Algorithmic complexity in the minority game

R. Mansilla
Department of Complex Systems, Physical Institute, National University of Mexico, 01000 Mexico D. F., Mexico and Department of Differential Equations, Faculty of Mathematics and Computer Science, University of Havana, Havana, Cuba
(Received 16 June 1999; revised manuscript received 19 November 1999)

In this paper, we present our approach for the study of the complexity of Minority Game using tools from thermodynamics and statistical physics. Previous attempts were based on the behavior of volatility, an observable of the financial markets. Our approach focuses on some properties of the binary stream of outcomes of the game. Physical complexity, a magnitude rooted in Kolmogorov-Chaitin theory, allows us to explain some properties of collective behavior of the agents. Mutual information function, a measure related to Shannon's information entropy, was useful to observe a kind of phase transition when applied to the binary string of the whole history of the game.

PACS number(s): 02.50.Le, 05.40.−a, 05.65.+b, 87.23.Ge

I. INTRODUCTION

In many natural and social systems agents establish among themselves a complex network of interactions. Often in such systems it is the case that successful agents are those which act in ways that are distinct from their competitors.

There have been many attempts to understand the general underlying dynamics of systems in which the agents seek to be different. Some of them have focused on the analysis of a class of simple games, which have come to be known as ''minority games'' [1–3].

The Minority Game was first introduced in the analysis of decision making by agents with bounded rationality, based on the ''El Farol'' bar problem [5], which allows one to address the question of how they react to public information–such as prices changes–and the feedback effects of these reactions.

The setup of the Minority Game is the following: N agents have to choose at each time step whether to go into room 0 or 1. The agents who have chosen the less crowded room (minority room) win, and the others lose. The agents have limited capabilities, and only ''remember'' the last m outcomes of the game. The number m is called memory size or brain size. In order to decide which room to enter, agents use strategies. A strategy is a device by which to process the outcomes of the winning room in the last m time steps, thus prescribing what room to enter next.

The agents randomly pick s strategies at the beginning of the game. After each turn, the agents assign one (virtual) point to each of their strategies, which would have predicted the correct outcome. At each turn of the game, they use whichever is the most successful strategy among the s in his possession, i.e., he chooses the one that has gained most virtual points.

As a dynamical system with many elements under mutual influence, the minority game is thought to underlie much of the phenomena associated with complexity. In order to understand this feature, particular emphasis has been devoted to study the mean square deviation of the number of agents making a given choice σ. In the financial context, this observable is called volatility. Although the great amount of paper devoted to study this quantity (see for example Refs. [3], [4], [6], [7]) A. Cavagna [8] proved that the behavior of the above mentioned measure does not depend of the real history of the game. According to Ref. [8] the only crucial requirement is that all the individuals must possess the same information irrespective of the fact that if this information is true or false. Starting from this, A. Cavagna reproduces the results obtained in Refs. [6] and [7]. N. F. Johnson *et al.* [9] proved that temporal correlation are relevant in the Minority Game in populations with agents having different values of m. They did not propose any alternating measure. D. Challet and Y. C. Zhang [10] also studied this problem giving a measure, which characterize the inefficiency of the system. As far as we know, a different measure taking into account the dynamical (i.e., time evolution) behavior of the model has not been proposed.

Our point of view is that volatility is not a good measure of the behavior of the model. All our claims are based in our belief that the binary string of successive outcome of the game contains all the relevant information about the model. Therefore, if volatility is insensitive to the real history of the game, better measures should be found.

In this paper, we introduce an approach for the study of the complex behavior of Minority Game borrowing tools from thermodynamics and statistical physics [11–14]. We will show that two measures, physical complexity [15] and mutual information function [16,17] strongly depend on brain size m of the agents and throw light on dynamics of the model. They are better measures than volatility.

II. PHYSICAL COMPLEXITY OF MINORITY GAME

The study of complex systems has enjoyed tremendous growth although the concept of complexity itself is vaguely defined. In searching for an adequate measure for complexity of binary string one could expect that the two limiting cases (e.g., regular strings and the random ones) have null complexity, while the ''intermediate'' strings that appears to have information encoded are thought to be complex. Besides, as remarked in Refs. [11] and [12] a classification of a string in absence of an environment within which it is to be interpreted is quite meaningless. In other words the complexity of a string should be determined by analyzing its corre-

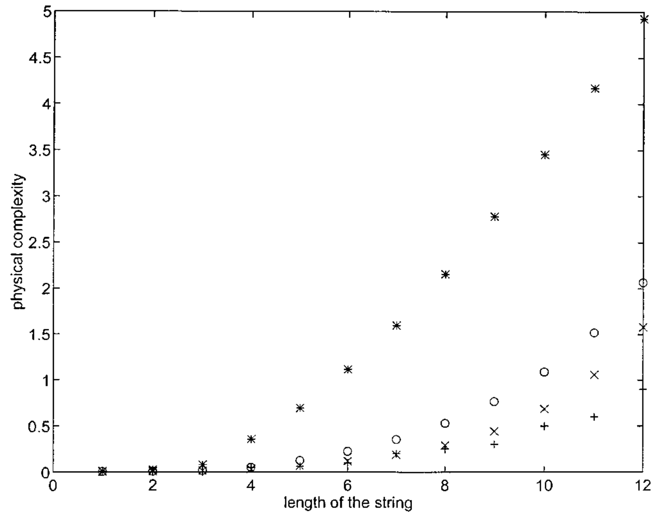

FIG. 1. Graphs of $C(l)$ versus l for different values of m. From above to below the plots correspond to $m=3,4,5$. The lowest plot is the mean value of $C(l)$ over 10 random sequences.

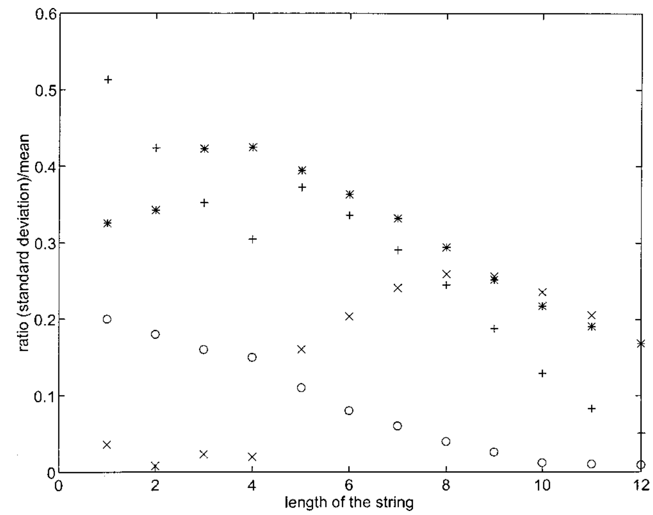

FIG. 3. The ratio (standard deviation)/mean for numerical simulations with different values of brain size and number of strategies per agent. The mean values were calculated over 10 runs with the same parameters. The curves intersect among them for $2 \leqslant l \leqslant 6$, but for $6 < l$ they are ordered. In the interval $6 < l$, from above to below in the graph: $m=3$ and $s=6$; $m=4$ and $s=6$; $m=3$ and $s=3$. The lowest curve corresponds to random sequences.

lation with a physical environment.

Physical complexity (first studied in Refs. [11] and [12]) is defined as the number of binary digits that are meaningful in a string η with respect to the environment. In reference to Minority Game the only physical record one gets is the binary string of the successive outcomes and we consider it as the environment ε. We study the physical complexity of substrings of ε. The comprehension of their complex features has high practical importance. Every agent on the game use only this kind of information to decide his next outcome which has some weight in the formation of future substrings to be used by the agents themselves in their future decision. We briefly review the above mentioned measures devoted to analyze the complexity of binary strings.

The Kolmogorov-Chaitin complexity [13,14] is defined as the length of the shortest program π producing the sequence η when run on universal Turing machine T

$$K(\eta)=\min\{|\pi|:\eta=T(\pi)\}, \tag{1}$$

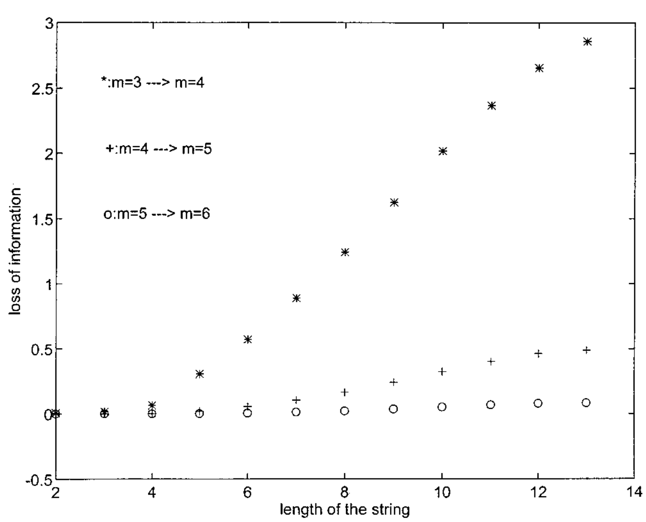

FIG. 2. Values of the loss of information versus l when the memory size m changes from 3 to 4 (upper plot) from 4 to 5 (middle plot) and from 5 to 6 (lower plot).

where $|\pi|$ represent the length of π in bits, $T(\pi)$ the result of running π on Turing machine T and $K(\eta)$ the Kolmogorov-Chaitin complexity of sequence π. In the framework of this theory, a string is said to be regular if $K(\eta) < \eta$. It means that η can be described by a program π with length smaller than η length.

As we have said, the interpretation of a string should be done in the framework of an environment. Hence, let imagine a Turing machine that takes an infinite string ε as input. We can define the conditional complexity $K(\eta/\varepsilon)$ [15] as the length of the smallest program that computes η in a Turing machine having ε as input:

$$K(\eta/\varepsilon)=\min\{|\pi|:\eta=C_T(\pi,\varepsilon)\}. \tag{2}$$

Finally, the physical complexity can be defined as the number of bits that are meaningful in η with respect to ε

$$K(\eta:\varepsilon)=|\eta|-K(\eta/\varepsilon). \tag{3}$$

Notice that $|\eta|$ also represent (see Refs. [12] and [15]) the unconditional complexity of string η, i.e., the value of complexity if the input would be $\varepsilon=\varnothing$. Of course, the measure $K(\eta:\varepsilon)$ as defined in Eq. (3) has few practical application, mainly because it is impossible to know the way in which information about ε is coded in η. However (as shown in Ref. [12] or Ref. [15]), if a statistical ensemble of strings is available to us, then the determination of complexity becomes an exercise in information theory. It can be proved that the average values $C(|\eta|)$ of the physical complexity $K(\eta:\varepsilon)$ taken over an ensemble Σ of strings of length $|\eta|$ can be approximated by

$$C(|\eta|)=\langle K(\eta:\varepsilon)\rangle_{\Sigma}\cong|\eta|-K(\Sigma/\varepsilon), \tag{4}$$

where

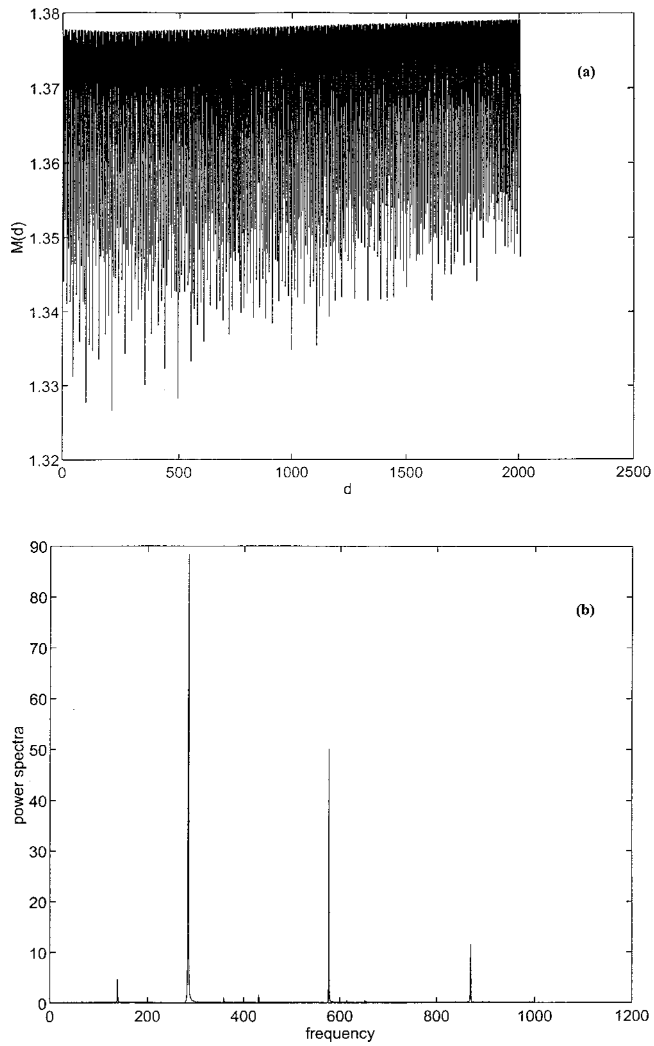

FIG. 4. Mutual information function (a) of stream ε and power spectrum (b) of that function. The value of brain size $m=3$ and number of strategies per agent $s=3$.

$$K(\Sigma/\varepsilon)=-\sum_{n\in\Sigma} p(\eta/\varepsilon)\log_2 p(\eta/\varepsilon) \qquad (5)$$

and the sum is taking over all the strings η in the ensemble Σ. In a population of N strings in environment ε, the quantity $n(\eta)/N$, where $n(s)$ denotes the number of strings equal to η in Σ, approximates $p(\eta/\varepsilon)$ as $N\to\infty$.

Let $\varepsilon=a_1a_2a_3\ldots a_n\ldots$; $a_i\in\{0,1\}$ be the stream of outcomes of the game and l a positive integer $l\geqslant 2$. Let Σ_l the ensemble of sequences of length l built up by a moving window of length l, i.e., if $\eta\in\Sigma_l$ then $\eta=a_ia_{i+1}\ldots a_{i+l-1}$ for some value of i.

We calculate the values of $C(l)$ using this kind of ensemble Σ_l. In Fig. 1 is shown the graph of $C(l)$ for different values of memory size m and for a fixed value of s. Notice that when m increase, the values of $C(l)$ for every fixed l decrease. The explanation of this fact is as follows: Consider the following two ''histories'' of the game

$$h_1=1010\ldots;\quad h_2=1011\ldots.$$

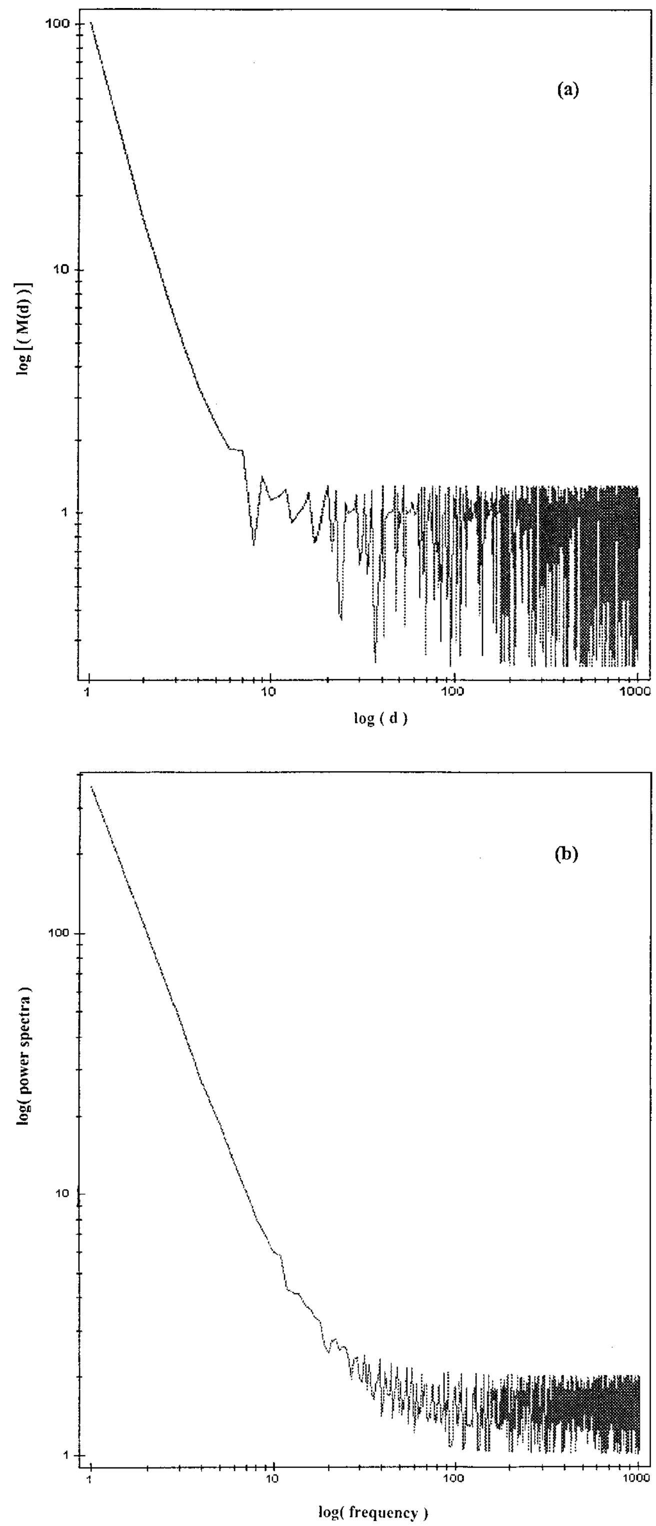

FIG. 5. Mutual information function (a) and power spectrum (b) for a random sequence.

If the brain size is $m=3$, then the agents cannot differentiate the above histories. Hence, they act in both cases as their best performing strategy suggests. If $m=4$ they can differentiate and in general have different responses to histories h_1 and h_2. Therefore, as m increases the perception of the agents become less ''coarsed'' and global response more unpredictable.

Then, there is an increase of entropy as m grows. More

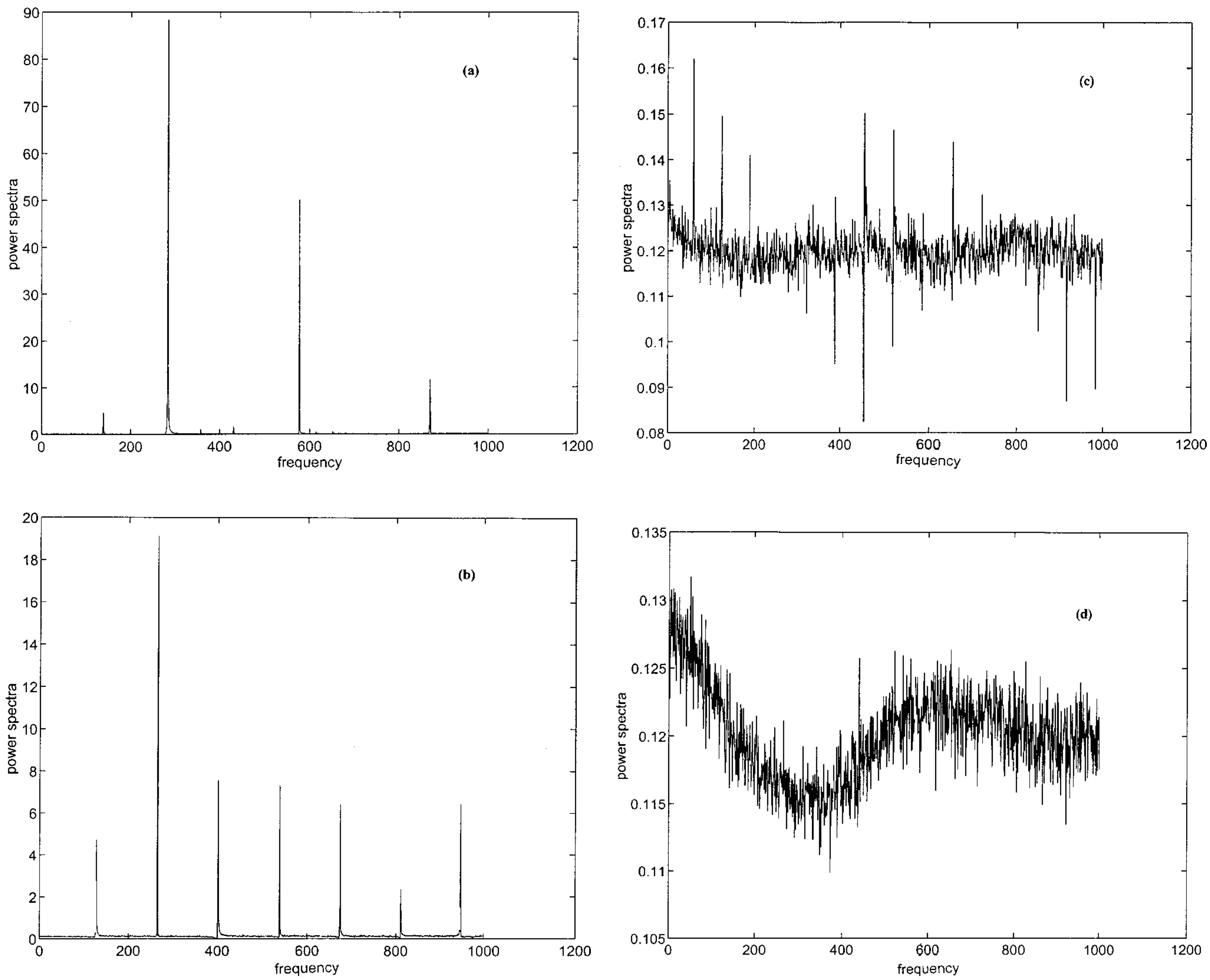

FIG. 6. Power spectra of mutual information function for several values of m. From (a) to (d) $m=3,4,5,6$. As the brain size increase, more and more frequencies enter to the signal. When m changes from 3 to 4 seems to appear a kind of phase transition. The number of strategies per agents in all simulations is $s=3$.

precisely, for every value of l the corresponding value of $C(l)$ decreases as m increases. In Fig. 2 is shown three curves concerning to that loss of information. Notice that as m increases the curves are more flat. It means that the speed of entropy growth decreases with m.

Besides, the calculated values of physical complexity are more ''stable'' as the length of the strings increase. In Fig. 3 are shown the ratio (standard deviation/mean) for several $C(l)$ curves for different values of the memory size m and number of strategies s. Notice that as the length l increase this ratio decrease, indicating that the standard deviation is a smaller fraction of the mean when the values of l grow. Interestingly, the lowest curve corresponds to physical complexity of random sequences. It confirms our above claim that physical complexity tends to be null in random sequences.

The above discussion show that the stream of binary data ε encodes relevant information about the game. The longer a substring of ε is, the larger the average number of binary digits that are meaningful in it. There is also a loss of information as m increases. The larger m, the smaller the number of binary digits that are meaningful in a substring of given length.

III. MUTUAL INFORMATION FUNCTION OF MINORITY GAME

In the last section, we show how the brain size of agents affects the degree of randomness of the stream of successive outcomes of the game. However, nothing has been said about the correlation of the outcomes along the time. Because the distance between two binary symbols represent the number of time iterations between them, a measure of the degree of correlation between elements in a symbolic string could yield information about time correlation.

The mutual information function is defined as

$$M(d)=\sum_{\alpha,\beta} P_{\alpha\beta}(d)\log_2\left[\frac{P_{\alpha\beta}(d)}{P_\alpha P_\beta}\right], \tag{6}$$

where $P_{\alpha\beta}(d)$ is the probability of having a symbol α followed d sites away by a symbol β and P_α the density of the symbol α. It can be proved [16] that mutual information function is a very sensitive measure of correlation.

Fourier spectra (see, e.g., Ref. [20]) is widely used in time series analysis, because the visual representation in the frequency domain can more easily reveal patterns that are harder to discern in the primary data, for example, intricate periodical behavior. We use here Fourier transform of mutual information function to detect some periodical features of that function when applied to the stream of outcomes of the game. From now on, we call power spectra of mutual information function to the product of Fourier transform of that function by its complex conjugate

$$\hat{S}(k)=\theta\left|\sum_{d=1}^{L} M(d)e^{-i2\pi(k/L)d}\right|^2, \tag{7}$$

where θ is a constant related with the sample frequency and L is the number of data available for $M(d)$, see Ref. [20] for details.

The most important feature of mutual information function of the string ε is his remarkable persistence of correlation at some large distances and his periodical behavior. In Fig. 4 are shown that function and his power spectra for a simulation of the game. There are abrupt changes in the correlation of symbols along the ε string for certain distances. Notice that a binary symbol belonging to ε could have high correlation with a second symbol far away of it and at the same time a very low correlation with some close neighbors of that second binary digit. More than that, this behavior is periodic as we could conclude from the power spectra of $M(d)$. Notice that this loss of correlation reflected in $M(d)$ is translated in loss of predictability of the agents of the game. Finally, we also stress that the above property strongly depends on the real history ε of the game. Mutual information function and his power spectrum for a random sequence are shown in Fig. 5. An extensive study of this fact can be found in Refs. [16] and [17]. A more structured symbolic sequence as those found in DNA molecule possess mutual information function very different of that shown in Fig. 4. See Refs. [18] and [19] for the details.

Another interesting fact is the behavior of the power spectra as the memory size m increase. In Fig. 6 this function is shown for several values of memory size. Notice that as m increases more and more frequencies enter to the spectra. It means that more often will appear abrupt changes in the mutual information function. When m changes from 3 to 4 the power spectrum become continous and behaves as $1/f^a$, where $\alpha=\alpha(m)$ reflecting an absence of characteristic time scale, typical in the behavior of financial index before crashes. As we understand it, this kind of phase transition has nothing to do with that reported in Ref. [3] or [4], because that transition strongly depends on the assumption that volatility does not depend on the history of the game (see for example the deduction of Eq. (5) in Ref. [4]). The above results are in perfect agreement with the result of the last section, because as we show there, the increase of m tends to decrease the predictability of the agents as the behavior of the averaged physical complexity shows.

IV. CONCLUSIONS

In this paper, we introduce a new approach for the study of the complex behavior of Minority Game using the tools of thermodynamics and statistical physics. We have shown that physical complexity, a magnitude rooted in Kolmogorov-Chaitin theory yields relevant information about the increase of entropy in the ensembles Σ_l when l increase. We show that as m increases, the average number of bits that are meaningful in a substring of length l of ε decreases. The mutual information function, a magnitude, which has his origin in Shannon's information entropy, throws light on the dynamics of the global time evolution of the model. The way in which the average loss of information impinges on the whole series of outcomes is yielding sudden changes of correlation in the series. As m increase these changes appear more often and for some values of m seems to arise the above mentioned phase transition and the power spectrum becomes continuos.

ACKNOWLEDGMENTS

The author is thankful to P. Miramontes, O. Miramontes, G. Cocho, D. Challet, and K. N. Ilinsky for their helpful comments. This work was supported by CONACYT, Mexico.

[1] D. Challet and Y. C. Zhang, Physica A **246**, 407 (1997).
[2] Y.-C. Zhang, Europhys. News **29**, 51 (1998).
[3] R. Savit *et al.*, Phys. Rev. Lett. **82**, 2203 (1999).
[4] D. Challet and M. Marsili, e-print cond-mat/9904071.
[5] W. B. Arthur, Am. Econ. Assoc. Papers Proc. **84**, 406 (1994).
[6] R. Savit *et al.*, e-print adap-org/9712006.
[7] N. F. Johnson *et al.*, Physica A **256**, 230 (1998).
[8] A. Cavagna, Phys. Rev. E **59**, R3787 (1999).
[9] N. F. Johnson *et al.*, e-print cond-mat/9903164.
[10] D. Challet and Y.-C. Zhang, Physica A **256**, 514 (1998).
[11] W. H. Zurek, Nature (London) **341**, 119 (1984).
[12] W. H. Zurek, Phys. Rev. A **40**, 4731 (1989).
[13] A. N. Kolmogorov, Rus. Math. Sur. **38**, 29–40 (1983).
[14] G. J. Chaitin, J. ACM **13**, 547 (1966).
[15] C. Adami and N. J. Carf, e-print adap-org/9605002.
[16] W. Li, J. Stat. Phys. **60**, 823 (1990).
[17] W. Li, Int. J. Bifurcation Chaos Appl. Sci. Eng. **2**, 137 (1992).
[18] R. Mansilla and R. Mateo-Reig, Int. J. Bifurcation Chaos Appl. Sci. Eng. **5**, 1235 (1995).
[19] R. Mansilla and G. Cocho, e-print chao-dyn/9809001.
[20] D. B. Percival and A. T. Walden, *Spectral Analysis for Physics Applications* (Cambridge University Press, Cambridge, England, 1993).