

Optimal sequence for Parrondo games

Luis Dinis*

*Grupo Interdisciplinar de Sistemas Complejos (GISC) and Departamento de Física Atómica, Molecular y Nuclear,
Universidad Complutense de Madrid, Ciudad Universitaria E-28040 Madrid, Spain*

(Received 14 November 2007; published 26 February 2008)

An algorithm based on backward induction is devised in order to compute the optimal sequence of games to be played in Parrondo games. The algorithm can be used to find the optimal sequence for any finite number of turns or in the steady state, showing that ABABB... is the sequence with the highest steady state average gain. The algorithm can also be generalized to find the optimal adaptive strategy in a multiplayer version of the games, where a finite number of players may choose, at every turn, the game the whole ensemble should play.

DOI: [10.1103/PhysRevE.77.021124](https://doi.org/10.1103/PhysRevE.77.021124)

PACS number(s): 05.40.-a, 02.30.Yy

I. INTRODUCTION

Rectification of thermal fluctuations has become a major topic in nonequilibrium statistical physics. Ajdari and Prost [1] discovered in 1993 a Brownian ratchet mechanism, afterwards named by Astumian and Bier the *flashing ratchet* [2]. In 1996, Parrondo [3] showed that this rectification mechanism also works when spatial and time degrees of freedom of the Brownian particle are discrete, as in the chance games thereafter known as Parrondo games [4]: two separate losing games that can be combined following a random or periodic strategy resulting in a winning game.

The games have received attention in several disciplines, ranging from quantum game theory [5,6], nonlinear or chaotic dynamics [7,8], economics [9,10], and biology [11,12]. However, the question on how to combine the games to get the highest increase in capital was still open. Sequences up to period 12 have been studied using symbolic manipulators [13], and the periodic sequence ABABB (or any of its permutations) has come up as the best in the sense that it provides the highest returns in the stationary state. In this paper we show that ABABB... is indeed the best sequence by applying Bellman's optimality criterion [14] and backward induction.

Recently, various multiplayer versions of the games have been proposed, giving rise to counterintuitive phenomena resembling those observed in game, control, and optimization theories or economics. For instance, it turns out that greedy algorithms or strategies may lead to suboptimal or even losing solutions [15–17].

However, it is worth noting that, contrary to what happens in many of the models used in economy or game theory, the behavior of Parrondo games is of a purely stochastic nature, therefore making them a good system to help understand the role of fluctuations and optimization in those systems where stochastic dynamics is relevant. As an example, it has also been shown that a related phenomenon may occur in a feedback controlled collective flashing ratchet [18,19]. Also in this context, the problem of finding an optimal protocol or strategy in a system where fluctuations have a major role has received attention lately in the field of finite-time thermodynamics and fluctuation theorems [20].

Finally, one can think of the problem of finding the best sequence in an alternative way. Imagine an infinite number of independent players who play against a casino with the only restriction that all of them must play the same game (A or B) at every turn t . That is, the decision to play A or B at t is taken collectively. If some information about the state of the system is known at t (we will later see which is the information needed), an optimal way of choosing A or B can be found so that the average capital is the maximum possible. This is an *adaptive strategy* in the sense that the choice taken can be adapted to the current state of the system; the interested reader may find a formal proof that backward induction can be used to calculate the best adaptive strategy explicitly in Ref. [21]. It then turns out that this optimal adaptive strategy makes the players use the sequence ABABB... in the long run. The average capital of an infinite ensemble of independent players is related, due to the law of large numbers, to the average of one player playing the same exact sequence of games. Hence this will allow us to state that ABABB... is the best periodic sequence for the Parrondo games.

We thus provide an example that an open-loop control problem (a control problem without information about the system) can be solved as a closed-loop optimization problem over an infinite collection of identical systems in which the information about their state may be used, an example that may be relevant for stochastic control theory.

The paper is organized as follows: we begin by briefly reviewing game rules and evolution equations in Sec. II. In Sec. III we state the problem and in Secs. IV–VI, we show how to find the best possible sequence to play in the long run for the original Parrondo games. In fact, the solution is more general and consists of finding the best sequence of games for any finite number of turns. The algorithm can be easily generalized to find the best way of choosing games for an arbitrary number of players (and number of turns to play), i.e., the best adaptive strategy for any number of players and any number of turns; this is done in Sec. VII. Section VIII is devoted to the application of the algorithm to the primary Parrondo paradox. Concluding remarks can be found in Sec. IX.

II. GAMES

Parrondo games can be stated as two simple coin tossing games, A and B. Game A is played with a coin slightly bi-

*ldinis@fis.ucm.es

ased so that the probability to win is less than one-half, that is $p_A=1/2-\epsilon$, with ϵ a small positive number. Let $X(t)$ be the capital of the player in turn t . The average capital $\langle X(t) \rangle$ evolves with the number of turns as

$$\langle X(t+1) \rangle = \langle X(t) \rangle + 2p_A - 1, \quad (1)$$

and therefore $\langle X(t) \rangle$ decreases with the number of turns. In this sense, we will call the game A a *losing* game. Analogously, a *winning* game will be one in which $\langle X(t) \rangle$ increases with t .

Game B is played with two biased coins, say the “good” and the “bad” coins. If the capital of the player is a multiple of 3, she must play the bad coin, which has a probability of winning $p_b=1/10-\epsilon$. Otherwise she tosses the good coin and wins with probability $p_g=3/4-\epsilon$. It can be shown [22] that these rules make B also a losing game in the long run. The paradox arises when alternating game A and B either in a random or periodic fashion, as this yields an average capital that increases with t , provided ϵ is small enough.

If game B is played at turn t , the capital of the player changes as

$$\langle X(t+1) \rangle = \langle X(t) \rangle + 2p_{\text{winB}}(t) - 1, \quad (2)$$

where p_{winB} is the probability to win in game B, which depends on the capital of the player in turn t . More precisely, it only depends on the probability $\pi_0(t)$ that the player has a capital multiple of three in the t th turn. With this definition, game B rules imply

$$p_{\text{winB}}(t) = \pi_0(t)p_b + (1 - \pi_0(t))p_g. \quad (3)$$

To compute $\pi_0(t)$, one can define $\pi_1(t)$ and $\pi_2(t)$ as the probabilities that the capital is a multiple of 3 plus 1 or plus 2, respectively, and $\boldsymbol{\pi}(t) \equiv (\pi_0(t), \pi_1(t), \pi_2(t))'$. Alternatively, $\langle X(t) \rangle$ can be interpreted as the average over the population of an infinite ensemble of independent players and $\boldsymbol{\pi}(t)$ as a fraction of players instead of probabilities. In either case, the following evolution equation applies:

$$\boldsymbol{\pi}(t+1) = \Pi_B \boldsymbol{\pi}(t), \quad \text{with} \quad \Pi_B = \begin{pmatrix} 0 & 1-p_g & p_g \\ p_b & 0 & 1-p_g \\ 1-p_b & p_g & 0 \end{pmatrix}. \quad (4)$$

Finally, game A can be expressed in the same terms and its evolution equation is

$$\boldsymbol{\pi}(t+1) = \Pi_A \boldsymbol{\pi}(t), \quad \text{where} \quad \Pi_A = \begin{pmatrix} 0 & 1-p_A & p_A \\ p_A & 0 & 1-p_A \\ 1-p_A & p_A & 0 \end{pmatrix}. \quad (5)$$

Due to normalization of probabilities, $\pi_0 + \pi_1 + \pi_2 = 1$, so the system state is fully determined by (π_0, π_1) . Hence the region accessible to the system can be represented as a rectangular triangle of side 1 in (π_0, π_1) space.

III. PROBLEM

The expected gain $g(\boldsymbol{\pi}(t))$ when playing game A or B in turn t is defined as

$$g(\boldsymbol{\pi}_0(t)) \equiv \langle X(t+1) \rangle - \langle X(t) \rangle, \quad (6)$$

and may have two different expressions,

$$g(\boldsymbol{\pi}_0(t)) = \begin{cases} g^A & \text{if A is played at } t, \\ g^B & \text{if B is played at } t \end{cases} \quad (7)$$

with

$$g^A \equiv 2p_A - 1, \quad (8)$$

$$g^B(\boldsymbol{\pi}_0(t)) \equiv 2[\pi_0(t)p_b + (1 - \pi_0(t))p_g] - 1, \quad (9)$$

as shown in the previous section. The total gain after T turns of the games is

$$G_T = \sum_{t=1}^T g(t) \quad (10)$$

and it is the target function we aim to maximize.

Let α_t be a parameter that may have values A or B to mark the game to be played at turn t , and $(\alpha_1, \alpha_2, \dots, \alpha_n)$ a sequence of decisions to play A or B. The problem we are faced with can be stated in the following way.

“Find the sequence of decisions $(\alpha_1, \alpha_2, \dots, \alpha_n)$ so that G_T attains its maximum value, with the restriction that $\boldsymbol{\pi}(t)$ evolves as

$$\boldsymbol{\pi}(t+1) = \Pi_A \boldsymbol{\pi}(t), \quad \text{if } \alpha_t = A \quad (11)$$

or

$$\boldsymbol{\pi}(t+1) = \Pi_B \boldsymbol{\pi}(t), \quad \text{if } \alpha_t = B, \quad (12)$$

provided $\boldsymbol{\pi}(t)$ is known.”

IV. FORMAL SOLUTION

Since any decision α_t affects the next, the best way to approach this problem is by proceeding backwards. The last decision will not affect any other, so it makes sense to start with that one. Once we know how we should proceed when we arrive at the last step, we can use that information to try and find out what is the best thing to do in the last but one turn of the games and so on.

Let us call $\hat{G}_n(\boldsymbol{\pi})$ the maximum possible value of the expected gain when there still are n turns left [26]. At this stage, we could choose to play game A or B. If we do the former, the state of the system changes to $\Pi_A \boldsymbol{\pi}$ and the gain obtained in that step is g^A . The highest expected gain attainable by choosing A is then

$$g^A + \hat{G}_{n-1}(\Pi_A \boldsymbol{\pi}), \quad (13)$$

because $\hat{G}_{n-1}(\Pi_A \boldsymbol{\pi})$ is by definition the maximum gain that can be attained provided there are $n-1$ turns left and the system is in state $\Pi_A \boldsymbol{\pi}$. This can be stated more formally using Bellman’s optimality criterion which assures that “an

optimal sequence has the property that, whatever the initial state and decision may be, the remaining decisions constitute an optimal sequence with respect to the state resulting from the first decision.” [14]

If on the other hand we choose to play B when there are n turns left, the maximum expected gain is

$$g^B(\pi_0) + \hat{G}_{n-1}(\Pi_B \pi). \quad (14)$$

Those are the only two possibilities and therefore

$$\hat{G}_n(\pi) = \max\{g^A + \hat{G}_{n-1}(\Pi_A \pi), g^B(\pi_0) + \hat{G}_{n-1}(\Pi_B \pi)\}. \quad (15)$$

It is now clear that information about the state π of the system is needed in order to maximize the gain, as stated in the Introduction.

Given the state π , the optimal decision at turn $t=T-n+1$ (n to the end) is to play A if the maximum corresponds to the first term in expression (15) and B otherwise, so the optimal decision α_{T-n+1} is A in the first case and B in the second. Thus each point in the state space can be related either to game A or B, creating a map that can be used at turn $T-n+1$ to decide which game to play in order to have the highest expected gain. In fact, due to linearity of the expressions involved, the state space is always divided in two connected regions, one for game A and the other for game B (see Fig. 2 for some examples).

Finally, when there is just one turn left, the best choice is to play game A if $g^A > g^B(\pi_0)$ and B otherwise [27]. Hence

$$\hat{G}_1(\pi) = \max\{g^A, g^B(\pi_0)\}, \quad (16)$$

which completes the induction algorithm.

V. NUMERICAL SOLUTION

The procedure explained in the preceding section can be readily turned into a recursive numerical algorithm to compute the maximum expected gain, given a starting state of the system $\pi(1)$. Though the algorithm is simple to program, it requires as many operations as a brute force approach (that is, systematic evaluation of the 2^T possible sequences) due to recursion, and what is even worse, does not provide any information about the solution for any other initial condition $\pi(1)$. To tackle this, we used a different approach.

(1) Define a grid in space (π_0, π_1) as shown in Fig. 1. Denote each point as (π_0^i, π_1^j) , or simply (i, j) , with $i, j \in \mathcal{Z}$.

(2) Set $n=1$. Evaluate $\hat{G}_1(i, j)$ in every point of the grid.

(3) Through evaluation of $\hat{G}_1(i, j)$, a map associating each point of the grid with the optimal choice (A or B) can be created.

(4) Increase n in one unit. Evaluate \hat{G}_n in every point of the grid. To do that, values of \hat{G}_{n-1} in points $\Pi_A(\pi_0^i, \pi_1^j, 1 - \pi_0^i - \pi_1^j)^t$ and $\Pi_B(\pi_0^i, \pi_1^j, 1 - \pi_0^i - \pi_1^j)^t$ which fall outside the grid would be needed. Approximate \hat{G}_{n-1} in each of those points for the value in the closest point down-left in the grid, whose value is already known.

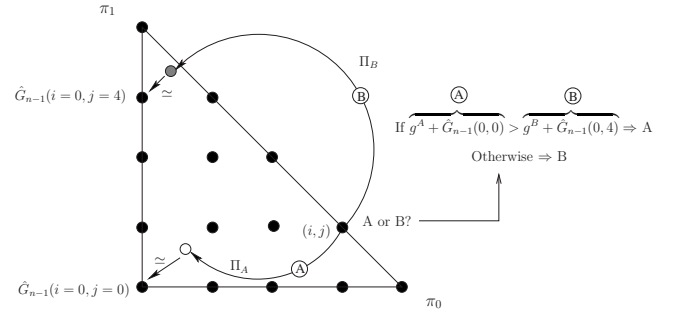


FIG. 1. Evaluation of $\hat{G}_n(\pi_0^i, \pi_1^j)$ and computation of the optimal choice of game for state (π_0^i, π_1^j) with n turns left.

(5) Evaluation of \hat{G}_n provides a map associating each point of the grid to the optimal choice of the game in that point, when there are n turns remaining.

(6) Repeat steps 4 and 5 until $n=T$, the total number of turns.

In step 4, \hat{G}_n is evaluated in every point of the grid for all n , although there is in principle no need to do so since the mappings defined by evolution equations (4) and (5) are contractive [21] and not every point can be reached by evolution after game A or B is played. However, by computing \hat{G}_n at every point, we will obtain the solution not only for the optimal sequence of length T but also of lengths $T-1, T-2, \dots$ in just one run of the algorithm. Moreover, all of these solutions are valid for any initial condition.

Steps 4 and 5 are represented schematically in Fig. 1. It is worth noting that the use of the approximation in step 4 implies the algorithm uses a time proportional to T to compute the solution.

VI. RESULTS: FROM THE MAPS TO THE SEQUENCE

As an example, imagine one player who is going to play the games four times ($T=4$) and whose initial capital is a multiple of 3 ($\pi_0=1, \pi_1=0$). Figure 2 shows the maps calculated using the aforementioned algorithm with a grid of 2000×2000 points and $\epsilon=0$. For the remaining part of the paper, I will take $\epsilon=0$ for simplicity. Choosing a small ϵ may shift the boundaries of the maps slightly. What is the sequence with the highest average gain provided the initial condition is $\pi_0=1$?

The map corresponding to $n=4$ indicates that for $(\pi_0, \pi_1)=(1,0)$ the best choice is game A. Game A takes (π_0, π_1) from its initial value to $\Pi_A(1,0,0)^t=(0,0.5,0.5)^t$. Now there are three turns left and map $n=3$ shows that game B should be chosen, because $(0,0.5)$ falls in the region B, as shown in Fig. 2. Game B takes the system to $(0.5, 0.125)$ in region A of map $n=2$, and finally A takes it to $(0.25, 0.4375)$ in region B of the last map. Therefore sequence ABAB is the one with the highest expected gain. Proceeding in this manner one can compute the optimal sequence for any number of turns T and any initial condition.

For a finite number of turns, some sprinting behavior can be observed that resembles the sprint effect at the beginning

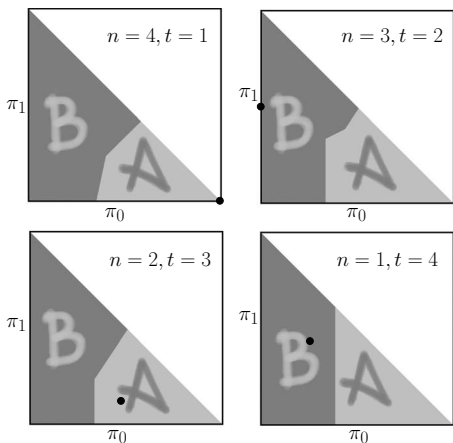


FIG. 2. Maps and evolution for $T=4$, for turns $t=1, 2, 3$, and 4 . $n=T-t+1$ denotes the number of turns left. The black circles mark the evolution of state (π_0, π_1) when playing ABAB, starting from initial condition $(1,0)$.

and the end of the time interval in general optimization problems [15,20]. The optimal sequence usually consists of several repetitions of the ABABB motif flanked by brief pieces of other sequences, as, for example, optimal sequence for 21 turns with initial condition $(1,0)$: AB ABABBABABBA-BABB ABBABB (the spaces have been added to help identify the different parts).

In order to find the sequence with the highest stationary average gain, the behavior of the maps for $n \gg 1$ must be analyzed. A regular pattern soon appears when increasing n , and the maps describing the regions in which to play A or B do not become completely independent of n but change periodically with n , converging to a cycle of the five different maps depicted in Fig. 3. This means that one should follow the prescription contained in these maps in a cyclic way and in order of decreasing n , as indicated by the arrows in Fig. 3.

After a number of runs and irrespective of the initial condition, a steady state is attained where the player ends up

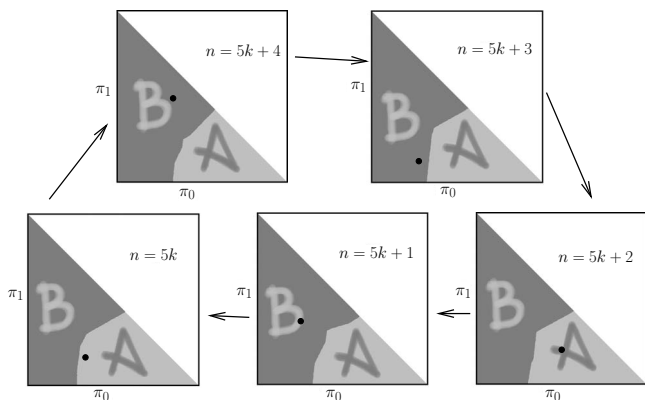


FIG. 3. Stationary state maps for the optimal gain and their corresponding limit cycle. These five maps repeat periodically for $n \gg 1$. To attain steady state, we set $k \gg 1 (\Rightarrow n \gg 1)$ (still many turns left) and $t \gg 1$ (many turns already played). To obtain the optimal sequence in the stationary state, the prescription indicated by the maps should be followed, using each map in the order indicated by the arrows.

playing the sequence ABABB..., and (π_0, π_1, π_2) follows the stationary five point cycle of this periodic sequence. The cycle is also shown in Fig. 3. Hence sequence ABABB... is the one with the highest average gain in the long run.

VII. FINITE NUMBER OF PLAYERS

In this section I turn to a different but related problem. Imagine now there is a single player which is allowed to decide whether to play A or B depending on her actual state 0, 1, or 2, as opposed to our previous problem in which the player chose the sequence of games beforehand and had to keep to it irrespective of the outcome of the games. Which is the way of choosing the games so that the average gain is the highest? The answer is quite trivial, the optimal choice can be expressed as “every time you play, choose A if your state is zero, B otherwise.” As stated in the Introduction, this kind of recipe to choose the game depending on the state is usually termed a strategy. By using the former strategy, the player completely avoids playing with the bad coin and her average gain is maximal.

The problem becomes more interesting if a finite ensemble of N players has to play independently against a casino with the constraint that all of them must play the same game. Since the players will be in different states in general, either game will be a good choice only for some part of them. Which is then the optimal strategy? To answer this question we can use the previous algorithm, with a slight modification.

For a finite number of players, the state of the whole system is given by the fraction of players in any of the three possible states, that is, $\pi(t)=(N_0, N_1, N_2)/N$, where N_i is the number of players with a capital multiple of 3 plus i , and N the total number of them. The main difference is that now evolution of $\pi(t)$ is stochastic, as opposed to the deterministic evolution described in Sec. II [28]. Due to normalization, the state is sufficiently determined by $(N_0(t), N_1(t))$.

The expected gains for one turn g^A and g^B , and also $\hat{G}_1(\pi_0)$, are defined as in the previous sections. However, to compute $\hat{G}_n(N_0/N, N_1/N)$, we must take into account the stochastic evolution of (N_0, N_1) . For example, if we start with a distribution of capitals $(N_0, N_1)=(i, j)$ and our first choice is A, we will in average obtain a gain

$$g^A + \sum_{(k,l)} p_{(i,j) \rightarrow (k,l)}^A \hat{G}_{n-1}(k/N, l/N), \tag{17}$$

where \hat{G}_{n-1} is weighted with the probability $p_{(i,j) \rightarrow (k,l)}^A$ to jump from state (i, j) to (k, l) in game A. An analogous expression can be written also for game B. Once the transition probabilities are computed using the rules of the games, the rest of the algorithm can be applied as described in Sec. V.

The results obtained for a finite number of players agree with the currently available analytical solutions for the steady state (up to $N=3$ in Ref. [23]). Furthermore, this algorithm can be used to compute the solution for more than 100 players in a PC. Finally, it is worth mentioning that the algorithm allowed us to state that it is impossible to devise a

TABLE I. Probabilities to win, stay with the same capital, or lose for games A and B in the primary Parrondo paradox.

| A | Win | Stay | Lose | B | Win | Stay | Lose |
|------|-----|------|------|------|-----|------|------|
| Odd | 1/4 | 1/2 | 1/4 | Odd | 1/9 | 2/3 | 2/9 |
| Even | 1/4 | 1/2 | 1/4 | Even | 4/9 | 1/3 | 2/9 |

strategy that gives the optimal average gain in the steady state irrespective of the number of players. The computation of the optimal ones for 25 and 100 players show that they differ in the game to be chosen at some points.

VIII. APPLICATION TO PRIMARY PARRONDO PARADOX

The algorithm can be also successfully applied to the primary Parrondo paradox (PPP) both to obtain the best *a priori* sequence for a single player or the optimal strategy in the multiplayer PPP version with “the freedom of choosing the common next game for the players.” [24]

PPP consists also of two games, A and B, with probabilities which depend in general on the capital of the player. Two possible states are defined, capital odd or even, and there is a probability to remain with the same capital after playing either of the games. The probability to win, stay with same capital, and lose are given in Table I for both games. The state of the system is sufficiently defined giving P_1 , the probability that the player has an even capital. With these probabilities one can show that $P_1=1/3$ always after playing B, and $P_1=1/2$ after playing A, irrespective of the previous state, a property known as superstability.

Due to the superstability of the PPP, which greatly simplifies the behavior of the system, the optimal sequence can be found analytically by applying Bellman’s optimality criterion. The average gains in one turn are $g_A=0$ and $g_B=1/3P_1-1/9$ for game A and B, respectively. One can easily show that

$$\hat{G}_1(P_1) = \begin{cases} 0 & \text{if } P_1 \leq \frac{1}{3} \text{ (play A)} \\ \frac{1}{3}P_1 - \frac{1}{9} & \text{if } P_1 > \frac{1}{3} \text{ (play B),} \end{cases} \quad (18)$$

$$\hat{G}_2(P_1) = \begin{cases} 0 & \text{if } P_1 < \frac{1}{2} \text{ (play A)} \\ \frac{1}{3}P_1 - \frac{1}{9} & \text{if } P_1 \geq \frac{1}{2} \text{ (play B).} \end{cases} \quad (19)$$

Then, \hat{G}_3 can be computed from \hat{G}_2 yielding

$$\hat{G}_3(P_1) = \begin{cases} \frac{1}{18} & \text{if } P_1 \leq \frac{1}{3} \text{ (A)} \\ \frac{1}{3}P_1 - \frac{1}{9} + \frac{1}{18} & \text{if } P_1 > \frac{1}{3} \text{ (B).} \end{cases} \quad (20)$$

Therefore $\hat{G}_3 = \hat{G}_1 + \frac{1}{18}$ and \hat{G}_4 necessarily has the same structure as \hat{G}_2 , prescribing also choose A if $P_1 < 1/2$ and B

otherwise; the computation of \hat{G}_5 yields the same prescription as \hat{G}_3 , etc. Consequently, the optimal choice will only depend on whether n is odd or even and can be expressed as

$$n \text{ odd} \Rightarrow \begin{cases} \text{play A} & \text{if } P_1 \leq \frac{1}{3} \\ \text{play B} & \text{if } P_1 > \frac{1}{3}, \end{cases} \quad (21)$$

$$n \text{ even} \Rightarrow \begin{cases} \text{play A} & \text{if } P_1 < \frac{1}{2} \\ \text{play B} & \text{if } P_1 \geq \frac{1}{2}. \end{cases} \quad (22)$$

These prescriptions together with the evolution that takes P_1 to $1/3$ when B is played and to $1/2$ when case A is played, yield the sequence ABAB... as the optimal sequence, in agreement with Ref. [24].

Regarding the multiplayer version of the PPP with strategy, the numerical algorithm was modified to take into account the different transition probabilities of the PPP and the fact that there are only two possible states instead of three. The optimal strategies provided by the modified algorithm have been checked proving identical to those reported in Ref. [25], that is up to $N=10$. Moreover, using the algorithm I was able to compute the optimal strategy up to $N=100$.

IX. CONCLUSIONS

Backward induction allowed us to compute the best sequence of games for any number of turns T in time proportional to T . The algorithm shows that ABABB... is the best periodic sequence in the long run in the original Parrondo games. It can also be generalized to multiplayer Parrondo games with strategy showing that the optimal strategy depends on the number of players. The solution provides an example that the optimal *a priori* protocol can be found by looking at a related problem in which the protocol or strategy may take into account the state of an infinitely large ensemble of copies of the original system.

Furthermore, the algorithm is quite general and may be applied to other Markov decision problems, as shown in the section about the primary Parrondo paradox. In fact, this algorithm can also be applied to a discretization of a stochastic system continuous in time and provides an approximation for the optimal control protocol [15]. The type of stochastic control problems relevant to stochastic thermodynamics [20] fit in this scheme.

ACKNOWLEDGMENTS

The author wishes to thank J.M.R. Parrondo for fruitful discussion and generous advice on the manuscript. The work was financially supported by Grants No. BFM2001-0291 from Dirección General de Enseñanza Superior and No. FIS2004-00271 (MCyT), MOSAICO (Consolider MEC), and No. PR27/05-13923 (Santander/Complutense).

- [1] A. Ajdari and J. Prost, *Compt. Rend. Acad. Sci. Paris (II)* **315**, 1635 (1993).
- [2] R. D. Astumian and M. Bier, *Phys. Rev. Lett.* **72**, 1766 (1994).
- [3] J. M. R. Parrondo, in *EEC HC&M Network on Complexity and Chaos* (#ERBCHRX-CT940546) (ISI, Torino, 1996).
- [4] G. P. Harmer and D. Abbott, *Nature (London)* **402**, 864 (1999).
- [5] A. P. Flitney, J. Ng, and D. Abbott, *Physica A* **314**, 35 (2002).
- [6] P. Gawron and J. A. Miszczak, *FNL* **5**, L471 (2005).
- [7] L. Kocarev and Z. Tasev, *Phys. Rev. E* **65**, 046215, (2002).
- [8] P. Arena, S. Fazzino, L. Fortuna, and P. Maniscalco, *Chaos, Solitons Fractals* **17**, 545 (2003).
- [9] M. Boman, S. J. Johansson, and D. Lyback, in *Intelligent Agent Technology*, edited by N. Zhong, J. Liu, S. Ohsuga, and J. Bradshaw (World Scientific, Singapore, 2001), p. 150.
- [10] R. Spurgin and M. Tamarin, *J. Behav. Finance* **6**, 15 (2005).
- [11] D. M. Wolf, V. V. Vazirani, and A. P. Arkin, *J. Theor. Biol.* **234**, 227 (2005).
- [12] P. C. W. Davies, in *The First Steps of Life in the Universe: Proceedings of the Sixth Trieste Conference on Chemical Evolution*, edited by J. Chela-Flores, T. C. Owen, and F. Raulin (Kluwer Academic, Dordrecht, 2001).
- [13] D. Vellerman and S. Wagon, *Math. Educ. Res* **9**, 85 (2000).
- [14] J. M. Layton, *Multivariable Control Theory* (Peter Peregrinus, Stevenage, 1976).
- [15] L. Dinis and J. M. R. Parrondo, *Europhys. Lett.* **63**, 319 (2003).
- [16] L. Dinis and J. M. R. Parrondo, *Physica A* **343**, 701 (2004).
- [17] J. M. R. Parrondo, L. Dinis, E. García-Toraño, and B. Sotillo, *Eur. Phys. J. Spec. Top.* **143**, 39 (2007).
- [18] F. J. Cao, L. Dinis, and J. M. R. Parrondo, *Phys. Rev. Lett.* **93**, 040603 (2004).
- [19] L. Dinis, J. M. R. Parrondo, and F. J. Cao, *Europhys. Lett.* **71**, 536 (2005).
- [20] T. Schmiedl and U. Seifert, *Phys. Rev. Lett.* **98**, 108301 (2007).
- [21] E. Behrends, *Proc. SPIE* **5471**, 510 (2004).
- [22] J. M. R. Parrondo and B. J. de Cisneros, *Rev. Esp. Fisica* **14**, 24 (2000).
- [23] B. Cleuren, Ph.D. thesis, Limburgs Universitair Centrum, 2004 (unpublished).
- [24] B. Cleuren and C. Van den Broeck, *Europhys. Lett.* **67**, 151 (2004).
- [25] B. Cleuren and C. Van den Broeck, *Phys. Rev. E* **70**, 067104 (2004).
- [26] We will later see that this only depends on the number of turns left and the state of the system π in that moment.
- [27] One could also choose to play game $g^A \geq g^B(\pi_0)$. This will affect which we consider the optimal sequence among various possible sequences with identical average gain in some special cases for a finite number of turns and initial conditions. It will nevertheless yield the same optimal steady state sequence since the points in the limit cycle of ABABB... lay far away from the boundary of the regions A and B.
- [28] If the ensemble were composed of an infinite number of players, the evolution could be described by the same Eqs. (4) and (5) and the system would behave exactly as a single player since all the players must choose the same game. The optimal strategy for an infinite ensemble would coincide then with the maps obtained in Sec. VI.