Chaotic signal detection and estimation based on attractor sets: Applications to secure communications

G. K. Rohde[a)]
NRC Postdoctoral Research Associate, U. S. Naval Research Laboratory, Optical Sciences Division, Washington, D.C. 20375, USA

J. M. Nichols and F. Bucholtz
U. S. Naval Research Laboratory, Optical Sciences Division, Washington, D.C. 20375, USA

(Received 21 September 2007; accepted 10 January 2008; published online 10 March 2008)

We consider the problem of detection and estimation of chaotic signals in the presence of white Gaussian noise. Traditionally this has been a difficult problem since generalized likelihood ratio tests are difficult to implement due to the chaotic nature of the signals of interest. Based on Poincare's recurrence theorem we derive an algorithm for approximating a chaotic time series with unknown initial conditions. The algorithm approximates signals using elements carefully chosen from a dictionary constructed based on the chaotic signal's attractor. We derive a detection approach based on the signal estimation algorithm and show, with simulated data, that the new approach can outperform other methods for chaotic signal detection. Finally, we describe how the attractor based detection scheme can be used in a secure binary digital communications protocol.
[DOI: 10.1063/1.2838853]

This paper offers a new detection scheme for chaotic signals buried in noise and compares this new detector to those previously used in the field. The approach takes a geometrical viewpoint of the detection problem. The nearness of an incoming signal to a known manifold serves as the detection statistic. Using bit error ratio (BER) curves, comparisons are made between this approach and the more commonly used "power detectors" and schemes based on digital chaos shift keying (DCSK). It is then demonstrated how this detection scheme can be used to develop a secure communication protocol using chaotic signals. Using this protocol, both "1s" and "0s" are coded as signals that are spectrally identical yet lie on different manifolds, thus are easily separable from a geometric point of view.

I. INTRODUCTION

We are interested in processing signals arising from nonlinear differential equations

$$\frac{d\mathbf{s}(t)}{dt} = F_\theta(\mathbf{s}(t)) \tag{1}$$

whose solutions $\mathbf{s}(t) \in \mathbb{R}^K$ are chaotic for some given set of initial conditions and where θ refers to a set of parameters describing the nonlinear function F. Chaotic signals $\mathbf{s}(t)$ are nonperiodic solutions which, after a transient period, stay confined to a bounded subset C (denoted as the attracting set) of $\mathbb{R}^K$, $\mathbf{s}(t) \in C \subset \mathbb{R}^K$. Therefore they possess a continuous, often broadband frequency spectrum. Another important property of such signals is that the distance between points in phase space increases exponentially with time. That is, given two initial conditions $\mathbf{s}_1(0)$ and $\mathbf{s}_2(0) = \mathbf{s}_1(0) + \Delta(0)$, $|\Delta(t)|/|\Delta(0)| \sim \exp(ht)$, with $h > 0$. Thus chaotic signals are difficult to predict over long time intervals as a slight disturbance of initial conditions $\Delta(0)$ in the nonlinear system will produce dramatically different results. A chaotic time series generated by using a Lorenz nonlinear system (see Appendix A for details) is shown in Fig. 1.

Chaotic behavior in dynamical systems has long been a subject of study in both theoretical and experimental physics.[1] More recently chaotic systems have also become the subject of study in the field of signal processing. Applications include modeling of radar signals[2] and communications,[3] amongst others. Throughout the past few years many researchers have noted that certain properties of chaotic time domain signals are advantageous for radar and communication protocols. Chaotic signals can be used to mitigate channel imperfections (such as frequency selective fading). In addition, these properties are useful for secure, private, low probability of intercept multiuser digital communications with good antijamming properties.[3] However, because of the complex structure and sensitivity to initial conditions, standard signal processing algorithms used in detection and estimation, such as the maximum likelihood method and generalized likelihood ratio, are difficult to implement. More specifically, performing detection and estimation of chaotic signals in noise by searching for the initial conditions that best match a given signal using a gradient based nonlinear optimization method, for example, is difficult.

Algorithms for signal detection and estimation using maximum likelihood approaches have been limited to very specific classes of chaotic signals (piecewise linear discrete maps).[4–9] Signal estimation methods for continuous chaotic

[a)]Present address: Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213. Electronic mail: gustavor@cmu.edu.

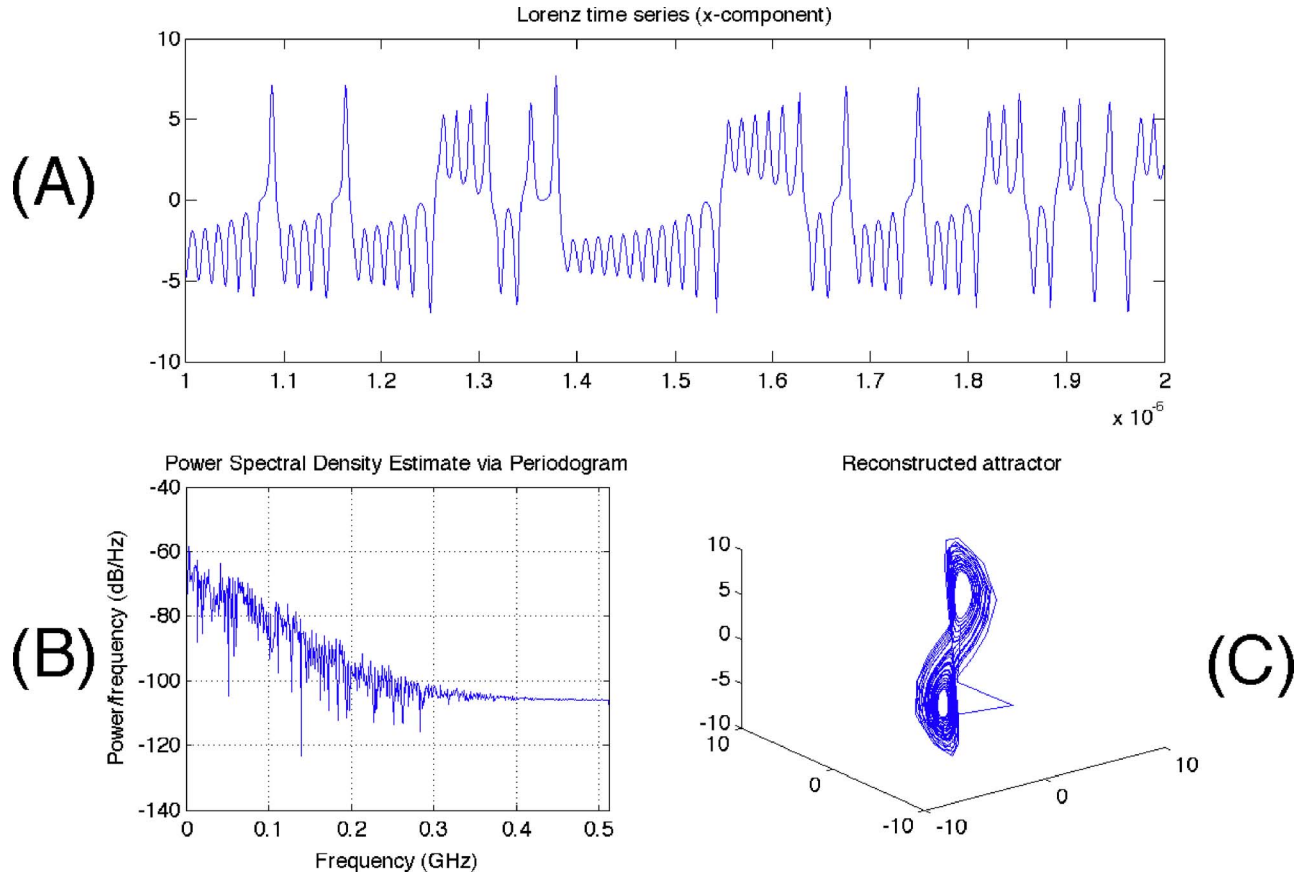

FIG. 1. (Color online) A sample chaotic signal generated from a Lorenz system. (A) time domain signal, (B) power spectral density estimate, (C) attractor reconstruction via delayed embedding.

dynamical systems have been few and offer poor robustness against noise.[10] Therefore implementation of receivers for continuous chaotic signals in binary digital communications, for example, have been hampered. An alternative for providing a reference signal with which to implement a coherent correlation-type receiver is to employ chaotic synchronization techniques.[11] Having two or more chaotic circuits synchronize, however, is a difficult task and because of these difficulties researchers have searched for nonoptimal detection algorithms based on the signal variance (bit energy).[3,12]

Here we describe a general method for detecting and estimating signals generated from chaotic signals (continuous or discrete) based on the idea that points in the phase space, after a transient period, lie on an attractor set C. With an estimate of the attractor set C computed via delay embedding, an estimate of a noisy version of a chaotic signal is found by searching for the phase space points in C closest to the embedded vectors of the noisy signal. A detection statistic is then derived based on the same principle. Finally, we show how these principles can be used to design a physical layer-secure binary digital communication protocol using the Lorenz nonlinear system.

The manuscript is organized as follows: In the next section we review the basic problem of signal detection and provide a geometric interpretation for it. We then use the geometric point of view as motivation in describing our approach for detection and estimation based on decompositions over dictionaries built based upon time-delayed reconstructions of the chaotic time series attractor. We then demonstrate the capabilities of detection and estimation of the algorithm on Monte Carlo-type simulations using pseudorandom numbers to simulate noise. An application to physical layer-secure binary digital communications is presented followed by discussion and concluding remarks.

II. THEORY

A classical signal detection problem can be posed as determining whether a scalar time domain signal $r(t)$, $0 \leqslant t \leqslant T$ is composed of noise alone or signal plus noise,

$$H_0: r(t) = n(t),$$
$$H_1: r(t) = s(t,\theta) + n(t), \tag{2}$$

where $n(t)$ refers to additive white Gaussian noise, and $s(t,\theta)$ refers to the uncorrupted target signal with unknown parameter θ. A typical solution to the problem in this case involves the generalized likelihood ratio,

$$\lambda(r;\hat{\theta}) = \frac{p_1(r|\hat{\theta})}{p_0(r)} \gtrless \eta, \tag{3}$$

where $\hat{\theta}$ is an estimate (e.g., maximum likelihood) of the parameter θ given the data, p_0 and p_1 represent the prior probabilities for the data given hypothesis 0 and 1, respectively, and η is a real valued threshold.

Consider the case when $s(t,\theta)=A\cos(wt+\theta)$, with θ unknown. Then it can be shown that the generalized likelihood ratio framework results in the following test:

$$q^2=|\langle r(t),\cos(wt)\rangle|^2+|\langle r(t),\sin(wt)\rangle|^2 \gtrless \gamma, \tag{4}$$

with $\langle a(t),b(t)\rangle=\int_0^T a(t)b^*(t)dt$ and γ is chosen based on some optimality criterion: minimum probability of error, Neyman–Pearson, etc. In this specific case, the generalized likelihood ratio and the average likelihood ratio yield the same test.[13]

Geometrically, $(4/T^2)q^2$ could be interpreted as the energy of the orthogonal projection of $r(t)$ onto the linear subspace defined by $V=\text{span}\{(2/T)\cos(wt),(2/T)\sin(wt)\}$,

$$q^2=\frac{T^2}{4}\|P_V r(t)\|^2, \tag{5}$$

where $\|a(t)\|=\sqrt{\langle a(t),a^*(t)\rangle}$. Stated another way, $P_V r(t)$ stands for the point on the linear space V which is closest, in the L_2 sense, to the signal $r(t)$. That is,

$$P_V r(t)=\arg\inf_{g(t)\in V}\|g(t)-r(t)\|, \tag{6}$$

where $\arg\inf_x f(x)$ denotes the argument for which there is no lower value of $f(x)$. Due to this interpretation, $P_V r(t)$ can also be used as an optimal estimate of the signal $r(t)$.

A. Chaotic signal detection and estimation

Let $s(t)$ represent one single component of a vector valued chaotic signal $\mathbf{s}(t)$, say the first component of a Lorenz time series (see Appendix A). In this case the attractor set can be reconstructed by embedding the time domain signal $s(t)$ onto a d-dimensional space through the method of delay reconstruction. The phase space reconstruction is given by

$$\mathbf{g}(t)=\{s(t),s(t+\tau),\ldots,s(t+(d-1)\tau)\} \tag{7}$$

for all time points. If $d\geqslant 2K+1$, then, with overwhelming probability, the reconstruction procedure yielding attractor $\tilde{C}$ is a diffeomorphism of C.[14] A time delayed embedding reconstruction of the Lorenz attractor is shown on the bottom of Fig. 1. Naturally, the same phase space plot reconstruction procedure can be applied to a chaotic time domain signal $r(t)$ originating from the same nonlinear system but with different initial conditions. Thus a natural procedure for approximating a chaotic signal $\mathbf{x}(t)$ [i.e., the incoming signal $r(t)$ at time points $t_i=i\tau+t$, for $i=0,\ldots,d-1$] is

$$\hat{\mathbf{x}}(t)=\arg\inf_{\mathbf{g}(t)\in\tilde{C}}\|\mathbf{g}(t)-\mathbf{x}(t)\|. \tag{8}$$

In the cases when $r(t)$ is corrupted by white Gaussian noise alone, for example, $\hat{\mathbf{x}}(t)$ is a maximum likelihood estimate.

The difficulty with this approach for chaotic signals is that a closed form parameterization for $\tilde{C}$ is normally not available. Sample functions $\mathbf{g}(t)\in\tilde{C}$ can be computed by initializing the nonlinear system (1) with different initial conditions. However, because of the exponential rate of divergence of phase space points, this is not a suitable optimization strategy. In the next section we discuss algorithms that can approximate the solution to problem in Eq. (8) with stored, sampled, digital data.

Before proceeding it will be useful to recall Poincaré's recurrence theorem, stated concisely in Ref. 15. Let Γ represent a measurable transformation $\Gamma:X\to X$ preserving a finite measure μ on metric space X with distance ρ. Then

Theorem II.1 (Poincarés Recurrence Theorem): *For almost every* $\mathbf{x}\in X$ *we have*

$$\lim_{n\to\infty}\inf\rho(\Gamma^n\mathbf{x},\mathbf{x})=0.$$

In other words, Poincaré's theorem tells us that, given enough time, the orbit of points $\mathbf{x}$ which belong to subsets X of finite measure returns arbitrarily close to the initial point. This result will be useful in guaranteeing convergence of the algorithm described in the next section.

III. ALGORITHMS

As mentioned earlier, a reconstruction $\tilde{C}$ of C can be obtained through the method of delayed embedding. This can be done by computing a solution $\mathbf{s}(t)$ for the nonlinear system (1) for time $0\leqslant t\leqslant L$ and sampling it at time points $t_i=i\tau+B$, $i=0,\ldots,N-1$, $N=(L-B)/\tau$ ($B>0$ is chosen so as to avoid transient phase space points). The attractor set is reconstructed by embedding the sampled component of interest $s(t_i)$ as in Eq. (7). Let $\mathcal{D}$ be a collection of $(N-d)$ vectors representing the embedded vectors $\mathbf{g}_i=\{s(i\tau+B),s((i+1)\tau+B),\ldots,s((i+d-1)\tau+B)\}$. Thus the dictionary $\mathcal{D}$ consists of a set of $(N-d)$ d-dimensional vectors extracted from a chaotic time series $s(t)$. The goal of the algorithm about to be described is to approximate a data vector (signal) $\mathbf{x}$ using the vectors in the collection $\mathcal{D}$. If the input vector (signal) originates from the same chaotic system, the algorithm will approximate $\mathbf{x}$ well. The initialization of the chaotic time series $s(t)$ is not important as the algorithm about to be described is not dependent on it.

Let $\mathbf{x}$ represent a sampled vector from a (possibly noise corrupted) chaotic time series $r(t)$. In fact, a chaotic time series $r(t)$ of length L will allow for multiple vectors $\mathbf{x}_m=\{r((m+1)\tau+B),\ldots,r((m+d-1)\tau+B)\}$ to be extracted from it. A simple approximation algorithm for each chaotic data vector $\mathbf{x}_m$ generated by using arbitrary initial conditions is then given by searching for the vector $\mathbf{g}$ in the collection $\mathcal{D}$ that matches the input vector $\mathbf{x}$ most closely. The notion of close or far is provided by the standard d-dimensional Euclidean distance. Thus the algorithm reduces to a search over the elements of $\mathcal{D}$ to find the vector $\mathbf{g}_i$ that is closest to $\mathbf{x}_m$,

$$\hat{\mathbf{x}}_m=\arg\min_{\mathbf{g}\in D}|\mathbf{g}-\mathbf{x}_m|, \tag{9}$$

where $|\mathbf{x}|=\sqrt{\Sigma_{i=0}^{d-1}x_i^2}$ and where the notation $\arg\min_x f(x)$ denotes the argument that minimizes function f. Poincaré's recurrence theorem guarantees that for almost every $\mathbf{x}_m$ belonging to the attractor set $\tilde{C}$,

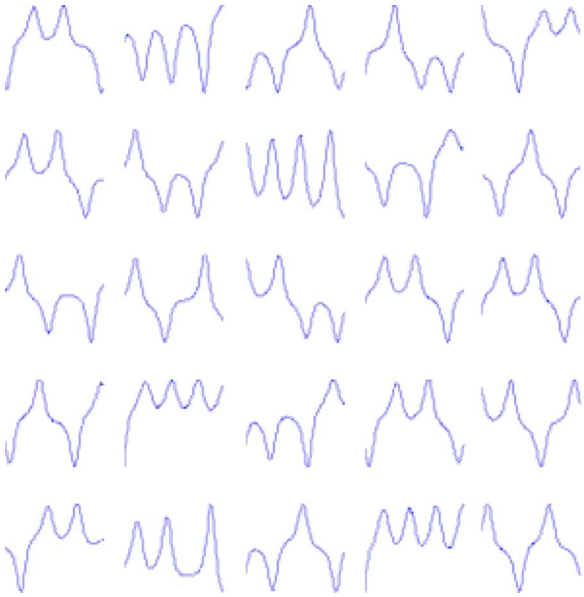

FIG. 2. (Color online) Sample (normalized) embedded vectors for a Lorenz dictionary.

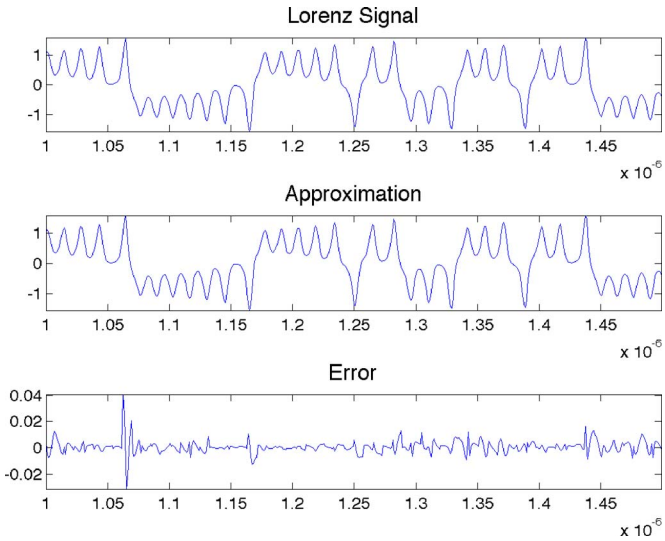

FIG. 3. (Color online) Approximate reconstruction (middle row) of a Lorenz time series (top) with a dictionary size of 20 472. The difference between the original and reconstructed time series is displayed on the bottom row.

$$\lim_{L\to\infty} |\hat{\mathbf{x}}_m - \mathbf{x}_m| = 0. \tag{10}$$

The chaotic signal estimation algorithm amounts then to a nearest neighbor search over the dictionary of vectors $\mathcal{D}$. Naturally, the solution to Eq. (9) only approximates the solution to Eq. (8) and for a fixed L the algorithm is not exact. However, the algorithm converges and the problem of optimizing over initial values is avoided.

A. Approximation over normalized attractor sets

The algorithm described above is not signal to noise ratio (SNR) independent. That is, a given dictionary $\tilde{C}$, constructed as described above, is not suitable for recovering a signal $\alpha r(t)$ for some arbitrary real valued α. To see that, simply consider the case when the attractor $\tilde{C}$ is built based upon the solution $s(t)$ of a specific nonlinear system. Now consider another solution of the same nonlinear system (computed using different initial conditions) $\alpha r(t)$ amplified by a constant coefficient $\alpha \gg 1$, for example. Then the vector $\mathbf{x}_m$ built based on a particular embedding of $\alpha r(t)$, in general, is not in $\tilde{C}$.

Because we would like to obtain a SNR independent method we modify the algorithm above by normalizing each column vector of $\mathcal{D}$, $\mathbf{g}_i \to \mathbf{g}_i/|\mathbf{g}_i|$. This has the effect of projecting each $\mathbf{g}_i$ vector onto a hypersphere of $\mathbb{R}^d$. For embedding dimensions greater than 3, a visual picture of the normalized dictionary is difficult to obtain. However, different normalized vectors comprising the dictionary can be displayed individually. This is done in Fig. 2 for a dictionary based on the Lorenz time series.

Since the vectors of $\mathcal{D}$ have norm 1, the search for the vector $\mathbf{g} \in \mathcal{D}$ which is closest to a particular vector $\mathbf{x}_m$ is equivalent to searching for the vector $\mathbf{g}$ whose dot product with $\mathbf{x}_m$ is highest. The approximation algorithm is then

$$\hat{\mathbf{x}}_m = (\mathbf{x}_m' \mathbf{g})\mathbf{g}, \tag{11}$$

where $\mathbf{x}'$ is the transpose of $\mathbf{x}$ and

$$\mathbf{g} = \arg\max_{\mathbf{h}\in\mathcal{D}} |\mathbf{h}'\mathbf{x}_m|. \tag{12}$$

To demonstrate this algorithm we compute approximations of a single instance $r(t)$ of the x component of a Lorenz chaotic time series over dictionaries $\mathcal{D}$ of increasing size. In this case the signal $s(t)$ from which the normalized dictionary was constructed was computed using $d=8$, while L (and consequently the size of the dictionary) was increased each time (see Appendix A for additional description of parameters used in this simulation). The signal being approximated was also sampled with the same rate as the dictionary, generating a time series of 1024 samples, and also partitioned in nonoverlapping windows of size 8. An approximation was performed in each window separately. For additional details about parameters, see Appendix A. The result of approximating the sampled signal $r(t)$ using a dictionary of 20 472 vectors is shown in Fig. 3. The total number of dictionary vectors in this simulation was chosen arbitrarily. The total error (summed across all windows) $\Sigma_{m=1}^{M} |\mathbf{x}_m - \hat{\mathbf{x}}_m|$ is displayed in Fig. 4.

B. Matching pursuits with chaotic attractor sets

The error in each window $\hat{\mathbf{x}}$ computed using the algorithm described in Eqs. (11) and (12) using a dictionary $\mathcal{D}$ of fixed size can be further reduced by using matching pursuit[16] type approaches. In fact Eqs. (11) and (12) comprise the first iteration of the matching pursuit algorithm with the discrete dictionary $\mathcal{D}$. Call the vector obtained with the first iteration of Eqs. (11) and (12) $\mathbf{x}^0$ and denote the residual $R^1\mathbf{x} = \mathbf{x}^0 - \mathbf{x}$. One can then proceed by replacing $R^1\mathbf{x}$ for $\mathbf{x}$ in Eqs. (11) and (12),

$$\mathbf{x}^1 = [(R^1\mathbf{x})'\mathbf{g}_1]\mathbf{g}_1, \tag{13}$$

where

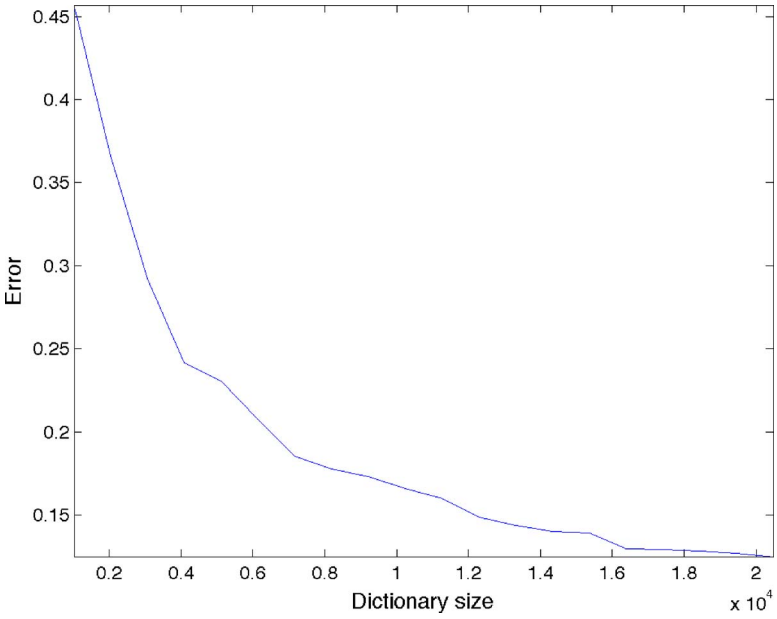

FIG. 4. (Color online) Error between the reconstructed and original Lorenz time series as a function of the size of the dictionary size.

$$\mathbf{g}_1 = \arg\max_{\mathbf{h}\in\mathcal{D}} |\mathbf{h}'(R^1\mathbf{x})|. \tag{14}$$

The procedure can be repeated n times and a data vector $\mathbf{x}$ can be written as

$$\mathbf{x} = \sum_{i=0}^{n-1} [(R^i\mathbf{x})'\mathbf{g}_i]\mathbf{g}_i + R^n\mathbf{x}, \tag{15}$$

where $R^0\mathbf{x}=\mathbf{x}$. We denote the n term approximation as

$$\hat{\mathbf{x}}^n = \sum_{i=0}^{n-1} [(R^i\mathbf{x})'\mathbf{g}_i]\mathbf{g}_i. \tag{16}$$

In other words, algorithm (16) recovers an approximation to data vector $\mathbf{x}$ based on element vectors $\mathbf{g}_i \in \mathcal{D}$. The approximation is not optimal in the sense that for a given n, Eq. (16) does not, for example, minimize $|\mathbf{x}-\hat{\mathbf{x}}^n|$. However, the approximation is efficient to compute (on the order of n nearest neighbor searches) and it does tend to provide sparse approximations; that is, most of the energy of the signals is contained in the first few coefficients $(R^i\mathbf{x})'\mathbf{g}_i$.

If $\mathcal{D}$ provides a complete representation of $\mathbb{R}^d$, then the algorithm in Eq. (16) recovers any vector in $\mathbb{R}^d$.[16] However, our goal here is not to provide an algorithm to reconstruct any data vector in $\mathbb{R}^d$ but only those which belong to the generalized cone defined by vectors $\alpha\mathbf{g}$, $\mathbf{g}\in\mathcal{D}$ and α an arbitrary constant. This could be especially useful for avoiding false positives in the problem of detection to be discussed in the next section. To that end the matching pursuit algorithm can be modified so that the reconstructed vector $\hat{\mathbf{x}}^n$ does not lie too far from the set defined by $\alpha\mathbf{g}$, $\mathbf{g}\in\mathcal{D}$. This can be achieved by using a "local" version of the matching pursuit algorithm. That is, the first iteration is computed normally using Eqs. (13) and (14). The algorithm proceeds by first "pruning" $\mathcal{D}$ so that it only contains some k nearest neighbors of $\mathbf{g}_1$. Denoting the pruned matrix as ${}^k\mathcal{D}$, the vector $\mathbf{g}_i$ in Eq. (15) is chosen from ${}^k\mathcal{D}$ instead of $\mathcal{D}$. Figure 5 displays the error per iteration of a local matching pursuit approximation of the same Lorenz signal, starting from the approximation displayed in Fig. 3.

C. Chaotic signal detection

Here we investigate the application of the chaotic signal estimation algorithms above to the solution of the classical signal detection problem (2). Following the approach outlined earlier for linear detection over hyperplanes, a natural detection statistic for chaotic signals is related to the energy of the vector $\alpha\mathbf{g}$, $\mathbf{g}\in\mathcal{D}$ closest to the signal data $\mathbf{x}$,

$$q^2 = |\hat{\mathbf{x}}|^2 \gtrless \gamma. \tag{17}$$

In the case of matching pursuit-based approximations,

$$q^2 = |\hat{\mathbf{x}}^n|^2 = \sum_{i=0}^{n-1} |(R^n\mathbf{x})'\mathbf{g}_i|^2 \gtrless \gamma. \tag{18}$$

As described above, a time series of length N can be divided in windows of size d and the detection statistic is simply the average of the detectors described above over all windows. We note that although in our work we have divided the signals into nonoverlapping windows, this is not a requirement. Of interest is the probability of error P_e [in communication problems P_e is referred to as bit error rate (BER)] as a function of L, d, n, and SNR. The probability of error is

$$P_e = P[D_1|H_0]P_0 + P[D_0|H_1]P_1, \tag{19}$$

where $P[D_1|H_0]$ stands for the probability of detecting a signal given that none was present, $P[D_0|H_1]$ vice versa, and the abbreviations $P_0=P[H_0]$, $P_1=P[H_1]$ were used. Assuming that a large number of data samples is available in a low SNR regime, the statistic $q^2=\Sigma_{m=1}^M|\hat{\mathbf{x}}_m^n|^2$ is approximately normally distributed. Thus we have, approximately,

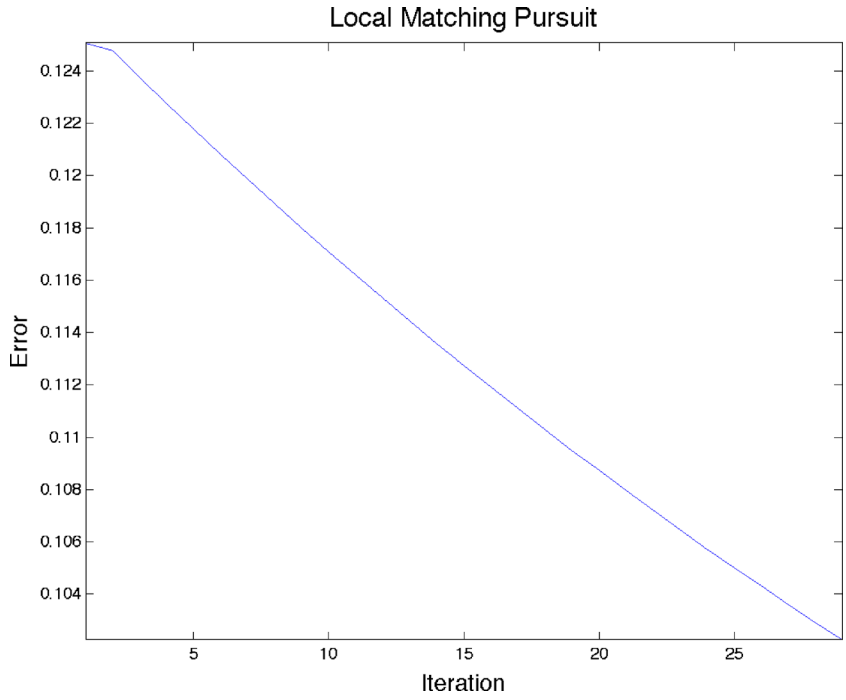

FIG. 5. (Color online) Error between reconstructed and original time series as a function of the number of elements in the "local" matching pursuit expansion.

$$P[D_1|H_0] = \int_{\gamma}^{\infty} \frac{1}{\sigma_0\sqrt{2\pi}} \exp\left[-\frac{(x-u_0)^2}{2\sigma_0^2}\right] dx \tag{20}$$

and

$$P[D_0|H_1] = \int_{-\infty}^{\gamma} \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[-\frac{(x-u_1)^2}{2\sigma_1^2}\right] dx \tag{21}$$

with u_0 (σ_0^2) and u_1 (σ_1^2) refer to the means (variances) of the detection statistic q^2 under the hypothesis H_0 and H_1, respectively. Solving for $dP_e/d\gamma=0$, when $\sigma_1 \neq \sigma_0$, we have

$$\gamma = \frac{-b-\sqrt{\Delta}}{2a} \tag{22}$$

with

$$a = \frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2}, \tag{23}$$

$$b = \frac{u_0}{\sigma_0^2} - \frac{u_1}{\sigma_1^2}, \tag{24}$$

$$c = \frac{u_1^2}{2\sigma_1^2} - \frac{u_0^2}{2\sigma_0^2} - \ln\left(\frac{P_1\sigma_0}{P_0\sigma_1}\right), \tag{25}$$

and $\Delta=b^2-4ac$. In the case where $\sigma_0=\sigma_1$ the threshold reduces to $\gamma=(u_0+u_1)/2$.

Describing the means and variances of q^2 under both hypotheses is difficult since the chaotic attractor sets, and their invariant measures, do not have a parameterized closed form. Thus, obtaining lower bounds (e.g., Rao–Cramer) on the variance of the estimators is a difficult task and is the subject of the current study. However, the performance of the detection algorithm above can be evaluated using Monte Carlo-type simulations. In the next section, the performance of the noncoherent chaotic signal detector described above will be compared to popular noncoherent detection methods such as the bit energy detector[17] and differentially coherent differential chaos shift keying (DCSK) methods[12] using several different chaotic time series.

IV. SIMULATION EXPERIMENTS

A. Denoising

The chaotic detection and estimation algorithm above can be used to extract signals from noisy samples. Figure 6 compares the matching pursuit decomposition of a noisy Lorenz time series with a Donoho's soft thresholding wavelet denoising algorithm[18] using the Symlet 8 wavelet.[19] Additive white Gaussian noise was added to the time series such that the ratio of the standard deviation of the signal over the standard deviation of the noise was 0.5. Here the "local" version of the matching pursuit algorithm described above was used, with the number of iterations set to 30, and $d=128$. As seen in this figure, the matching pursuit algorithm produces a superior estimate of the time series as compared to the wavelet one. This is to be expected, since the pursuit algorithm uses *a priori* information and searches for the signal in the correct space (normalized attractor set). Decomposing the noisy signal using basis functions chosen arbitrarily may result in distorted signals, as this example shows.

B. Detection

The performance of the chaotic signal detection algorithms described earlier for solving the detection problem summarized in Eq. (2) were evaluated using simulations. The simulations were computed in the Matlab programming language using the "randn" function for simulating additive white Gaussian noise. The simulations were based on the baseband discrete model described in Ref. 3, for example, and summarized in Appendix B.

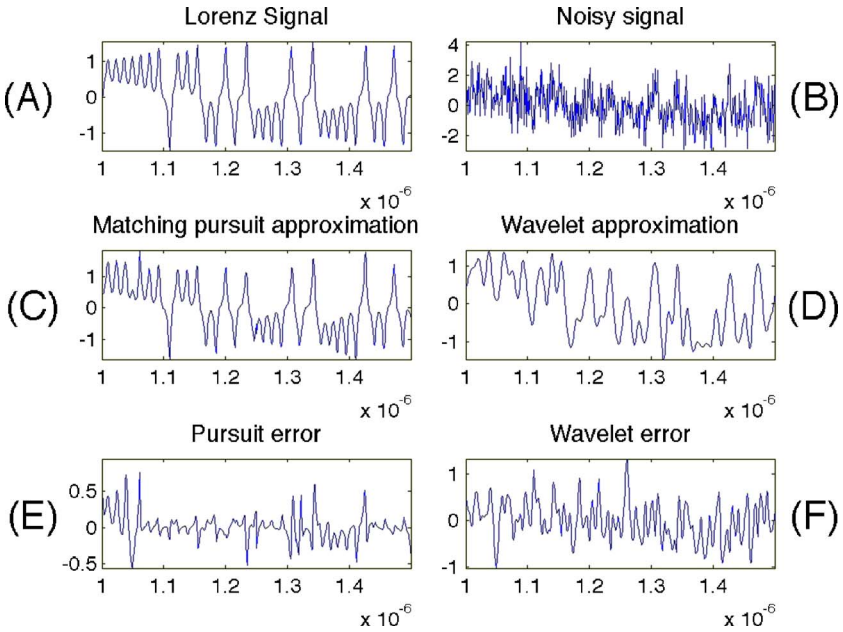

FIG. 6. (Color online) The matching pursuit algorithm can be used for extracting a Lorenz time series from a noisy observation. (A) the original Lorenz signal, (C) matching pursuit approximation of the noisy signal, (E) the difference between the original time signal and the matching pursuit approximation, (B) the noisy Lorenz time series, (D) wavelet approximation using the Donoho's soft thresholding approximation, and (F) error between the wavelet approximation and original time series.

The bit error ratio curve for the on-off chaotic communication system described above is shown in Fig. 7 for the Lorenz, Rossler, and Henon systems. Here $d=128$ and the number of iterations in the matching pursuit algorithm was set to $n=1$. In our simulations for the detection problem, the effect of increasing the number of elements in the signal expansion was minimal, so for this and all the other simulations the number of elements was set to 1. In this simulation, the size of each dictionary $\mathcal{D}$ was 16 384. The bit error ratio (BER) was computed using Eq. (19). 20 000 simulation signals were used to compute the mean and the variance of the detection statistics, with the length of each signal set to $N=1024$, and $P_1=P_0=0.5$.

The detection algorithm above was compared to the bit energy detector described in Ref. 17. In addition, we also include simulations depicting the performance of the differentially coherent DCSK. The idea behind differentially coherent DCSK is to sacrifice half of the transmission time (half the samples in a discrete signal) to send the key to decoding each bit. Thus, the first half of each signal sent is composed of a chaotic time series. The second half is a copy of the chaotic time series in the case of a 1 being sent, and the negative of the chaotic time series in case of a 0 being sent. The detector decodes the bits by taking the inner product between the first and second halves of each signal sent. If this inner product is less than zero, a zero bit is declared. If it is greater than zero, a one bit is declared. For more details, see Refs. 3 and 12.

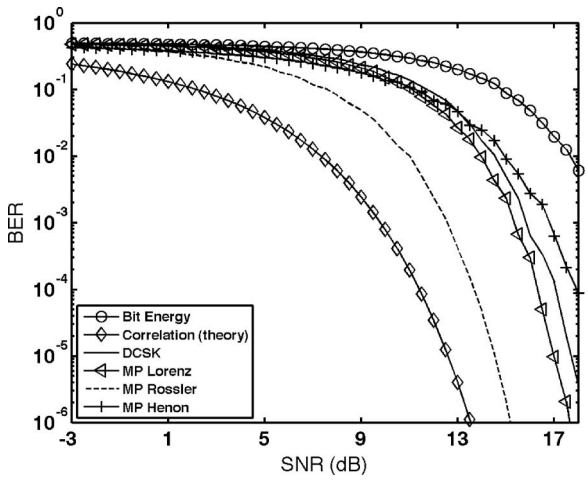

FIG. 7. Bit error ratio curves for different chaotic signal detection schemes as a function of the signal to noise ratio.

As shown in Fig. 7, the performance of the matching pursuit (MP) detectors for the Lorenz and Henon systems is comparable to that of DCSK for the system parameters used. When a Rossler system is used instead, a significant improvement in BER is possible. Finally, note that the figure also includes the performance of the correlation detector (the optimal receiver in the likelihood ratio framework).[13] However, the correlation receiver was included here for comparison purposes only. Implementing a correlation receiver for communication systems based on chaotic signals involves chaos synchronization, which, as mentioned earlier, is a difficult task in practice.[3]

Bit error ratios as a function of the size of the dictionary $\mathcal{D}$ used in the detection scheme were also computed for the same three chaotic systems. Results are shown in Fig. 8. For the system parameters used in these simulations, it seems that relatively few dictionary vectors are necessary for obtaining the BER performance shown in Fig. 7. The simulations were computed for SNR=15 dB.

The performance of the detector as a function of embedding dimension d was also tested at SNR=15 dB, whereas the size of each library was 16 384. Results are plotted in Fig. 9. Embedding dimension had a clear effect on the continuous systems, especially the Rossler system; the higher the embedding dimension, in general, the better the detection performance.

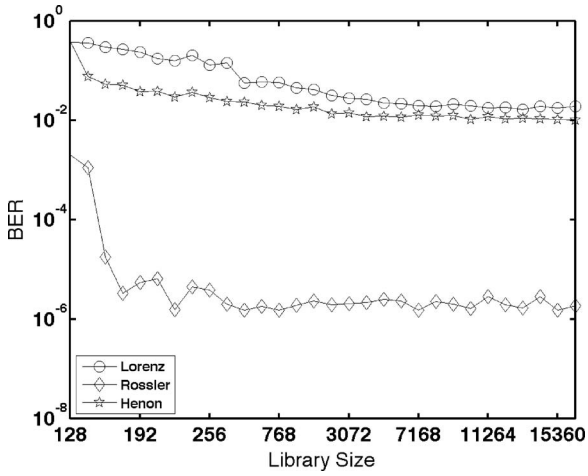

FIG. 8. Bit error ratio curves for matching pursuit-based chaotic signal detection for different systems as a function of the dictionary (library) size for SNR= +15 dB.

V. APPLICATIONS TO SECURE BINARY DIGITAL COMMUNICATIONS

The binary digital communication protocols described above are not secure in the sense that well known techniques can be used to detect and demodulate such signals. The on-off communications protocols can be easily decoded using a power detector (albeit at a lower BER performance). In addition, differentially coherent DCSK techniques are relatively well known and decoders for them are easy to implement. We now describe a novel binary digital communications protocol based on the ideas presented above. The protocol has the advantage that signals cannot be demodulated using standard spectral analysis methods, thus offering physical layer security, reducing the risk of eavesdropping. The proposed protocol utilizes a noncoherent receiver, and does not depend on synchronization of different chaotic systems, nor on the exact knowledge of initial conditions of the nonlinear system.

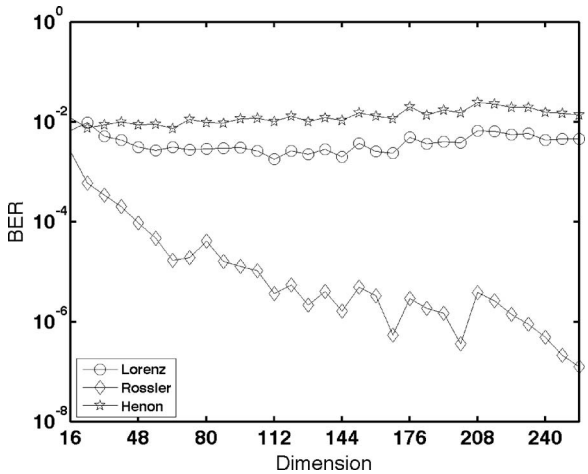

FIG. 9. Bit error ratio curves for matching pursuit-based chaotic signal detection for different systems as a function of the embedding dimension for SNR= +15 dB.

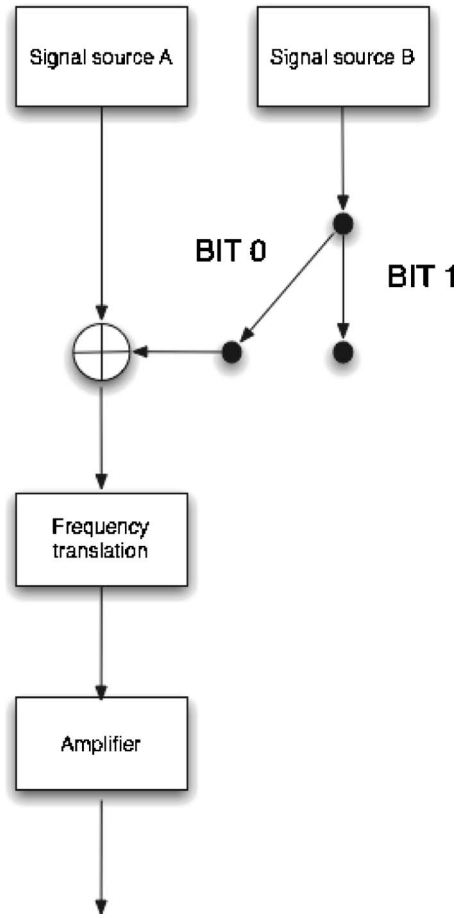

FIG. 10. Schematic diagram depicting implementation of bits 0 and 1.

We study the following encoding scheme: Given a specific nonlinear system F and associated cone $\alpha\mathbf{g}$, $\mathbf{g} \in \widetilde{C}$, the bit zero can be represented by a signal whose embedded vectors do not belong to the cone while bit one is represented by a signal whose embedded vectors do belong to it. Bit one can be easily generated by using a signal arising from the nonlinear system F (by initializing the system with random initial conditions, for example). In order to generate a signal with a similar Fourier spectrum for encoding bit 0 one may take advantage of the fact that the cone $\alpha\mathbf{g}$, $\mathbf{g} \in \widetilde{C}$, in general, is not a linear space or convex set. Thus a chaotic signal whose embedded vectors do not belong to cone $\alpha\mathbf{g}$, $\mathbf{g} \in \widetilde{C}$ may be generated by simply adding together two chaotic signals also arising from nonlinear system F, initialized with different initial conditions.

A system diagram illustrating this concept is shown in Fig. 10. Figure 11 shows two sample signals prior to frequency translation generated using computer simulated Lorenz time series. Note that the two signal components were appropriately weighted so that the energy of the resulting signal was the same as the energy of the signal encoding bit one. The bit one is encoded by a single signal arising from the same Lorenz system (bottom of the figure). Note that no noise was added to these figures. For details about the Lorenz nonlinear system, see Appendix A.

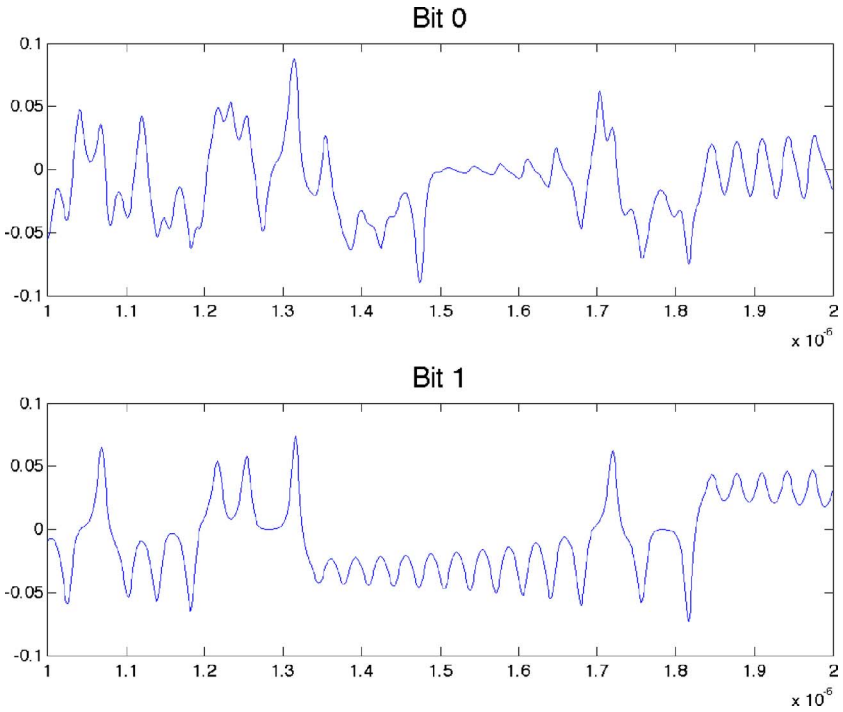

FIG. 11. (Color online) Chaotic signals arising from the Lorenz nonlinear system. Top: the addition of two Lorenz signals encoding bit 0. Bottom: a signal arising from the Lorenz system encoding bit one.

The primary task of the receiver is to decode the received signals into bits zero and one. In this case, due to the fact that the chaotic signals in use are ergodic, and uncorrelated, the Fourier spectra for bits 0 and 1 are nearly identical. This is demonstrated in Fig. 12 which contains an estimate of the Fourier spectra of the signals shown in Fig. 11 computed using the periodogram method from the average of 100 realizations of the signals. Therefore, the signals cannot be decoded using standard Fourier analysis techniques. In fact, any decomposition of the chaotic signals onto linear subspaces is not likely to be informative. Time frequency transforms are also uninformative when it comes to decoding the information present in the different signals. Figure 13 displays the Wigner–Ville time-frequency distribution (computed as described in Ref. 20 from the same 100 realization of the signals above). The time-frequency distributions for the signals encoding bits 0 and 1 are nearly identical.

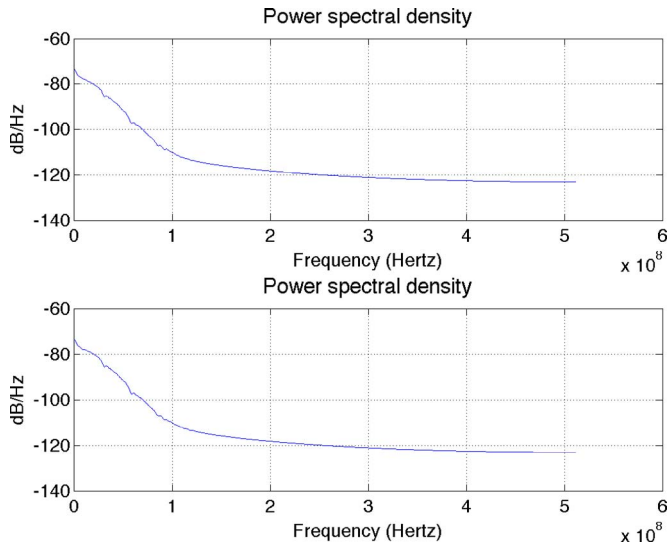

FIG. 12. (Color online) Power spectral densities for chaotic signals encoding bits 0 (top) and 1 (bottom).

In this case demodulation requires learning from the received signal $r(t)$ whether or not its embedded vectors belong to the cone $\alpha\mathbf{g}$, $\mathbf{g} \in \widetilde{C}$. To that end we use the receiver described in Eq. (17) and test it against a threshold γ chosen so as to minimize the probability of making an error P_e (declaring a zero received when in fact a one was sent, or vice versa). The BER performance for this digital communications protocol was simulated as described above using the Lorenz time series. Results are shown in Fig. 14.

VI. DISCUSSION AND CONCLUSIONS

We have presented a method for detecting and estimating signals arising from chaotic systems. The nonlinear, chaotic signal detection and estimation algorithm exploits the intrinsic geometry of chaotic attractor sets to provide a reconstruction of a given chaotic signal generated with arbitrary initial conditions. Naturally, for any given application, the dynamics of the nonlinear system (as well as the sampling rate) must be known so that an appropriate attractor set can be reconstructed and used by the algorithm. The algorithm can be thought of as a matching pursuit[16] over the chaotic attractor set. However, the goal is not to provide an errorless representation of any discrete signal in $\mathbb{R}^d$ but only an accurate representation of signals that belong to a particular class (i.e., some specific chaotic signals).

The algorithm was used in experimental simulations in denoising and detection problems. In the denoising problem, the importance of obtaining decompositions over the appropriate set was demonstrated. Though the wavelet approximation is capable of producing an errorless expansion of any

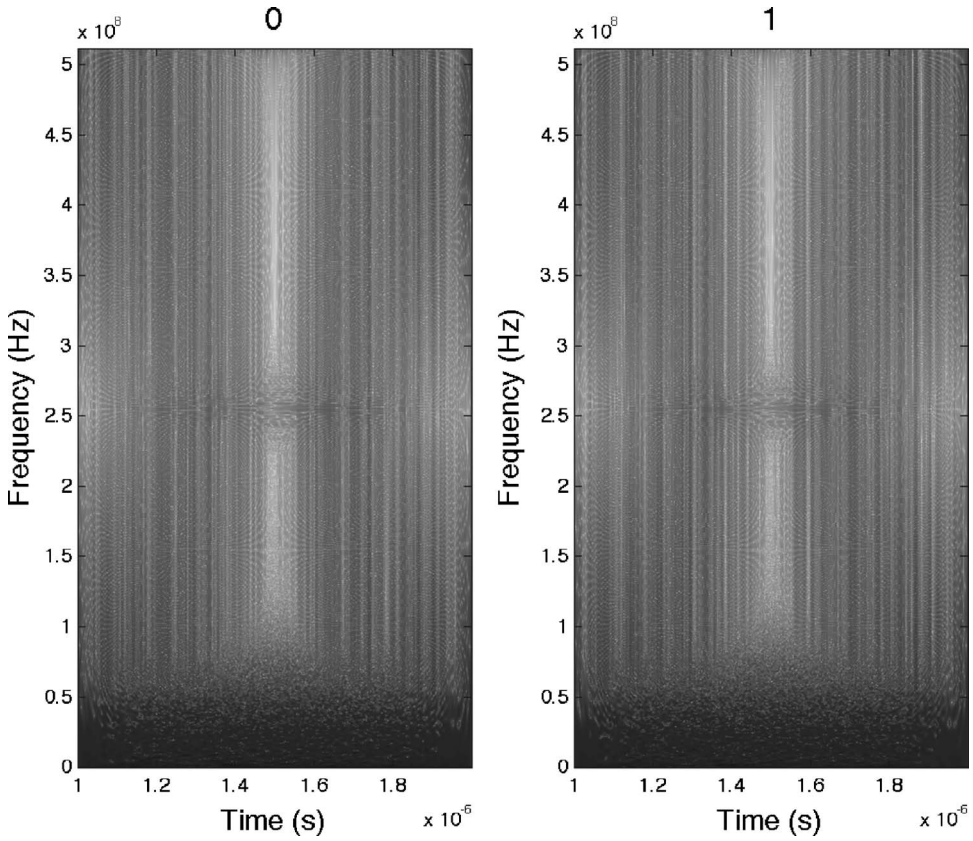

FIG. 13. Wigner–Ville time-frequency distributions of signals encoding bits 0 (left) and 1 (right).

discrete signal, it was not nearly as effective as the denoising procedure using the matching pursuit algorithm.

In the detection problem, the matching pursuit computed over the appropriately chosen attractor set also improved on the bit energy detector used for chaotic communications by others.[3] In addition, the new detection scheme seems to perform comparably to the performance of DCSK. In the case of the Rossler system, the increase in performance in comparison to the bit energy and DCSK detector is significant.

Though based on geometric observations with close analogs in more classical signal detection approaches, the attractor-based algorithm we propose contains several parameters: embedding dimension, dictionary size, and number of iterations in the matching pursuit. The effect of these parameters on the detection problem was studied empirically. While dictionary size and embedding dimension had noticeable effects in the BER performance of detection problems (see Figs. 8 and 9) their precise effect cannot be generally described at this point, as simulations showed them to have different effects for different chaotic systems. The number of iterations in the matching pursuit, however, did not have a noticeable effect in the experiments performed. That is, results with two or more iterations produced results no better than one single iteration.

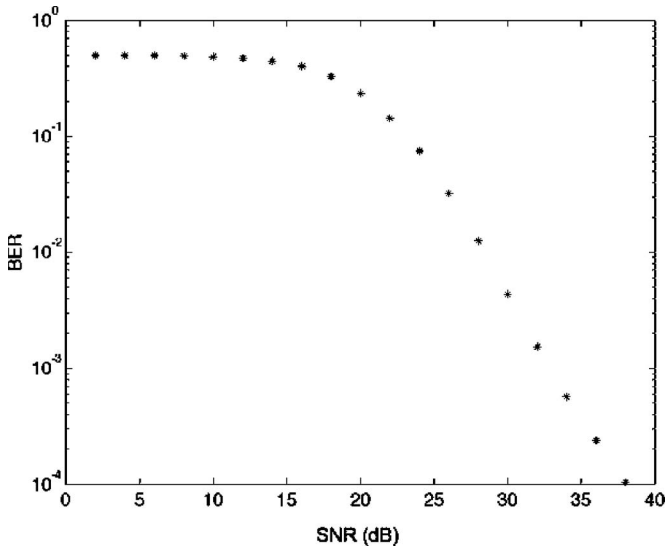

FIG. 14. BER curve for secure communications based on chaotic signals.

Several unresolved issues remain. Perhaps none more important than understanding the limitations of the detection and estimation algorithm as a function of the several parameters involved: embedding dimension, size of the dictionary, number of elements in the pursuit algorithm, etc. Performance bounds on the variance of the estimators would lead to a better understanding of optimal bit error ratio that could be achieved by the algorithm.

Finally, the algorithm could be used by many signal processing applications using chaotic signals. One important example could be low probability of intercept digital communications. We demonstrated the application of the method to the design of a physical layer-secure binary digital communications scheme using a Lorenz nonlinear system. The advantages of the secure communications protocol proposed here over previous works describing communication systems are that (1) it does not depend on synchronization (it uses an incoherent receiver) and (2) it does not depend on exact

knowledge of the initial conditions used to generate the signals. In addition, each signal representing bit 1, for example, is different from the previous signals that encoded the same bit. In fact, they are nearly uncorrelated.

A disadvantage of the algorithm proposed here is that the computational complexity at the receiver's end is higher than many other methods. More specifically, given a dictionary $\tilde{C}$ of size $N \times d$, the computational cost of the detection and estimation algorithms described above is $O(Nd)$. In our implementation, using a signal length of 256 samples, and dictionary size of 512 vectors, decoding 1000 bits takes, on average, 0.87 s on a 2.4 GHz PowerPC processor using standard C-based software libraries. However, these operations can be easily parallelized in the presence of multiple processors and may not be overly restrictive even for real-time implementations.

ACKNOWLEDGMENTS

The authors would like to thank Dr. L. Pecora and Dr. T. Carroll of the Naval Research Laboratory, Washington, D.C., USA, for sharing their knowledge of chaotic systems.

APPENDIX A: NONLINEAR SYSTEM DEFINITIONS

The Lorenz system is comprised of three nonlinear, ordinary differential equations

$$dx/dt = \xi[a(y-x)], \quad dy/dt = \xi(bx - y - xz), \quad dz/dt = \xi(cz + xy), \tag{A1}$$

with constant parameters ξ, a, b, c. In this study the parameters were fixed to the values $\xi = 5 \times 10^7$, $a = 10$, $b = 28$, $c = -8/3$ resulting in chaotic dynamics. Time series were obtained by numerically integrating Eqs. (A1) using a fifth order Runge-Kutta algorithm with a fixed time step of $\Delta t = 9.76 \times 10^{-10}$. The transient integration time B was set to 9.76×10^{-7} s, and each signal was composed of $N = 1024$ samples. $s(t) = x(t)$ was taken as the signal of interest.

The Rossler system is given by

$$dx/dt = \xi(-y - z), \quad dy/dt = \xi(x + ay), \quad dz/dt = \xi[b + z(x - c)], \tag{A2}$$

where we have chosen the constants $a = 0.15$, $b = 0.2$, $c = 10.0$. All other parameters associated with the integration were the same as with the Lorenz system.

The final system explored is the discrete Henon map

$$x(n+1) = y(n) + 1 - ax(n)^2, \quad y(n+1) = bx(n), \tag{A3}$$

where $a = 1.4$, $b = 0.3$ are constants. Initial conditions for this system were chosen in the range $0 < x(0), y(0) < 0.25$ in order to ensure the solution converged to a stable attractor (large initial conditions lead to an unbounded system response). As with the Lorenz and Rossler systems we discard the first 1024 iterations as transients.

APPENDIX B: BASE BAND SIMULATION MODEL

Here a summary of the discrete time baseband model assumed to compute the simulations discussed above. Let a chaotic time domain signal $s(t)$ have a bandwidth of $\pm W$. Here we assume that the signal $s(t)$ is effectively band-limited. In an amplitude modulation communications link, the frequency content of $s(t)$ is translated and centered at f_0, with $f_0 \gg W$, via multiplication with a complex exponential $\exp(\iota f_0 t)$ prior to transmission over the channel. At the receiving end, the signal is demixed and filtered so its bandwidth is again $\pm W$. Assuming additive white Gaussian noise distortion in the channel, denote the two sided power spectral density of the noise $N_0/2$. The signal is then discretized using the Nyquist time step $\tau = 1/W$. Thus the (band-limited) noise in the signal is assumed uncorrelated and its stationary variance is given by $\sigma^2 = N_0 W = N_0/2\tau$. The simulations above were computed starting from the discretization of the baseband signal $s(t)$.

[1] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, 1993).
[2] S. Haykin and S. Puthusserypady, *Chaotic Dynamics of Sea Clutter* (Wiley, New York, 1999).
[3] F. Lau and C. Tse, *Chaos-Based Digital Communication Systems* (Springer, Berlin, 2003).
[4] H. Papadopoulos and G. Wornell, IEEE Trans. Inf. Theory **41**, 312 (1995).
[5] S. M. Kay, IEEE Trans. Signal Process. **43**, 2013 (1995).
[6] C. Pantaleon, D. Luengo, and I. Santamaria, IEEE Signal Process. Lett. **7**, 235 (2000).
[7] J. Schweizer and T. Schimming, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **48**, 1269 (2001).
[8] Y. Hwang and H. C. Papadopoulos, IEEE Trans. Signal Process. **52**, 2637 (2004).
[9] H. Leung, S. Shanmugam, N. Xie, and S. Wang, IEEE Trans. Signal Process. **54**, 1091 (2006).
[10] M. Davies, Physica D **79**, 174 (1994).
[11] L. Pecora, T. Carrol, G. Johnson, D. Mar, and J. Heagy, Chaos **7**, 520 (1997).
[12] G. Kolumbán, G. Vizvari, W. Scharz, and A. Abel, in *International Workshop and Nonlinear Dynamics of Electronic Systems* (Seville, Spain, 1996), pp. 92–97.
[13] A. Whalen, *Detection of Signals in Noise* (Academic, New York, 1971).
[14] F. Takens, in *Lecture Notes in Mathematics* (Springer-Verlag, Berlin, 1981).
[15] L. Barreira, in *XIV International Congress on Mathematical Physics (Lisboa, 2003)* (World Scientific, Singapore, 2005), pp. 415–422.
[16] S. Mallat and Z. Zhang, IEEE Trans. Signal Process. **41**, 3397 (1993).
[17] G. Kolumbán, M. Kennedy, and L. Chua, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **45**, 1129 (1998).
[18] D. L. Donoho, IEEE Trans. Inf. Theory **41**, 613 (1995).
[19] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).
[20] S. Mallat, *A Wavelet Tour of Signal Processing* (Academic, London, 1999).