful to uncouple the two variables in (1) and write a single equation for the aggressor population frequency  $x_1$ :

$$\begin{aligned} \dot{x}_1 &= x_1(1-x_1)[(A\bar{x})_1 - (A\bar{x})_2] \\ &= x_1(1-x_1)\{(a_{12} - a_{22}) + [(a_{11} - a_{21}) - (a_{12} - a_{22})]x_1\}. \end{aligned} \quad (3)$$

Note also that a single equation for the passive population  $x_2$  is easily obtained using the change of variable  $x_1 = 1 - x_2$  in Eq. (3).

The question we address in the paper is whether it is possible to alter the entries in the payoff matrix  $A$  in a time-dependent fashion (dynamic incentives) in order to optimally achieve some predetermined goal (such as minimizing aggression) at the end of fixed time  $T$ . Dynamically altering the entries of a payoff matrix in an evolutionary game setting has only recently been studied by coupling the entries, for example, to a system that represents an external environment [12,13]. One can easily imagine different scenarios in which this might be useful. In the context of nuclear brinkmanship, for example, is it possible to alter the payoff incentives dynamically in order to achieve a goal [6] that would not be achievable with fixed payoffs? Is it possible to offer dynamic economic incentives that optimize some desired outcome across a population of participants [14,15]? In this context, one can fruitfully think of the problem as one in optimal control theory where in real time, the government either turns on or turns off various economic incentives, such as adjusting interest rate policies or implementing targeted changes to control inflationary pressure. In the context of evolutionary game theory, this translates into introducing time-dependent control schedules that can alter the punishment-reward structure of the system by explicitly controlling the entries of the payoff matrix.

Can one optimally design time-dependent incentive schedules of rewards/punishments to compel groups of people to get vaccinated [16]? For coevolving microbial populations, is it possible to dynamically schedule selective antibiotic agents in order to *steer* the evolutionary trajectory in an advantageous direction [17,18], or even reverse antibiotic resistance, or in the context of scheduling chemotherapy treatments, is it possible to design schedules optimally that make best use of the chemotherapy agents administered in order to delay chemotherapeutic resistance [8,9,19–21]? Control theory is increasingly being used in a wide range of biological applications [21–27] but to date has not been systematically implemented in the context of evolutionary games as far as we know, aside from Refs. [8,9,21,28].

One interesting evolutionary context where an apparent Hawk-Dove scenario may require attainment of a quasistable equilibrium condition is during the evolution of symbiotic relationships in which one partner is aggressive or predatory. For example, hostile colonies of eusocial insects, such as ants and termites, are plagued by a diversity of solitary arthropods that have evolved to infiltrate the social system and parasitize the nest [30,31]. The majority of such parasitic species evolved from free-living ancestors without any behavioral specialization [32,33]. It follows that the initial steps in establishing the symbiosis were contingent on these free-living species (the Doves) entering into equilibrium with their aggressive eusocial hosts (the Hawks). This equilibrium, once attained, may have provided an essential, permissive stepping stone to evolving the essential adaptive traits—such as social behaviors and pheromonal mimicry—that facilitate social parasitism [32].

To address these and related types of settings, we develop a mathematical framework to determine time-dependent incentive schedules for altering the payoff entries of a Hawk-Dove evolutionary game in such a way as to (i) maximize aggression at the end of time  $T$ , and (ii) minimize aggression at the end of time  $T$ . By considering the bang-bang schedules that produce these upper and lower bounds on the competing frequencies, we can conclude that any alternative payoff schedule will produce a result that lies somewhere between the two bounds as each are extremizers of a cost function associated with the Pontryagin maximum (minimum) principle. We then extend the time period to time  $nT$  ( $n = 1, \dots, 5$ ) by using an adaptive control method that adjusts the schedule in the  $(n+1)$ st window based on the ending frequency values from the  $n$ th window. The schedules produced drive aggression down to an absolute minimum ( $x_1^{\min}$ ), or drive it up to an absolute maximum ( $x_1^{\max}$ ), which are functions of the cycle time  $T$ . These values provide absolute lower and upper bounds on opposing behavior strategies in an evolutionary setting.

## II. OPTIMAL CONTROL THEORY FOR THE REPLICATOR DYNAMICAL SYSTEM

To implement an optimal dynamic incentive strategy, we consider the time-dependent system

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = A_0 + A_1(t) \quad (4)$$

$$= \begin{bmatrix} 3 & 1 \\ 5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 4u_2(t) \\ 6u_1(t) & 0 \end{bmatrix} \quad (5)$$

$$= \begin{bmatrix} 3 & 1 + 4u_2(t) \\ 5 + 6u_1(t) & 0 \end{bmatrix}, \quad (6)$$

where  $A_1(t)$  represents our control with entries in the off-diagonal terms, and  $A_0$  is the baseline Hawk-Dove payoff matrix. The time-dependent functions  $\vec{u}(t) = [u_1(t), u_2(t)] \in \mathbb{R}^2$  are bounded above and below,  $-1 \leq u_1(t) \leq 1$ ,  $-1 \leq u_2(t) \leq 1$  and have a range ( $-3 \leq a_{12} \leq 5$ ;  $-1 \leq a_{21} \leq 11$ ) that allows us to traverse the plane along any path depicted in red in Fig. 1, starting in the Hawk-Dove zone in the uncontrolled ( $u_1 = 0; u_2 = 0$ ) case which is shown in Fig. 2 in the phase plane [Fig. 2(a)] and the frequency plane [Fig. 2(b)]. Traversing this plane amounts to dynamically sampling the five reciprocity mechanisms allowable in  $2 \times 2$  games, discussed in the context of social dilemmas in Refs. [34,35]. Dynamically altering the entries of the payoff matrix allows the freedom to choose different mechanisms to gain advantage over some prescribed finite time. Note that we have chosen the coefficients and bounds on the controllers to ensure that all regions of the plane in Fig. 1 are accessible. The ESS for the uncontrolled case is  $x_1 = 1/3$ . The control path chosen, and the time parametrization  $0 \leq t \leq T$  determines both the sequence of games being played as well as the switching times (the times at which the path crosses over from one region to the next) between games. We denote the total control output

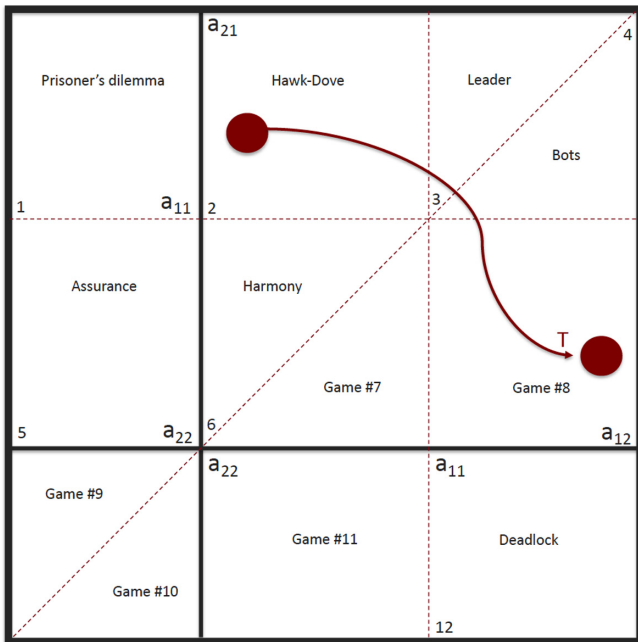


FIG. 1. Twelve regions in the off-diagonal ( $a_{12}, a_{21}$ ) plane [29] (delineated by dashed lines) define which game is being played. We choose  $a_{22}$  at the origin (without loss). Starting at  $t = 0$  in the Hawk-Dove square, indicated by the initial large (red) dot, what are the paths that minimize and maximize aggression at time  $t = T$ ? The solid (red) curve is one example of a path that crosses into four different regions of the plane.

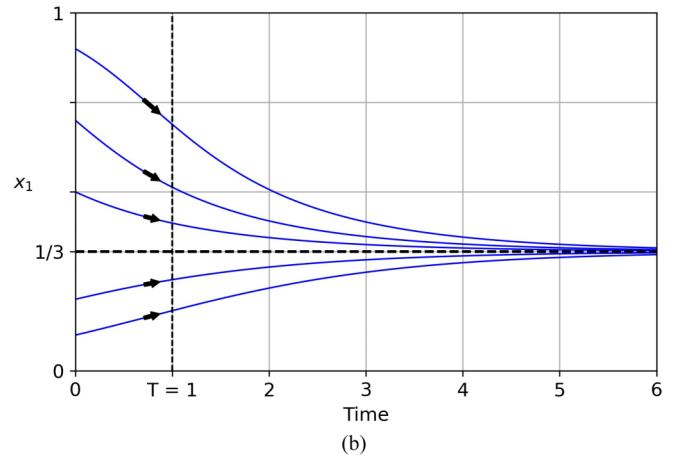
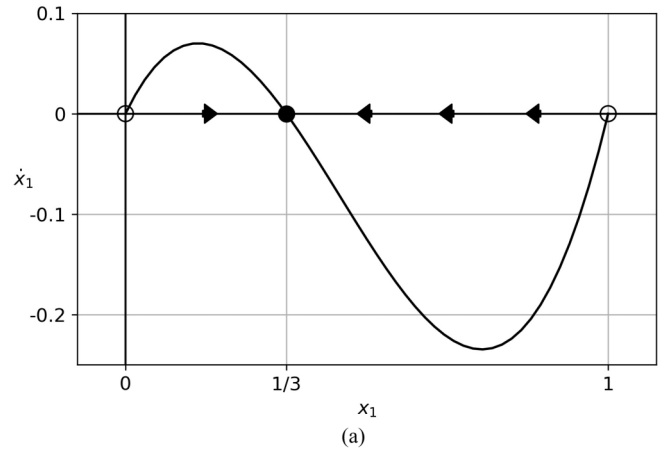


FIG. 2. Dynamics of the uncontrolled ( $u_1 = 0, u_2 = 0$ ) Hawk-Dove evolutionary game. (a) Phase portrait associated with the aggressor population  $x_1$ . Both Hawk and Dove dominance ( $x_1 = 1, 0$ ) are unstable fixed points, while the mixed state  $x_1 = 1/3$  is the evolutionarily stable strategy (ESS). (b) Hawk dynamics for various initial conditions.  $T = 1$  is the end of one control cycle and also the linear growth rate of the Hawk-Dove system.

$$\vec{C} \in \mathbb{R}^2,$$

$$\vec{C}(t) = [C_1(t), C_2(t)] = \int_0^t \vec{u}(t) dt, \quad (7)$$

with total output delivered in time  $t$ , then

$$\dot{\vec{C}}(t) = \vec{u}(t) \quad (8)$$

and

$$\vec{C}(0) = 0, \quad (9)$$

$$\vec{C}(T) = \int_0^T \vec{u}(t) dt = \vec{C}_T, \quad (10)$$

where  $T$  denotes a final time in which we implement the control over one cycle. We consider  $\vec{C}_T$  as a constraint on the optimization problem, with  $\vec{C}_T = (0, 0)$ , and our goal is to first find schedules that minimize and maximize aggression ( $x_1$ ) at the end of one cycle  $t = T$  subject to this constraint. For the uncontrolled case, we know  $x_1 \rightarrow 1/3$  as  $t \rightarrow \infty$  and we compare the controlled cases with the uncontrolled

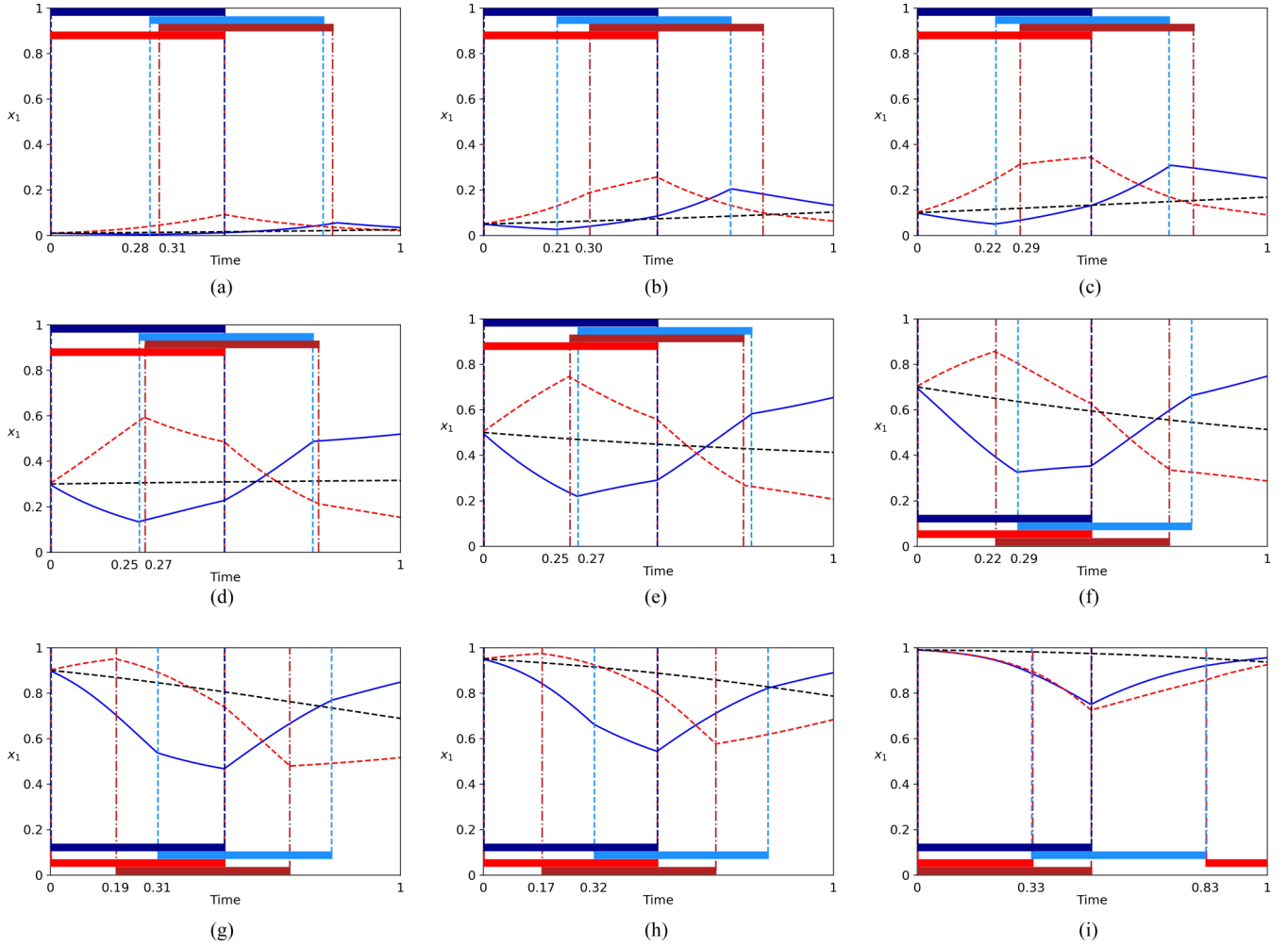


FIG. 3. Maximizing (solid blue) and minimizing (dashed red) trajectories for nine initial conditions. Also shown is the (dashed black) curve for the uncontrolled Hawk-Dove trajectory which lands in between the maximum and minimum at  $T = 1$ . The top bar (dark blue) corresponds to  $u_1 = 1$ , with white indicating  $u_1 = -1$ . The second bar (light blue) corresponds to  $u_2 = 1$ , with white indicating  $u_2 = -1$  associated with the maximizing control schedule. The third bar (dark red) corresponds to  $u_1 = 1$ , with the white bar indicating  $u_1 = -1$ . The fourth bar (light red) corresponds to  $u_2 = 1$ , with white indicating  $u_2 = -1$  associated with the minimizing control schedule. All schedules are bang-bang. (a)  $x_1(0) = 0.01$ , (b)  $x_1(0) = 0.05$ , (c)  $x_1(0) = 0.1$ , (d)  $x_1(0) = 0.3$ , (e)  $x_1(0) = 0.5$ , (f)  $x_1(0) = 0.7$ , (g)  $x_1(0) = 0.9$ , (h)  $x_1(0) = 0.95$ , and (i)  $x_1(0) = 0.99$ .

case, both satisfying the constraint. Also notice that the linear growth rate in (3) is  $(a_{12} - a_{22}) = 1 - 0 = 1$ , so we scale  $T$  the same way in our computations, as  $T = 1$ . We then perform the optimization adaptively over multiple cycles  $nT$  using the end value of cycle  $nT$  as the initial condition to compute the optimal schedule for the  $(n + 1)$ st cycle. Using this method, we are able to identify absolute maximizers and minimizers as a function of the cycle time  $T$ .

#### A. Optimal control formulation

A standard form for implementing the Pontryagin maximum (minimum) principle with boundary value constraints is

$$\vec{X} = [\vec{x}(t), \vec{C}(t)]^T, \quad \vec{X} \in \mathbb{R}^4, \quad (11)$$

$$\dot{\vec{X}} = \vec{F}(\vec{X}) = [\dot{\vec{x}}, \dot{\vec{C}}(t)]^T, \quad \vec{F} : \mathbb{R}^4 \rightarrow \mathbb{R}^4, \quad (12)$$

where one would like to minimize or maximize a general cost function:

$$\int_0^T L(\vec{x}(t), \vec{u}(t), t) dt + \phi(\vec{x}(T)). \quad (13)$$

In our case, we are only interested in minimizing the terminal value  $\phi(\vec{x}(T)) \equiv x_1(T)$ , which is called a classical Mayer problem (specifically choosing  $L = 0$  and optimizing the terminal value, developed in the context of missile guidance problems where the final distance from the target is optimized) discussed in detail in Sec. 13.3, and together with the Pontryagin principle in Sec. 14.6 of Ref. [36]. The controllers, of course, still play an important role as they enter both the payoff matrix through Eq. (6) as well as the constraint Eq. (7).

Since the method is standard, we will just briefly describe the basic framework and refer readers to Refs. [37–41] for more details on how to implement the approach. Following



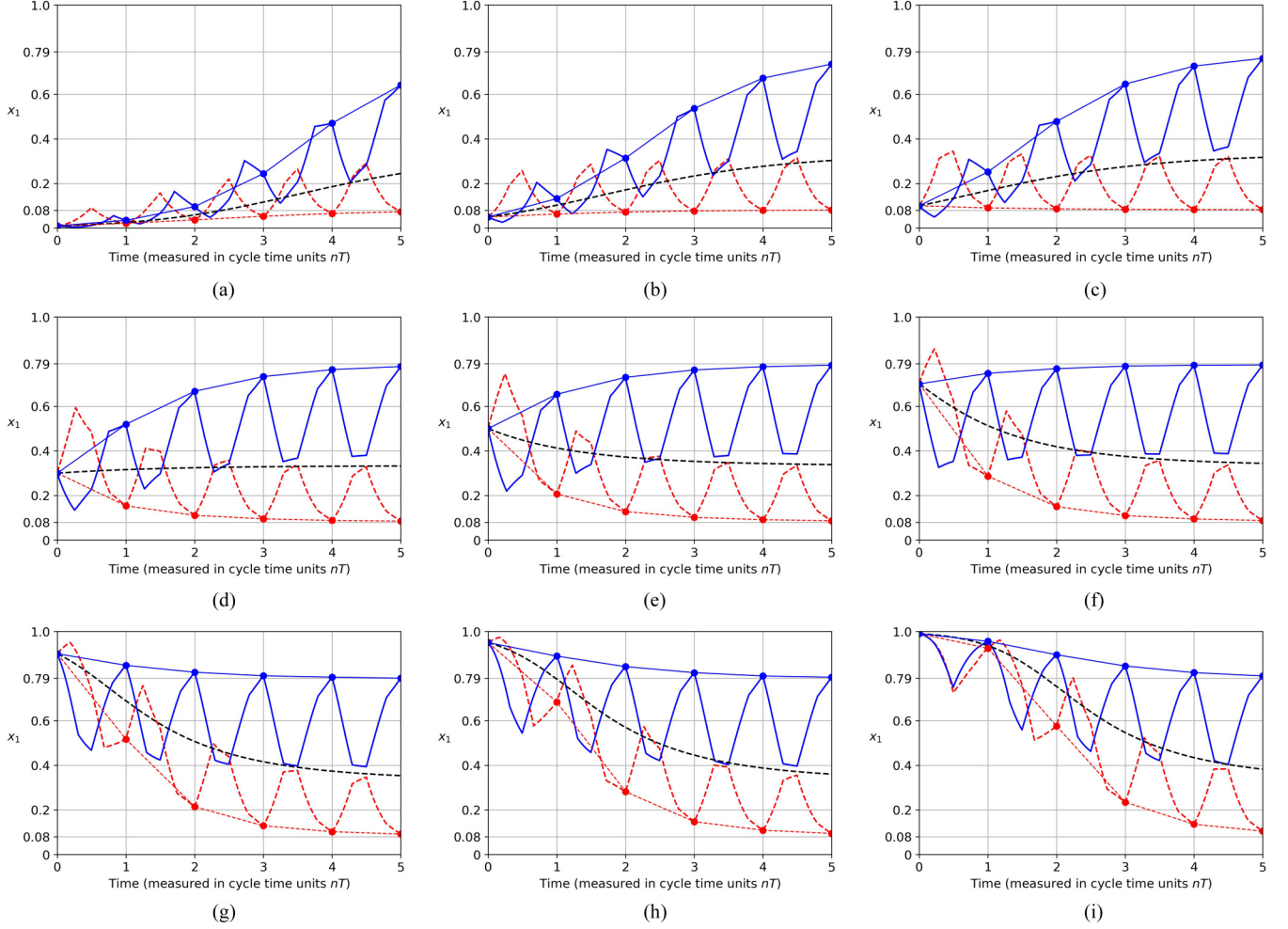


FIG. 4. Maximizing (solid blue) and minimizing (dashed red) trajectories for nine initial conditions over  $n = 5$  cycles, with solid (blue)/dashed (red) curves joining the end values after each cycle. The adaptive schedule for the  $(n + 1)$ st cycle is calculated based on the endpoint of the  $n$ th cycle. The dashed black curve shows the uncontrolled Hawk-Dove trajectory. (a)  $x_1(0) = 0.01$ , (b)  $x_1(0) = 0.05$ , (c)  $x_1(0) = 0.1$ , (d)  $x_1(0) = 0.3$ , (e)  $x_1(0) = 0.5$ , (f)  $x_1(0) = 0.7$ , (g)  $x_1(0) = 0.9$ , (h)  $x_1(0) = 0.95$ , and (i)  $x_1(0) = 0.99$ .

Ref. [40] in particular (see p. 62, Theorem 4.2.1), we construct the control theory Hamiltonian,

$$H(\vec{x}(t), \vec{C}(t), \vec{\lambda}, \vec{u}(t)) = \vec{\lambda}^T \vec{F}(\vec{x}), \quad (14)$$

where  $\vec{\lambda} = [\lambda_1, \lambda_2, \mu_1, \mu_2]^T$  are the costate functions (i.e., momenta) associated with  $\vec{x}$  and  $\vec{C}$ , respectively. Assuming that  $\vec{u}^*(t)$  is the optimal control for this problem, with corresponding trajectory  $\vec{x}^*(t)$ ,  $\vec{C}^*(t)$ , the canonical equations satisfy

$$\dot{x}_i^*(t) = \frac{\partial H}{\partial \lambda_i^*}, \quad (15)$$

$$\dot{C}_i^*(t) = \frac{\partial H}{\partial \mu_i^*}, \quad (16)$$

$$\dot{\lambda}_i^*(t) = -\frac{\partial H}{\partial x_i^*}, \quad (17)$$

$$\dot{\mu}_i^*(t) = -\frac{\partial H}{\partial C_i^*}, \quad (18)$$

where  $i = (1, 2)$ . The corresponding boundary conditions are

$$\vec{x}^*(0) = \vec{x}_0, \quad (19)$$

$$\vec{C}^*(0) = 0, \quad \vec{C}^*(T) = \vec{C}_T^*, \quad (20)$$

$$\lambda_i^*(T) = \frac{\partial \varphi(\vec{x}(T))}{\partial x_i^*(T)}. \quad (21)$$

Then, at any point in time, the optimal control  $\vec{u}^*(t)$  will minimize the control theory Hamiltonian:

$$\vec{u}^*(t) = \arg \min_{\vec{u}(t)} H(\vec{x}^*(t), \vec{C}^*(t), \vec{\lambda}^*(t), \vec{u}(t)). \quad (22)$$

The optimization problem becomes a two-point boundary value problem [using (19)–(21)] with unknowns  $[\lambda_2^*(0), x_2^*(T)]$  whose solution gives rise to the optimal trajectory  $\vec{x}^*(t)$  [from (15)] and the corresponding control  $\vec{u}^*(t)$  that produces it [37–40]. We solve this problem by standard numerical shooting-type methods [40]. The result is that the optimal controllers follow a bang-bang schedule, taking on

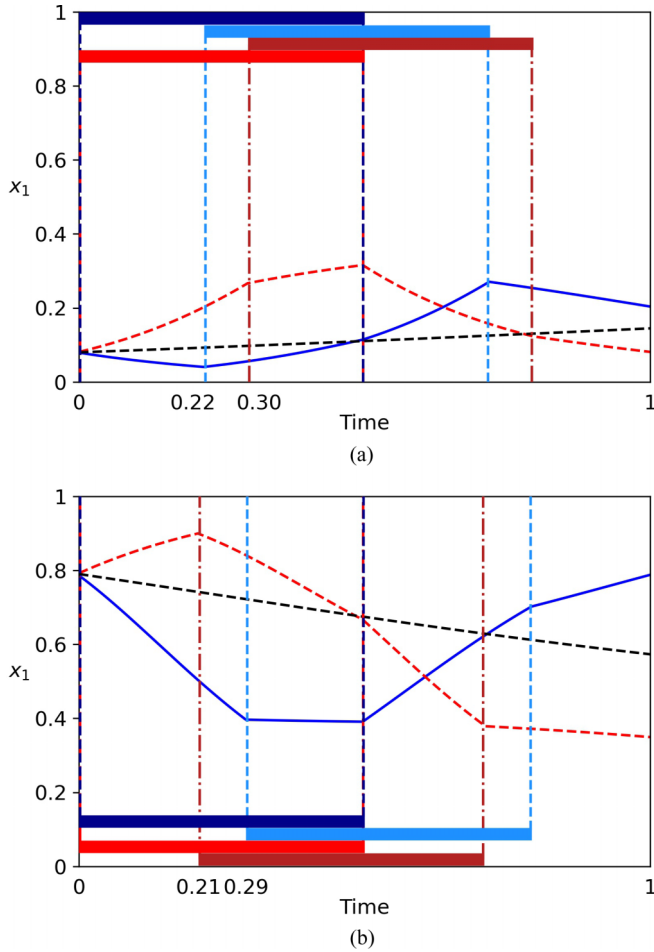


FIG. 5. Maximizing (solid blue) and minimizing (dashed red) trajectories for the two special initial conditions  $x_1(0) = 0.08, 0.79$ . For the larger initial condition, the maximizing schedule (top two bars in blue) produces an absolute maximum  $x_1^{\max}$  for  $T = 1$ . For the smaller initial condition, the minimizing schedule (bottom two bars in red) produces an absolute minimum  $x_1^{\min}$  for  $T = 1$ . The dashed (black) curve shows the uncontrolled Hawk-Dove trajectory. The top bar (dark blue) shows  $u_1 = 1$ , and the white bar shows  $u_1 = -1$ . The second bar (light blue) shows  $u_2 = 1$ , and the white bar shows  $u_2 = -1$  associated with the maximizing control schedule. The third (dark red) bar shows  $u_1 = 1$ , and the white bar shows  $u_1 = -1$ . The fourth (light red) bar shows  $u_2 = 1$ , and the white bar shows  $u_2 = -1$  associated with the minimizing control schedule. (a)  $x_1(0) = 0.08$ , and (b)  $x_1(0) = 0.79$ .

only the extreme values  $+1$  or  $-1$ , and not values throughout the interval  $[-1, 1]$ .

### III. RESULTS

In this section we show the results of solving the adaptive optimal control method to minimize and maximize aggression at time  $T = 1$ , and then further at the end of multiple cycles  $t = nT$ . Figures 3(a)–3(i) show the maximizing (blue) and minimizing (red) trajectories for nine initial conditions. The corresponding bang-bang schedules that produce these trajectories are also shown in each case. It is straightforward to prove that the optimal schedules must be bang-bang since the

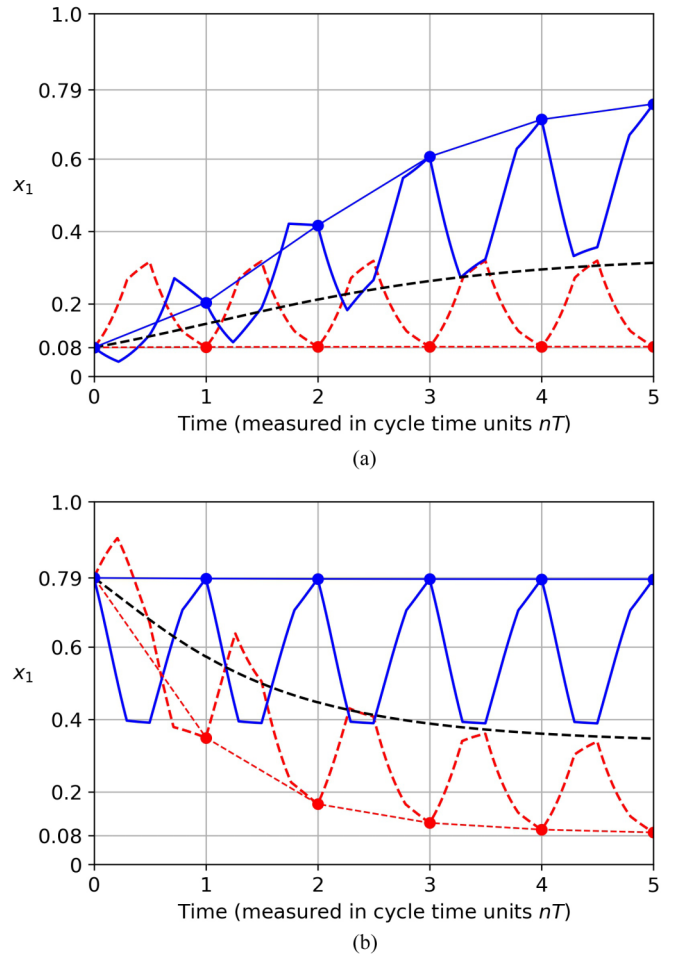


FIG. 6. Maximizing (solid blue) and minimizing (dashed red) trajectories for the two special initial conditions  $x_1(0) = 0.08, 0.79$  over  $n = 5$  cycles. Notice that the minimizing trajectory (dashed red) shown in (a) exactly repeats for each cycle [ $x_1(0) = x_1(T)$ ] since the schedule is an absolute minimizer, while the maximizing trajectory (solid blue) shown in (b) exactly repeats for each cycle [ $x_1(0) = x_1(T)$ ] since the schedule is an absolute maximizer. The dashed (black) curve shows the uncontrolled Hawk-Dove trajectory. (a)  $x_1(0) = 0.08$ . The dashed (red) horizontal line indicates an absolute minimizer at  $x_1^{\min}$ . (b)  $x_1(0) = 0.079$ . The solid (blue) horizontal line indicates an absolute maximizer at  $x_1^{\max}$ .

controllers are linear in the governing equations. In each case, we show the uncontrolled (dashed curve) Hawk-Dove trajectory, which ends in between the maximizer and minimizer as expected.

Figure 4 shows the maximizing (blue) and minimizing trajectories over  $n = 5$  cycles. We obtain these adaptively, using the endpoint from the  $n$ th cycle to compute the optimal schedule for the following  $(n + 1)$ st cycle. Two special initial conditions are shown in Fig. 5. For  $x_1(0) = 0.08$ , the minimizing (red) trajectory shown in Fig. 5(a) ends at  $x_1(1) = 0.08$ , hence is periodic. This value (and corresponding schedule) corresponds to an absolute minimizer  $x_1^{\min}$  for aggression  $x_1$ . By contrast, for  $x_1(0) = 0.79$  shown in Fig. 5(b), the maximizing (blue) trajectory ends at  $x_1(1) = 0.79$ , hence is periodic. This value (and the corresponding schedule) corresponds to an absolute maximizer  $x_1^{\max}$  for aggression. These two

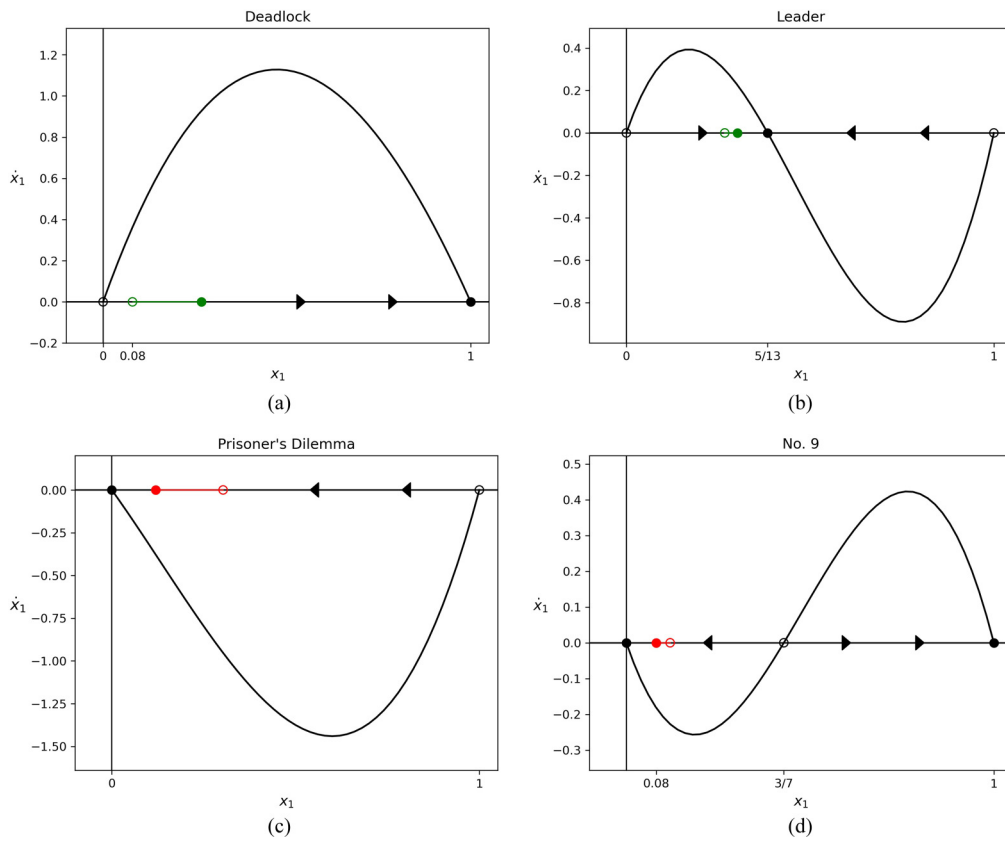


FIG. 7. Minimizing sequence of four games (Deadlock; Leader; Prisoner’s Dilemma; Game 9) associated with initial condition  $x_1(0) = 0.08$  that produces the absolute minimizer. Open dots show the starting point, and the solid dot shows the ending point.

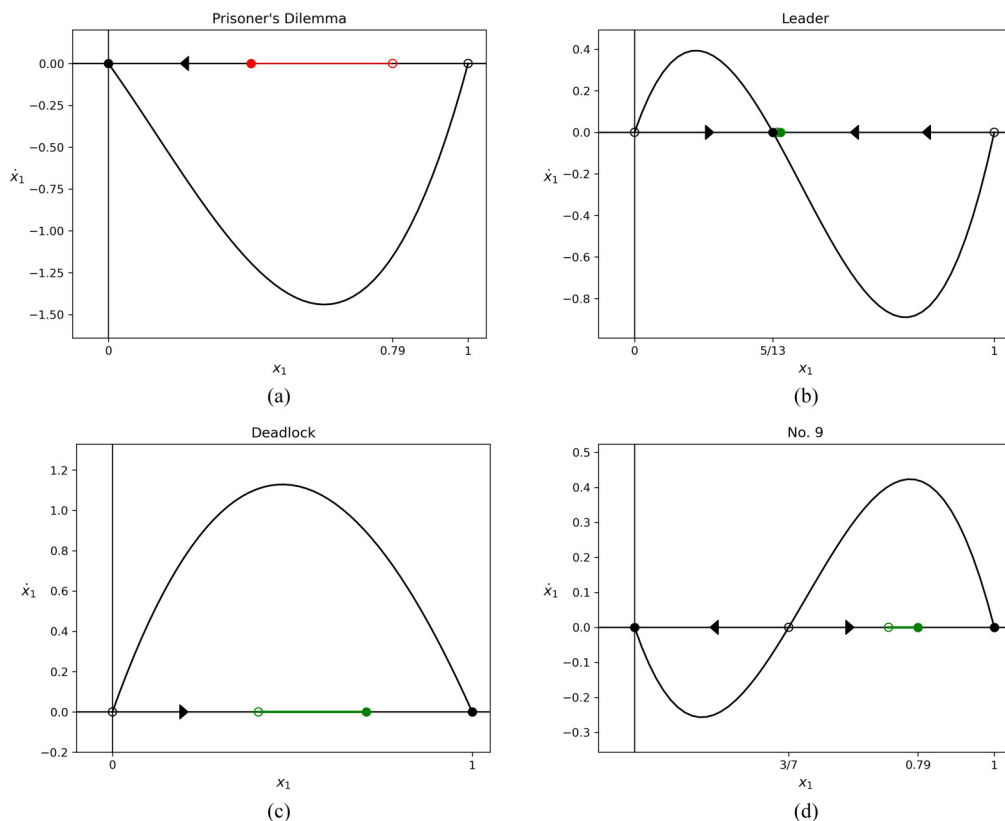


FIG. 8. Maximizing sequence of four games (Prisoner’s Dilemma; Leader; Deadlock; Game 9) associated with initial condition  $x_1(0) = 0.79$  that produces the absolute maximizer. Open dots show the starting point, and the solid dot shows the ending point.

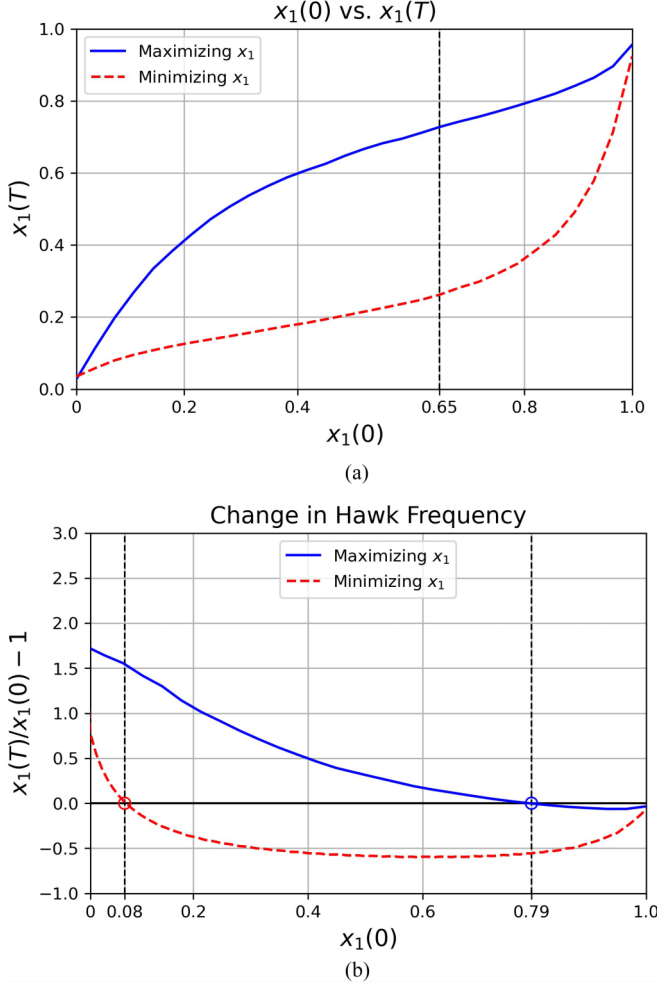


FIG. 9. (a) Hawk initial condition  $x_1(0)$  vs Hawk frequency at final time  $x_1(T)$  for maximizing (solid blue) and minimizing (dashed red) schedules. The vertical dashed line at  $x_1(0) = 0.65$  marks the maximum difference between the minimizer and the maximizer. (b) Change in Hawk frequency as a function of initial condition. Points above the line  $x_1(T)/x_1(0) - 1 = 0$  represent an increase over time and points below this line represent a decrease over time. The two intersection points  $x_1(0) = 0.08$  and  $x_1(0) = 0.79$  mark the absolute minimizer and maximizer initial conditions for  $T = 1$ .

special initial conditions are shown in Fig. 6 over  $n = 5$  cycles, confirming the periodicity of the minimizing trajectory (red) in Fig. 6(a) and the periodicity of the maximizing (blue) trajectory in Fig. 6(b). The sequence of games that the system cycles through to achieve the minimizing sequence is shown in Fig. 7, while the maximizing sequence is shown in Fig. 8. These are obtained from Eq. (3) and the following four equations:

- (1)  $u_1 = 1, u_2 = 1: \dot{x}_1 = x_1(1 - x_1)(5 - 13x_1)$ ;
- (2)  $u_1 = 1, u_2 = -1: \dot{x}_1 = x_1(1 - x_1)(-3 - 5x_1)$ ;
- (3)  $u_1 = -1, u_2 = 1: \dot{x}_1 = x_1(1 - x_1)(5 - x_1)$ ;
- (4)  $u_1 = -1, u_2 = -1: \dot{x}_1 = x_1(1 - x_1)(-3 + 7x_1)$ .

In Fig. 9 we show the minimizing values and maximizing values of  $x_1(T)$  vs  $x_1(0)$  through the full range  $0 \leq x_1(0) \leq 1$ . Notice that at the endpoints, the two values converge (since for the linear system, the schedule does

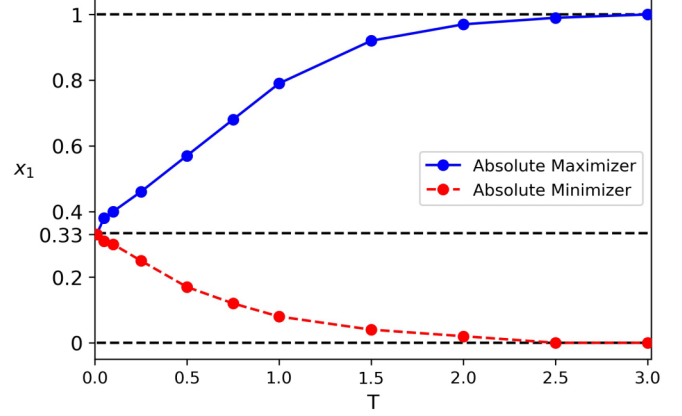


FIG. 10.  $x_1^{\max}$  (solid blue) and  $x_1^{\min}$  (dashed red) as a function of cycle time  $T$ . The dashed horizontal line at  $x_1 = 1/3$  is the ESS for the uncontrolled Hawk-Dove system where the two curves meet as  $T \rightarrow 0$ .

not matter, only the total  $\vec{C}_T$ ). Figure 9(b) shows the ratio  $x_1(T)/x_1(0) - 1$  (percentage increase or decrease) versus the initial condition  $x_1(0)$  through the full range  $0 \leq x_1(0) \leq 1$ . When the maximizing (blue) curve crosses  $x_1(T)/x_1(0) - 1 = 0$ , [i.e.,  $x_1(0) = x_1(T)$ ] an absolute maximizer is achieved [for  $x_1(0) = x_1(T) = x_1^{\max} = 0.79$ ], while when the minimizing (red) curve crosses  $x_1(T)/x_1(0) - 1 = 0$ , an absolute minimizer is achieved [for  $x_1(0) = x_1(T) = x_1^{\min} = 0.08$ ]. In Fig. 10 we show how  $x_1^{\max}$  and  $x_1^{\min}$  depend on the cycle time  $T$ . Interestingly, as  $T \rightarrow 0$ ,  $x_1^{\max}, x_1^{\min} \rightarrow \frac{(a_{12} - a_{22})}{(a_{12} + a_{21}) - (a_{11} + a_{22})} = 1/3$  which is the ESS for the uncontrolled Hawk-Dove system. For  $T \geq \frac{(a_{12} + a_{21}) - (a_{11} + a_{22})}{(a_{12} - a_{22})} = 3$ ,  $x_1^{\max} \rightarrow 1$  and  $x_1^{\min} \rightarrow 0$  showing that for large enough cycle times we can drive either of the subpopulations to extinction or to fixation.

#### IV. DISCUSSION

Our goal in this paper is to lay out a mathematical framework for determining optimal dynamic incentive schedules (time-dependent payoff schedules) that maximize/minimize certain behaviors in an evolutionary game theory setting using the  $2 \times 2$  replicator dynamical system with a Hawk-Dove payoff matrix as our baseline. By changing the payoff entries in a time-dependent manner, subject to constraints, we are altering the payoff-reward structure of the Hawk-Dove interaction as the populations evolve, which is equivalent to selecting a sequence of  $2 \times 2$  evolutionary games in such a way that an optimum is achieved after a fixed passage of time. The determination of these schedules requires a balance between the timescale on which the payoffs change, and the timescale of the underlying replicator dynamical system in such a way that the Pontryagin maximum/minimum principle is satisfied.

As mentioned earlier, there are many settings in which dynamic payoffs can be used to achieve a certain outcome (developing chemotherapeutic schedules that manage chemoresistance, antibiotic scheduling to avoid and even reverse antibiotic resistance in microbial populations, or the introduction of economic incentive packages to guide behavior). One of the more compelling potential applications of the methods developed in this paper in the context of social



dilemmas, is to frame people's attitudes towards vaccination acceptance as a social contract [42–44] and to devise dynamic incentive methods to encourage vaccination acceptance as well as to explore their theoretical limitations. Our method uses the Pontryagin maximum/minimum principle along with the  $2 \times 2$  replicator dynamical system, with constraints, to determine schedules over one cycle time  $T$ , then we extend the results adaptively over multiple cycles  $nT$ . We show this leads to the identification of an absolute maximizer and minimizer  $(x_1^{\max}, x_1^{\min})$  for the aggressor population, both of which are

functions of the cycle time  $T$ . We believe the framework laid out in the paper can be extended to  $N \times N$  replicator systems, as well as discrete (stochastic) models for the interaction of a finite number of participants using a frequency-dependent Moran process.

#### ACKNOWLEDGMENT

We gratefully acknowledge support from the Army Research Office MURI Award No. W911NF1910269.

- 
- [1] J. M. Smith, *On Evolution* (Edinburgh University Press, Edinburgh, 1972).
- [2] J. M. Smith and G. Price, *Nature (London)* **246**, 15 (1973).
- [3] J. M. Smith, *J. Theor. Biol.* **47**, 209 (1974).
- [4] J. M. Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).
- [5] M. A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life* (Harvard University Press, Cambridge, MA, 2006).
- [6] S. Brams, *Game Theory and Politics* (Dover, New York, 2004).
- [7] R. Cressman and Y. Tao, *Proc. Natl. Acad. Sci. USA* **111**, 10810 (2014).
- [8] P. K. Newton and Y. Ma, *Phys. Rev. E* **99**, 022404 (2019).
- [9] Y. Ma and P. K. Newton, *Phys. Rev. E* **103**, 032408 (2021).
- [10] T. Börgers and R. Sarin, *J. Econ. Theory* **77**, 1 (1997).
- [11] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, UK, 1998).
- [12] J. S. Weitz, C. Eksin, K. Paarporn, S. P. Brown, and W. C. Ratcliff, *Proc. Natl. Acad. Sci. USA* **113**, E7518 (2016).
- [13] A. Tilman, J. Plotkin, and E. Akçay, *Nat. Commun.* **11**, 915 (2020).
- [14] D. Friedman, *Econometrica* **59**, 637 (1991).
- [15] L. Hurwicz and S. Reiter, *Designing Economic Mechanisms* (Cambridge University Press, Cambridge, UK, 2006).
- [16] S. Higgins, E. Klemperer, and S. Coleman, *Prev. Med.* **145**, 106421 (2021).
- [17] P. Yeh, M. Hegrennes, A. Aiden, and R. Kishony, *Nat. Rev. Microbiol.* **7**, 460 (2009).
- [18] R. E. Lenski, *Int. Microbiol.* **1**(4), 265 (1998).
- [19] J. West, Z. Hasnain, J. Mason, and P. K. Newton, *Converg. Sci. Phys. Oncol.* **2**, 035002 (2016).
- [20] J. West, L. You, J. Zhang, R. Gatenby, J. Brown, P. K. Newton, and A. Anderson, *Cancer Res.* **80**, 1578 (2020).
- [21] M. Gluzman, J. Scott, and A. Vladimirovsky, *Proc. R. Soc. B* **287**, 20192454 (2020).
- [22] B. San Goh, G. Leitmann, and T. L. Vincent, *Math. Biosci.* **19**, 263 (1974).
- [23] S. Bewick, R. Yang, and M. Zhang, in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, New York, 2009), pp. 6026–6029.
- [24] H. R. Joshi, *Optim. Control Appl. Methods* **23**, 199 (2002).
- [25] R. B. Martin, *Automatica* **28**, 1113 (1992).
- [26] G. W. Swan, *Math. Biosci.* **101**, 237 (1990).
- [27] A. J. Coldman and J. Murray, *Math. Biosci.* **168**, 187 (2000).
- [28] P. K. Newton and Y. Ma, *Phys. Rev. E* **103**, 012304 (2021).
- [29] A. Kaznatcheev, in *Complex Adaptive Systems: Resilience, Robustness, and Evolvability*, Technical Report No. FS-10-03 (AAAI Press, Menlo Park, CA, 2010), p. 71.
- [30] B. Hölldobler and E. O. Wilson, *The Ants* (Harvard University Press, Cambridge, MA, 1990).
- [31] D. H. Kistner, in *Social Insects*, edited by H. R. Hermann, Vol. 3 (Academic Press, New York, 1982), pp. 1–244.
- [32] J. Parker, *Myrmecol. News* **22**, 65 (2016).
- [33] M. Maruyama and J. Parker, *Curr. Biol.* **27**, 920 (2017).
- [34] Z. Wang, S. Kokubo, M. Jusup, and J. Tanimoto, *Phys. Life Rev.* **14**, 1 (2015).
- [35] H. Ito and J. Tanimoto, *R. Soc. Open Sci.* **5**, 181085 (2018).
- [36] J. Hsu and A. Meyer, *Modern Control Principles and Applications* (McGraw-Hill, New York, 1968).
- [37] L. S. Pontryagin, *Mathematical Theory of Optimal Processes* (CRC Press, Boca Raton, FL, 1987).
- [38] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory* (Wiley, New York, 1967).
- [39] I. Ross, *A Primer on Pontryagin's Principle in Optimal Control*, 2nd ed. (Collegiate Press, Boston, 2015).
- [40] K. L. Teo, C. Goh, and K. Wong, *A Unified Computational Approach to Optimal Control Problems* (Academic Press, New York, 1991).
- [41] P. Newton and Y. Ma, *Am. J. Phys.* **89**, 134 (2021).
- [42] C. Bauch and D. Earn, *Proc. Natl. Acad. Sci. USA* **101**, 13391 (2004).
- [43] L. Korn, R. Bohm, N. Meier, and C. Betsch, *Proc. Natl. Acad. Sci. USA* **117**, 14890 (2020).
- [44] M. R. Arefin, K. M. Ariful Kabir, M. Jusup, H. Ito, and J. Tanimoto, *Sci. Rep.* **10**, 16092 (2020).