Complete and flexible replacement of chaotic uncertainty with transmitted information

Matthew B. Kennel*
Institute for Nonlinear Science, University of California, San Diego, La Jolla, California 92093-0402, USA

Shawn D. Pethel
U. S. Army Aviation and Missile Command, AMSAM-RD-WS-ST, Redstone Arsenal, Alabama 35898, USA

(Received 30 August 2004; published 19 October 2005)

Natural chaos can be described as an information source emitting symbolic sequences with positive entropy. We use two algorithmic techniques from data compression in a nonstandard way along with a control scheme to replace the natural uncertainty in chaotic systems with an arbitrary digital message. Unlike previous targeting-based control, the controlled, deterministic, transmission appears statistically identical to natural chaos, with a message modulated on it at the intrinsic Kolmogorov-Sinai information generation rate of the chaotic oscillator. Thus, chaotic communication by targeting need not consume any additional channel capacity beyond that required by the message itself and the message-bearing signal may appear identical to the uncontrolled oscillator. We also demonstrate control and data transmission at the channel capacity of the oscillator, the maximum possible data rate compatible with the grammar.

DOI: 10.1103/PhysRevE.72.046212 PACS number(s): 05.45.Vx, 05.45.Gg

I. INTRODUCTION

Chaotic systems are natural information sources. The sensitivity inherent in these processes creates macroscopic uncertainty by amplifying microscopic fluctuations from the surrounding environment. But unlike pure noise, chaos is highly structured and often low dimensional and thus easily modeled and controlled. Such systems may further be designed to convey useful information and thus function as a communication system, transmitting information on a chaotic carrier. One class of techniques applies infinitesimal perturbations to use the dynamics to target specific orbit sequences that encode a message [1–4]. As these orbits are indistinguishable from solutions of the natural, deterministic chaotic dynamical system, one can use the techniques of *symbolic dynamics* that provide a natural bridge between dynamics and digital communication. Orbits correspond to sequences of symbols from a finite alphabet, $s_1, s_2, \ldots$. Each symbol encodes a state-space region intersected by the trajectory as it passes through a Poincaré section. In the symbolic domain, the distinction between the discrete stochastic sources traditionally considered in information theory and deterministic chaos is nearly erased. Deterministic equations of motion translate to conditional probability laws, e.g., $p(s_i|s_{i-1}, s_{i-2}, \ldots)$, and the evolution of an orbit is considered as a stochastic information source. The topology of the chaotic attractor in continuous space corresponds, in symbolic language, to the list of specific allowed and disallowed symbolic words, and the density of orbits corresponds to the varying relative probabilities of sequences of symbols.

To communicate, the state of a transmitter must be recreated at a receiver to ensure decoding. A theoretical result [5] states that to synchronize to a chaotic transmitter the minimum channel capacity necessary is equal to the Kolmogorov-Sinai (KS) entropy rate (assuming, of course, that the entropy rate of the message R is less than this). That number (e.g., in bits/s) quantifies the rate of *spontaneous* information generation in natural chaos due to intrinsic instability of orbits in an unperturbed attractor.

Our purpose in this work is to explicitly demonstrate the algorithmic technology necessary to produce a channel code for chaotic oscillators with empirically observed symbolic dynamics. This channel code takes streams of white iid bits (considered without a loss of generality to be compressed or encyphered messages with statistics akin to white bits, but not "random" to the intended recipient) to structured streams of symbols (not necessarily in a binary alphabet), which inform the targeting apparatus of the transmitting oscillator about the desired symbolic dynamics of the orbits that the oscillator ought to execute. This code is reversible and the original message may be retrieved. The code construction is general purpose, not constrained to a specific type of oscillator or parameters: this method can drive any chaotic dynamical system provided a control scheme that can execute any dynamically allowable symbolic orbit on a finite alphabet.

Our experimental transmitter is a chaotic electronic circuit [2] whose orbits can be targeted [4] with very small perturbations to follow any desired symbolic sequence allowed by the intrinsic dynamics. We adapt two technologies originally invented for lossless data compression, Markov-model source estimation and arithmetic coding, but use them in a nonstandard configuration to drive the oscillator to follow a specially constructed symbol sequence.

One result is that we show how to transmit arbitrary streams of human-supplied binary information so that the encoded signal behaves identically to the natural chaotic oscillations, and the message is transmitted at the same rate as information that would be generated by natural chaos. Here, we have substituted all the natural chaotic uncertainty, quantified by the KS rate, with a desired information stream with the same Shannon rate and simultaneously maintain the sta-

*Electronic address: mkennel@ucsd.edu

tistics of the natural symbolic dynamics, and thus constructively show how to achieve the theoretical limit [5], using all the information required to synchronize to transmit a useful message as well. This is "chaotic stegeanography" (hiding messages in what appears to be something else), as the transmission appears statistically nearly identical to a free-running chaotic oscillator, though we make no security claims.

Our procedure furthermore allows channel codes with different transition probabilities than the unperturbed (natural) chaos, as long as the symbolic grammar is maintained. In the class of control by infinitesimal perturbations, the capacity of the chaotic oscillator—the maximum possible data rate—is the topological entropy rate, greater than or equal to the KS entropy rate. We recall the explicit solution for the transition probabilities for this and demonstrate transmission at the channel capacity as well.

II. CONSTRAINED CODES

The first task is to design and implement a constrained code for the symbolic dynamics of the chaotic system. In other words, a procedure to transform a sequence of arbitrary message bits—in this case assumed to be independent, equiprobable, binary bits—into a sequence of symbols that matches the symbolic dynamics of a certain dynamical system.

Constrained codes have a significant history, in particular, as channel codes for magnetic recording. The most famous of these are the (d,k) codes, implemented in nearly every commercial magnetic disk drive as a data modulation code (see, e.g., Refs. [6,7] for the mathematical theory surrounding them). For physical engineering reasons, it is desirable to ensure that on the magnetic medium there is at least some number, d, of 0 symbols between every 1 symbol encoded, but, at most, k consecutive zeroes (meaning no magnetic domain boundary) between every occurrence of a 1. The input message bits, which have no constraints, are transformed into a longer sequence that obeys the given constraints. These codes, in symbolic language, correspond to constraining the grammar of the sequences—which are allowed or disallowed—but impose no other targets on the relative probabilities of words in the encoded sequence. Hence, the key criterion is that the *topological* entropy rate of the shift corresponding to the constraint must be equal to or exceed the entropy rate of the transmitted sequences of alphabet A (assumed equiprobable), i.e., $\log A$.

Some data compression algorithms—source codes instead of channel codes—operate by estimating an explicit probabilistic model of the structured input sequence, and, using an "arithmetic coder," then emit a sequence of compressed bits that are nearly white and independent (i.e., maximal entropy), assuming the model accurately models the source. The decompresser uses an arithmetic decoder along with the identical model to reverse the compression. Suppose in this setup that the compressed bits (being nearly white) were replaced by a new bit stream that is also white and independent. The decompresser would then create a structured sequence that is *statistically* equivalent to a sample generated from the the model of the source.

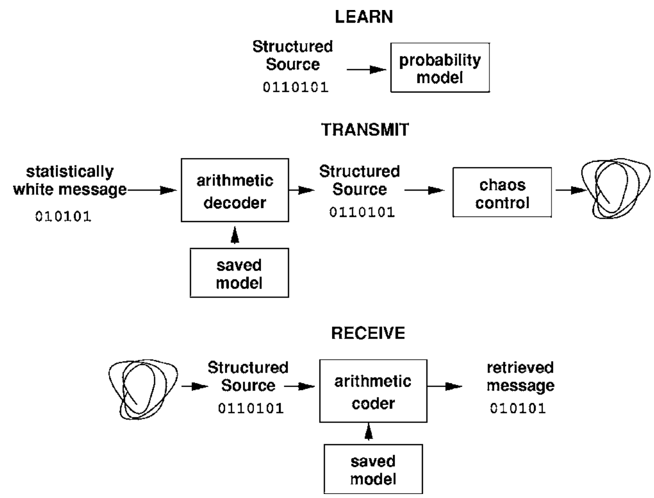

FIG. 1. Schematic diagram of the control scheme.

To communicate through the chaotic process, we reverse the usual position of decoder and encoder, and use an arithmetic decoder as a channel encoder. We estimate a probabilistic model of the source's symbolic dynamics, $P(s_{t+1}|s_t, s_{t-1}, \ldots)$ from an observation of the natural, unperturbed, chaotic oscillator and applying a "context-tree"-based modeling algorithm [8] adapted from source coding methods. The modeling technique is based on universal compression so that the model's metric entropy rate is guaranteed to approach the source's, given enough training data. We assume without loss of generality that the desired message has already been adequately source coded (compressed) or enciphered by conventional means so that it is statistically like random bits, but in truth, was still was formed deterministically. We set up the arithmetic decoder with preknowledge of the model structure and feed this whitened message into it, producing a stream of structured symbols. We control the transmitter to follow this symbolic sequence. By construction, it reflects the natural chaotic structure; hence the transitions requested will be experimentally allowable and hence fully controllable. The resulting pseudochaos can be transmitted over some link, and when its symbolic sequence is recreated (which may require only coarse measurement precision), the message can be retrieved by the reverse of the transmitter, feeding the symbols into a conventional arithmetic coder with the same model.

Figure 1 shows a diagram outlining the steps of the encoding process. First, the symbolic output from an uncontrolled oscillator is the source for learning a predictive probability model, in our case, a context tree. To transmit, this model along with the statistically white message is fed into an arithmetic decoder, producing a stream of symbols with the same probability structure and grammar as the saved model, which was estimated to be very similar to the original source. This symbolic sequence drives the controller of the analog state of the oscillator. Figure 2 shows a "symbologram" [8] of an uncontrolled and a controlled symbol sequence. Consider some point along the sequence. Take the future binary symbols as digits of the binary fraction that forms the y-axis value, and similarly, in reverse time, for the x axis. The figure is a graphical illustration of the fact that

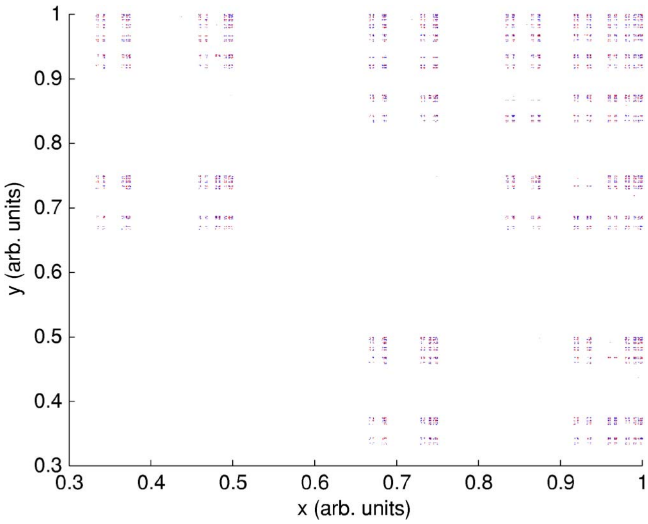

FIG. 2. (Color online) Comparing symbolic sequences of uncontrolled and controlled configurations. The x axis is the arithmetic equivalent of binary symbols going into the past, the y axis of symbols going into the future. Hence, sequences of binary symbols may be conveniently projected into the unit square, conveniently summarizing the symbolic dynamics. The fact the points arising from the controlled (red, color online) and uncontrolled (blue, color online) time series have nearly the same density in symbolic space reflects the success in designing a channel code that transforms random-appearing white message bits to symbolic sequences that closely match the unperturbed oscillator's grammar and metric properties.

statistically the two sequences are essentially indistinguishable: the location and approximate density of points in the symbol plane, like the oscillators in the continuous time representation, are the same. Of course, if the input bits had not been statistically white, this would not be so.

To receive a message, the process is reversed. We have not examined any particular noisy channel model, as this is not a work on communication engineering yet. We assume that the state can be reconstructed sufficiently to produce the same symbolic sequence. Chaotic synchronization is a potential technique that may assist here but it is not essential. The process is reversed, and along with the same saved model, the original message is reconstructed.

The arithmetic coder and decoder software that we employed was based on work by Wheeler [10] that was a reimplementation of the original work in Ref. [11]. In practice, there needs to be a known initialization state for the arithmetic coder and decoder so that all data are properly reconstructed. For our particular example, we chose the string "0101," which was a terminal node in the context tree model as the initial state. In a more realistic example, one might choose to designate one moderately rare symbolic sequence in the structured source as the "synchronization word" that would signal the transmitter and receiver to initialize their coders in a prescribed way. Given ergodicity, a receiver that happens to listen in on a sequence of output from the transmitter starting from an unknown state will eventually encounter a synchronization word and be able to decode the message from that point forward.

III. CONTROLLING THE OSCILLATOR

We show a proof-of-principle experimental demonstration in an electronic circuit. The oscillator and its control system have been described previously in [2,4]. Briefly, the transmitter is a nonlinear oscillator with a fold providing chaos of the Roessler type. The control perturbations are achieved with "dynamic limiting," a certain maximum in voltage is expected for each oscillation. The control is such that sufficient current is applied to prevent the voltage from exceeding this limit. Proper targeting involves judiciously choosing the desired target voltage individually for each oscillation. When the control system is calibrated, as described below, once the oscillator is synchronized to the controller (which happens very rapidly), the desired target will be very close to the actual naturally achieved maximum, and the control force (as measured by the control current applied) is very small, often approaching the noise level of the oscillator and observational electronics. The key to this is that there is a buffer of B (usually 12 to 16) symbols as a symbolic "look-ahead," so that the new symbolic control input at each step is fed in as the least significant symbol, and the controller may be aware of the whole length of symbols to make a control input. As shown in [4], this allows the perturbations to be very small because the system has, in the previous time steps, already been controlled to follow an orbit corresponding to all but one of the symbols.

Training the symbolic control scheme requires only observing a long trajectory of uncontrolled orbits. An appropriate Poincaré section is defined (which does not occur necessarily at the maximum value used for control), and on this plane the return map is practically one dimensional so a generating partition can be defined by dividing at the obvious location, where the derivative is zero. To train the control, one correlates the observed maximum value on the subsequent oscillation with the future sequence of B symbols observed to occur from this point, and finds the appropriate ensemble average. This yields the target limiter value to apply in order to control the next B symbols to the given sequence. To control to a given long sequence, one feeds in the new symbols successively into the least significant bit of a shift register, reads out the B-bit symbolic word, and looks up the target limiter value from the training data. In sum, using the method in [4] without alteration, we control the transmitter to any given symbolic stream as long as that stream is compatible with the grammar of the underlying system. The coding construction we describe here ensures that this is always the case.

Figure 3 shows time-delay plots of uncontrolled and message-transmitting oscillators. Statistical methods [9] do not distinguish them in any way. We have demonstrated by explicit experimental transformation how the dynamical "information" generated by natural chaotic instability is not in any fundamental way different from the usual information considered in Shannon's theory and communication engineering.

IV. TRANSMITTING ON CHAOTIC SADDLES

By altering transition probabilities, we may transmit on various "chaotic saddles" that are contained in the natural

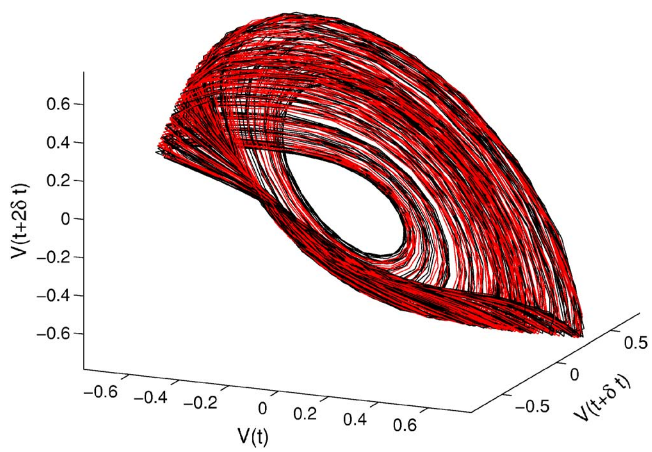

FIG. 3. (Color online) Time delay plot of voltage from transmitting circuit. The dark lines (black) indicate a sample of an uncontrolled orbit, and the light lines (red) indicate a sample from an orbit transmitting a message. They are statistically indistinguishable.

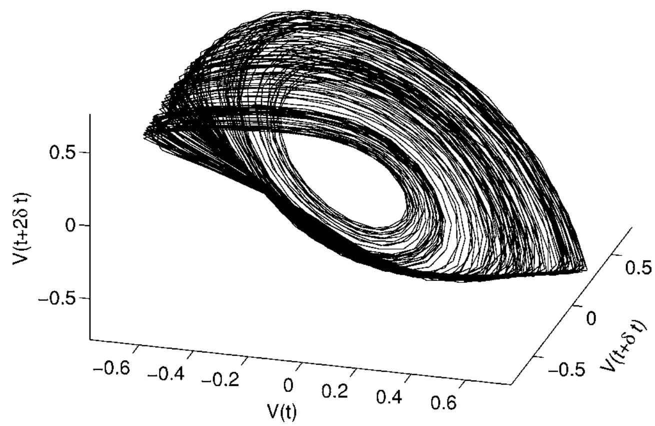

FIG. 4. Time-delay plot of transmitter voltage with a channel-capacity-achieving control signal.

attractor. By this we mean transmitting with the same grammar or a subset of the grammar (i.e., which symbol words are allowed or not), but possibly with different probabilities than the natural chaotic dynamics. These are all orbits of the dynamical system and hence their density an invariant set of the evolution operator. These would be statistically distinguishable [9] from the natural attractor, of course, but this might have some practical advantages. For example, one may prefer certain statistical characteristics for the best communication over some physical channel (see, for instance, [3] and references therein). Or one may be interested in the maximum bit rate that can be transmitted given the grammar. No change to the controller is necessary since all sequences that would be controlled to now would exist in the natural system and control to a pseudonatural orbit. Furthermore, an oscillator identical to the transmitter could possibly synchronize to all of these signals, as they are all orbits of the same dynamical system, though stable synchronization is not guaranteed [10,11].

Imagine the case where we learn a context tree model from a long series of symbols taken from the natural uncontrolled dynamics. Using the methods of [8] one can always deterministically convert the estimated context tree model into such a finite-state first-order Markov chain. Each of the N_s states corresponds to some particular history of recent symbols. There is a transition matrix $\mathbf{P}_{ij}$, giving transition probabilities from state i to state j, with each transition emitting one symbol. If the size of the alphabet (number of symbols in the symbolic dynamics) is A, then there are, at most, A nonzero entries per row, and the sum across rows is identically 1. The metric (Shannon) entropy rate of this Markov chain (assuming one strongly connected component in the graph as expected for ergodic dynamics) is

$$h=\sum_{ij}-\mathbf{P}_{ij}\log\mathbf{P}_{ij}\mu_i,$$

where μ is the stationary density given by $\mu\mathbf{P}=\mu$. This entropy h is the rate at which information can be transmitted in the standard case where $\mathbf{P}$ is the model used in the arithmetic coder.

This is not necessarily the best possible rate, however. The optimum solution is a Markov chain with the same grammar (pattern of allowed and disallowed transitions) but one whose Shannon entropy equals the *channel capacity*, the maximum possible entropy rate consistent with the grammar. To find the channel capacity, form the connectivity matrix $\mathbf{T}_{ij}=\delta(\mathbf{P}_{ij}>0)$, which gives a 1 for all allowed and a 0 for all disallowed transitions. Let λ be the largest eigenvalue of $\mathbf{T}$, $\mathbf{T}b=\lambda b$ for some eigenvector b. The channel capacity C, also known as the topological entropy, is the logarithm of the largest eigenvalue λ of $\mathbf{T}$: $C=h_T=\log\lambda$. The various chaotic saddles corresponding to the same grammar as the attractor have probabilistic transition operators $\mathbf{P}'$ that may differ from $\mathbf{P}$, but they must have the same pattern of allowed and disallowed transitions, i.e., $\mathbf{T}'=\mathbf{T}$. In other words, we may modify any entry of $\mathbf{P}$ that is neither identically 0 nor 1 to form $\mathbf{P}'$. Shannon derived [12] the explicit solution for transmitting at the maximum rate when the symbolic dynamics is a finite-state first-order Markov chain, i.e., when the probability distribution of transitions to any subsequent state, emitting a symbol in the process, depends only on the present state. This transition matrix, using the eigenvector b found before,

$$\mathbf{P}_{ij}^{MAX}=\lambda^{-1}\frac{b_j}{b_i}\mathbf{T}_{ij}, \tag{1}$$

does maximize the entropy rate while staying consistent with the grammar. The entropy rate, channel capacity, and topological entropy are all equal with $\mathbf{P}^{MAX}$.

Figure 4 shows a time-delay plot using $\mathbf{P}^{MAX}$ as the transmitting model. $\mathbf{P}^{MAX}$ was derived from the Markov chain estimated from the uncontrolled oscillation of the experimental circuit. Visually the controlled attractor is still somewhat similar to the natural chaos, but the two can be distinguished with statistical methods. The same type of trajectories will occur in both, but their relative probabilities are altered in the channel-capacity-achieving solution. In fact, to achieve channel capacity, all allowed symbolic words of asymptotically long length ought to occur with equal probability.

Therefore, compared to natural chaos, the rate-maximizing solution will have more occurrences of previously rare sequences.

V. CONCLUSION

In summary, we have described a general-purpose constructive algorithm and performed an experimental demonstration showing that arbitrary messages may be modulated onto the symbolic dynamics of general chaotic systems whose symbolic dynamics are known only via observation. The context-tree estimator can make excellent finite-state tree-machine approximation models given observed symbolic dynamics, and the arithmetic coder can modulate arbitrary white bits onto this proxy dynamical system. The result is chaotic modulation of information at virtually the identical rate as the Kolmogorov-Sinai entropy rate of the underlying chaotic system, resulting in a controlled chaotic oscillation.

In conventional communication with a sinusoidal carrier (e.g., FM or AM radio), the carrier has a zero entropy rate and does not effectively use up any channel bandwidth, leaving it all to the desired message. Stojanovski *et al.* [5] investigated the theoretical requirements for chaotic communication, namely the channel capacity when there is a chaotic carrier, and found that the carrier alone requires a channel capacity equal to its KS entropy. This might suggest naively that chaotic communication would have more overhead than conventional communication by using up the channel bandwith that could otherwise be devoted to the message. Here we have explicitly demonstrated that this need not be the case, by devising a general-purpose approach to transmit a message embedded entirely within the natural KS entropy rate of the uncontrolled system, achieving the best theoretical bound of [5]. When this happens, the controlled system must necessarily appear statistically just like the uncontrolled system, and our result is hence the ideal limit of modulation-based (vs targeting-based) chaotic communication.

Thus we have shown, by explicit experimental demonstration, that the information commonly said to be "created" in a free-running chaotic dynamical system is fungible, interchangeable with the ordinary information of communication theory. All work using the targeting of symbolic dynamics, starting from [1], suggest this, but we think ours is the most compelling demonstration since the result can be made virtually identical to the free-running case. This work significantly broadens the generality of small-perturbation-based approaches to the design and control of chaotic systems. Starting from [1], various *ad hoc* modulation schemes used control inputs to fully determine evolution in a desired way, but the resulting dynamics under modulation were arbitrary in their statistics. In a modern analysis of communication schemes, the information-theoretical aspects of the channel code are divorced from the particular physics and engineering of the modulation hardware; our methods here do the same for targeting-based chaotic communication.

The central prerequisite for our method is to construct a chaotic transmitter whose symbolic dynamics are practically targetable. With our approach, this may be designed with experimental engineering practicality being paramount, as the symbolic end of constructing the transmitted sequence is solved with substantially general and nearly optimal algorithms (with the possible exception of expending channel capacity for error correction). In other words, the symbolic dynamics of the transmitter does not need to be specially tuned to the desired message statistics or engineered for particularly simple symbolic dynamics. Furthermore, one can design a dynamically synchronizing receiver with respect to the unperturbed natural oscillator: finding stable synchronization in its natural condition is easier than ensuring the stability remains under modulation.

Alternatively, one might intentionally want to alter the relative probabilities on the transition matrix used for transmission, e.g., to achieve the channel capacity maximizing solution, to tune the spectral characteristics or noise resistance in continuous space, or to create a system that has a robust synchronizing receiver. Our use of the automatic context-tree modeling method provides a good and sufficiently compact and yet accurate estimate of the unperturbed dynamics and a finite subshift approximation to its grammar. This is a surprisingly nontrivial problem in engineering practice, as there the dynamics of the oscillator will likely be determined by physical implementation necessities, which will not necessarily yield a particularly simple symbolic dynamics. The symbolic dynamics must be estimated entirely empirically, and that calls for the kind of algorithms we employ.

Bollt and Dolnik [13] previously showed an example of inserting binary streams of information into a subshift of the finite type on binary symbols from a model of a chemical reaction. Their estimated subshift had transitions with one or two possible futures—when there was but one future, it was transmitted to respect the grammar, and when there were two, the bit from the message was inserted literally. This procedure, more like data modulation codes, preserves the topological structure (list of forbidden words) but not the metric properties, unlike our present method. It is surprising but true that applying a message bit that has equal probability at every controllable binary transition does *not* necessarily result in the maximum rate solution, i.e., that given by (1), even though instantaneously, such a white bit has the highest entropy pointwise. The complication is that such a transition choice would generically result in executing orbit segments with zero-data rate deterministic transitions more frequently than the optimal solution would. This reduces the overall data rate to be below the maximum achievable (the channel capacity C) even though the per-symbol data rate—at non-deterministic transitions—would be locally maximized. A little introspection reveals the solution to be unobvious and we commend the reader to review Shannon's derivation [12].

Our use of the arithmetic coding procedure permits one to modulate a binary stream onto a symbolic dynamics that is representable with a three-symbol alphabet with more than one bit per symbol entropy, for instance. Baptista *et al.* [14] investigated a somewhat complementary scenario, whereby a chaotic system and control was synthesized to represent the grammar of a message that was assumed to have a certain, known, grammatical, and probabilistic structure. Our method performs "impedance matching" between arbitrary white bits (which can come from a compressed structured source) and arbitrary symbolic dynamics with a finite entropy rate. Our use of the arithmetic decoder as a synthesis tool is most similar to procedures used for efficient production of pseudorandom numbers with a certain structure, given an underlying uniform source of random variates or random bits, e.g., [15,16]. Merhav and Weinberger [17] investigated the theoretical bounds on algorithms for resimulating new data from an information source given the original data, which absent our insertion of nearly white true message bits instead of truly random bits, is similar to the overall architecture of our problem, as we estimate a structured source from the symbolized observed chaotic data, and then resimulate a similarly structured symbolic stream.

[1] S. Hayes, C. Grebogi, and E. Ott, Phys. Rev. Lett. **70**, 3031 (1993).
[2] N. J. Corron and S. D. Pethel, Chaos **12**, 1 (2002).
[3] E. M. Bollt, Int. J. Bifurcation Chaos Appl. Sci. Eng. **13**, 269 (2003).
[4] N. J. Corron and S. D. Pethel, Phys. Lett. A **313**, 192 (2003).
[5] T. Stojanovski, L. Kocarev, and R. Harris, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **44**, 1014 (1997).
[6] B. H. Marcus, P. H. Seigel, and J. K. Wolf, IEEE J. Sel. Areas Commun. **10**, 5 (1992).
[7] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding* (Cambridge University Press, Cambridge, UK, 1995).
[8] M. B. Kennel and A. I. Mees, Phys. Rev. E **66**, 056209 (2002).
[9] M. B. Kennel, Phys. Rev. E **56**, 316 (1997).
[10] http://www.cipr.rpi.edu/~wheeler/ac.
[11] I. H. Witten, R. M. Neal, and J. G. Cleary, Commun. ACM **30**, 520 (1987).
[12] C. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[13] E. M. Bollt and M. Dolnik, Phys. Rev. E **55**, 6404 (1997).
[14] M. S. Baptista, E. Rosa, and C. Grebogi, Phys. Rev. E **61**, 3590 (2000).
[15] T. S. Han and M. Hoshi, IEEE Trans. Inf. Theory **43**, 599–611 (1997).
[16] K. Visweswariah, S. R. Kulkarni, and S. Verdu, IEEE Trans. Inf. Theory **44**, 462 (1998).
[17] N. Merhav and M. J. Weinberger, IEEE Trans. Inf. Theory **50**, 5 (2004).