

Influence of observational noise on the recurrence quantification analysis

Marco Thiel^{a,*}, M. Carmen Romano^a, Jürgen Kurths^a, Riccardo Meucci^b,
Enrico Allaria^b, F. Tito Arcelli^b

^a *Nonlinear Dynamics, University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany*

^b *Instituto Nazionale di Ottica Applicata, Largo E. Fermi 6, I-50135 Florence, Italy*

Received 25 October 2001; received in revised form 29 May 2002; accepted 13 June 2002

Communicated by E. Kostelich

Abstract

In this paper, we estimate the errors due to observational noise on the recurrence quantification analysis (RQA). Based on this estimation, we present ways to minimize these errors. We give a criterion to choose the threshold ε needed for the optimal computation of the recurrence plot (RP). One important point is to show the limits of interpretability of the results of the RQA if it is applied to measured time series. We show that even though the RQA is very susceptible to observational noise, it can yield reliable results for an optimal choice of ε if the noise level is not too high. We apply the results to typical models, such as white noise, the logistic map and the Lorenz system, and to experimental laser data.

© 2002 Elsevier Science B.V. All rights reserved.

PACS: 07.05.Rm; 07.05.Kf

Keywords: Recurrence quantification analysis; Recurrence plots; Stochastic processes; Dynamical processes; Observational noise

1. Introduction

Recurrence plots (RPs) successfully provide a qualitative impression of the dynamics of a given system [2]. To obtain some quantitative information also, the recurrence quantification analysis (RQA) was invented [12]. RQA is based on the distributions of the diagonal, horizontal and vertical lines that are found in the RP. The connection between RPs and RQA forms a powerful tool; both methods have been widely applied in many fields such as physics, medicine, earth science and economics [3,6,7,9,12]. However, the structures that are found in the RPs are not substantiated mathematically. The interpretation is, to a high degree, a matter of experience and is somewhat subjective. There is no well-founded criterion to choose the parameters used in the methods [4]. In spite of their numerous applications to measured data, the influence of noise on RPs and the RQA is still not sufficiently studied.

* Corresponding author.

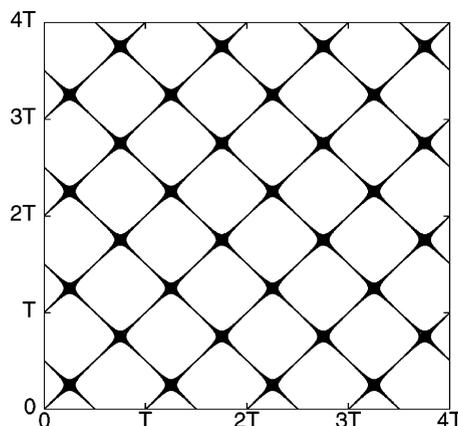


Fig. 1. RP of a sine function.

The main point of this paper is to determine the changes that noise induces in the RPs and in the results of the RQA of a given time series. The results allow one to specify the necessary parameters optimally. Our results show that a careful choice of the parameters is crucial to assure the relevance of the analysis of measured data based on the RQA.

1.1. Recurrence plots

In the analysis of time series of dynamical systems, it is often convenient to describe the system's state in phase space [5]. RPs were introduced to visualize the behavior of trajectories in phase space [2]. The representation is based on the matrix

$$\mathbf{R}_{i,j} = \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), \quad i, j = 1, \dots, N, \quad (1)$$

where \vec{x}_i stands for the point in phase space at which the system is situated at time i , and ε is a predefined threshold. $\Theta(x)$ is the Heaviside function. The matrix consists of the values 1 and 0 only. The graphical representation is an $N \times N$ grid of points, which are encoded as black for 1 and white for 0. A black point in the RP means that the system returns to an ε -neighborhood of the corresponding point in phase space. This recurrence gives the name to the method.

The RP of the sine function (Fig. 1) is computed with an embedding dimension of 1. The periodicity of the sine function is obviously reflected in the RP. The main diagonal occurs in all RPs and corresponds to the fact that the distance $x_i - x_i$ is 0. On the other hand, a homogeneous plot consisting of mostly isolated points may indicate mainly stochastic system. A decreasing recurrence rate that is perpendicular to the main diagonal may indicate a drift, i.e., non-stationarity of the time series. Long diagonal lines mean that the system “recurs” to a state in phase space, so that the future development of the system is similar to its past.

Zbilut et al. [14] have expanded the method of RP by considering two different time series. The cross-recurrence between two series $\{x_i\}$ and $\{y_i\}$ is defined as

$$\mathbf{CR}_{i,j} = \Theta(\varepsilon - \|\vec{x}_i - \vec{y}_j\|). \quad (2)$$

Obviously, the two time series have to be embedded in the same phase space to be able to define a distance. The representation is analogous to the RP, and it is called a cross-recurrence plot (CRP).

1.2. Recurrence quantification analysis

To quantify the structures that are found in RPs and CRPs, the RQA was proposed [12]. There are different measures that can be considered. One crucial point for many of these measures, such as the determinism

$$\text{DET} = \frac{\sum_{l=l_{\min}}^{l_{\max}} l \cdot P(l)}{\sum_{l=1}^{l_{\max}} l \cdot P(l)} \quad (3)$$

is the distribution of the lengths of the diagonal lines $P(l)$ of length l that are found in the plot [8]. Interest focuses on these lines because they are linked to the largest Lyapunov exponent if there is an underlying dynamical system. Eckmann et al. [2] claim:

The length of the lines is thus related to the inverse of the largest positive Lyapunov exponent. If the $x(i)$ were randomly chosen rather than coming from a dynamical system, there would be no such lines.

We agree with the first statement in the case of chaotic systems, but we show that the second does not strictly hold. Even in the case of white noise, a rather large number of short diagonal lines exists. We focus our attention in this paper on the structures that appear in RPs and CRPs of random data, i.e., basic stochastic processes, and we compute analytically the corresponding distribution of the diagonal lines.

It is crucial to consider randomness, as the methods of RP and CRP are frequently applied to experimental data sets [12]. To give some insight into the importance of the problem, we compute the distribution of diagonal lengths for three realizations of a Gaussian stochastic process $\mathcal{WN}(0, \sigma^2)$. Each time series contains 3000 points; we use no embedding and set $\varepsilon = 0.05$. The results are obtained simply by counting the number of diagonal lines of length l in the RP (Table 1). Lines of length 1 occur with high frequency, as expected. But as Table 1 shows, there are also lines of length greater than 1. The maximal observed length of the lines is 4. (The main diagonal of the RP is not considered as it is trivial.)

RPs are frequently applied to embedded time series, which are corrupted by observational noise. We stress that the concept of phase space is not valid in the case of stochastic systems. Hence, it is also problematic to use embedding. There are two ways to handle this problem. First, one can use RPs without embedding. Then, one may interpret the diagonal lines only in the sense of a recurrence of some structure in time. These RPs contain useful information, as we will see in the next three sections. Second, to maintain the successful concept of embedding for dynamical systems, one can use the following argument. The aim is to analyze the RP of the embedded time series (without noise). The embedding procedure for time series with observational noise yields deviations from the RP of the underlying process. If the magnitudes of these deviations are known, then conclusions about the underlying process can be drawn. This paper presents estimates for these deviations and so offers the opportunity to apply both methods.

The outline of this paper is as follows. First, we calculate analytically the distribution of the diagonal lines for basic stochastic processes (Gaussian and uniformly distributed noise). Second, we estimate the influence of observational noise on the distribution of diagonal lines for an arbitrary underlying process and develop a criterion to choose ε optimally.

Table 1

Numerical results for the distribution of diagonal lengths for three time series (3000 points) of Gaussian white noise, $\mathcal{WN}(0, 1)$, with $\varepsilon = 0.05$ and $l_{\min} = 2$

Run	Length (l)	1	2	3	4	5	6	7	8	DET
1	Number of diagonals	234416	6550	170	12	0	0	0	0	0.055
2	Number of diagonals	233938	6440	148	4	0	0	0	0	0.054
3	Number of diagonals	239390	6966	194	0	0	0	0	0	0.057

2. Analytical calculation of the distribution of lengths of the diagonals for Gaussian white noise

We first calculate the distribution $P(l)$ for a CRP obtained from the two time series $\{x_i\}$ and $\{y_i\}$ without embedding; each is independent Gaussian noise $\mathcal{WN}(0, \sigma^2)$. The probability to find a point in the interval $[x, x + \Delta x]$ is given by

$$p(x + \Delta x) = \int_x^{x+\Delta x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-y^2}{2\sigma^2}\right) dy.$$

To find a recurrence point at the coordinates (i, j) in the CRP, the condition $|x_i - y_j| < \varepsilon$ must hold. Hence, the probability for the occurrence of a recurrence point is

$$P_{\text{RP}} = \frac{1}{(\sqrt{2\pi}\sigma)^2} \int_{-\infty}^{\infty} \exp\left(\frac{-y^2}{2\sigma^2}\right) \int_{y-\varepsilon}^{y+\varepsilon} \exp\left(\frac{-x^2}{2\sigma^2}\right) dx dy = \text{erf}\left(\frac{\varepsilon}{2\sigma}\right). \quad (4)$$

When $\varepsilon/\sigma = 0.1$, we obtain $P_{\text{RP}} = 0.056$. To calculate the distribution of diagonal lengths, we compute first the probability to find l recurrence points in a row. As the time series are independent noise, this probability is the product

$$P_{l\text{-RP}'s}(l) = P_{\text{RP}}^l.$$

To be sure that the line has exactly the length l , the point before the first recurrence point and the point after the last recurrence point must be white, i.e., both must not be recurrence points. The probability to find such a white point is $(1 - P_{\text{RP}})$. Thus, the probability to find a line of length l is given by

$$P(l) = P_{\text{RP}}^l (1 - P_{\text{RP}})^2. \quad (5)$$

This calculation let us to compute the absolute number $N(l)$ of diagonal lines of lengths l in a CRP or RP of total length L using the relation

$$N(l) = L^2 P_{\text{RP}}^l (1 - P_{\text{RP}})^2 \quad (6)$$

for a CRP and

$$N(l) = (L^2 - L) P_{\text{RP}}^l (1 - P_{\text{RP}})^2 \quad (7)$$

for a RP. Numerically, one does not count points on the main diagonal. We take this fact into consideration by ignoring all L points on the main diagonal and by subtracting L from the total number of points in the RP. Simulations confirm Eq. (7).

3. RP of a function with observational noise

In this section, we discuss the influence of independent Gaussian observational noise on RPs. (The analysis for uniformly distributed noise is given in [Appendix B](#).) We suppose that a given time series $\{x_i\}$ is corrupted by some independent Gaussian observational noise η_i ,

$$y_i = x_i + \eta_i, \quad (8)$$

where $\eta_i \sim \mathcal{WN}(0, \sigma^2)$. The recurrence matrix $R_{i,j}$ of the underlying time series $\{x_i\}$ is

$$R_{i,j} = \Theta(\varepsilon - |D_{i,j}|) \quad (9)$$

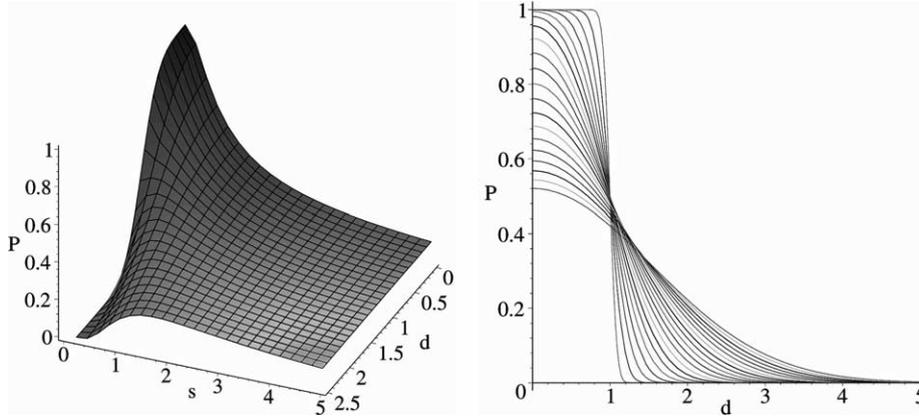


Fig. 2. Left panel: dependence of P on $d = D/\varepsilon$ and $s = \sigma/\varepsilon$ in Eq. (11). Right panel: dependence of P on d for $s \in [0.05, 1]$ in steps of 0.05.

with $D_{i,j} = x_i - x_j$. Due to the observational noise that enters the measured time series $\{y_i\}$, $R_{i,j}$ is replaced by

$$\tilde{R}_{i,j} = \Theta(\varepsilon - |\tilde{D}_{i,j}|) \quad (10)$$

with $\tilde{D}_{i,j} = y_i - y_j = x_i - x_j + \eta_i - \eta_j$. The question that we address is how this observational noise changes the structures that are found in $\tilde{R}_{i,j}$ with respect to $R_{i,j}$. Therefore, we compute the probability $P_{i,j}$ to find a recurrence point at the coordinates (i, j) if the underlying time series x_i has a distance $D_{i,j}$ at these coordinates.

Given a fixed underlying time series $\{x_i\}$, we measure the same process with different realizations of additive noise $\mathcal{WN}(0, \sigma^2)$. Then we obtain for each time index i an ensemble of measurements $\{y_n\}_i$, i.e., Gaussian distributed with mean x_i and variance σ^2 . As we are interested in the distances $D_{i,j}$, we can assume without loss of generality that $x_i = 0$ and $x_j = -D_{i,j}$. Then the probability to find a recurrence point at the coordinates i, j is given by

$$\begin{aligned} P_{i,j} &= \frac{1}{(\sqrt{2\pi}\sigma)^2} \int_{-\infty}^{\infty} \exp\left(-\frac{\eta_i^2}{2\sigma^2}\right) \int_{\eta_i-\varepsilon}^{\eta_i+\varepsilon} \exp\left(-\frac{(\eta_j - D_{i,j})^2}{2\sigma^2}\right) d\eta_j d\eta_i \\ &= \frac{1}{8} \left\{ \operatorname{erfc}^2\left(\frac{D_{i,j} - \varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(-\frac{D_{i,j} - \varepsilon}{2\sigma}\right) + \operatorname{erfc}^2\left(-\frac{D_{i,j} + \varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(\frac{D_{i,j} + \varepsilon}{2\sigma}\right) \right\}, \quad (11) \end{aligned}$$

where $\operatorname{erfc}(\cdot) = 1 - \operatorname{erf}(\cdot)$. Eq. (11) is one of the key results of this paper. It relates the distance matrix $D_{i,j}$ to a probability matrix $P_{i,j}$ that gives the probability to find a recurrence point in the RP of $y_i = x_i + \eta_i$ at the coordinates i, j for a fixed underlying process and an ensemble of realizations of the observational noise.

The matrix $P_{i,j}$ takes values in the interval $[0, 1]$ in contrast to the usual RP, which relates the distance matrix $D_{i,j}$ to the “binary” matrix $R_{i,j}$, i.e., the matrix consisting of only the two symbols 0 (if $|x_i - x_j| > \varepsilon$) for large distances and 1 (if $|x_i - x_j| \leq \varepsilon$) for small distances.¹

It is also possible to treat continuous functions. In that case, the matrix $D_{i,j}$ is replaced by a function $D(t, d)$, which depends on the two continuous variables t and d . Then the probability matrix $P(t, d)$ is continuous. From now on, we write P and D instead of $P_{i,j}$ and $D_{i,j}$ for convenience.

The left panel of Fig. 2 shows the dependence of P in Eq. (11) on $d = D/\varepsilon$ and $s = \sigma/\varepsilon$. We show only the part with $d > 0$, as the picture is symmetric for $d < 0$. An equivalent representation is given in the right panel of Fig. 2, which shows slices parallel to the d axis for different values of s , where s goes from 0.05 to 1 in steps of 0.05.

¹ There are also “Unthresholded Recurrence Plots”. They encode different distances with different colors [4].

The smaller the values of s , the closer P comes to a Heaviside function. In the presence of noise, the probabilities differ from this ideal. Due to the noise, points that are recurrence points for the series x_i are recognized as such with probability less than 1. Similarly, points that are not recurrence points in the absence of noise are recognized as such with probability less than 1. Hence, the probability plot for systems with observational noise appears gray (the main diagonal is not considered).

One striking result is that only a small amount of noise reduces the reliability of the calculation of the distribution of the lengths of the diagonals. To give an example: for a noise level of about 4%, the probability to find a black point near the distance threshold $d = 1$ is reduced from 1 to less than 0.6; see Eq. (11). Therefore, more than 40% of the recurrence points are not recognized. The distribution of the diagonals changes enormously. This result has crucial consequences for the application of RQA to observed data.

There are three ways to overcome this problem:

- (i) We can measure the data with very little observational noise. However, this is usually impossible, because the noise often cannot be controlled.
- (ii) We can develop a statistic other than the distribution of the diagonals. The approach presented in this paper suggests some alternatives; we will treat this problem in a future paper.
- (iii) We can choose an optimal ε , so that these misclassifications are reduced as much as possible. We have developed a criterion based on Eq. (11) in Section 4 to calculate such an ε_{opt} .

Various suggestions appear in the literature for the choice of ε ,² but they are heuristic.³ Next we discuss a prototypical case to demonstrate the application of Eq. (11).

3.1. The logistic map

We analyze as an example a chaotic time series that is generated by the logistic map

$$x_{n+1} = 4x_n(1 - x_n). \quad (12)$$

The left panel of Fig. 3 represents the probability to find a recurrence point in the time series without observational noise. As each “realization” is identical in this deterministic case, there are only two probabilities: 1 (black) and 0 (white). There is a pronounced structure in the plot. Obviously, there are not only diagonal lines but also vertical and horizontal ones. These structures indicate a retention period in a small volume of the phase space (intermittency also can cause these patterns). The right panel shows the same plot for the time series with observational noise. The black lines of the left plot appear to be smeared out, as expected. There is no point that has probability 1 or 0 (except on the main diagonal, which always has probability 1). We obtain different results when calculating the distribution of the diagonal lines in the two cases. If we consider the logistic map and take the usually recommended ratio $\varepsilon/\sigma = 1$, a signal/noise ratio of 10 decreases the probability to correctly recognize a recurrence point to about 50%. We will see in Section 4 that, in this case, the result for the distribution of the diagonals fails to detect the structures of the underlying process. We will also discuss the magnitudes of the changes due to noise and then propose an optimal choice for ε that allows for correct recognition of recurrence points with much higher probability.

² In practice, we fix r as small as possible (typically not greater than 10% of the normalized mean distance of the first embedding) relative to the noise level [13]. We define the diameter, $d \in \mathcal{R}^+$, of a reconstruction to be $d = \max\{\|\vec{v}_i - \vec{v}_j\|, i, j \in \{1, \dots, N_v\} \text{ and } i \neq j\}$. 10% of the diameter is a typical cut-off value, c [6]. For the radius ε , one uses usually 10% of the maximum diameter ϕ_{max} of the phase space. If there are some outliers in the time series, this ansatz may yield values of ε that are not reasonable. The values should be of the order of magnitude of the standard deviation [8].

³ Besides being critically important, the selection of threshold corridor is also difficult to systematize in any sensible way. Solutions in the literature are unsatisfying: Webber and Zbilut, without comment, prescribe a threshold corridor corresponding to the lower 10% of the entire distance range present in the corresponding UTRP (unthreshold recurrence plot) [4].

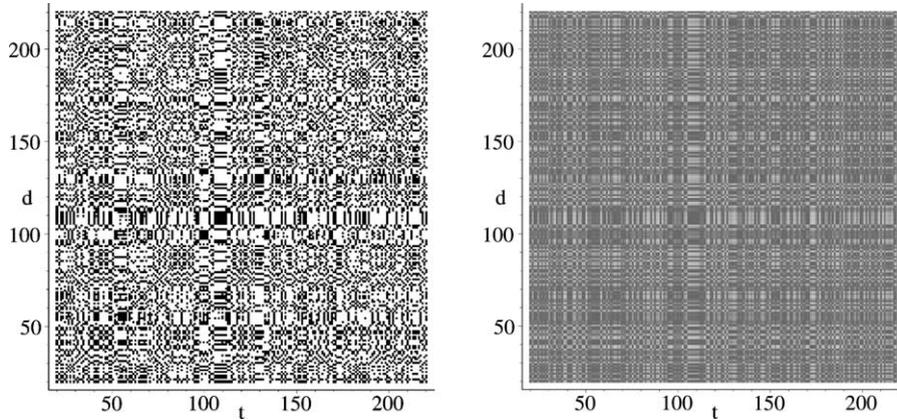


Fig. 3. Left panel: RP of a time series from the logistic map, Eq. (12). Right panel: probability plot of a time series of the logistic map, Eq. (12), to which observational Gaussian white noise has been added with standard deviation $\sigma = 0.1\Sigma$, where Σ is the standard deviation of x_n .

4. Estimation of the errors due to observational noise

In this section, we compute analytically the percentage of recurrence and non-recurrence points that are properly recognized. This result will be used to derive a procedure to determine the optimal ε . We show that the results of RQA improve substantially with this optimal choice ε_{opt} .

Fig. 2 represents the probabilities to find recurrence points in the presence of noise if the underlying process leads to a difference d . Without noise, the probability is 1 for $|d| < 1$ and 0 otherwise. If the density $\rho(\cdot)$ of the increments $x_i - x_j$ is given, we can compute the percentage of recurrence points $p_b(\varepsilon, \sigma)$ that are properly recognized in the presence of observational noise by

$$p_b(\varepsilon, \sigma) = \frac{\int_{-\varepsilon}^{\varepsilon} P(D, \varepsilon, \sigma)\rho(D) dD}{\int_{-\varepsilon}^{\varepsilon} \rho(D) dD}, \tag{13}$$

where

$$P(D, \varepsilon, \sigma) = \frac{1}{8} \left\{ \operatorname{erfc}^2\left(\frac{D - \varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(-\frac{D - \varepsilon}{2\sigma}\right) + \operatorname{erfc}^2\left(-\frac{D + \varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(\frac{D + \varepsilon}{2\sigma}\right) \right\} \tag{14}$$

is the solution of Eq. (11). Analogously, the percentage of properly recognized non-recurrence points is

$$p_w(\varepsilon, \sigma) = \frac{\int_{-\infty}^{-\varepsilon} [1 - P(D, \varepsilon, \sigma)]\rho(D) dD + \int_{\varepsilon}^{\infty} [1 - P(D, \varepsilon, \sigma)]\rho(D) dD}{1 - \int_{-\varepsilon}^{\varepsilon} \rho(D) dD}. \tag{15}$$

We distinguish two types of errors that can occur (Table 2). To calculate the optimal ε that reduces both types of errors, we need the percentage (density) of recurrence points in the presence of noise, which is the sum of the

Table 2
Errors and properly recognized recurrence points

Point in the plot	Recognized as recurrence point	Not recognized as recurrence point
Recurrence point	p_b	$1 - p_b$
Not recurrence point	$1 - p_w$	p_w

properly and improperly recognized recurrence points in the plot, divided by the total number of points:

$$p_b^{\text{density}} = \int_{-\infty}^{\infty} P(D)\rho(D) dD. \quad (16)$$

For the non-recurrence points, $p_w^{\text{density}} = 1 - p_b^{\text{density}}$.

We compute p_b and p_w numerically for two different sets of data: a shot noise process and a logistic process. See [Appendix A](#) for a detailed discussion. Here we summarize the main results.

The main consequence of our approach is that we can optimize the choice of ε if the standard deviations σ of the Gaussian noise and Σ of the underlying process are given, so that the bias due to the noise is minimized. Otherwise, we can compute the influence of noise on a given process when only its distribution is known. Simulations show that, in many cases, it is possible to estimate the influence of noise even for a measured time series, i.e., a time series plus noise, without knowing the underlying process. This works if the observational noise does not change the distribution of the time series too much.

Although the results depend on the distribution of the time series, there are some general statements that we can make (see [Appendix A](#)).

1. For a wide class of processes, the choice

$$\varepsilon \approx 5\sigma \quad (17)$$

is appropriate. If ε is smaller, the effects of the observational noise have a dominant influence on the detection of recurrence points. On the other hand, if $\varepsilon \approx \Sigma$, where Σ is the standard deviation of the underlying process, then the density of recurrence points is too high to detect the detailed structures of the underlying process.

2. In the literature, a frequently applied choice for ε is $\varepsilon \approx \frac{1}{10}\Sigma$ (see [8] and Footnote 2). Note, however, that this choice is appropriate only if $\sigma < \frac{1}{50}\Sigma$, as the condition [Eq. \(17\)](#) must be satisfied.
3. Some processes, e.g., discrete processes with only a few states, may require choices of ε other than [Eq. \(17\)](#). In these cases, one must compute the quantities in [Eqs. \(13\) and \(15\)](#).

These considerations allow for the estimation of errors that occur in the computation of the lengths of the diagonal, vertical and horizontal lines. This estimation is of special interest, as these distributions are the skeleton of the RQA. The idea is to calculate a correction factor for the number of lines of the length l . Analogous to the calculus of the distribution of the diagonal lengths for stochastic processes, we have to consider not only the l black points that form the line, but also the white points at the beginning and the end of the lines. Hence, for a given process and ratio ε/σ , we compute the percentage of properly recognized recurrence points (p_b) and non-recurrence points (p_w). The two white points are properly recognized with probability p_w and the l black points with probability p_b . This leads to the correction factor

$$K(l) = p_w^2 p_b^l. \quad (18)$$

The meaning of this factor is the following. Given a process x_i and the corresponding distribution of diagonal (horizontal, vertical) lines $P_{\text{lines}}(l)$, then

$$P_{\text{lines}}^{\text{prop}}(l) = K(l) P_{\text{lines}}(l) \quad (19)$$

is the fraction of diagonals of length l that one properly recognizes, i.e., the lines that are found in the RP without noise and are still recognized in the plot with noise. Note that $P_{\text{lines}}^{\text{prop}}(l)$ is not the fraction that one actually counts in a RP with noise: there is also a spurious detection of diagonals of length l that does not occur in the plot without noise. Although the distribution $P_{\text{lines}}(l)$ of the plot without noise obviously depends on the dynamics in the time series (e.g., that is why it reflects the Lyapunov exponent in the case of chaotic systems), the correction factor $K(l)$ only

Table 3

Comparison of the probabilities to correctly recognize recurrence lines for different ratios of ε/σ in a time series from the logistic map. Top row: $\varepsilon/\sigma = 1$ (usual choice), bottom row: $\varepsilon/\sigma = 5$ (optimal choice)

ε/σ	$K, \% (l = 1)$	$K, \% (l = 2)$	$K, \% (l = 3)$	$K, \% (l = 4)$	$K, \% (l = 5)$	$K, \% (l = 6)$
1	46.4	23.0	11.4	5.66	2.81	1.39
5	88.4	82.2	76.5	71.1	66.1	61.5

depends on the distribution of the time series. It would be the same for a chaotic logistic process, a sine function, or noise, if they all had the same distribution.

We illustrate how to apply $K(l)$ for the case of the logistic map ($r = 4$; see also [Appendix A](#)). We first estimate the error that occurs when computing the distribution of diagonal lines in the presence of 10% noise for two different ratios of ε/σ . [Table 3](#) gives the percentage of properly recognized diagonals of length l . The first row corresponds to the frequently applied choice of $\varepsilon = \sigma$, and the second row corresponds to the optimal threshold, $\varepsilon = 5\sigma$.

In the case of the usual choice, $\varepsilon/\sigma = 1$, the percentage of properly recognized diagonals is too low to draw conclusions about the dynamics. For the optimal choice, $\varepsilon/\sigma = 5$, the results are much better. Although we cannot determine the distribution of diagonals perfectly, we can hope to conserve enough structure to recognize the main features of the underlying dynamics.

To illustrate the result of [Eq. \(19\)](#) and [Table 3](#), we compute the distribution of diagonals. The embedding dimension is 1 and the level of noise is 10%. The left panel in [Fig. 4](#) shows the distribution for the usual choice, $\varepsilon/\sigma = 1$. In this case, the distribution of the underlying process (solid line) is biased if observational noise is present (dashed line). The right panel shows the distributions for the optimal choice, $\varepsilon/\sigma = 5$. The distributions in this case coincide for a large interval of l .

These results are also strongly reflected in the computed value of the determinism, DET ([Eq. \(3\)](#)), which is a basic quantity in the RQA ([Table 4](#)). The difference in the values of DET for the case without noise is due to the different choices of ε . The results show that, for the optimal choice of ε , the computed value of DET in the case of observational noise is nearly identical to that of the underlying process, whereas the usual choice, $\varepsilon/\sigma = 1$, leads to a strongly biased estimate: the relative error in DET is about 67%.

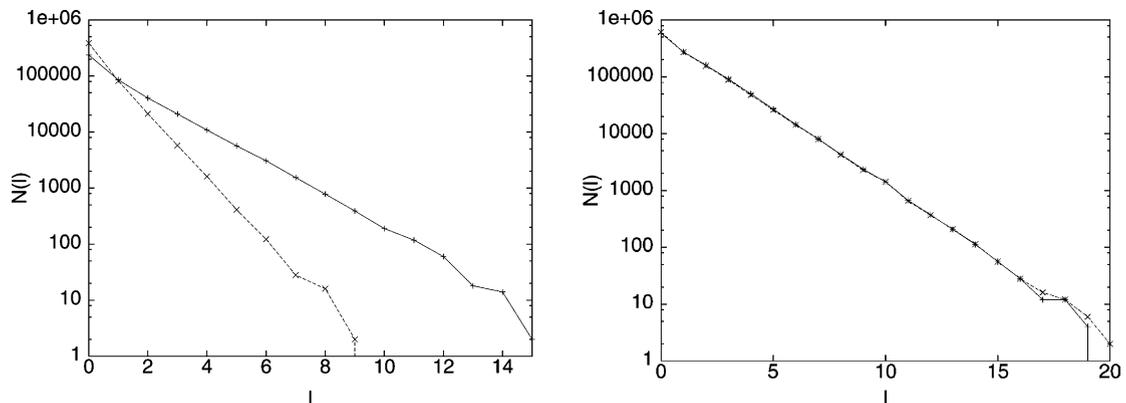


Fig. 4. Left panel: number of diagonals, $N(l) = L^2 P(l)$, for the logistic map, analogous to [Eq. \(7\)](#), with the usual choice of ε . Solid line: without noise, dashed line: with 10% noise. Right panel: as at left, but with the optimal choice of ε .

Table 4

Computed values of the determinism, DET, for the usual and optimal choices of ε for time series of the logistic map with and without noise. Parameters: length of the time series, $L = 3000$; embedding dimension, $m = 1$; minimal length, $l_{\min} = 3$

Choice of ε	DET without noise	DET with noise
Usual, $\varepsilon/\sigma = 1$	0.46	0.15
Optimal, $\varepsilon/\sigma = 5$	0.57	0.56

5. Results for embedded time series

In this section, the results of Section 4 are extended to embedded measurements

$$\vec{y}_i = \vec{x}_i + \vec{\eta}_i.$$

Eq. (11) has to be extended to d dimensions. We apply the usual method [5,11] to reconstruct the vector

$$\vec{D} = (x_i - x_j, x_{i+\tau} - x_{j+\tau}, \dots, x_{i+(d-1)\tau} - x_{j+(d-1)\tau}).$$

Here, \vec{D} is the vector that corresponds to the point (i, j) in the CRP.⁴ It is not necessary to consider the coordinates explicitly in the calculations, so we write D instead of $D_{i,j}$. Now

$$P(\vec{D}) = \frac{1}{(\sqrt{2\pi}\sigma)^{2d}} \int_{\mathcal{R}^d} \exp\left(-\frac{\vec{\eta}}{2\sigma^2}\right) \int_{\mathcal{U}_\varepsilon(\vec{\eta})} \exp\left(-\frac{(\vec{\vartheta} - \vec{D})^2}{2\sigma^2}\right) d\vec{\vartheta} d\vec{\eta}. \quad (20)$$

The outer integral must be evaluated over the entire space \mathcal{R}^d . The inner integral is evaluated in the ε -neighborhood of $\vec{\eta}$. The integral is more difficult to evaluate if we consider a sphere around $\vec{\eta}$. It is easier to consider a box whose sides have length 2ε . Let D_k denote the k th component of \vec{D} ; then the solution of Eq. (20) is given by

$$P(\vec{D}) = \prod_{k=1}^d \frac{1}{8} \left\{ \operatorname{erfc}^2\left(\frac{D_k - \varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(-\frac{D_k - \varepsilon}{2\sigma}\right) + \operatorname{erfc}^2\left(-\frac{D_k + \varepsilon}{2\sigma}\right) - \operatorname{erfc}^2\left(\frac{D_k + \varepsilon}{2\sigma}\right) \right\} \quad (21)$$

if the D_k are considered to be independent. The probability to find a recurrence point at the coordinates (i, j) is $P(\vec{D}) = P(\vec{D}_{i,j})$ if the underlying process y_i is embedded in a space of dimension d .

It is problematic to apply embedding methods to random time series. However, the optimal choice of ε in the case with embedding is the same as in the case without embedding, as the components of \vec{D} enter Eq. (21) separately. The recurrence point is properly recognized if the product in Eq. (21) is maximal. In a first approximation, this is the case if each factor is maximal. This condition leads to the same procedure as described for the case without embedding (see Section 4).

We illustrate this result for the Lorenz system with the parameters $\sigma = 16$, $r = 45.92$ and $b = 4$. The step size for the integration is $h = 0.001$ and the sampling rate is $\delta t = 10h$. We use an embedding dimension of $m = 3$ and a delay of $\tau = 8$. Fig. 5 shows the number, $N(l)$, of diagonal lines of length l , and Table 5 shows the values of the determinism, DET, for a time series of length $N = 5000$ for the case without and with observational noise. The noise is white and Gaussian with $\sigma = 0.1\Sigma$. The left panel shows the results for the usual choice of ε ; the right, for the optimal one.

The value of DET ($l_{\min} = 2$) can be computed from this distribution (Table 5). For the usual choice of ε , the computed values of DET for the case with and without noise differ by a factor of about 7, whereas for the optimal

⁴ We do not analyze a RP as we would have to consider spurious correlations due to the embedding. The CRP considers the same underlying process but different realizations of the noise in the two axis of the plot.

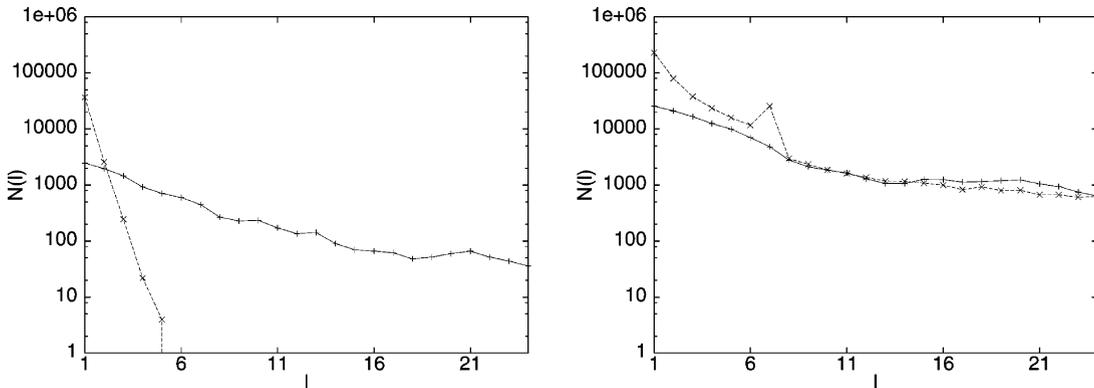


Fig. 5. Left panel: the number $N(l)$ of the diagonals for the Lorenz system with the usual choice of ε . Solid line: without noise, dashed line: with noise, $\sigma = 0.1\Sigma$. Right panel: as at left, but with the optimal choice of ε .

Table 5

Determinism for usual and optimal choice of ε for the Lorenz system in the case with and without noise. Parameters: length of the time series, $L = 5000$; embedding dimension, $m = 3$; delay, $\tau = 8$; minimal length, $l_{\min} = 2$

Choice of ε	DET without noise	DET with noise
Usual, $\varepsilon/\sigma = 1$	0.98	0.14
Optimal, $\varepsilon/\sigma = 5$	0.99	0.87

choice, the factor is 1.14. The optimal choice reduces the factor by nearly one order of magnitude. This difference is crucial: in the case of noisy time series with the usual choice of ε , the computed value of DET is so small that it is indistinguishable from that of a pure stochastic process. The optimal choice of the threshold, on the other hand, reveals the determinism in the underlying dynamics.

6. Application to experimental data

Next, we apply our modified RQA to data from an experimental system consisting of a CO₂ laser with sinusoidal modulation of the cavity losses. More precisely, the apparatus is a conventional low-speed longitudinal gas flow CO₂ laser with a Ge intracavity acousto-optic modulator (Mod. AGM-406B1 IntraAction). The optical cavity, 1.35 m long, is defined by a diffraction grating selecting the $P(20)$ line at 10.6 mm and an outcoupler mirror with a reflectivity of 90%. By applying a sinusoidal signal to the acousto-optic modulator at 100 kHz, which is close to the relaxation oscillation of the system, the laser reaches a chaotic condition after a sequence of subharmonic bifurcations [1,10]. The chaotic output intensity is detected by means of a high speed uncooled Hg–Cd–Zn–Te photodetector with a detectivity of 2.3×10^7 cm Hz^{1/2}/W and a bandwidth of 20 MHz (PD-10.6-8 Vigo System). Sequences of 50,000 points at a time resolution of 200 ns are recorded on a digital scope (LT 342L Lecroy). Fig. 6 shows a portion (20 periods) of the data.

We perform the RQA to compute the distribution of diagonal lines by using the embedding dimension $m = 3$ and the delay $\tau = 10$. The results of the original time series and the time series corrupted by 10% observational Gaussian white noise are shown in Fig. 7.

Here, the ratio between the computed values for the determinism for the case without and with noise and the usual choice of ε is 3.8. For the optimal choice of ε , the factor is 1.04, and, hence, very close to unity (Table 6).

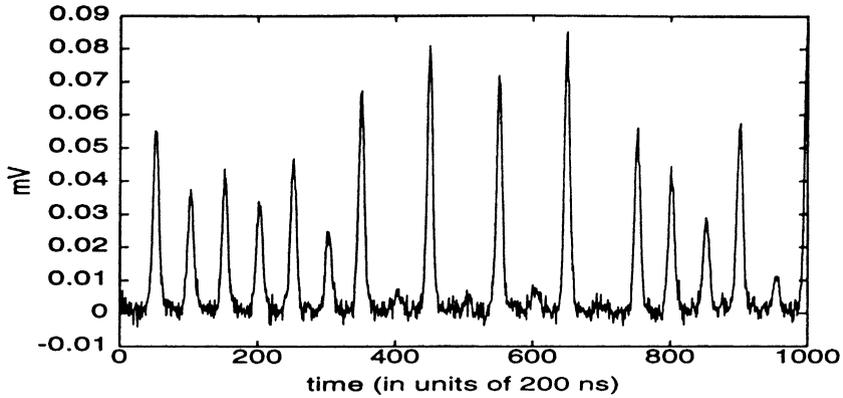


Fig. 6. Output of the CO₂ laser corrupted by 10% observational Gaussian noise.

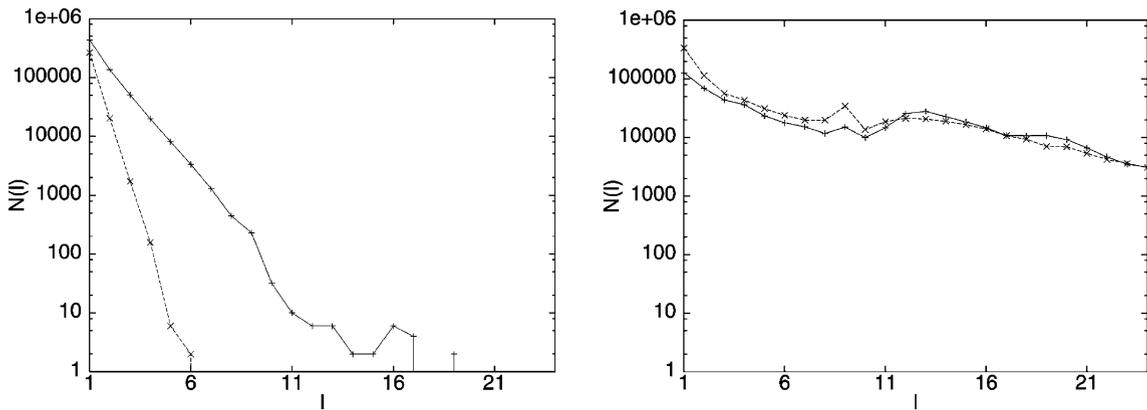


Fig. 7. Left panel: the number $N(l)$ of the diagonals for the laser system, computed with the usual choice of ϵ . Solid line: without noise, dashed line: with 10% noise. Right panel: as at left, but with the optimal choice of ϵ .

Table 6

Computed values of the determinism, DET, for the usual and optimal choice of ϵ for the laser system, with and without noise. Parameters: time series length, $L = 5000$; embedding dimension, $m = 3$; delay, $\tau = 10$; minimal length, $l_{\min} = 2$

Choice of ϵ	DET without noise	DET with noise
Usual, $\epsilon/\sigma = 1$	0.57	0.15
Optimal, $\epsilon/\sigma = 5$	0.98	0.94

7. Conclusions

The main point of this paper is that there are severe problems applying RQA to observed time series. Even low levels of observational noise change the statistics considerably. To solve this problem, we have proposed, based on our analytical results, a threshold ϵ that is at least five times the standard deviation of the observational Gaussian noise σ . This choice is appropriate in most cases. Furthermore, we have presented expressions for the deviations we have to take into account in the case of observational noise, and that help to evaluate the reliability of the results of

the RQA. However, if the level of observational noise is too high ($\sim 20\%$ of the standard deviation of the underlying process or more), then the application of the RQA to the data can lead to pitfalls.

The results motivate the development of new statistics for the RQA, different from simple distributions of diagonal lengths. Such statistics have to take the deviations due to noise—e.g., interruptions of the diagonal lines—into consideration.

Our results can easily be extended to other classes of noise and hold for all underlying processes. We have already mentioned the problems that stochasticity causes for embedding. The results of this paper show that it is possible to consider embedding, too.

A further advantage of such a mathematical description arises if RPs and RQA are used for hypothesis testing. Frequently, the null hypothesis that is tested for is based on surrogate data [12]. For shuffled surrogates, we can compute the distribution of diagonals analytically.

Acknowledgements

We thank Udo Schwarz, Norbert Marwan and Kai-Uwe Thiel, for fruitful discussions. The project was supported by the DFG-Schwerpunktprogramm 1114, the DFG-Projekt: KU 837/11-1, the DFG Research Group Conflicting Rules in Cognitive Systems and the EU HPRN-CT-2000-00158.

Appendix A. On the optimal choice of ε

Programs to compute probabilities to recognize a recurrence point p_b or a non-recurrence point p_w in the presence of noise are available on request. We apply the programs to two time series (logistic map and shot noise process) and present the results in Figs. 8 and 9.

A.1. Logistic map

First we consider the logistic map

$$x_n = 4x_n(1 - x_n). \quad (\text{A.1})$$

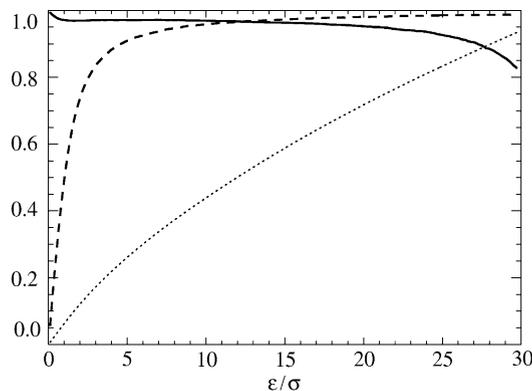


Fig. 8. Rates of properly recognized recurrence and non-recurrence points in the case of the logistic map for different ratios of ε/σ (solid line, respectively, dashed line). Dotted line: density of recurrence points.

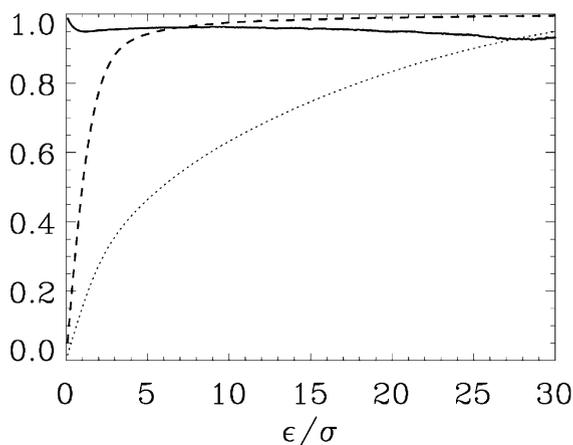


Fig. 9. Rates of properly recognized recurrence and non-recurrence points in the case of a shot noise process for different ratios of ϵ/σ (solid line, respectively, dashed line). Dotted line: density of recurrence points.

The standard deviation of the process is $\Sigma = 3/10$, and we use a noise level of 10%, i.e., $\sigma = 0.1\Sigma$. The length of the time series is 65,000.

Fig. 8 shows that p_b increases as ϵ/σ grows. This is plausible, as for small ϵ/σ , the noise makes it difficult to recognize properly small differences $D_{i,j} = x_i - x_j$. If $D_{i,j}$ is small, it is probable to find $\tilde{D}_{i,j} > D_{i,j}$, where $\tilde{D}_{i,j} = y_i - y_j = x_i - x_j + \eta_i - \eta_j$ (see Section 3). As ϵ/σ grows, this effect becomes less important for the discrimination of recurrence points. Hence, p_b increases as ϵ/σ increases. Analogously, p_w decreases as ϵ/σ . The bigger this ratio, the more difficult it becomes to determine non-recurrence points. Therefore, it is reasonable to maximize p_b and p_w : p_b increases until $\epsilon/\sigma \sim 5$ and then saturates; p_w is nearly constant for $\epsilon/\sigma < 20$ and then decreases. In the interval $\epsilon/\sigma \in [5, 20]$, both probabilities are relatively large; consequently, all the values of ϵ/σ in this interval are suitable. If we take into consideration that the percentage of recurrence points is very high for large fractions of ϵ/σ , then we tend to choose optimally $\epsilon/\sigma \approx 5$, because smaller ratios allow us to distinguish smaller changes in the dynamics of the underlying process.

A.2. Shot noise process

Next, we consider a shot noise process with an exponential decay. It consists of 65,000 points and 300 shots. The amplitude of the shots is 1 and the decay is given by $\lambda = 0.02$. The standard deviation of the time series is $\Sigma = 1/3$, and the standard deviation of the Gaussian noise is $\sigma = 0.1\Sigma = 1/30$.

Obviously, Fig. 8 is very similar to Fig. 9, i.e., the behavior of the probabilities with respect to changes of ϵ/σ is essentially the same.

The magnitudes of the probabilities are very similar, too. This is very interesting as the distribution of the time series (and, hence, of the increments) is quite different; p_b saturates at $\epsilon/\sigma \approx 5$, and p_w is nearly constant until $\epsilon/\sigma \approx 15$. As in the case of the logistic map, we choose the smallest ratio of the suitable interval, i.e., $\epsilon/\sigma \approx 5$.

Appendix B. Uniformly distributed noise

In this section, we present results for another class of random numbers (see Section 3). We consider a uniform distribution in $[-a, a]$. The probability to find a point at x is given by

$$p(x) = \frac{1}{2a} [\Theta(x+a) - \Theta(x-a)]. \quad (\text{B.1})$$

Analogous to the case of the Gaussian distribution, the probability to yield a recurrence point given the distance D of the underlying process is

$$\begin{aligned}
 P_{\text{RP}} &= \int_{-\infty}^{\infty} p(\eta) \int_{\eta-D-\varepsilon}^{\eta-D+\varepsilon} p(\vartheta) d\vartheta d\eta \\
 &= \frac{1}{8a^2} [2\Theta(D-\varepsilon)(D-\varepsilon)^2 - 2\Theta(D+\varepsilon)(D+\varepsilon)^2 + \Theta(D-2a+\varepsilon)\{(D+\varepsilon)^2 + 4(a^2-da-\varepsilon a)\} \\
 &\quad - \Theta(D-2a-\varepsilon)\{(D-\varepsilon)^2 + 4(a^2-da+\varepsilon a)\} + \Theta(D+2a+\varepsilon)\{(D+\varepsilon)^2 + 4(a^2+da+\varepsilon a)\} \\
 &\quad - \Theta(D+2a-\varepsilon)\{(D-\varepsilon)^2 + 4(a^2+da-\varepsilon a)\}] \tag{B.2}
 \end{aligned}$$

The following procedure is completely analogous to that for the Gaussian distribution. One can simply substitute Eq. (B.2) into the corresponding expressions in the latter case. We do not discuss it in more detail here.

By applying the first line of Eq. (B.2) to distributions $p(x)$ different from Eq. (B.1), one can compute P_{RP} for arbitrary probability distributions. Even an extension to embedded time series is analogous to the case of the Gaussian distribution.

References

- [1] F.T. Arecchi, R. Meucci, G. Puccioni, J. Tredicce, Experimental evidence of subharmonic bifurcations, multistability and turbulence in a Q-switched gas laser, *Phys. Rev. Lett.* 49 (1982) 1217–1220.
- [2] J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems, *Europhys. Lett.* 4 (1987) 973–977.
- [3] J.A. Holyst, M. Zebrowska, K. Urbanowicz, Observations of deterministic chaos in financial time series by recurrence plots, can one control chaotic economy? *Eur. Phys. J. B* 20 (2001) 531–535.
- [4] J.S. Iwanski, E. Bradley, Recurrence plots of experimental data: to embed or not to embed? *Chaos* 8 (4) (1998) 861–871.
- [5] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 1997.
- [6] M. Koebbe, G. Mayer-Kress, Use of recurrence plots in the analysis of time-series data, nonlinear modeling and forecasting, in: *Proceedings of SFI Studies in the Science of Complexity*, Vol. XXI, Addison-Wesley, Santa Fe, 1992, pp. 361–378.
- [7] J. Kurths, U. Schwarz, C.P. Sonett, U. Parlitz, Testing for nonlinearity in radiocarbon data, *Nonlinear Process. Geophys.* 1 (1994) 72–75.
- [8] N. Marwan, Study of climate variability in NW Argentina with a quantitative analysis of recurrence plots, Diploma Thesis, Dresden University of Technology, 1999.
- [9] N. Marwan, M. Thiel, N.R. Nowaczyk, Cross recurrence plot based synchronization of time series, *Nonlinear Process. Geophys.* 9 (2002) 325–332.
- [10] A.N. Pisarchik, R. Meucci, F.T. Arecchi, Theoretical and experimental study of discrete behavior of Shilnikov chaos in a CO₂ laser, *Eur. Phys. J. D* 13 (2001) 385–391.
- [11] F. Takens, Detecting strange attractors in turbulence, in: D.A. Rand, L.-S. Young (Eds.), *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, Vol. 898, Springer, Berlin, 1980.
- [12] C.L. Webber Jr., J.P. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies, *J. Appl. Physiol.* 76 (1994) 965–973.
- [13] J.P. Zbilut, C.L. Webber Jr., Embeddings and delays as derived from quantification of recurrence plots, *Phys. Lett. A* 171 (1992) 199–203.
- [14] J.P. Zbilut, A. Giuliani, C.L. Webber Jr., Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification, *Phys. Lett. A* 246 (1998) 122–128.