Chapter 9
Detection of Determinism

In Chap. 2 we have illustrated the applications of ordinal patterns with four examples. In this chapter we present a further application, this time to the detection of determinism in noisy time series. Following the common usage of the term in applied science, "determinism" is meant here as the opposite to statistical independence, hence it includes colored noise as well. This application hinges on two basic properties of ordinal patterns: existence of forbidden patterns in the orbits of maps (Sects. 1.2, 3.3, and 7.7) and robustness to observational noise (Sects. 3.4.3, and 9.1). We shall actually present two detection methods.

Method I is based on the number of missing ordinal patterns. It proceeds by (i) counting the number of missing ordinal patterns in sliding, overlapping windows of size L along the data sequence, (ii) randomizing the sequence, and (iii) repeating (i) with the randomized sequence. Is the result of step (iii) clearly greater than the result of step (i), so may we conclude that the original noisy sequence has a deterministic component.

Method II is based on the distribution of the visible ordinal patterns. This method proceeds by (i) counting the number of ordinal patterns in sliding, non-overlapping windows of size L along the data sequence and (ii) performing a χ^2 test based on the results of (i), the null hypothesis being that the data are white noise. Hold the null hypothesis, so should all possible ordinal L-patterns be visible and evenly distributed over sufficiently many windows, at variance with what happens in the case of noisy deterministic data. In the latter case, the number of missing ordinal patterns is higher, its decay rate with L is slower, and the distribution of patterns is not necessarily uniform.

Both methods, as other applications of permutation entropy, are conceptually simple and computationally fast for moderate values of L. But not only this: Method II compares favorably to the popular Brock–Dechert–Scheinkman (BDS) independence test when applied to time series projected from the attractors of the Lorenz map and the time-delayed Hénon map. The bottom line is that determinism in noisy multivariate time series can be detected by observing a single component, a possibility that can come in handy in experimental situations.

Noisy univariate and multivariate time series have been intensively studied in the last few decades [1, 112]. Depending on the noise level of the data, one can expect to recover the full deterministic dynamics, to reconstruct the geometry of the noise-free

signal in some appropriate space, or just to ascertain the existence of an underlying determinism. The ordinal pattern-based methods described in this chapter falls in the third category. As a compensation for such a seemingly modest accomplishment, it has a remarkable success even with very high levels of noise. Besides the BDS method, which is based on the correlation dimension, other detection methods for determinism use the smoothness of the measure along reconstructed trajectories [164], functionals of probabilistic distributions [176], or the Higuchi fractal dimension on Poincaré sections [85].

9.1 Dynamical Robustness Against Observational Noise

Ordinal patterns are robust against small additive perturbations on account of being defined by inequalities. This property was called conditional robustness in Sect. 3.4.3. Yet, this property alone would not explain the persistence of forbidden patterns in the very noisy deterministic sequences that we are going to study in the next section. It turns out that, in deterministic sequences, there is a second mechanism for robustness, also in case of multi-dimensional maps—the dynamics itself. The result is an enhancement of the robustness of ordinal patterns against additive noise, which we call *dynamical robustness*. A simple explanation follows.

In the sequel we deal with a time series of the form

$$\xi_n = f^n(x_0) + w_n = x_n + w_n \tag{9.1}$$

($n \in \mathbb{N}_0$, or in practice $0 \leq n \leq N - 1$), where f is a self-map of the interval $[a, b] \subset \mathbb{R}$ and w_n are independent and uniformly distributed random variables (i.e., uniform white noise) in the interval $[-\eta, \eta]$. In order that the noise destroys a given allowed or forbidden pattern $\pi = \langle \pi_0, \ldots, \pi_{L-1} \rangle$ of the noise-free sequence $(x_n)_{n \in \mathbb{N}_0}$, it must happen that

$$x_{\pi_i} < x_{\pi_{i+1}}$$

but

$$x_{\pi_i} + w_{\pi_i} > x_{\pi_{i+1}} + w_{\pi_{i+1}}$$

for some $0 \leq i \leq L - 2$ and $w_{\pi_i}, w_{\pi_{i+1}} \in [-\eta, \eta]$. If η is small, this will be only possible if $x_{\pi_i} \approx x_{\pi_{i+1}}$, i.e., if $f^{\min\{\pi_i, \pi_{i+1}\}}(x_0)$ is an "approximately" periodic point with period $|\pi_i - \pi_{i+1}|$. We conclude that, indeed, the dynamics imposes an extra condition on $x_{\pi_i}, x_{\pi_{i+1}}$ so that a small amplitude perturbation can reverse their order.

To put some numbers on this argument, take $f(x) = 4x(1 - x)$, $0 \leq x \leq 1$, the logistic map. We know that for $\eta = 0$ this map has one forbidden 3 -pattern, namely, $\langle 2, 1, 0 \rangle$ (Fig. 1.6). In other words, there exists no $x \in [0, 1]$ such that $f^2(x) < f(x) < x$. The pattern $\langle 2, 1, 0 \rangle$ can appear in the noisy sequence (9.1) by

a single order reversal if the noise changes the order of x_n, x_{n+1} or the order of x_{n+1}, x_{n+2} in the allowed patterns

$$x_{n+2} < x_n < x_{n+1} \quad \text{or} \quad x_{n+1} < x_{n+2} < x_n,$$

respectively. In the first case, this requires $x_n \approx x_{n+1} = f(x_n)$, i.e., x_n must be close to any of the two fixed points of the map: $x = 0$ or $x = \frac{3}{4}$ (see Fig. 1.5). In the second case, the same applies to x_{n+1} and $x_{n+2} = f(x_{n+1})$. Therefore, it suffices to discuss the first case.

Consider the fixed point $x = 0$ and take $x_n = \delta > 0$. Then $x_{n+1} = f'(0)\delta + R\delta^2$, where R can be estimated with the remainder of the Taylor series. Since $\xi_n \in [x_n - \eta, \, x_n + \eta] =: I_n$, the inequality $\xi_{n+1} < \xi_n$ can be fulfilled only if the intervals I_n and I_{n+1} overlap, i.e., if

$$\delta \le \delta_0(\eta) = \frac{1 - f'(0) + \sqrt{(1 - f'(0))^2 + 8R\eta}}{2R}. \tag{9.2}$$

One can analogously estimate $\delta_+(\eta) > 0$ and $\delta_-(\eta) > 0$ such that if $x_n \in [\frac{3}{4} - \delta_-(\eta), \frac{3}{4} + \delta_+(\eta)]$, then x_n is sufficiently close to $x = \frac{3}{4}$ again in the sense that the inequality $\xi_{n+1} < \xi_n$ can hold for η small.

Thus, the probability $\Pr(\eta)$ for two consecutive orbit points $(x_n, x_{n+1}$ or $x_{n+1}, x_{n+2})$ to lie sufficiently close to either fixed point so as the pattern $\langle 2, 1, 0 \rangle$ becomes observable in a noisy orbit of the logistic map by means of a single order reversal is

$$\Pr(\eta) = \mu([0, \delta_0(\eta)]) + \mu([\tfrac{3}{4} - \delta_-(\eta), \tfrac{3}{4} + \delta_+(\eta)]),$$

where μ is the natural invariant measure for the logistic map,

$$\mu([c, d]) = \int_c^d \frac{dx}{\pi \sqrt{x(1 - x)}}$$

(see (1.20)). To make the argument even simpler, observe that once two consecutive orbit points in x_n, x_{n+1}, x_{n+2} are close to a fixed point, we may assume that the third one is around as well. In this case, the type of $\xi_n, \xi_{n+1}, \xi_{n+2}$ is going to depend basically on the type of w_n, w_{n+1}, w_{n+2}.

Consider now a string of length N, $\xi_0^{N-1} = \xi_0, \xi_1, \ldots, \xi_{N-1}$, along with the $\lfloor \frac{N}{3} \rfloor$ independent random vectors $\xi_n^{n+2} = \xi_n, \xi_{n+1}, \xi_{n+2}, n = 0, 3, 6, \ldots$. If we pick one of those vectors, the probability $\Pr(\langle 2, 1, 0 \rangle)$ that $\xi_{n+2} < \xi_{n+1} < \xi_n$ holds is then

$$\Pr(\langle 2, 1, 0 \rangle) \approx \Pr(\eta)\Pr\{w_{n+2} < w_{n+1} < w_n\})$$

$$= \Pr(\eta) \cdot \frac{1}{6}.$$

In order to verify these results, the probability P of finding at least once the pattern $\langle 2, 1, 0 \rangle$ in any of the $\lfloor \frac{N}{3} \rfloor$ windows $\xi_{3n}, \xi_{3n+1}, \xi_{3n+2}$ of the noisy time

series $(\xi_n)_{n=0}^{N-1}$, (9.1), was calculated numerically. From the reasoning above, this probability should be close to $1 - (1 - \Pr(\eta)/6)^{\lfloor \frac{N}{3} \rfloor}$ for the logistic map contaminated with additive, uniform white noise of small amplitude η, whereas it should be $1 - (1 - 1/6)^{\lfloor \frac{N}{3} \rfloor}$ for uniform white noise only (i.e., $\xi_n = w_n$ in (9.1)). Clearly, the former probability is greater than the latter because $\Pr(\eta)$ is going to be very small. This is confirmed by Fig. 9.1.
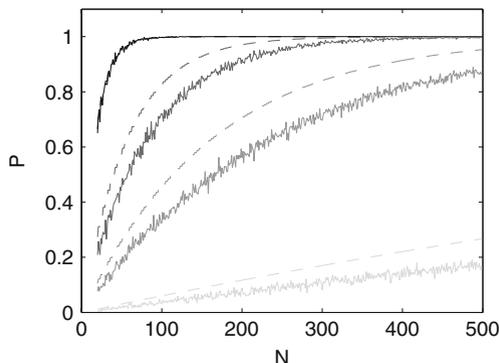
Fig. 9.1 Numerical computation (*continuous line*) and analytical estimation (*dashed*) of the probability P of finding the pattern $\langle 2, 1, 0 \rangle$ in any of the $\lfloor \frac{N}{3} \rfloor$ windows ξ_{3n}, ξ_{3n+1}, ξ_{3n+2} of a time series of length N generated with the logistic map. The noise amplitude is $\eta = 0.0001$ (*light gray*), $\eta = 0.01$ (*gray*), $\eta = 0.1$ (*dark gray*). The *top curve* corresponds to uniform white noise. Clearly the probability P is smaller for a noisy, deterministic time series than for uniform white noise

9.2 Detection of Determinism I: Number of Missing Ordinal Patterns

We already know (Sect. 1.2) that if $(x_n)_{n \in \mathbb{N}_0}$ is a univariate time series generated by a piecewise monotone interval map f, then there exist ordinal patterns which are forbidden for f. The theoretical situation in higher dimensions is less satisfactory in that the existence of forbidden patterns has been proved so far only under the somewhat restrictive condition of expansiveness (Sect. 7.6). There is nevertheless numerical evidence that forbidden ordinal patterns are also a general feature of higher dimensional dynamics. Since, on the other hand, univariate and multivariate random sequences have no forbidden patterns with probability 1, we conclude that the existence of forbidden patterns can be used as a fingerprint of deterministic orbit generation. Here "random sequence" means generated by an unconstrained, stochastic process taking on values in an interval. In summary, the difference between deterministic and random time series is clear-cut from an ordinal-theoretical point of view: the former have forbidden patterns while the latter have not.

However, when it comes to exploit this forbidden pattern-based strategy to detect determinism, two important practical issues arise: finiteness and noise contamination. Finiteness produces *false forbidden patterns*, that is, ordinal patterns which are

missing in a finite (segment of a) random sequence without constraints. Noise destroys forbidden patterns; for instance, a forbidden pattern of the "clean" sequence can turn visible because of additive random fluctuations. Let us mention in passing that were not for the observational noise, determinism could be easily ascertained, for example, with graphical methods. It is therefore interesting that ordinal patterns themselves provide the remedy to the two said issues. First of all, the number of false forbidden patterns of a fixed length always decreases with the length of the time series. Second, "true" forbidden patterns (i.e., forbidden patterns for an underlying deterministic dynamics) possess an additional dynamical robustness against additive noise (Sect. 9.1). This translates into a greater number of missing ordinal patterns in a noisy deterministic sequence than in a random one, and also to a slower decay rate with the length of the sequence. We shall shortly present numerical evidence that forbidden patterns persist in very noisy deterministic data—so noisy that the traditional methods [1, 112, 152] fail to uncover the underlying deterministic dynamics. But before coming to this point, let us dwell on some practical issues.

In practice one uses sliding windows of size L to comb a finite sequence $(x_n)_{n=0}^{N}$ for visible ordinal L-patterns. Note that a sequence of length N allows $N - L + 1$ windows of size L, for $2 \leq L \leq N$. Thus, in order to allow every possible ordinal pattern of length L to occur in a time series of length N, the condition $L! \leq N - L + 1$ must hold. Moreover, in cases where undersampling might occur, $N \gg L! + L - 1$ should also hold. As a rule of thumb we chose $(L + 1)! \leq N$ in the numerical simulations below, although $L! \leq N$ would do also in our case (very noisy data). Furthermore, $(x_n)_{n=0}^{N}$ will be initial segments of variable length $N \leq N_{\max} = 8000$, taken from a sequence $(x_n)_{n=0}^{N_{\max}}$. All these constraints leave $L = 4, 5, 6$ as interesting choices for L. In general one takes also moderate values for L, not least because of the sharp increase of the function $L!$.

Under these provisos, suppose now that the ordinal pattern $\pi \in S_L$ is missing in a finite noise-free time series. Of course, the odds that a *false* forbidden pattern persists in a random or deterministic sequence (or sample of sequences) will decrease exponentially with the number of data (see, e.g., Sect. 9.1). As a result, the number of false forbidden patterns in $(x_n)_{n=0}^{N}$ will decay as N increases up to $N_{\max}$, the number of data at our disposal. Otherwise, if $(x_n)_{n=0}^{N_{\max}}$ is a deterministic *noise-free* time series and π is a forbidden pattern, then π will be missing in $(x_n)_{n=0}^{N}$ for all $N \leq N_{\max}$. In other words, the number of *true* forbidden patterns in $(x_n)_{n=0}^{N}$ does not depend on N.

Consider a fixed initial condition x and suppose that $\pi_{\text{forb}} = \langle \pi_0, \ldots, \pi_{L-1} \rangle$ is a forbidden pattern for f. Suppose furthermore that we switch on a discrete-time random perturbation w_k, $|w_k| \leq w_{\max}$, such that π_{forb} is still missing in the finite sequence $\left(f^k(x) + w_k \right)_{k=0}^{N-1}$ (due to robustness). Observe that the *noisy* time series $\xi_k = f^k(x) + w_k$ can be viewed both as a perturbation of an underlying deterministic dynamics and as a random process correlated with the deterministic dynamics[1] f.

---

[1] Sometimes *colored noise* (i.e., a random process whose variables are statistically dependent) is numerically simulated in this way. For other methods, see, e.g., [113, 83].

If the orbit of x would be infinitely long, then the noisy time series had no missing patterns and π_{forb} would be visible with probability 1. In the finite-length case we are considering, this is in general not the case; rather, there is a threshold $\theta = \theta(N)$ (the greater N, the smaller θ) such that π_{forb} will do appear in $(\xi_k)_{k=0}^{N-1}$ only if $w_{\max} > \theta$. We conclude that amplifying a random perturbation destroys progressively the forbidden patterns of the underlying deterministic dynamic.

In the following we are going to test numerically one of the properties discussed above, namely, the robustness of *true* forbidden patterns against additive random perturbations. In order to estimate the average number $\langle n(L, N) \rangle$ of missing ordinal L-patterns in a finite, noisy sequence of length N,

$$\xi_k = x_k + w_k, \quad 0 \le k \le N - 1,$$

with $x_{k+1} = f(x_k)$ and w_k a random process, we generate 100 samples of length $N_{\max} = 8000$ and normalize the corresponding count of missing patterns of lengths $4 \le L \le 6$. To check the decay of $\langle n(L, N) \rangle$ with N, this parameter is allowed to vary in the range $(L + 1)! \le N \le N_{\max}$. We highlight next a few results obtained with f being the logistic map and w_k being white noise uniformly distributed in the interval $[-w_{\max}, w_{\max}], 0 \le w_{\max} \le 1$.

Figure 9.2 shows $\langle n(L, N) \rangle$ when (a) $w_{\max} = 0.25$, (b) $w_{\max} = 0.50$, and (c) $w_{\max} = 1$ and $f^k(x) = 0$ (noise only), respectively. Note the different orders of magnitude of the vertical scales. Needless to say, $\langle n(L, N) \rangle$ decays with increasing N because the greater the N, the more unlikely that an L-pattern is missing in a noisy or random sequence of length N; this is a statistical effect. The important features for us are the magnitude of $\langle n(L, N) \rangle$ and its decay rate with N, since these two properties are tightly related to the forbidden patterns of the underlying deterministic dynamic via robustness: the smaller the $w_{\max}$, the closer we are to the deterministic case, therefore, the more missing ordinal patterns and the slower their decrease with N.
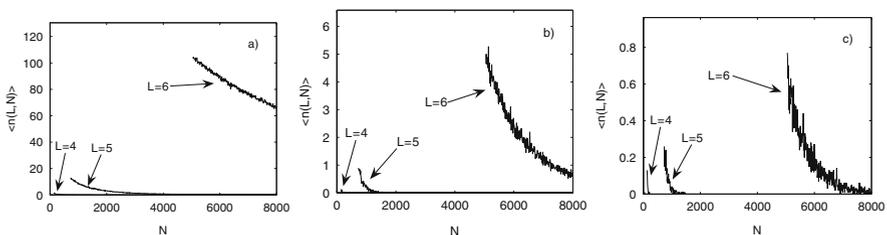
Fig. 9.2 Average number of missing ordinal patterns of length L found in a time series of length N, $\langle n(L, N) \rangle$, for noisy series of the logistic map with $w_{\max} = 0.25$ (**a**), $w_{\max} = 0.5$ (**b**), and for a series of uniformly distributed noise (**c**)

Figure 9.3 depicts ξ_{k+1} vs ξ_k in the previous cases (a) and (b). The higher order of magnitude of, e.g., $\langle n(6, N) \rangle$ in Fig. 9.2(b) as compared to Fig. 9.2(c) signalizes an underlying deterministic law, in spite of the fact that Fig. 9.3(b) hardly gives any clue about this.
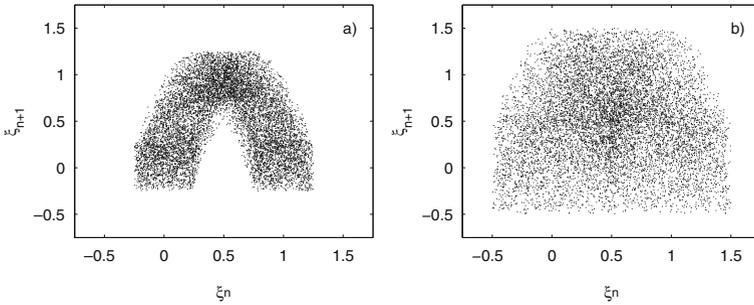
Fig. 9.3 Return map for noisy time series from the logistic map with $w_{max} = 0.25$ (**a**) and $w_{max} = 0.5$ (**b**). In the latter case, the high noise level does not allow to recognize the underlying deterministic dynamics. However, the number of missing ordinal patterns is sensibly higher than in the purely random case
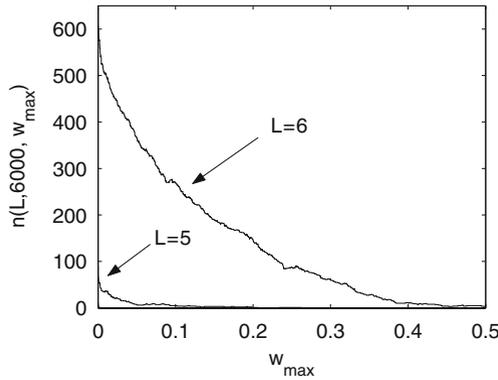
Fig. 9.4 Number of missing ordinal patterns of length L found in a noisy time series of the logistic map with length 6000 vs the uniform noise amplitude w_{max}

Finally, Fig. 9.4 nicely illustrates the resistance of the true forbidden patterns to disappear with increasing noise levels. In this figure, $N = 6000$, $L = 5, 6$, and $0 \le w_{max} \le 0.5$.

These numerical simulations suggest the following simple-minded, three-step method to discriminate noisy, deterministic, finite time series from random ones, at least when the noise is white.

(a) Compute the number of missing ordinal L-patterns of adequate length (say $(L + 1)! \le N$) in sliding windows along the sequence. It is convenient to use segments of variable length N and to draw the corresponding curves, as in Fig. 9.2.

(b) Randomize the sequence, i.e., change the temporal structure of the data in a random way.

(c) Proceed as in step (a) with the randomized sequence.

If the results of (a) and (c) are about the same, the sequence is very likely not deterministic (or the observational noise is so strong as compared to the deterministic signal that the latter has been completely masked). Otherwise, the sequence stems from a deterministic one. Needless to say, the method is more reliable if a statistically significant sample of sequences can be obtained, for instance, by cutting a long sequence into shorter pieces. In the next section we discuss a more quantitative method.

9.3 Detection of Determinism II: Distribution of Visible Ordinal Patterns

Consider once more a univariate or multivariate time series of the form

$$\xi_n = f^n(x_0) + w_n, \tag{9.3}$$

$(0 \le n \le N - 1)$ where w_n is white noise, i.e., outcomes of an independent and identically distributed (i.i.d.) random process. In order to differentiate white noise from a noisy deterministic time series of form (9.3), the perhaps simplest tool consists in counting visible ordinal patterns before and after randomizing the time series under scrutiny; depending on whether the number of visible patterns remains about the same or decreases significantly, we may conclude that the series is random or deterministic, respectively. This is the method discussed in Sect. 9.2.

A more quantitative method calls for performing a chi-square test based on the count of visible ordinal patterns. The *null hypothesis* reads

$$H_0: \quad \text{the } \xi_n \text{ are i.i.d.} \tag{9.4}$$

From a statistical point of view, this method is going to be a test of independence since the alternative to H_0 includes also colored noise.

The method goes as follows. Take sliding windows of size $L \ge 2$, overlapping at a single point (i.e., the last point of a window is the first point of the next one) down the sequence $\xi_0^{N-1} = \xi_0, \ldots, \xi_{N-1}$. For brevity, we call them "non-overlapping" windows. The number of such windows is

$$K = \left\lfloor \frac{N-1}{L-1} \right\rfloor, \tag{9.5}$$

each comprising the entries

$$\mathbf{e}_k = \xi_{kL-k}, \ldots, \xi_{(k+1)L-(k+1)}, \quad 0 \le k \le K - 1.$$

Notice that if the variables $\xi_0, \xi_1, \ldots, \xi_{N-1}$ are independently drawn from the same probability distribution, then the ordinal L-patterns defined by the components of $\mathbf{e}_k \in \mathbb{R}^L$, which we denote by $\pi(\mathbf{e}_k) \in \mathcal{S}_L$, will also be independent and, moreover,

uniformly distributed random variables. Therefore, if one or several ordinal patterns are missing in a sample obtained using non-overlapping windows, this might be a statistically significant signal that independence and/or the equality of the distribution are/is not fulfilled.

Given the non-overlapping windows $\{\mathbf{e}_k \in \mathbb{R}^L : k \geq 0\}$ corresponding to an arbitrarily long time series $\{\xi_n : n \geq 0\}$, suppose that some ordinal patterns of length L are missing in the initial segment $\xi_0, \xi_1, \ldots, \xi_{N-1}$. Let ν_π be the number of $\mathbf{e}_k$'s such that $\mathbf{e}_k$ is of type $\pi \in \mathcal{S}_L$ (i.e., $\pi(\mathbf{e}_k) = \pi$). Thus, $\nu_\pi = 0$ means that the L-pattern π has not been observed.

In order to accept or reject the null hypothesis H_0, (9.4), based on our observations, we apply a chi-square goodness-of-fit hypothesis test with statistic [135]

$$
\begin{aligned}
\chi^2(L) &= \sum_{\pi \in \mathcal{S}_L} \frac{(\nu_\pi - K/L!)^2}{K/L!} \\
&= \frac{L!}{K} \left(\sum_{\pi \in \mathcal{S}_L} \nu_\pi^2 - 2\frac{K}{L!} \sum_{\pi \in \mathcal{S}_L} \nu_\pi + \left(\frac{K}{L!}\right)^2 \sum_{\pi \in \mathcal{S}_L} 1 \right) \\
&= \frac{L!}{K} \sum_{\pi \in \mathcal{S}_L} \nu_\pi^2 - 2K + K \\
&= \frac{L!}{K} \sum_{\pi \in \mathcal{S}_L : \text{visible}} \nu_\pi^2 - K,
\end{aligned}
\tag{9.6}
$$

since (i) $\sum_{\pi \in \mathcal{S}_L} \nu_\pi = K$ and (ii) $\nu_\pi = 0$ if π is missing. Here $K/L!$ is the expected relative frequency of an ordinal L-pattern, if H_0 holds true. In the affirmative case, $\chi^2 = \chi^2(L)$ converges in distribution (as $K \to \infty$) to a chi-square distribution with $L! - 1$ degrees of freedom. Thus, for large K, a test with approximate level α is obtained by rejecting H_0 if $\chi^2 > \chi^2_{L!-1, 1-\alpha}$, where $\chi^2_{L!-1, 1-\alpha}$ is the upper $1 - \alpha$ critical point for the chi-square distribution with $L! - 1$ degrees of freedom [135]. In our case, the hypothetical convergence of χ^2 to the corresponding chi-square distribution may be considered sufficiently good if $\nu_\pi > 10$ for all visible L-patterns π, and

$$
\frac{K}{L!} > 5.
\tag{9.7}
$$

Notice that since this test is based on distributions, it could happen that a deterministic map has no forbidden L-patterns, thus $\nu_\pi \neq 0$ for all $\pi \in \mathcal{S}_L$; however, the null hypothesis be rejected because those ν_π's are not evenly distributed.

9.4 A Benchmark

A well-known benchmark for independence in time series is the Brock–Dechert–Scheinkman (BDS) test [38, 193], which is based on the correlation dimension. Since the numerical simulations below use the algorithm provided in [136], we follow this reference for the basics of the BDS test.

Let X_t, $t \geq 1$, be i.i.d. random variables, and

$$
I_\epsilon(x, y) = \begin{cases} 1 & \text{if } |x - y| < \epsilon, \\ 0 & \text{otherwise.} \end{cases}
$$

The probability that two length-m vectors are within ϵ can be estimated by the correlation sum

$$
C_{m,n}(\epsilon) = \frac{2}{n(n-1)} \sum_{s=1}^{n} \sum_{t=s+1}^{n} \prod_{j=0}^{m-1} I_\epsilon(X_{s-j}, X_{t-j}).
$$

It is shown in [38] that

$$
W_{m,n}(\epsilon) = \sqrt{n}\, \frac{C_{m,n}(\epsilon) - C_{1,n}^m(\epsilon)}{\sigma_{m,n}(\epsilon)}
$$

converges in distribution to a standard normal distribution. The normalization $\sigma_{m,n}(\epsilon)$ is given by

$$
\sigma_{m,n}^2(\epsilon) = 4 \left[B^m + 2 \sum_{j=1}^{m-1} B^{m-j} C^{2j} + (m-1)^2 C^{2m} - m^2 B C^{2m-2} \right],
$$

where C is consistently estimated by $C_{1,n}(\epsilon)$ and B can be estimated by

$$
B_n(\epsilon) = \frac{6}{n(n-1)(n-2)} \sum_{t=1}^{n} \sum_{s=t+1}^{n} \sum_{r=s+1}^{n} h_\epsilon(X_t, X_s, X_r),
$$

$$
h_\epsilon(i, j, k) = \frac{1}{3} \left[I_\epsilon(i, j) I_\epsilon(j, k) + I_\epsilon(i, k) I_\epsilon(k, j) + I_\epsilon(j, i) I_\epsilon(i, k) \right].
$$

A statistically significant non-zero value of $W_{m,n}(\epsilon)$ is evidence for determinism in the univariate time series $\{X_t : t \geq 1\}$.

This method relies on the selection of the parameters m and ϵ. Following the usual procedure [140], we take $\epsilon = 0.9^j$ with $j = 0, 1, 2, \ldots$. The criterion to say whether a combination of m and ϵ is "adequate" call for evaluating if a random time series is accepted as deterministic using this test the number of cases prescribed by the significance level of the test α.

9.5 Numerical Simulations

As underlying deterministic time series we use projections on the first coordinate of orbits generated by the Lorenz and time-delayed Hénon maps (this amounts in practice to using the standard lexicographical order). The additive noise w_n is modeled as Gaussian white noise,

$$\mathbb{E}(w_m \cdot w_n) = \sigma^2 \delta_{mn}$$

($\mathbb{E}$ stands for expectation value), with different standard deviations σ. Simulations with uniformly distributed noise yield similar results.

Two kinds of results are going to be presented in the two next sections: (i) Plots of the number of missing ordinal patterns as in Sect. 9.2 and (ii) plots of the distribution of the χ^2 statistic. Although the first ones provide only qualitative information, they can eventually complement the information provided by the second ones, as we shall see in the case of the Lorenz map. The specifics of plots (i) and (ii) are as follows.

(i) Let $N_{\max}$ denote the length of the data sequence under scrutiny and let $n(L, N)$ be the number of missing L-patterns in the initial segment $\xi_0, \xi_1, \ldots, \xi_{N-1}$ of variable length $N \leq N_{\max}$. The numbers $n(L, N)$ are determined with *overlapping* sliding windows of sizes $4 \leq L \leq 7$. In order to make the most of sequences of length $N_{\max} = 8000$, we take this time

$$L! \lesssim N \leq N_{\max}.$$

An average number $\langle n(L, N) \rangle$ is then estimated from 100 sequences.

(ii) *Non-overlapping* windows are used for the chi-square test of independence based on the distributions of ordinal L-patterns, with statistic (9.6)

$$\chi^2 = \chi^2(L) = \frac{L!}{K} \sum_{\pi \in \mathcal{S}_L : \text{visible}} v_\pi^2 - K. \tag{9.8}$$

Here, $K = \left\lfloor \frac{N-1}{L-1} \right\rfloor$ is the number of non-overlapping windows of size L in a data sequence of length N, (9.5). The window sizes in the simulations are $L = 4, 5$. For $L = 4$, the acceptance/rejection thresholds of the null hypothesis (9.5) at levels $\alpha = 0.10, 0.05$ are

$$\chi^2_{23, 0.90} = 32.01, \quad \chi^2_{23, 0.95} = 35.17, \tag{9.9}$$

respectively. For $L \geq 5$, corresponding to degrees of freedom over 100, the following approximation for the thresholds $\chi^2_{L!-1, 1-\alpha}$ is used [135]:

$$\chi^2_{L!-1, 1-\alpha} \approx (L! - 1) \left(1 - \frac{2}{9(L! - 1)} + z_{1-\alpha} \sqrt{\frac{2}{9(L! - 1)}} \right)^3,$$

where $z_{1-\alpha}$ is the upper $1 - \alpha$ critical point for the standard normal distribution, $\mathcal{N}(0, 1)$; in particular, $z_{0.90} = 1.282$ and $z_{0.95} = 1.645$. Thus,

$$\chi^2_{119, 0.90} = 139.15, \quad \chi^2_{119, 0.95} = 145.46. \tag{9.10}$$

Remember from (9.7) that $5L! \lesssim K$ should hold for the chi-square test to be statistically significant. Therefore

$$5L! \lesssim \frac{N}{L - 1},$$

i.e., $N \gtrsim 5(L - 1)L!$. In consequence we take sequences of length $N = 1000$ for $L = 4$ and $N = 8000$ for $L = 5$. To plot the χ^2-value distribution, a sample of $10,000$ sequences was used.

The numerical results are summarized in the following two sections.

9.5.1 The Lorenz Map

The *Lorenz map* [193] is defined as

$$x_{n+1} = x_n y_n - z_n, \quad y_{n+1} = x_n, \quad z_{n+1} = y_n. \tag{9.11}$$

It has an attractor with *Kaplan–Yorke dimension* $D_{KY} = 2$ [193]. Assuming the well-tested Kaplan–Yorke conjecture $D_{KY} = D_1$, where D_1 is the *information dimension*, then the *fractal dimension* D_0 satisfies

$$D_0 \geq D_1 = 2.$$

Figure 9.5 shows the return map $\xi_{n+1} = x_{n+1} + w_{n+1}$ vs $\xi_n = x_n + w_n$ for a typical orbit of the Lorenz map on its attractor and additive Gaussian white noise w_n with $\sigma = 0.25$ ($SNR^2 \simeq 10$ dB). The geometry of the attractor has been completely washed out by the noise, but the underlying determinism can still be detected because of the different count of missing ordinal patterns before (Fig. 9.6) and after (Fig. 9.7) switching off the deterministic signal. Not only the count of missing ordinal patterns is different in these two cases, but also their decay rate with N. The different behavior in Fig. 9.6 of the curve $L = 4$, on the one hand, and the curves $L \geq 5$, on the other hand, strongly indicates that the Lorenz map has no forbidden 4-patterns.

Figure 9.8 shows the distribution of the statistic χ^2, (9.8), obtained from 10,000 projections x_0^{N-1} of orbits of the Lorenz map, contaminated with additive Gaussian noise with $\sigma = 0.25, 0.50$ ($SNR \simeq 10, 4.0$ dB, respectively). Since the rejection

---

[2] SNR is short for "signal-to-noise ratio" and dB is short for "decibel."
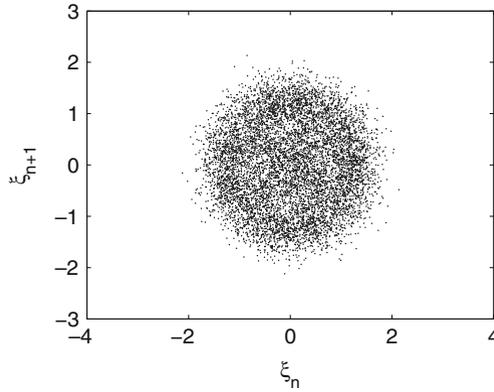
Fig. 9.5 Return map for a time series of the Lorenz map contaminated with Gaussian white noise with $\sigma = 0.25$ (SNR $\simeq 10$ dB). The structure of the underlying chaotic attractor has been totally blurred. However, the count of missing ordinal patterns is sensibly higher than in the purely random case
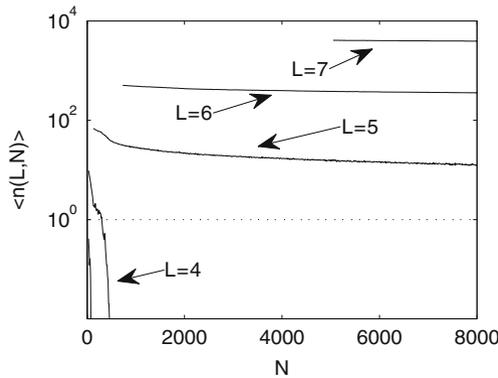
Fig. 9.6 Average number of missing ordinal patterns of length L found in a time series of length N, $\langle n(L, N) \rangle$ (in logarithmic scale), for a noisy series of the Lorenz map with $\sigma = 0.25$ (SNR $\simeq 10$ dB)

threshold of the null hypothesis H_0 (9.4) at level $\alpha = 0.05$ is $\chi^2_{23, 0.95} = 35.17$ in (a) and $\chi^2_{119, 0.95} = 145.46$ in (b), see (9.9), the χ^2 test clearly detects determinism. It is worth noticing that the rejection of H_0 in case (a) is due to the non-uniform distribution of ν_π since, according to Fig. 9.6, all 4-patterns are visible in noisy time series generated by the Lorenz map with $N \gtrsim 500$ and $\sigma = 0.25$.

Finally, the comparison with the BDS test is shown in Fig. 9.9. There we show the probability P of rejecting the null hypothesis (9.4) for the 27 possible adequate BDS tests on a time series $\xi_0^{N-1} = (x_n + w_n)_{n=0}^{N-1}$ of length $N = 1000$, where now w_n is Gaussian white noise with $0 \leq \sigma \leq 2$. In the same figure it is also
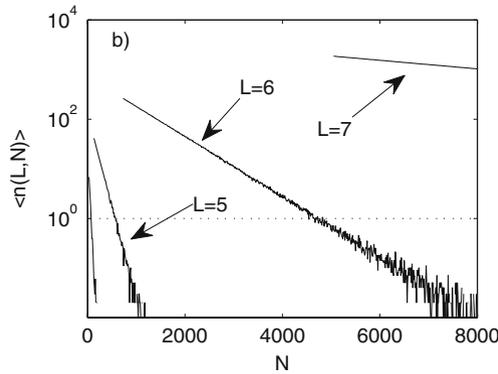
Fig. 9.7 Average number of missing ordinal patterns of length L found in a time series of length N, $\langle n(L,N) \rangle$ (in logarithmic scale), for time series of Gaussian white noise with $\sigma = 0.25$
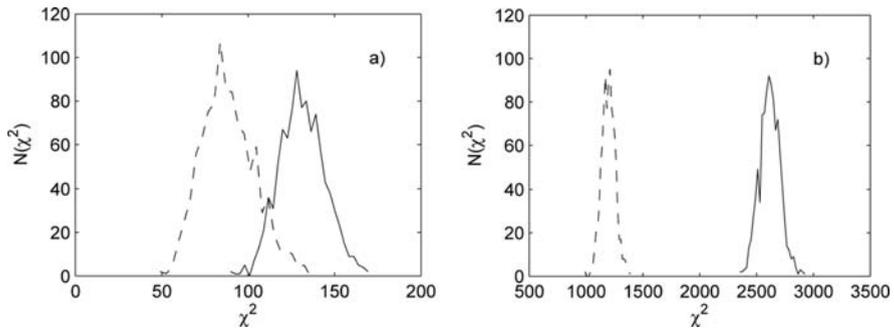
Fig. 9.8 Distribution $N(\chi^2)$ of χ^2 for 10, 000 noisy sequences generated with the Lorenz map, for $L = 4$, $N = 1000$, $\sigma = 0.25$ (*continuous line*) and $\sigma = 0.50$ (*dashed line*) (SNR $\simeq$ 10, 4.0 dB, respectively) (**a**) and for $L = 5$, $N = 8000$, $\sigma = 0.25$ (*continuous line*) and $\sigma = 0.50$ (*dashed line*) (SNR $\simeq$ 10, 4.0 dB, respectively) (**b**)

plotted the probability P of rejecting the null hypothesis using the chi-square test with the same level $\alpha = 0.05$. Notice that the chi-square test correctly rejects the null hypothesis with higher probability than the BDS test in the high-noise regime ($\sigma \geq 1$), and its performance is comparable to the best one of the BDS test in the low-noise regime ($\sigma \leq 1$). Put in a different way, the probability of a false positive is higher with the BDS test. We conclude also from Fig. 9.9 that the BDS test performance strongly depends on the combinations of ϵ and m; for some combinations, this method wrongly accepts the null hypothesis even for small values of σ.

9.5.2 The Delayed Hénon Map

The *time-delayed Hénon map* [194] is defined as

$$x_n = 1 - ax_{n-1}^2 + bx_{n-d}, \tag{9.12}$$
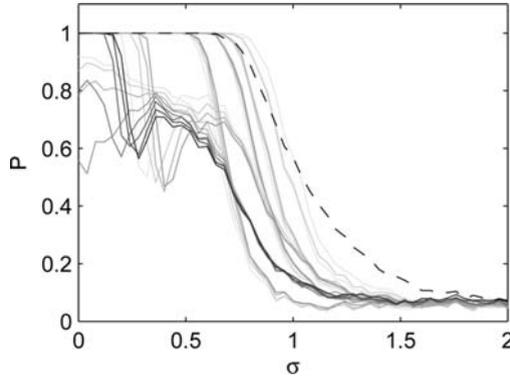
Fig. 9.9 The *continuous lines* indicate the probability of rejecting the null hypothesis H_0 ("the time series is i.i.d.") for a time series projected from the Lorenz map's attractor, contaminated with Gaussian white noise with σ up to $\sigma = 2$, when applying the BDS test with level $\alpha = 0.05$. In total, 27 tests for different combinations of ϵ and m were performed. The lighter the *gray color* is, the bigger is the value of ϵ used (see text for details). The *dashed line* indicates the probability of rejecting H_0 when using the chi-square test based on missing ordinal patterns, with the same level $\alpha = 0.05$. The chi-square test correctly rejects the null hypothesis more often than the BDS test

where a, b are real constants and $d \geq 1$. For $d = 1$, the time-delayed Hénon map is equivalent to the logistic map $x_{n+1} = Ax_{n-1}(1 - x_{n-1})$, with [194]

$$A = \frac{b-1}{2a} \pm \frac{1}{2a}\sqrt{(b-1)^2 + 4a}.$$

For $d = 2$ and $a = 1.4, b = 0.3$, we recover the familiar two-dimensional dissipative Hénon map.

For $a = 1.6$ and $b = 0.1$, Sprott [194] finds the following linear relation between D_{KY} and d over the range $1 \leq d \leq 100$:

$$D_{KY} \cong 0.192d + 0.699.$$

The Kaplan–Yorke conjecture implies now

$$D_0 \geq D_1 = D_{KY} \cong 0.192d + 0.699$$

for the fractal dimension D_0 of the attractor, $1 \leq d \leq 100$. In particular, $D_0 \geq 1.083$ for $d = 2$, $D_0 \geq 10.299$ for $d = 50$, and $D_0 \geq 19.899$ for $d = 100$. Thus, this family of maps provides attractors with a wide range of fractal dimensions.

Figure 9.10 shows the return map ξ_{n+1} vs ξ_n for a typical orbit on the attractor of the time-delayed Hénon map with $d = 50$, both in the absence of noise, $\xi_n = x_n$ (a) and corrupted with Gaussian white noise, $\xi_n = x_n + w_n$, with $\sigma = 0.5$ (SNR $\simeq$ 1.3 dB) (b). Again, the geometry of the attractor has been completely blurred by the
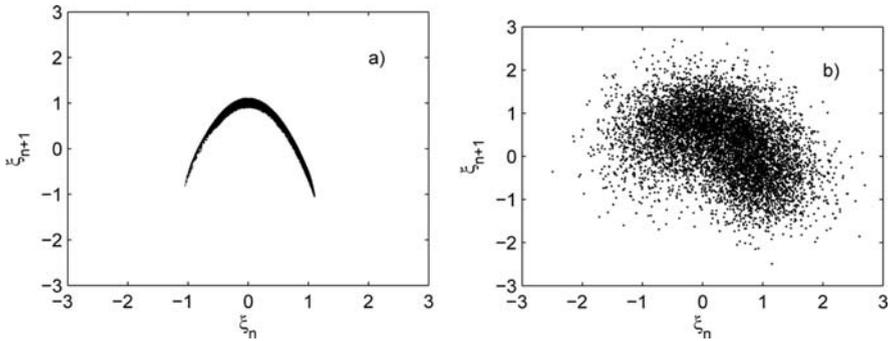
Fig. 9.10 Return map for a time series of the time-delayed Hénon map with $d = 50$ in the absence of noise (**a**) and contaminated with Gaussian white noise with $\sigma = 0.5$ (SNR $\simeq 1.3$ dB) (**b**). The structure of the underlying chaotic attractor has been totally blurred. Here again the count of missing ordinal patterns is sensibly higher than in the purely random case
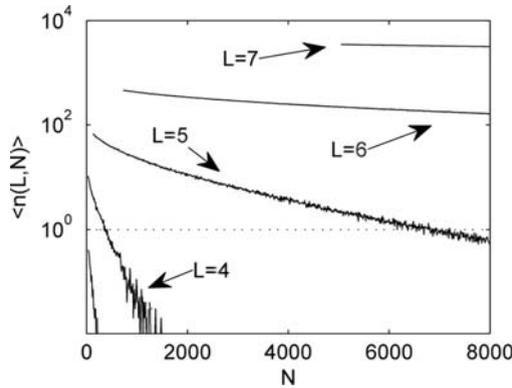
Fig. 9.11 Average number of missing ordinal patterns of length L found in a time series of length N, $\langle n(L,N) \rangle$ (in logarithmic scale), for a noisy series of the time-delayed Hénon map with $\sigma = 0.5$ (SNR $\simeq 1.3$ dB)

presence of the noise. However, it can be seen in Fig. 9.11 that also in this case, the number of missing ordinal L-patterns found in a time series of length N, $\langle n(L,N) \rangle$, is sensibly larger than in the white noise-only case, Fig. 9.7.

Figure 9.12(a)–(c) depicts the comparison of the chi-square test with the BDS test for $d = 2$, $d = 50$, and $d = 100$, respectively. Again, the probability of a false positive is higher with the BDS test. Since we are interested in the detection of determinism, we may conclude that the chi-square test, based on the distribution of visible ordinal patterns, is more reliable.

In conclusion, the (conditional + dynamical) robustness against additive noise of the forbidden patterns makes them a practical tool to distinguish deterministic, noisy time series from white noise. It is in this sense that we claim that forbidden patterns can be used to detect determinism in noisy time series—determinism as opposite to
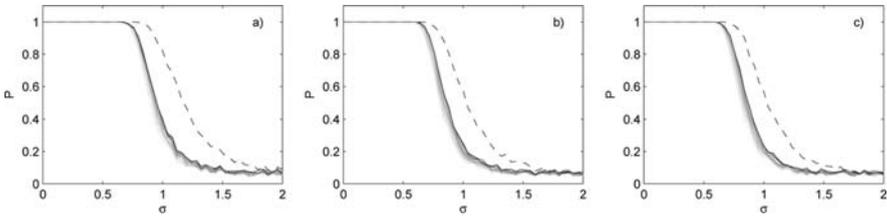
Fig. 9.12 Comparison of the chi-square test and the BDS test applied to projections of the time-delayed Hénon map with $d = 2$ (**a**), $d = 50$ (**b**), $d = 100$ (**c**), and Gaussian white noise with $0 \leq \sigma \leq 2$. The *continuous lines* indicate the probability of rejecting the null hypothesis H_0 ("the time series is i.i.d.") when applying the BDS test with level $\alpha = 0.05$. In total, 27 tests with different combinations of ϵ and m were performed. The lighter the *gray color* is, the bigger is the value of ϵ. The *dashed line* indicates the probability of rejecting H_0 when using the chi-square test with the same level $\alpha = 0.05$. Clearly, the chi-square test rejects the null hypothesis more often than the BDS for all noise values and for the three values of d

statistical independence. In fact, determinism is usually equated to statistical dependence among the observations in applications. On the other hand, the discrimination of deterministic, noisy time series from colored noise seems problematic, although some interesting methods have been proposed; see, e.g., [119] for a method based on nonlinear predictability.