Stochastic analysis of recurrence plots with applications to the detection of deterministic signals

Gustavo K. Rohde,*†Jonathan M. Nichols, Bryan M. Dissinger, Frank Bucholtz

Naval Research Laboratory, Washington, DC

September 17, 2007

Abstract

Recurrence plots have been widely used for a variety of purposes such as analyzing dynamical systems, denoising, as well as detection of deterministic signals embedded in noise. Though it has been postulated previously that recurrence plots contain time correlation information here we make the relationship between unthresholded recurrence plots and the covariance of a random process more precise. Computations using examples from harmonic processes, autoregressive models, and outputs from nonlinear systems are shown to illustrate this relationship. Finally, the use of recurrence plots for detection of deterministic signals in the presence of noise is investigated and compared to traditional signal detection methods based on the likelihood ratio test. Results using simulated data show that detectors based on certain statistics derived from recurrence plots are sub-optimal when compared to well-known detectors based on the likelihood ratio.

Keywords: Recurrence plots, stochastic processes, detection

PACS codes: 05.45.Tp, 02.50.Fz

*Currently at Center for Bioimage Informatics, Biomedical Engineering Department, Carnegie Mellon University.

†Corresponding author: HH C 122, Center for Bioimage Informatics, Biomedical Engineering Department, Carnegie Mellon University. 5000 Forbes Ave., Pittsburgh, PA. 15213. Email: gustavor@cmu.edu

1 Introduction

Since its introduction by Eckmann and Ruelle [1] the recurrence plot has emerged as a useful tool in the analysis of nonlinear, non-stationary time series. As the name suggests, a recurrence plot provides a graphical picture of the times at which a process will return (recur) to a given state. Given a real-valued time function $x(t)$ a recurrence plot is built by first 'embedding' the function in a multi-dimensional space by constructing a vector

$$\mathbf{u}(t) = \{x(t), x(t+\tau), \cdots, x(t+(d-1)\tau)\}^T, \tag{1}$$

where d and τ are the embedding dimension and delay, respectively [1]. The recurrence plot (RP) is then usually defined as a matrix $R(t,s) = H(\epsilon - \|\mathbf{u}(t) - \mathbf{u}(s)\|)$, where $H(x)$ is the Heaviside function, ϵ is a chosen threshold, and $\|\mathbf{a}\|^2 = \mathbf{a}^T\mathbf{a}$ is the standard 2 norm for vectors. The definition described above implies that the recurrence plot is a binary plot, where a dot is placed in the coordinate t, s for which $\mathbf{u}(t)$ and $\mathbf{u}(s)$ are relatively close to each other.

The original motivation for such a plot was to gain insight into the time scales involved in dynamical processes and to provide a technique for detecting nonstationarities (e.g. parameter drift). Subsequently, recurrence plots and the various metrics derived from them have been used for a variety of purposes such as the study of nonlinear chaotic systems [2, 3, 4], data denoising [5], as well as analysis of sound signals [6]. The emergence of Recurrence Quantification Analysis (RQA)[7] has furthered the use of recurrence plots and has resulted in a number of recurrence-based studies in the literature. RQA consists of computing various statistics that reflect the composition of recurrence plots by quantifying both the number of recurrences and the length of time a process will remain correlated (remain in the same state) with itself. Studies making use of RQA include the analysis of heart-rate-variability data [8], analysis of protein sequences [9], as well as detection of deterministic signals buried in noise [10, 11, 12] among many others. Other works relating RQA metrics to estimates of invariant measures include [13, 14]. In this work, however, we are not interested in the problem of estimating such measures, but rather in the problem of detecting a signal buried

in noise. For this different application a different type of analysis is required. Rather than treat the signals as the output of a deterministic, dynamical system we explore the statistical properties of the recurrence plot.

Though it has often been postulated that recurrence plots contain some kind of 'time correlation information' (see [1] for example) here we make this relationship more precise. Many researchers recognize that uncertainty due to noise and other causes has some effect on the computation of recurrence plots. However, not a great deal of information about the stochastic properties of recurrence plots is available. Thiel *et al.* [15] study the influence of observational noise on recurrence plots and the effect it has on quantities derived from them. Casdagli [2] proposed so called *meta* recurrence plots (recurrence plots modified by local averaging of the thresholded recurrence plot) to reconstruct the driving source of a dynamical system.

Here we describe the stochastic properties of recurrence plots by drawing analogies from the second order theory of random processes (which includes observational white noise) with the ultimate goal of characterizing the performance of RP-based statistics for standard deterministic signal detection problems. We show that, in many practical circumstances described in detail below, the expected value of the entries of the unthresholded, squared RP (USRP) can be expressed in terms of the variance and covariance of an underlying random process. In addition, we show that in two specific cases (instances of harmonic processes with specific embedding, or instances of stationary, zero mean, processes with large embedding dimension) the entries of USRPs can be written as functions of the estimates of the covariance matrix for the random process.

More specific information about the settings under which the equivalence between covariance and unthresholded recurrence plots is given below. We study unthresholded recurrence plots of several kinds of random processes, stationary and non-stationary, and arising from linear as well as nonlinear systems with the end goal of characterizing the performance of signal detection methods based on recurrence plots. Finally, we study the application of techniques derived from recurrence plot analysis to the problem of detection of deterministic signals buried in noise in an effort to characterize the performance of recurrence-based

detectors proposed recently elsewhere [10, 11, 12]. We show that the performance of many recurrence-based detectors compares unfavorably with more traditional approaches.

2 Recurrence plots of stochastic processes

There are several good reasons for investigating the stochastic properties of recurrence plots due to the fact that many real-world signals (say a voltage, displacement, or price index) are at least partially described by random phenomena, with noise being the primary example. First consider the autoregressive model of order 1 (AR1) (see subsection 2.1 below), a widely used model. Recurrence plots of two instances of the same random process are displayed in Figure 1. By looking at this figure it is easy to understand why so many researchers have become enamored with the technique. It produces interesting graphical descriptions of a data series (made even more interesting by the fact that recurrence plots are usually constructed to be symmetric). However, recurrence plots are highly sensitive to several features such as embedding dimension d, the time delay used in the embedding τ, and choice of threshold ϵ. A small change in one of these parameters can change the appearance of recurrence plots significantly. In addition to the strong dependence on parameters d, τ, and ϵ, recurrence plots tend to be highly sensitive to randomness in the signals. For example, the recurrence plots shown in Figure 1 are very different from one another, even though the data used in both came from the same process. Another good reason to investigate the stochastic properties of recurrence plots is related to the application of detection of deterministic signals in noise, the main objective of this paper. In this application one is usually interested in constructing detectors (receivers) that are at the same time powerful and significant (these concepts are explained in more detail below) in a statistical sense. Therefore expectations (in the sense of ensemble averages) of detection statistics become necessary.

For the reasons delineated above (and also because they are theoretically easier to analyze) in this article we will study, mostly, USRPs. That is, instead of analyzing $R(t,s) = H(\epsilon - \|\mathbf{u}(t) - \mathbf{u}(s)\|)$ we analyze

$$D(t,s) = \|\mathbf{u}(t) - \mathbf{u}(s)\|^2. \tag{2}$$

We note that unthresholded recurrence plots have been used previously in the literature [3].

We begin by letting a real-valued time function $x(t)$ represent a continuous random process. Note that we use the same notation to denote a random process and a realization of the random process. More precisely, we consider each $x(t)$ to be a random variable with mean $\mu(t) = \mathrm{E}\{x(t)\}$, where the expectation $\mathrm{E}\{g(\xi)\} = \int g(\xi)\mathrm{pr}[\xi]d\xi$, with $\mathrm{pr}[\xi]$ representing the probability density function of random variable ξ. The variance of $x(t)$ is $\sigma^2(t) = \mathrm{E}\{(x(t) - \mu(t))^2\}$ while the covariance between $x(t)$ and $x(s)$ is given by:

$$C(t,s) = \mathrm{E}\{(x(t) - \mu(t))(x(s) - \mu(s))\} = \mathrm{E}\{x(t)x(s)\} - \mu(t)\mu(s). \tag{3}$$

The expectation value of the USRP can be written as:

$$\mathrm{E}\{\|\mathbf{u}(t) - \mathbf{u}(s)\|^2\} = \mathrm{E}\{\|\mathbf{u}(t)\|^2\} - 2\mathrm{E}\{\mathbf{u}^T(t)\mathbf{u}(s)\} + \mathrm{E}\{\|\mathbf{u}(s)\|^2\}. \tag{4}$$

For the case $d = 1$, and if $\mu(t) = 0$ for all t (i.e., zero mean random process), then it is easy to see that:

$$\frac{1}{d}\mathrm{E}\{D(t,s)\} = \sigma^2(t) - 2C(t,s) + \sigma^2(s). \tag{5}$$

Clearly, the same holds for zero mean, second order stationary random processes (stationary in mean and covariance):

$$\frac{1}{d}\mathrm{E}\{D(t,s)\} = 2(\sigma^2(t) - C(t,s)), \tag{6}$$

regardless of the choice of d and τ. Finally, equation (5) also holds, approximately, for random processes that are zero mean and locally stationary, so long as d and τ are chosen such that each vector $\mathbf{u}(t)$ contains data from a finite region whose covariance matrix is approximately stationary.

The analysis above links unthresholded recurrence plots $D(t,s)$ to variance and covariance of random processes. That is, if we had access to multiple, statistically independent, sample functions $x(t)$, the ensemble average of the USRP, under the conditions described above, would be a simple function of the ensemble variance and covariance of the data series. We

should point out that typically RPs are used to analyze single realizations of a signal (as opposed to an ensemble). In this case, RP analysis does not necessarily equate to second order statistics and can in fact be used to explore higher-order correlations among time series data.

2.1 Linear Autoregressive Processes

Consider a first order linear autoregressive discrete random process (AR1) given by

$$x(n) = ax(n-1) + e(n), \tag{7}$$

where $n = 1, 2, \cdots$ and $e(n)$ are zero mean, uncorrelated, normally distributed random variables with variance σ_e^2. If $|a| < 1$, then the random process above is asymptotically stationary to second order and its covariance is given, approximately, by [16]:

$$C(i,j) = \sigma_e^2 \frac{a^{|i-j|}}{(1-a^2)}. \tag{8}$$

Naturally, the variance of this random process is given by, approximately, $\sigma_x^2 = \sigma_e^2/(1-a^2)$. Figure 2 (top) displays a sample function of the random process above with $x_0 = 0$, $a = 0.8$, and $\sigma_e^2 = 1$. Note that, unless noted otherwise, all axis in the figures contain time information. The time index in this figure was chosen so that $t = n/10^3$ and $n = 1, \cdots, 2^{10}$. The expectation of the unthresholded recurrence plot $\mathrm{E}\{D(i,j)\}$ was computed by generating an ensemble of 1000 replicates of the random process above and then computing its average. This plot is displayed on the bottom left panel of figure 2. For these computations we set $d = 3$ and $\tau = 2$ ($2/10^3$ in real time coordinates). Note that in regions where $|i-j|$ is large (regions which correspond to the regions far away from the diagonal in the recurrence plot), according to equation (8) one must have that $\sigma^2(i) - 2C(i,j) + \sigma^2(j) = 5.56$, which, by looking at the bottom left panel plot, agrees fairly well with the results obtained. The absolute error between $\mathrm{E}\{D(i,j)\}/d$ and $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$ (estimated from the same ensemble) is shown on the bottom right panel. The mean value of $\mathrm{E}\{D(i,j)\}/d$ (sum of

$\mathrm{E}\{D(i,j)\}/d$ over all indices i,j divided by N^2) is 5.49 while the mean value of the error plot is 0.2. It is possible to reduce the magnitude of the error by increasing the number of replicates in the ensemble.

Note that, because the random process above is asymptotically stationary and ergodic, one may substitute ensemble averages with time averages. Therefore it is possible to compare $D(i,j)/d$ (without taking expectations) and $\sigma^2(i)-2C(i,j)+\sigma^2(j)$ directly simply by choosing a large embedding dimension d. Since $D(i,j)/d = \|\mathbf{u}(i)\|^2/d - 2\mathbf{u}^T(i)\mathbf{u}(j)/d + \|\mathbf{u}(j)\|^2/d$ and (noting the mean of the random process is zero)

$$\lim_{d\to\infty} \|\mathbf{u}(i)\|^2/d = \lim_{d\to\infty} \frac{1}{d}\sum_{k=1}^{d} |u_k(i)|^2 \sim \sigma^2(i) = \sigma_e^2/(1-a^2), \tag{9}$$

$$\lim_{d\to\infty} -2\mathbf{u}(i)^T\mathbf{u}(j)/d = \lim_{d\to\infty} -2\frac{1}{d}\sum_{k=1}^{d} u_k(i)u_k(j) \sim -2C(i,j) = -2\sigma_e^2\frac{a^{|i-j|}}{(1-a^2)}. \tag{10}$$

In Figure 3 we repeat the experiment using $d = 400$, $\tau = 1$. As before, the error between $D(i,j)/d$ and $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$ can be reduced by increasing d. Here the mean value of the recurrence plot was 4.6 while the mean value of the error 0.6.

The autoregressive process above can be modified so that it is no longer stationary by having a vary with time:

$$x(n) = a(n)x(n-1) + e(n). \tag{11}$$

If the coefficients $a(n)$ do not vary rapidly, the covariance matrix for this random process can be seen as locally stationary. Using a sample function $a(n)$ that varies smoothly between 0 and 0.9 we have computed the expected value of $D(i,j)$ and $\sigma^2(i)-2C(i,j)+\sigma^2(j)$ estimated from an ensemble of 1000 realizations. Results are shown in figure 4. $\sigma^2(i)-2C(i,j)+\sigma^2(j)$ is shown on the left and $\mathrm{E}\{D(i,j)\}$ on the middle panel. The error is shown on the right panel. Care was taken to choose $d = 3$ and $\tau = 2$ so that each vector $\mathbf{u}(j)$ contained data from a small region so that the local stationarity assumption is approximately satisfied.

2.2 Harmonic Processes

Harmonic models have traditionally been used to model cyclic phenomena which can be well approximated as sums of sines and cosines. They have been extensively used, for example, in modeling electromagnetic waves used in radar and communications applications. One harmonic random process often studied is

$$x(t) = \sum_{k=1}^{K} A_k \cos(w_k t + \phi_k) \tag{12}$$

where A_k, w_k are constants and ϕ_k are independent random variables uniformly distributed in the interval $(-\pi, \pi)$. This is a stationary random process. $\mathrm{E}\{x(t)\} = 0$ for all t while the covariance of this random process is

$$C(t,s) = \sum_{k=1}^{K} \frac{1}{2} A_k^2 \cos(w_k |t-s|). \tag{13}$$

Hence,

$$\sigma^2 = \sum_{k=1}^{K} \frac{1}{2} A_k^2. \tag{14}$$

Clearly equation (6) holds as before regardless of the choice of d and τ because the process is stationary. The simulation experiment described above was repeated this time using $K = 3$, $A_1 = 2$, $A_2 = 3$, $A_3 = 0.3$, $\Delta t = 1/10^3$, $d = 3$, $\tau = 2$, $w_1 = 20$, $w_2 = 36$, $w_3 = 48$, $t = n\Delta t$, $n = 1 \cdots 2^{10}$, and 1000 replicates as before. Each ϕ_k was chosen to be uniformly distributed between $-\pi$ and π. Results are shown in figure 5. A sample signal is shown on top, $\sigma^2(t) - 2C(t,s) + \sigma^2(s)$ is shown on the bottom left panel, and the expected value of the unthresholded RP is shown on the bottom center panel. The error is shown on the right panel. The average error was 0.27 while the average value of $\mathrm{E}\{D(i,j)\}/d$ was 13.1. Note that, since in this case the process is stationary, the choices for the parameters described above is arbitrary. That is, the equivalence holds in general settings.

As before, it is not necessary to take expectations to relate recurrence plots and covariance. For example, consider the process given in equation (12) with $w_k = k\beta$. One may check by substitution (see appendix A) that if $d > 2K$ and $\tau = \frac{2\pi}{\beta d}$ then unthresholded recurrence

plots equate to a simple function of the covariance structure for the random process, without taking expectations:

$$\frac{1}{d}D(t,s) = \sigma^2(t) - 2C(t,s) + \sigma^2(s). \tag{15}$$

2.3 Duffing System

In the literature recurrence plots are often used to study the correlation structure in nonlinear (often chaotic) signals. For this reason we use recurrence plots to analyze the response of a chaotic Duffing system

$$\frac{d^2x}{dt^2} - c\frac{dx}{dt} + x(1+x^2) = f(t) \tag{16}$$

where $f(t)$ is the simple harmonic random process:

$$f(t) = A\cos(wt + \phi), \tag{17}$$

with ϕ being a random variable uniformly distributed between $-\pi$ and π. The Duffing system was solved numerically using $c = 0.25$, $A = 0.4$, and $w = 1$ in the time interval between 0 and 100, taking only samples after time 10 to allow for transients to fade away. The time samples were evenly distributed between 10 and 100 while the data length again was 1024. Again we wish to compare the quantities $\mathrm{E}\{D(t,s)\}$ defined in equation (2) and $\sigma^2(t) - 2C(t,s) + \sigma^2(s)$ by generating an ensemble of 1000 realizations. Results are shown in figure 6. The top part contains a sample signal $x(t)$, while the bottom left panel displays $\mathrm{E}\{D(t,s)\}$. The bottom right panel contains $\sigma^2(t) - 2C(t,s) + \sigma^2(s)$ as estimated from the ensemble. The average error was 0.1 while the average value of $\mathrm{E}\{D(i,j)\}/d$ was 1.5.

In this simulation the embedding dimension was 3 while the time delay was 1. As shown in this figure, both quantities contain very similar information. More precisely, unlike the AR1 random process for example, the output of a Duffing system when excited by a harmonic process contains time correlations which are not simple decaying functions, but contain fluctuations. However, as shown in this experiment, the information content of the covariance of the random process and the expectation value of the unthresholded recurrence plot is very

similar.

3 Detection of Deterministic Signals in Noise

Detecting whether an incoming signal contains information from a deterministic physical source (as opposed to random thermal noise, for example) is an important task with numerous applications. Recurrence plots have been used for such a task in Zbilut and others [10, 11, 12]. By drawing from the theory developed above we now compare a few detectors based on recurrence plots with more traditional detectors based on likelihood ratio tests. We note that in our study we do not perform an exhaustive comparison between detectors based on likelihood ratios and recurrence-based time series analysis methods. Many works describing advanced methods for recurrence-based time series analysis exist (see [13] for an example). In this section our goal is to characterize the performance of recurrence-based methods for detecting deterministic signals already proposed elsewhere, such as in the work of Zbilut and colleagues [10, 11, 12], in comparison with more traditional methods developed in the electrical engineering literature.

In its simplest form, the detection problem can be cast as follows. Given a data series $\mathbf{x} = \{x(1), \cdots, x(N)\}^T$ (note that without loss of generality we take Δt to be 1) we ask which hypothesis is most likely to have occurred:

$$\begin{aligned} H_0 &: \text{signal absent } [\mathbf{x} = \mathbf{e}] \\ H_1 &: \text{signal present } [\mathbf{x} = \mathbf{s} + \mathbf{e}] \end{aligned} \tag{18}$$

where $\mathbf{e} = \{e(1), \cdots, e(N)\}^T$ represents a noise vector while $\mathbf{s} = \{s(1), \cdots, s(N)\}^T$ represents a deterministic signal from a physical source. If enough is known about the deterministic and stochastic parts of the data series $\mathbf{x}$ under both hypotheses a likelihood ratio test can be performed

$$\lambda(\mathbf{x}) = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} \gtrless \eta \tag{19}$$

where $p_1(\mathbf{x})$ is the likelihood function for the data series under hypothesis H_1 and $p_0(\mathbf{x})$ is the

likelihood function for the data series under H_0. From (19) is is possible to derive statistics $\lambda(\mathbf{x})$ which can be chosen to be optimal in several senses. The threshold η can be chosen according to a variety of criteria. One may choose to minimize the probability of making an error, or, as in a Neyman-Pearson test for example, a threshold can be chosen to satisfy some probability of false alarm. The probability of false alarm P_{fa} (in statistics known as the significance of a test) is defined as the probability of deciding that a signal is present when in fact it is not. It can be calculated directly from the test statistic $\lambda(\mathbf{x})$:

$$P_{\text{fa}} = \int_{\eta}^{\infty} p_0(\lambda(\mathbf{x}))d\lambda(\mathbf{x}) \tag{20}$$

where $p_0(\lambda(\mathbf{x}))$ is the probability density function, under H_0, of the statistic $\lambda(\mathbf{x})$. In essence, it computes the probability that, given hypothesis zero is true, the statistic λ would be greater than the threshold η. If η represents the decision boundary, this would be an error (false alarm), and hence the terminology false alarm.

Similarly, the probability of detection given a deterministic signal is present P_{d} (in statistics known as the power of a test) is defined as the area under $p_1(\lambda(\mathbf{x}))$ to the right of the decision boundary:

$$P_{\text{d}} = \int_{\eta}^{\infty} p_1(\lambda(\mathbf{x}))d\lambda(\mathbf{x}) \tag{21}$$

where $p_1(\lambda(\mathbf{x}))$ is the probability density function of the statistic $\lambda(\mathbf{x})$ under H_1. This defines the probability that, given that hypothesis one is true, the detection statistic would correctly determine that a signal is present.

A curve plotting P_{fa} versus P_{d}, known as a Receiver Operating Characteristic (ROC) curve, is a good measure of the performance of a particular receiver. The ROC curve is, in fact, one standard platform for comparing different detectors. For a fixed P_{fa}, a higher probability of detection P_{d} represents a better detector. Thus an optimal detector is the one for which, given some probability of false alarm P_{fa}, the probability of detection P_{d} is maximized. The problem of optimal signal detection has been extensively developed in the engineering literature. For comprehensive references, including plots of $p_1(\lambda(\mathbf{x}))$ and $p_0(\lambda(\mathbf{x}))$ for different experimental setups, see [17, 18].

We will compare detectors based on unthresholded recurrence plots with detectors based on the likelihood ratio framework described above under two general circumstances: when the signal $\mathbf{s}$ is known and when it is not known. We will use several types of signals derived from harmonic processes (simple sinusoids and linear chirps) as well as signals arising from the Duffing nonlinear system.

When the deterministic signal $\mathbf{s}$ one is trying to detect is known, and assuming that $\mathbf{e}$ is a sequence of identically and independently distributed (iid) normal random variables (white noise), the likelihood ratio test reduces to the well-known correlation receiver (see [17] for a derivation)

$$\frac{1}{\sigma^2}\mathbf{x}^T\mathbf{s} \gtrless \ln(\eta) + \frac{1}{2\sigma^2}\|\mathbf{s}\|^2. \tag{22}$$

where σ^2 is the variance of the noise. In cases where not enough information is available to characterize the probability density function of an incoming deterministic signal alternative approaches must be used. For example, for detecting a sampled sinusoid, $n = 1, \cdots, N$ with unknown frequency and phase an average likelihood ratio test (assuming a uniform distribution of frequencies and phases) is often used (see [17], for example)

$$\sum_{m=1}^{N} q_m^2 \gtrless V \tag{23}$$

where V (to be chosen given some optimality criterion) is a scalar threshold and

$$q_m = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \exp\left(\frac{\imath 2\pi n m}{N}\right). \tag{24}$$

If $x(n)$ is zero mean, testing for $\sum_{m=1}^{N} q_m^2$ is equivalent to testing for the variance (power) of the signal.

3.1 Detection of unknown signals

Let us start by looking at the problem of detecting a cosine signal sampled at $n = 1, \cdots, N$ with unknown frequency and phase. Zbilut and colleagues suggest using quantities computed

based on thresholded recurrence plots such as the percent recurrence, defined as:

$$\%\text{REC} = \frac{1}{N^2} \sum_{i,j=1}^{N} H\left(\epsilon - \sqrt{D(i,j)}\right), \tag{25}$$

where $H(\cdot)$ is the Heaviside function, as well as the percent determinism, defined as:

$$\%\text{DET} = \frac{\sum_{l=l_{min}}^{l_{max}} l \cdot P(l)}{\sum_{l=1}^{l_{max}} l \cdot P(l)} \tag{26}$$

where $P(l)$ is the distribution (usually computed via histogram binning) of diagonal lines of length l in a thresholded recurrence plot. l_{min} was set to 2 and l_{max} was set to the maximum possible, given the finite size of the data series, and consequently the recurrence plot. In addition to these, we have also tested the average line length (ALL) in the recurrence plot as a statistic for detection. Finally, note that these detectors can be built based on diagonal lines of the recurrence plot [10, 12] or vertical lines [8].

In Figure 7 we plot ROC curves computed based on the statistics described above, as well as the ROC curve based on the average value of the unthresholded recurrence plot. In this simulation a simple cosine random process (12) was used with $K = 1$, $w_1 = 20$, $\Delta t = 1$, $d = 4$, $\tau = 12$, and $\epsilon = 1.5$. Two experiments are shown. In the first (top part of the figure) the %DET and ALL detectors were based on diagonal lines, while on the bottom figure they were based on vertical lines. We note that results obtained using this specific setting for the parameters are representative of a larger series of experiments where we have varied the parameters d, τ and ϵ. We omit the plot of additional results for brevity. For all experiments tried the detector based on the average, unthresholded, recurrence plot outperforms the detectors based on %REC, %DET and ALL. Therefore, from here on we focus exclusively on the average recurrence as a statistic to be used in detecting signals.

A few notes about the selection of parameters d, τ, and ϵ are appropriate. For the purpose of attractor reconstruction, and estimation of physical quantities related to it, general guidelines for the settings of the parameters d, τ are are provided by Takens [19]. However, for some applications such as the estimation of dynamical invariants embedding appears to play a diminished role [20, 21]. The role of a "proper" embedding with regard to the

signal detection problem has not yet been explored. Stark et. al. [22] suggest embedding stochastic systems is still appropriate in some cases but there exists few guidelines in this case. In the absence of guidelines for choosing these parameters for the problem of detecting signals buried in noise we tested several settings experimentally. As reported above, in the cases we tried, we did not find settings so that detectors based on binary recurrence plots would outperform the simple average, unthresholded, squared recurrence. In addition, the performance of the average, unthresholded, squared recurrence detector remains constant for the choices of embedding we tested.

The detector based on the average (unthresholded) recurrence plot is defined as:

$$\lambda(\mathbf{x}) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} D(i,j). \tag{27}$$

Taking the expectation of the expression above and using the result in equation (5) we have:

$$\mathrm{E}\left\{\lambda(\mathbf{x})\right\} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathrm{E}\left\{D(i,j)\right\} = \frac{d}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{\sigma^2(i) - 2C(i,j) + \sigma^2(j)\right\}. \tag{28}$$

The sufficient conditions for the equation above to hold were explained earlier. To summarize, the equation above holds for any kind of zero mean random process if $d = 1$. For stationary processes, zero mean, (such as the harmonic random process), the result holds for any choice of d and τ. For non-stationary, zero mean, processes, the result above holds, approximately, if $d\tau$ is smaller than the length of time over which the covariance $C(i,j)$ changes appreciably.

Equation (28) above links the unthresholded average recurrence detector to the power detector described earlier. Therefore we should expect that both detectors should perform similarly when it comes to detecting unknown signals. In fact this is exactly the result obtained for the three signals investigated here (cosine, linear chirp, and Duffing). The results are shown in Figure 8. The cosine wave was computed using the random process definition in (12). The parameters used were $w = 20$, $K = 1$, while $d = 4$, $\tau = 12$, and $\epsilon = 1.5$. The chirp wave was computed such that the frequency varied linearly between 5 Hz at time 0 and 40 Hz at time 1. The phase was randomly selected between $(-\pi, \pi)$. In

this experiment the embedding dimension was set to 4, and the delay to 6, while $\epsilon = 1.5$. Finally, the Duffing system was solved as explained earlier. Embedding dimension was set to 4, $\tau = 15$, and $\epsilon = 3.5$. In all cases the signal to noise ratio (variance of signal over variance of noise) was 0.1. The results in this figure demonstrate, in these specific cases, that there is no advantage, from a statistical point of view, to using detectors based on %REC, %DET and ALL over the simplest of detectors such as the power detector, whether it is implemented in the time or frequency domain.

3.2 Detection of known signals

When the form of the deterministic signal $\mathbf{s}$ one is trying to detect is known Zbilut and colleagues [10, 11, 12] have proposed using cross-recurrence plot (CRP) as a tool for detecting signals. A CRP follows basically the same construction as presented earlier with the difference that now:

$$R_{i,j} = H(\epsilon - \|\mathbf{u}(i) - \mathbf{v}(j)\|) \tag{29}$$

where $\mathbf{u}(i)$ is the embedded signal $\mathbf{x}$ while $\mathbf{v}(j)$ is the embedded probe (the known signal) $\mathbf{y}$. As before we compare the detector based on the statistic

$$\lambda(\mathbf{x}) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \|\mathbf{u}(i) - \mathbf{v}(j)\|^2 \tag{30}$$

to the correlation detector given in equation (22). The comparison is given in Figure 9. The signal here was the same cosine wave described earlier. The phase term, however, was set to zero so that the signal was always in phase with the probe, as done in [11]. Clearly, the CRP receiver is outperformed by the traditional correlation receiver by a significant amount. In fact this is to be expected since the correlation receiver, in this experiment, can be shown to be optimum in the sense that it maximizes the probability of detection for a given probability of false alarm [17].

4 Summary and Conclusions

We have used well-known tools from the second order theory of random processes to investigate the stochastic properties of recurrence plots. We have shown that for some types of random processes, the expected value of the entries of USRPs can be expressed in terms of the variance and covariance of the random process. This relationship is expressed in equation (5), which is appropriate when any the following holds:

- if the embedding dimension is set to 1 (no embedding);
- for any embedding dimension d and delay τ so long as the random process is stationary to second order;
- when processes are non-stationary but locally stationary, the relationship holds, approximately, if d and τ are chosen so that $d\tau$ is smaller than the period of time over which the covariance of the random process changes appreciably.

We have also shown that the relationship (5) holds exactly for some choices of d and τ for stationary harmonic processes, even without taking expectations (equation (15)). For that one must choose the embedding dimension to be greater than twice the number of sinusoids in the random process ($d > 2K$), and the delay must be $\tau = 2\pi/(\beta d)$.

Note that when the conditions delineated above are not met, unthresholded recurrence plots, even in expectation, may not equate to second order analysis. In fact it is likely that these two quantities can be significantly different. Consider the case of a linear chirp with random phase uniformly distributed between 0 and 2π, an instance of which is shown in Figure 10. The signal consists of a single chirp varying from 1 Hz to 100 Hz in the span of 1 second. The signal was sampled with $\Delta t = 1/10^3$. The covariance function of the random process is shown in the bottom left of Figure 10. The expected value of the squared unthresholded recurrence (with $d = 3$ and $\tau = 30$) plot is shown in the middle of the bottom row. The error between these two quantities is shown on the right bottom panel in the same figure. As shown here the error between these quantities is significantly larger than the error between the same quantities when the stationarity assumption holds (cosine example shown

earlier). The average error shown is 0.5, while the average value of the squared, unthresholded recurrence plot is 1.

Note also that, even though in the settings we described above, unthresholded RP statistics equate to second order statistics of random process this does not mean that quantities derived from RP in different settings (in the study of dynamical systems, for example) equate to second order analysis. For example, the concept of correlation dimension often used to characterize time series arising from dynamical systems (which can be computed using a recurrence plot [23]) is not the same as the correlation (defined through ensemble expectations) between random variables $x(t_1)$ and $x(t_2)$. Naturally these quantities may be related through the fact that they can often be computed from the same data series. However, the concepts are different and one must be careful not to allow the similar nomenclature to cause confusion.

We have used relationship (5) to analyze the performance of the recurrence plot as a tool for deterministic signal detection. For detection of unknown signals (signals for which a model is not available) the detector based on the average unthresholded recurrence plot (which performed better than detectors based on other thresholded recurrence statistics) is essentially equivalent to the variance (power) detector. This was confirmed by simulations using harmonic signals (sinusoids, linear chirps) as well as outputs from the Duffing nonlinear system. In addition, as a statistical test for detection of deterministic signals, the performance of the CRP detector proposed by Zbilut and colleagues [11] falls significantly below that of the traditional correlation receiver. We conclude that, while recurrence plots remain a useful tool in certain estimation problems, their performance in classical signal detection problems does not compare well with traditional approaches based on the likelihood ratio framework.

Appendix A

Here we show that for certain choices of embedding dimension d and delay τ, the recurrence plot (unthresholded) of a harmonic process is exactly equivalent to the formula (15) even without taking (ensemble) averages. Again, let the random process be defined as:

$$x(t) = \sum_{k=1}^{K} A_k \cos(k\beta t + \phi_k), \tag{31}$$

while the embedding $\mathbf{u}(t)$ is defined by equation (1). Naturally, if one wishes that

$$\begin{aligned} \|\mathbf{u}(t) - \mathbf{u}(s)\|^2/d &= \|\mathbf{u}(t)\|^2/d - 2\mathbf{u}^T(t)\mathbf{u}(s)/d + \|\mathbf{u}(s)\|^2/d \\ &= \sigma^2(t) - 2C(t,s) + \sigma^2(s), \end{aligned} \tag{32}$$

the following conditions suffice:

$$\|\mathbf{u}(t)\|^2/d = \sigma^2(t) \tag{33}$$

$$\mathbf{u}^T(t)\mathbf{u}(s)/d = C(t,s). \tag{34}$$

Using $\tau = \frac{2\pi}{\beta d}$, each component of vector $\mathbf{u}(t)$ is given by:

$$(\mathbf{u}(t))_i = x(t + (i-1)\tau) = \sum_{k=1}^{K} A_k \cos\left(k\beta t + k\frac{2\pi}{d}(i-1) + \phi_k\right).$$

First, we are interested in computing $\|\mathbf{u}(t)\|^2$:

$$\|\mathbf{u}(t)\|^2 = \sum_{i=1}^{d}\sum_{k=1}^{K} A_k^2 \cos^2\left(k\beta t + k\frac{2\pi}{d}(i-1) + \phi_k\right).$$

Using the identity:

$$\cos(u)\cos(v) = \frac{1}{2}[\cos(u-v) + \cos(u+v)], \tag{35}$$

one can show that

$$\|\mathbf{u}(t)\|^2 = \frac{1}{2}\sum_{k=1}^{K} A_k^2 \sum_{i=1}^{d}\left[1 + \cos\left(2k\beta t + 2\phi_k + 2k\frac{2\pi}{d}(i-1)\right)\right].$$

Using the fact that, if $d > 2k$,

$$\sum_{i=1}^{d} \cos\left(2k\beta t + 2\phi_k + 2k\frac{2\pi}{d}(i-1)\right) = 0$$

(this result can be shown using the identity $\sum_{m=0}^{L-1} e^{-\imath\omega m} = \frac{\sin(\omega L/2)}{\sin(\omega/2)} e^{-\imath\omega(L-1)/2}$) and equation (14) one arrives at the condition (33).

Now consider

$$\mathbf{u}^T(t)\mathbf{u}(s) = \sum_{k=1}^{K}\sum_{j=1}^{K} A_k A_j \sum_{i=1}^{d} \cos\left(k\beta s + k(i-1)\frac{2\pi}{d} + \phi_k\right) \times \cos\left(j\beta t + j(i-1)\frac{2\pi}{d} + \phi_j\right).$$

Using (35) again and re-arranging

$$\mathbf{u}^T(t)\mathbf{u}(s) = \sum_{k=1}^{K}\sum_{j=1}^{K} \frac{1}{2} A_k A_j \sum_{i=1}^{d} \cos\left(\beta(kt - js) + \phi_k - \phi_j + (i-1)\frac{2\pi}{d}(k-j)\right) + \sum_{k=1}^{K}\sum_{j=1}^{K} \frac{1}{2} A_k A_j \sum_{i=1}^{d} \cos\left(\beta(kt + js) + \phi_k + \phi_j + (i-1)\frac{2\pi}{d}(k+j)\right).$$

The first cosine sum of the equation above is only nonzero when $k = j$. If $d > 2K$ again, the second cosine sum is always zero. Therefore (taking advantage that the cosine function is even)

$$\mathbf{u}^T(t)\mathbf{u}(s) = \frac{d}{2}\sum_{k=1}^{K} A_k^2 \cos\left(k\beta|t-s|\right).$$

Dividing by d we arrive at the result (34).

To summarize, given an instance of a generic stationary harmonic random process defined in equation (31) one can compute a recurrence plot so that it is a simple function (given in equation (32)) of the covariance structure of the random process. To that end one must first choose an embedding dimension so that $d > 2K$, and then set the embedding delay to $\tau = \frac{2\pi}{\beta d}$. These are sufficient conditions for (32) to hold.

References

[1] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhysics Letters*, vol. 4, pp. 973–977, 1987.

[2] M.C. Casdagli, "Recurrence plots revisited," *Physica D*, vol. 108, pp. 12–44, 1997.

[3] G. McGuire, N.B. Azar, and M. Shelhamer, "Recurrence matrices and the preservation of dynamical properties," *Physics Letters A*, vol. 237, pp. 43–47, 1997.

[4] L.L. Trulla, A. Giuliani, J.P. Zbilut, and C.L. Webber Jr., "Recurrence quantification analysis of the logistic equation with transients," *Physics Letters A*, vol. 223, pp. 255–260, 1996.

[5] L. Matassini, H. Kantz, J. Holyst, and R. Hegger, "Optimizing of recurrence plots for noise reduction," *Physical Review E*, vol. 65, pp. 021102, 2002.

[6] A. Facchini, H. Kantz, and E. Tiezzi, "Recurrence plot analysis of nonstationary data: The understanding of curved patterns," *Physical Review E*, vol. 72, pp. 021915–1:021915–06, 2005.

[7] C.L. Webber Jr. and J. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Modeling in Physiology*, vol. 94, pp. 965–973, 1994.

[8] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, "Recurrence-plot-based measures of complexity and their application to heart-rate-variablility data," *Physical Review E*, vol. 66, pp. 026702, 2002.

[9] P. Babinec and J. Leszczynski, "Recurrence plot analysis of nonlinear vibrational dynamics in H_3^+ molecule," *Chaos Solutions and Fractals*, vol. 17, pp. 981–984, 2003.

[10] J.P. Zbilut, A. Giuliani, and Jr. C.L. Webber, "Recurrence quantification analysis and principal components in the detection of short complex signals," *Physics Letters A*, vol. 237, pp. 131–135, 1998.

[11] J.P. Zbilut, A. Giuliani, and Jr. C.L. Webber, "Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification," *Physics Letters A*, vol. 246, pp. 122–128, 1998.

[12] J.P. Zbilut, A. Giuliani, and Jr. C.L. Webber, "Recurrence quantification analysis as an empirical test to distinguish relatively short deterministic versus random number series," *Physics Letters A*, vol. 267, pp. 174–178, 2000.

[13] K Urbanowicz and J.A. Holyst, "Noise-level estimation of time series using coarse-grained entropy," *Physical Review E*, vol. 67, pp. 046218, 2003.

[14] P. Faure and H. Korn, "A new method to estimate the kolmogorov entropy from recurrence plots: its application to neuronal signals," *Physica D*, vol. 122, pp. 265–279, 1998.

[15] M. Thiel, M.C. Romano, J. Kurths, R. Meucci, E. Allaria, and F.T. Arecchi, "Influence of observational noise on the recurrence quantification analysis," *Physica D*, pp. 138–152, 2002.

[16] M.B. Priestley, *Spectral Analysis and Time Series*, Academic Press, New York, 1981.

[17] A.D. Whalen, *Detection of Signals in Noise*, Academic Press, New York, 1971.

[18] H.L. Van Trees, *Detection, Estimation, and Modulation Theory - Part III*, Wiley, New York, 1971.

[19] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, D. Rand and L.-S Young, Eds., p. 366. Springer, 1981.

[20] J. S. Iwanski and E. Bradley, "Recurrence plots of experimental data: To embed or not to embed?," *Chaos*, vol. 8, pp. 861–871, 1998.

[21] M. Thiel, M. C. Romano, P. L. Read, and J. Kurths, "Estimation of dynamical invariants without embedding by recurrence plots," *Chaos*, vol. 14, pp. 234–243, 2004.

[22] J. Stark, D. S. Broomhead, M. E. Davies, and J. Huke, "Takens embedding theorems for forced and stochastic systems," *Nonlinear Analysis Theory Methods Applications*, vol. 30, pp. 5303–5314, 1997.

[23] J. Gao and H. Cai, "On the structures and quantification of recurrence plots," *Physics Letters A*, vol. 270, pp. 75–87, 2000.

Figure captions

Figure 1

Thresholded recurrence plots of two instances of the same (AR) random process. Even though the random process parameters are the same the recurrence plots can look dramatically different.

Figure 2

Stochastic analysis of AR1 using unthresholded recurrence plots. The top part shows a sample signal from an AR1 random process. The bottom left panel shows the expected value of the squared, unthresholded, recurrence plot computing from an ensemble of 1000 repetitions. The bottom right panel shows the error between the expected value of the recurrence plot and $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$ (estimated from the same ensemble). See text for a description of the parameters used in this simulation.

Figure 3

Analysis of a single realization of the AR1 process using recurrence plots. A sample function is shown on top while the bottom left panel shows a single unthresholded, squared, recurrence plot. The bottom right panel shows the error between the unthresholded, squared, recurrence plot and $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$.

Figure 4

Stochastic analysis of nonstationary AR1 using unthresholded recurrence plots. The panel on the left shows $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$, computed from an ensemble of 1000 realizations. $\mathrm{E}\{D(i,j)\}/d$ is shown in the middle panel and the error is shown on the right panel.

Figure 5

Stochastic analysis of cosine harmonic process. A sample function is shown on top. $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$, computed from an ensemble of 1000 realizations, is shown on the bottom left, while $\mathrm{E}\{D(i,j)\}/d$ is shown in the middle panel. The error is shown in the bottom right panel.

Figure 6

Stochastic analysis of Duffing system. A sample function is shown on the top of the figure. $\sigma^2(i) - 2C(i,j) + \sigma^2(j)$, computed from an ensemble of 1000 realizations, is shown on the bottom left, while $\mathrm{E}\{D(i,j)\}/d$ is shown in the middle panel. The error is shown in the bottom right panel.

Figure 7

Detector statistics for simple cosine wave (SNR = 0.1).

Figure 8

ROC curves comparing power detector with average recurrence detector for cosine (top), chirp (middle) and Duffing (bottom) signals.

Figure 9

ROC curves comparing CRP receiver with the well-known correlation receiver.

Figure 10

Expectation of unthresholded recurrence plot (middle of bottom row) for a linear chirp random process (top). In this case the error between the recurrence plot approximated using the covariance of the random process can be large.

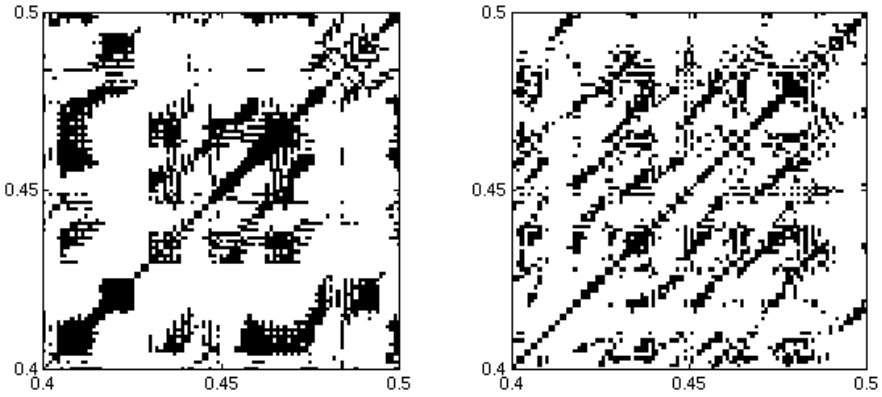

Figure 1:

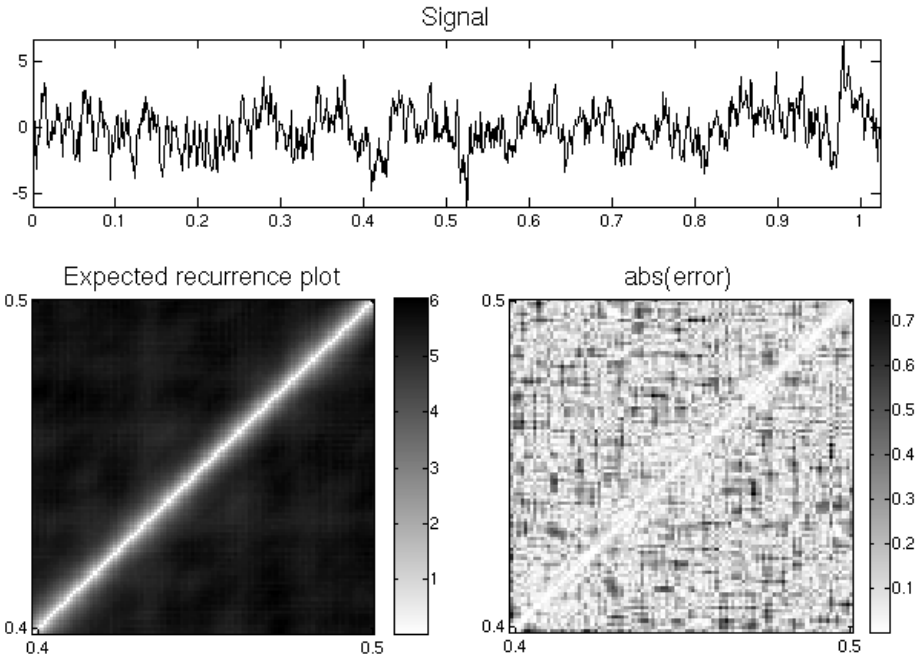

Figure 2:

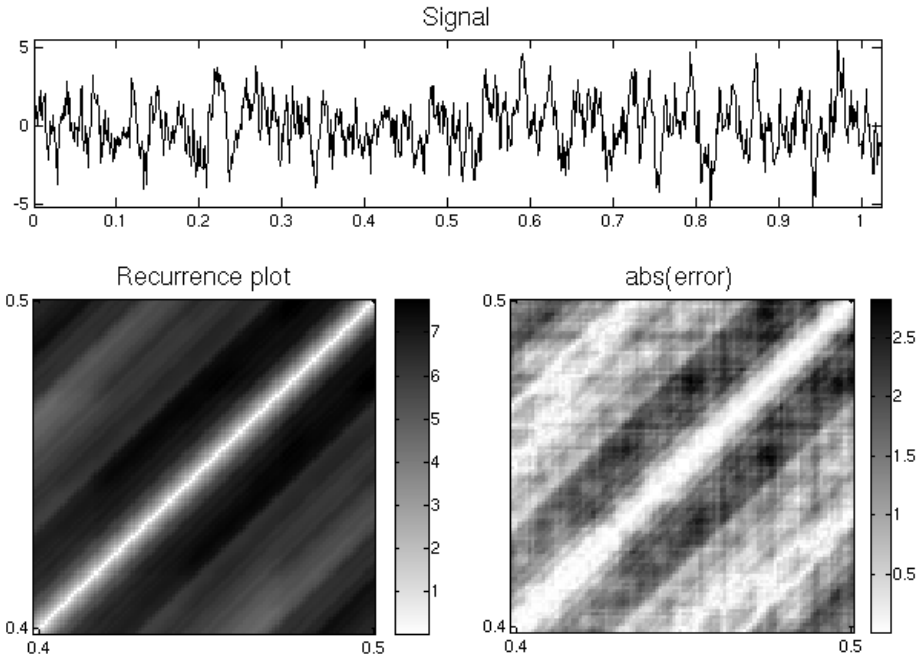

Figure 3:

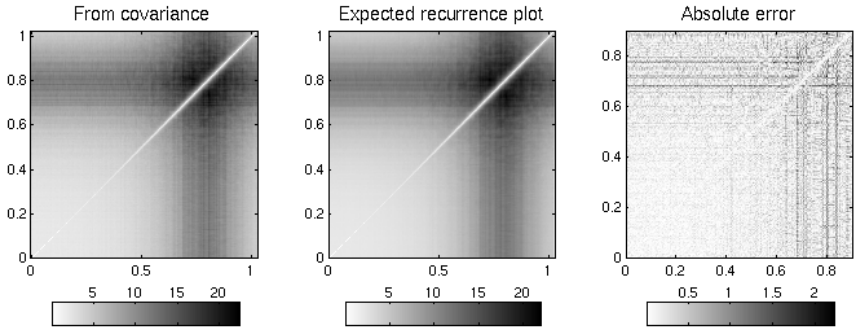

Figure 4:

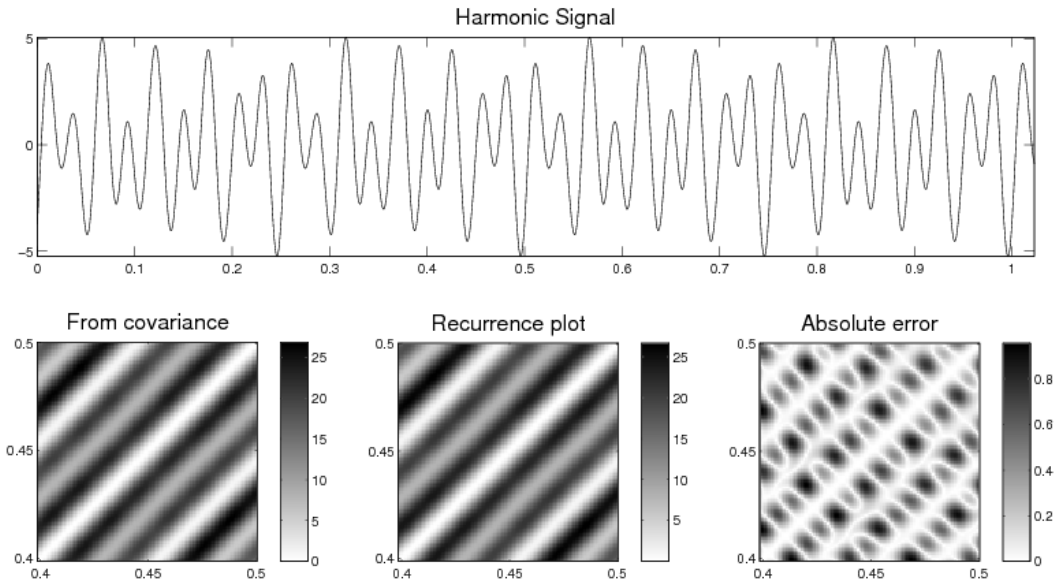

Figure 5:

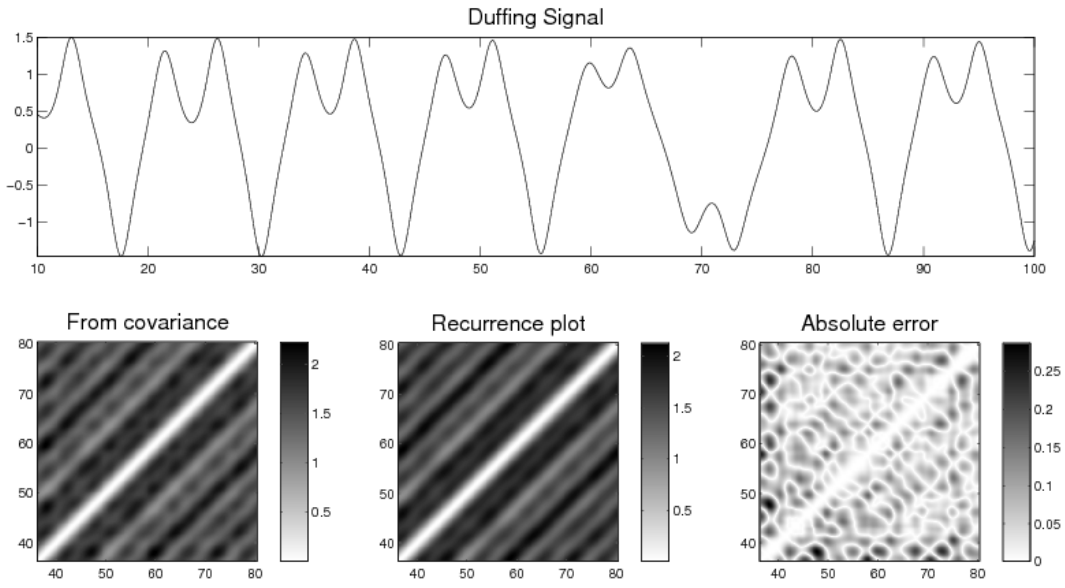

Figure 6:

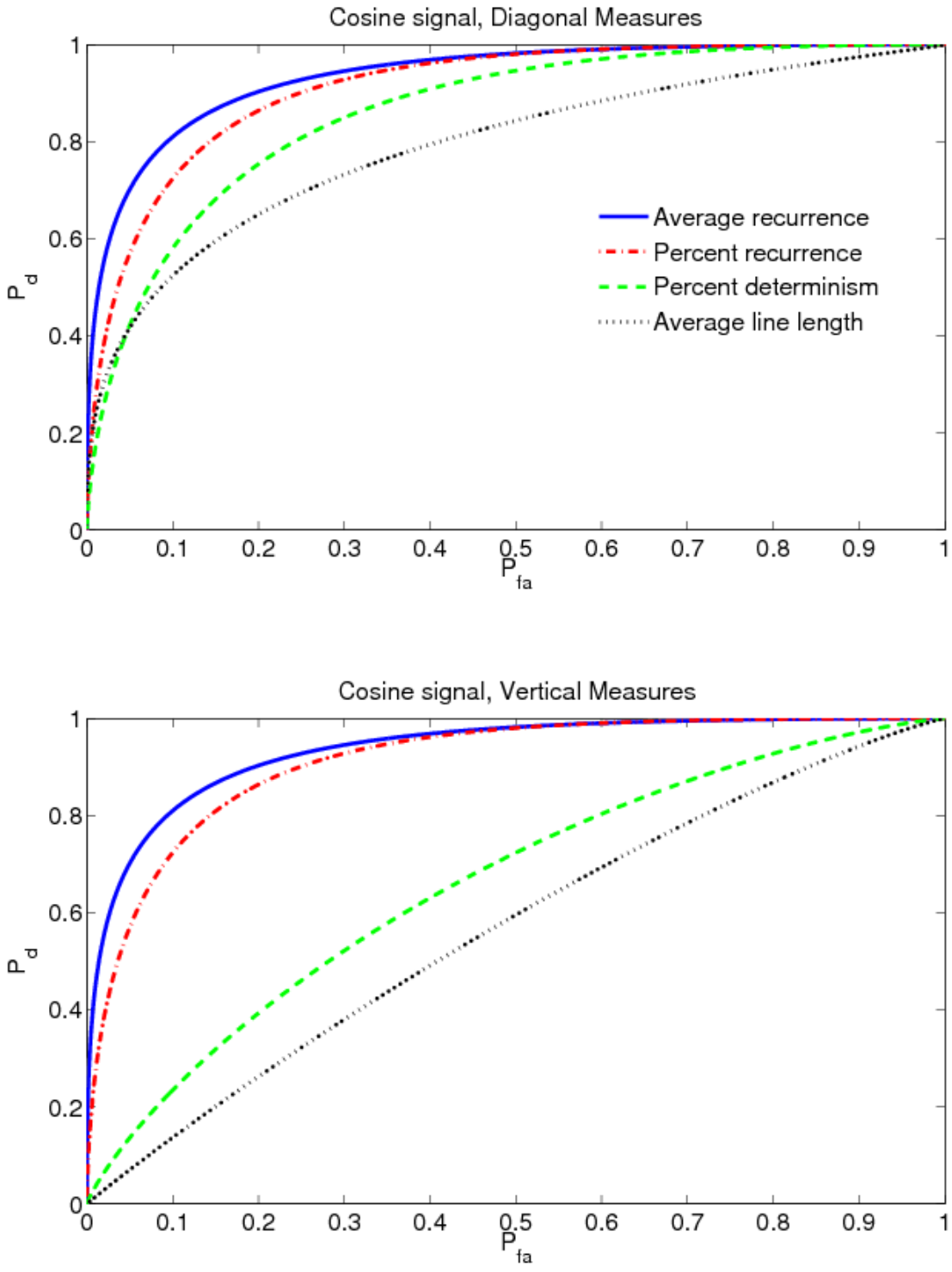
Cosine signal, Diagonal Measures
Average recurrence
Percent recurrence
Percent determinism
Average line length
P_d
P_fa
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5
0.7
0.9
Cosine signal, Vertical Measures
P_d
P_fa

Figure 7:

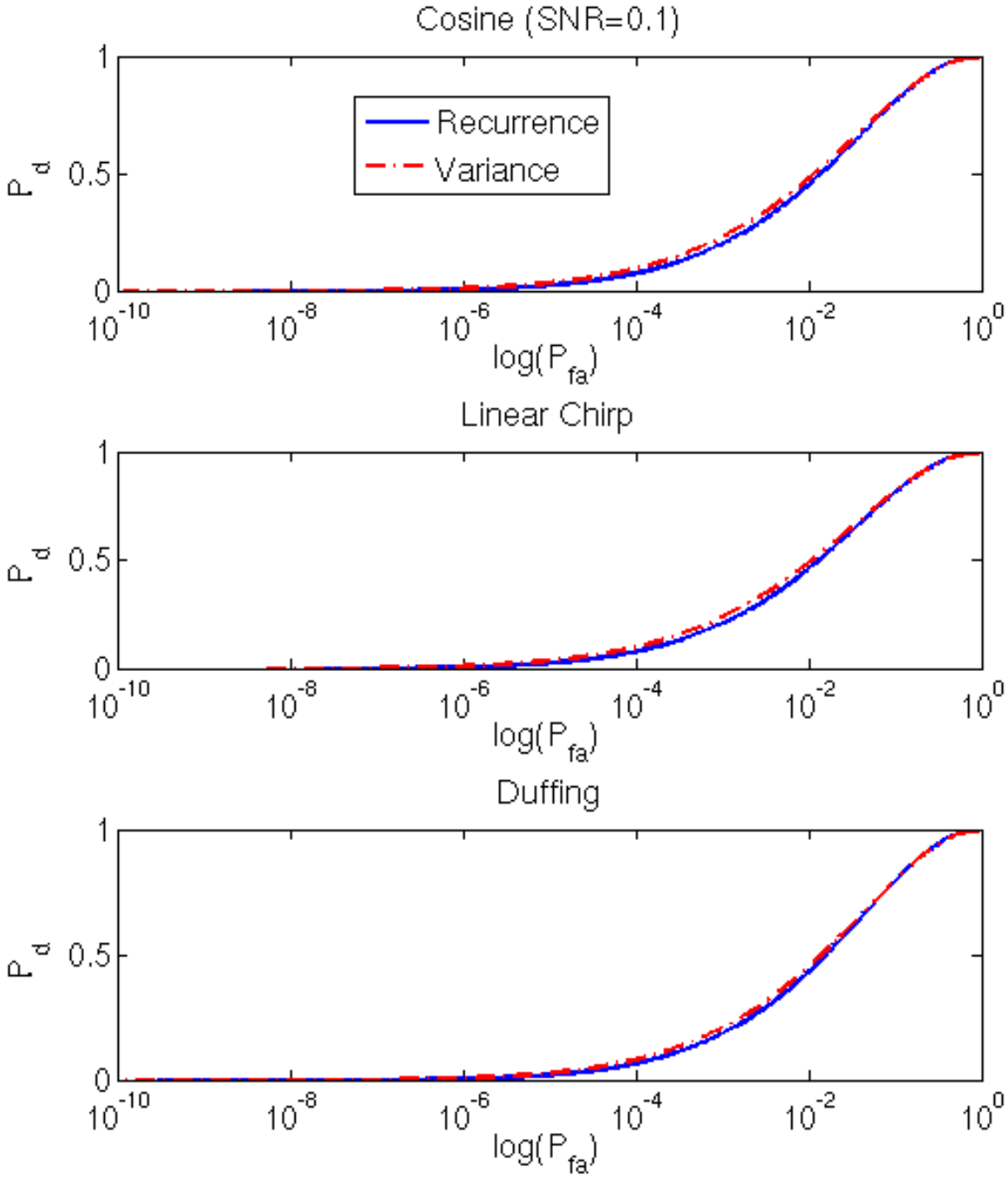
Cosine (SNR=0.1)
Recurrence
Variance
P_d
log(P_fa)
Linear Chirp
Duffing

Figure 8:

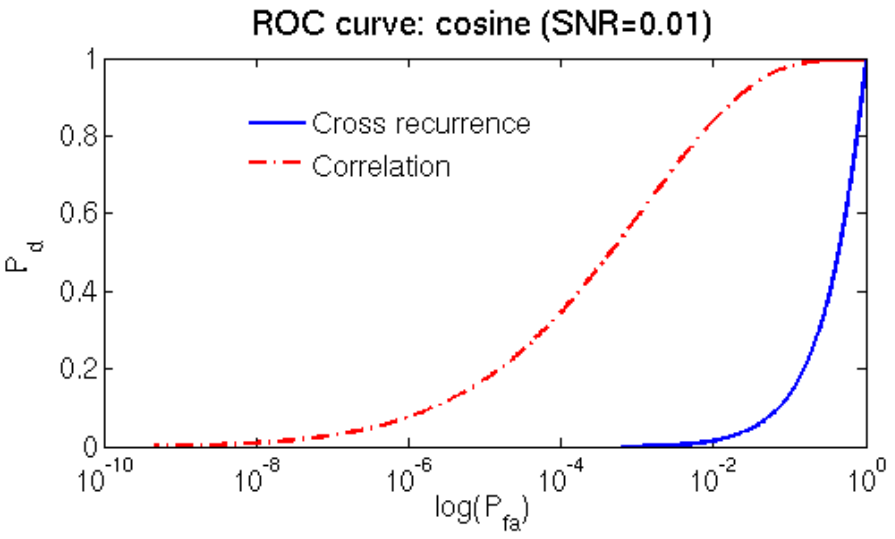

Figure 9:

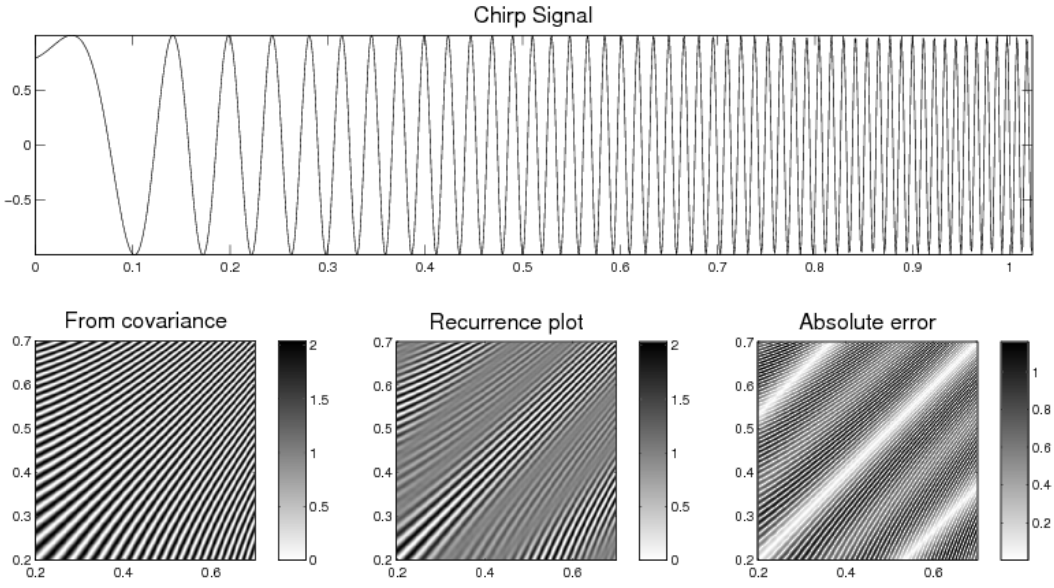

Figure 10: