PHYSICAL REVIEW E, VOLUME 65, 055201(R)

Information transfer in chaos-based communication

Murilo S. Baptista
Instituto de Física, Universidade de São Paulo, Caixa Postal 66318, CEP 05315-970 São Paulo, São Paulo, Brazil

Luis López
Escuela de Ciencias Experimentales e Ingeniería, Universidad Rey Juan Carlos, Caixa Postal 28933 Móstoles, Madrid, Spain
(Received 11 May 2001; revised manuscript received 1 February 2002; published 2 May 2002)

This paper presents fundamentals of a theory to characterize chaos-based communication. We describe the amount of information a dynamical system is able to transmit, the dynamical channel capacity, which takes into account the information that a dynamical system generates.

DOI: 10.1103/PhysRevE.65.055201 PACS number(s): 05.45.Vx, 05.45.Gg, 01.20.+x

Among the works on chaos-based communication, a few [1–5], have suggested the use of chaos to enhance communication efficiency: security, transmission rate, and problem solving. The idea of using the capability of chaotic systems to code some source of information was experimentally demonstrated in Ref. [1]. In that work, the authors made arbitrarily small time-dependent perturbations manipulating the chaotic system, in order to generate a desired encoding signal which codes some information to be transmitted.

We show that the dynamical channel capacity, which measures the amount of information transfered in chaos-based communication, assumes for very small Gaussian noise variance values multiples of the H_{KS} [6].

Recent works [2,4], showed, among other things, that capability dynamical systems have to filter noise. That is, given a chaotic orbit corrupted by additive Gaussian noise, there are ways to transform this noisy orbit into an orbit much closer to the non-noisy trajectory. We show that the dynamical filter is able to act only on the noisy, preserving the real transmitted trajectory. In addition, for very small noise variances, the dynamical filter succeeds in completely reducing the noise.

Chaos-based communication can be summarized in a three-step process. *Sampling* the input signal, *filtering* the output signal, and finally, *predicting* with the filtered signal. By *sampling* we mean that one has to find an appropriate partition of the phase space with which encoding of the message is possible. For example, one could use a coarse-grained partition as suggested in [4] in order to have the symbolic encoded stream and the message with a similar transition statistic. We do not treat this problem here, but we assume a partition is known by both the transmitter and the receiver. So, from now on, a chaotic wave signal is, in fact, a set of points obtained through a discreting process (a mapping) of the higher-dimensional continuous trajectory, a trajectory which is the wave-signal used to transmit information over a channel. By *filtering* we mean that one has to recover the wave signal after it is transmitted and corrupted by noise and distorted by the strictly band-limited frequency, physical limitations imposed by the channel. In this paper, we only deal with noise filtering because we consider that the problem of eliminating the noise from the output signal is similar to recovering the input signal from a distorted output. Finally, to transmit a large amount of information, we only transmit pieces of the signal that after the *filtering* process enables the receiver to fully recover the trajectory through *prediction*, what leads to recovering of the message.

We assume that our chaotic signal is a one-dimensional Bernoulli shift mapping, the Baker's map F defined as, $x_{n+1}=2x_n+\xi_0$ if $x_n \leqslant 0.5$ or $x_{n+1}=2x_n-1+\xi_0$ if $x_n>0.5$, where ξ_0 is a Gaussian noise perturbation with variance $\eta_0=[10^{-12},0.001]$, and zero mean. The reason for the noisy term is to maintain the orbit of the mapping onto a set very close to the chaotic set. Also, this perturbation is responsible to affect the long-term evolution of the map. With this, we simulate a real control application (as done in [1]) which drives the trajectory such that it encodes arbitrary messages.

We encode binary information using this mapping. For that we partition the phase-space domain $J=[0,1]$, into a generating partition w composed of two partitions, $\omega_1=[0,0.5]$ and $\omega_2=]0.5,1]$. Trajectory points falling in partition ω_0 encodes for the symbol $R_0=$ ''0,'' and points falling in ω_1 encodes for the symbol $R_1=$ ''1,'' The probability with which the orbit stay in the partition ω_k is $\sigma(\omega_k)=0.5$. The message to be transmitted, $X=X_1,X_2,\ldots$, the channel input, is composed by a symmetric discrete binary alphabet $\mathcal{S}$, with components R_0,R_1, appearing in the message with frequencies $p_0=0.5$ and $p_1=0.5$. The choice of the Baker's map is due to the fact that it represents a Bernoulli shift, and therefore, its trajectory encodes any sequence of symbols.

To measure the amount of information transferred using chaotic signals, we introduce the Shannon entropy, H_s [7], the Kolmogorov-Sinai (KS) entropy [6], H_{KS}, and the topological entropy, H_T [8]. To quantify the amount of information of the message, we use the Shannon's entropy, H_s, with the natural logarithm. This is so, because we want to compare the information carried by the message, with the information produced by the dynamical trajectory, which is calculated using the natural logarithm. So, $H_s(\mathcal{S})=\sum_{k=0}^{k=1}p_k\ln(1/p_k)$. One important property of the entropy $H_s(\mathcal{S})$ is that $0\leqslant H_s(\mathcal{S})\leqslant \ln 2$, where the upper limit is reached if and only if $p_k=1/K$ for all k. Note that a typical message (with large enough symbols) encoded by the Baker's map will have a Shannon entropy equal to the Shannon entropy of the message. If the source has an entropy of ln(2), the typical encoded message should have an entropy close to ln(2). If ω is a generation partition then H_{KS} can be represented by $H_{\mathrm{KS}}=\sum_{k=0}^{k=1}\sigma(\omega_k)\ln(1/\sigma(\omega_k))$. The H_{KS} is also

connected to metric characteristics of the dynamical system. It was shown by Ruelle [9] that $H_{\mathrm{KS}} \leqslant \Sigma_{\lambda_i>0}\lambda_i$, where λ_i's are the positive Lyapunov exponents of the dynamical system. Our chaotic system is one dimensional, therefore, it has only one positive Lyapunov exponent $\lambda = \ln(2)$. For this map, $H_{\mathrm{KS}} = \ln(2)$ [9].

In communication with chaos another important entropy is the topological entropy H_T, which is based on the fact that the number $E(P)$ of periodic orbits grows exponentially with the period P, as $E(P) \sim \exp^{H_T P}$. This entropy measures the information the dynamical system can encode with very small manipulations on the trajectory. Such manipulations are required in order to have the message X coded by the system's trajectory. What is important here is that $H_T \geqslant H_{\mathrm{KS}}(\omega_k)$. In order to better understand the relation between H_{KS}, H_T, and H_s, with the amount of noise and the amount of information received, we study a particular case, where the message (input symbol stream) is optimized for the maximum information transfer, i.e., the message X is actually the natural symbol sequence of the nonperturbed Baker's map. This choice is also made in order to simulate the use of vanishingly small controlling perturbations (very small η) what would thus lead to the following equality: $H_T = H_{\mathrm{KS}} = H_s = \ln(2)$, in case ω_K is a generating partition. This equality also means that the Baker's map trajectory can encode as much information as a random independent discrete source of information.

We now consider a noisy channel. For that, an important variable is the signal-to-noise ratio $\zeta = P/\eta$, where P represents the power of the signal, and η the variance of the Gaussian noise with zero mean (which is also the power of the signal). In a noisy channel, the channel capacity C_s, i.e., the average maximal amount of information an independent discrete source is able to transmit *per transmission* is given by Shannon's channel theorem [7], which states that

$$C_s = 0.5 \ln(1+\zeta). \tag{1}$$

To show the limits in transfering information in noisy channels, using chaos, we add to the trajectory x, noise with variance η, creating a noisy trajectory y. Placement of y in the partitions w_k decodes y into the channel output Y. The quantification of how much information is lost when noise is introduced is measured by $H_e = -p_{10}\ln(p_{10}) - p_{01}\ln(p_{01})$ (this is so, for $p_{11} = p_{00}$). The error probability p_{ij} represents the probability with which one sends the symbol "i" and decodes the symbol "j." Once, $\sigma(\omega_1) = \sigma(\omega_2)$, and the noise is Gaussian, $p_{10} \cong p_{01}$. In a noisy channel, the amount of information that reaches the receiver, *per transmission*, is given by the mutual information $I(\eta) = H_s(\mathcal{S}) - H_e$.

The trajectory x has length $l = 19\,000$ iterations. From this trajectory, we construct the noisy trajectory y. For different values of η, varying from 0 up to 1/3 (note that the power of x is 1/3), we show in Fig. 1 the mutual information between channel input and the channel output. We see that for a very small η, the mutual information is $I(\eta) = \ln(2)$.

Now, we use the sensibility to initial conditions to filter the noisy trajectory. Consequently, this filtered trajectory is used to guess (predict) orbit points that were not transmitted.

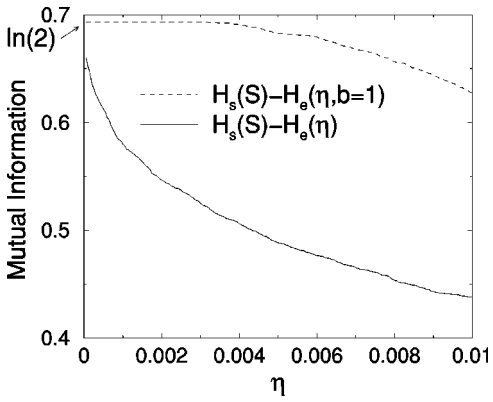

FIG. 1. The solid line represents the mutual information without using any information of the dynamical system, and the dashed line represents the mutual information using the dynamics, filtering the noise trajectory y by applying one backward iteration ($b=1$). η is the noise level.

It is known that two neighboring trajectories, ϵ distant from each other, when iterated forward in time, n iterations, diverge exponentially proportionally to $\epsilon \exp^{\lambda n}$. On the other hand, applying backward iterations, two neighboring trajectories, δ distant from each other, converge exponentially proportionally to $\epsilon \exp^{-\lambda n}$. So, we iterate backward points of the noisy trajectory y obtaining a filtered trajectory z with points closer to points of the noiseless trajectory x. This backward process is not so easy because the Baker's map is invertible, and each backward iteration has two solutions. Therefore, for b backward iterations, 2^b backward trajectories are possible. The task is identifying from all these trajectories, the one that is closer to the y. We proceed as follows. Suppose we want to obtain z, applying b backward iterations. One backward iteration of y_{n+b}, generates two solutions: w^j_{n+b-1}, with $j = 1, \ldots, 2$: $w^1_{n+b-1} = y_{n+b}/2$ and $w^2_{n+b-1} = y_{n+b}/2 + 1$. We calculate the distance between both points w^1_{n+b-1} and w^2_{n+b-1}, with the noisy point y_{n+b-1}. Two backward iteration of y_{n+b}, generates four solutions: w^j_{n+b-2}, with $j = 1, \ldots, 2^2$: $w^1_{n+b-2} = w^1_{n+b-1}/2$ and $w^2_{n+b-2} = w^1_{n+b-1}/2 + 1$, $w^3_{n+b-2} = w^2_{n+b-1}/2$ and $w^4_{n+b-2} = w^2_{n+b-1}/2 + 1$. We calculate the distances between these four solutions with the noisy point y_{n+b-2}. In this way, b backward iteration of y_{n+b}, generates 2^b solutions, w^j_n, with $j = \{1, \ldots, 2^b\}$. We have to define from all these 2^b backward trajectories w, which one is closer to the noisy trajectory y. Let us say that the chosen backward trajectory is w^4_n, w^2_{n+1}. Thus, $z_n = w^4_n$ and $z_{n+1} = w^2_{n+1}$. We expect that the points z_n and z_{n+1} decodes for the same symbol of the noiseless points x_n and x_{n+1}. This filtering process is an improvement of the method used in [2,4].

Doing this filtering process, the mutual information is calculated considering the filtered trajectory w, and not the noisy trajectory y. So, it is appropriate to name the uncertainty of the dynamical channel, by applying b backward iterations, by $H_e(\eta, b, x, z)$. Therefore, the mutual informations is given by $I_e(\eta, b, x, z) = H_s(\mathcal{S}) - H_e(\eta, b, x, z)$. In Fig. 1, we show by dashed line the mutual information of the dynamical channel for $b = 1$. We see that up to a certain

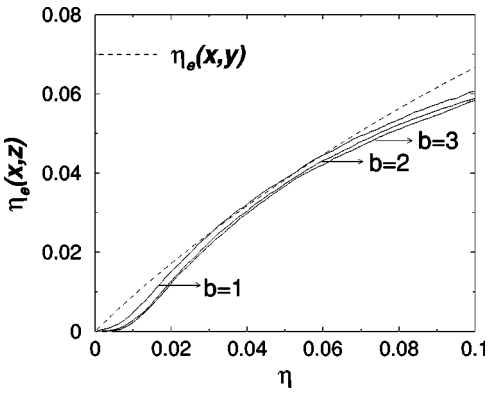

FIG. 2. By the dotted line we show the average distance, $\eta_e(x,y)$, between the noiseless trajectory x and the noisy trajectory y. By the solid lines we show $\eta_e(x,z)$, with the filtered trajectory z obtained for different backward iterations b.

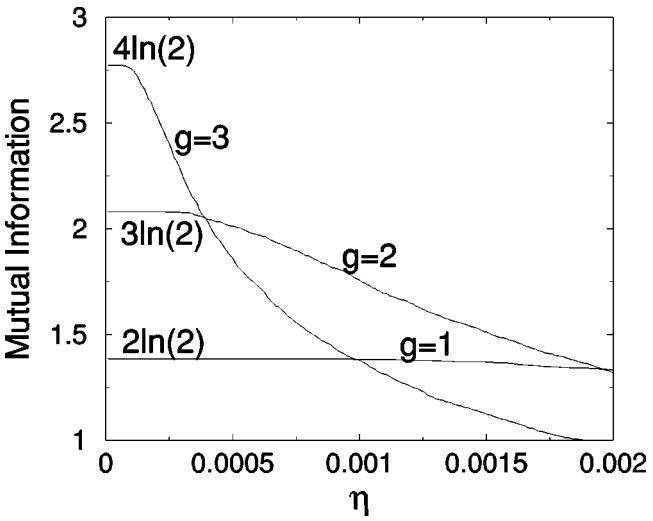

FIG. 3. The mutual information for many configurations of the parameter g that represents the length of the gap in the transmitted trajectory. C_d represents the dynamical channel capacity.

noise level, the mutual information is kept constant. As we increase the value of b, we get an increase in the mutual information. However, the average distance between the trajectory x and the filtered trajectory $z(b)$, defined as $\eta_e(x,z)=[1/(l-1)]\Sigma_{n=1}^{l}(x_n-z_n)^2$ decreases. This distance is in fact the effective noise of the dynamical channel. In Fig. 2, we show by dashed line the average amount of noise that arrives in the receiver, and by solid lines the amount of noise $\eta_e(x,z)$, after the filtering process for different backward iterations b.

After the filtering process, we proceed with the prediction process, to recover nontransmitted trajectory points. Different from random variables, whose elements are independent, dynamical variables are dependent. Therefore, one element contains information of the one that generated it. This property can be explored such that not all the trajectory is transmitted, what results in an increase of the mutual information per transmission. Another use of the dynamical variables in communication is that they offer a natural way to overcome dropouts in the transmission once the missing information would be recovered by looking at the received information. What we do is withdrawing g points out of $g+1$ points. So, we introduce another parameter in the calculation of the mutual information, that is the length of the gap g of the transmitted signal. For example, if $g=1$, we transmit x_n, x_{n+2}, and x_{n+4}, and so on. So, for a gap of length g, we transmit x_n, $x_{n+(g+1)}$, $x_{n+2(g+1)}, \ldots, x_{n+q(g+1)}$, with $q(g+1)+n \leqslant l$. The backward trajectories are calculated considering the received points y_n, $y_{n+(g+1)}$, $y_{n+2(g+1)}$, ..., and $y_{n+q(g+1)}$. From these points, we calculate the filtered trajectory, z_n, $z_{n+(g+1)}$, $z_{n+2(g+1)}$, ..., and $z_{n+q(g+1)}$. Thus, every point of the gap-filtered trajectory, is used to predict the nontransmitted points. So, assuming that $g=1$ and $b=2$, we reconstruct the trajectory by doing $z_{n+1}=F(z_n)$ and $z_{n+3}=F(z_{n+2})$. Note that, if $g \neq 0$, b cannot assume any value, otherwise the reconstruction of the trajectory might create a distorted signal. Therefore, for $g \neq 0$, b should assume the values $k(g+1)]$, for $k=1,2,3,\ldots$. In this paper, we limit our analysis to $b=g+1$. Using the fact that every iteration generates H_{KS} of information, the mutual information for $g>0$, named $I_g(\eta,b,g)$ is calculated by $I_g(\eta,b,g)=I_e(\eta,b)(g+1)$.

When we start using the prediction property of dynamical systems, the mutual information between the channel output, obtained from the filtered and predicted trajectory increases for an interval of the noise variance values. In Fig. 3, we show the mutual information I_g for different values of the parameter g. For an interval of η values, we see that there are no errors in the decoding of the transmitted signal, as we increase the noise variance within a small length interval.

In order to explain this step function in the mutual information, which means that for intervals of η the mutual information is constant, we have to understand the effect of noise in the reconstruction of the transmitted trajectory. Up to some maximum noise level $\eta_m(g)$, there might be no errors in transmission, i.e., $H_e=0$. The reason for the rare occurrence of errors is that we bound the received trajectory to be within the domain of the Baker's map, i.e., [0,1], what has the effect of making the noise to be bound when applied to a point close to the boundaries of the map domain. To be more specific, if the point 1.01 is received, we consider that the point 1.00 was received before the filtering process is performed. Even though there is a small amount of noise (even for the filtered trajectory), that amount is very unlikely to make the receiver decode a wrong message. In general, it should be expected that for very small noise, there are no errors and $I_g=(g+1)H_{\mathrm{KS}}$. For this noise level, the signal-to-noise rate is $\zeta_m(g)=P/\eta_m(g)$.

We define the *dynamical channel capacity* C_d as the maximum amount of mutual information between the channel input X and the filtered channel output Z over all the parameters b and g, and for a given interval of the signal-to-noise rate $\Delta\zeta=[\zeta_m(g),\zeta_m(g+1)]$,

$$C_d(g,\Delta\zeta)=(g+1)H_{\mathrm{KS}}. \tag{2}$$

Doing $C_d=I_g$, we obtain that $\zeta_m(g)=\exp^{(g+3)H_{\mathrm{KS}}-1}$. Using

$\zeta_m(g)$ in Eq. (2) we see that the dynamical channel capacity for an specific dynamical system, characterized by H_{KS}, is a step function which assumes the values given by Eq. (2), for the interval of $[\zeta_m(g),\zeta_m(g+1)]$. In fact, the finding of $\zeta_m(g)$ is the most important information to define the dynamical channel capacity of a dynamical system. With that one learns how much robust to noise the specific dynamical system is. Note that $\zeta_m(g+1)=[\zeta_m(g)+1]\exp^{H_{KS}}-1$, which is approximately given by $\zeta(g+1)=2\zeta(g)+1$. We may also express the dynamical information capacity per seconds, using the proposed sampling. If f_c is giving in hertz, $C_d(snr)=2f_c(g+1)H_{\mathrm{KS}}$ per seconds. One could say that we can have g as great as we want and then, have a communication system with infinite capacity for information transfer. But g should be bounded according to the size of the controlling perturbation ξ. Note that while Eq. (1) shows that the channel capacity decreases with the increasing of the noise amplitude, in the chaos-based communication proposed, noise up to a maximum variance value η_m, does not affect the dynamical channel capacity introduced in Eq. (2). In Fig. 3, we also show that as the noise variance increases, for $\eta > \eta_m$, there is still an interval of values for which the dynamical channel capacity is constant, although smaller than $C_d(g,\Delta\zeta)$.

Increasing of the noise η_0, and the usage of higher g values, does not change the results described by Eq. (2). Our results are also not affected by the slight change in the power of the channel output, due to use of noise η_0 with higher variance.

Thus, the concept of dynamical channel capacity may be a key component for the description of communication processes that, like the biological ones, present this particular type of behavior.

In conclusion, with the introduction of chaos in communication, the channel should be defined by not only the power of the signal, the frequency bandwidth, and the noise level, but also by the Kolmogorov-Sinai entropy of the dynamical system. However, the dynamical channel capacity does not violate the Shannon capacity, which means its value is always smaller than the upper bound imposed by Eq. (1).

This work is partially supported by FAPESP. The first author thanks A. Rodrigues for useful discussions. The authors thank the Max-Planck-Institute Für Physik Komplexer Systeme at Dresden for their hospitality and financial support.

[1] S. Hayes, C. Grebogi, E. Ott, and A. Mark, Phys. Rev. Lett. **73**, 1781 (1994).

[2] E. Rosa, Jr., S. Hayes, and C. Grebogi, Phys. Rev. Lett. **78**, 1247 (1997).

[3] M. Hasler, Int. J. of Bifurcation and Chaos **8**, 647 1998).

[4] M. S. Baptista, E. E. Macau, C. Grebogi, Y.-C. Lai, and E. Rosa, Phys. Rev. E **62**, 4835 (2000).

[5] M. S. Baptista, E. Rosa, Jr., C. Grebogi, Phys. Rev. E **61**, 3590 (2000).

[6] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **119**, 861 (1958).

[7] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (The University of Illinois Press, Illinois 1949).

[8] R. C. Adler, A. C. Konheim, and M. H. McAndrew, Trans. Am. Math. Soc. **114**, 309 (1965).

[9] D. Ruelle, *Chaotic Evolution and Strange Attractors* (Cambridge University Press, New York, 1989).