NONLINEAR TIME SERIES ANALYSIS, WITH APPLICATIONS TO MEDICINE

José María Amigó

Centro de Investigación Operativa, Universidad Miguel Hernández, Elche (Spain)

LECTURE 3

SYMBOLIC DYNAMICS

OUTLINE

Lecture 1 **RANDOM PROCESSES**	Lecture 2 **DYNAMICAL SYSTEMS**
Random sequence	**Orbit**
Stationary probability distribution	**Invariant measure**
Stationary random process	**Measure-preserving DS**
Shannon entropy	**???**

Questions.

Q1 Can RP be formulated as DS?

A1 Yes, via *shift systems* (Sect. 1)

Q2 Can a DS generate a random sequence?

A2 Yes, via *partitions* (Sects. 2 & 4)

1. Shift systems

Fact. *Every stationary, finite-state* $\mathbf{X} = \{X_n\}$ *can be associated with a measure-preserving DS* $(\mathcal{X}^\infty, \sigma, m)$

- the state space $\mathcal{X}^\infty$ is a *sequence space*,

$$\mathcal{X}^\infty = \begin{cases} \{x_0^\infty = (x_0, x_1, ..., x_n, ...) : x_n \in \mathcal{X}\} & \text{if } \mathbf{X} \text{ is one-sided} \\ \{x_{-\infty}^\infty = (..., x_{-n}, ..., x_0, ..., x_n, ...) : x_n \in \mathcal{X}\} & \text{if } \mathbf{X} \text{ is two-sided} \end{cases}$$

- the map σ is the (left) *shift transformation*,

$$\sigma(..., x_{-n}, ..., x_0, x_1, ..., x_n, ...) = (..., x_{-n+1}, ..., x_1, x_2, ..., x_{n+1}, ...),$$

- the *shift invariant measure* m is

$$m\{x_0^\infty \text{ or } x_{-\infty}^\infty : x_{i_1} = a_1, ..., x_{i_n} = a_n\} = \Pr\{X_{i_1} = a_1, ..., X_{i_n} = a_n\}$$

$\Longrightarrow$ $(\mathcal{X}^\infty, \sigma, m)$ is called the *shift space model* of $\mathbf{X}$.

Remarks.

- Shift space models allow to deal RP as DS
- The states are infinite sequences.
- The "transported" measure m is invariant because $\mathbf{X}$ is stationary.
- The shift transformation σ models time passing.
- $\mathbf{X}$ is ergodic iff $(\mathcal{X}^{\infty}, \sigma, \mu)$ is ergodic.

1. Shift systems

Example. (*Coin tossing*) $X_n \in \{0,1\}$, $n \geq 0$, with

$$\Pr\{X_n = 0\} = p_0, \Pr\{X_n = 1\} = p_1 = 1 - p_0.$$

- $\mathcal{X}^\infty = \{x_0^\infty = (x_0, x_1, ..., x_n, ...) : x_n = 0, 1\} = \{\text{binary sequences}\}$
- $m\{x_0^\infty : x_n = i_n, ..., x_{n+l} = i_{n+l}\} = p_{i_n} ... p_{i_{n+l}}$.

This shift space is called the (p_0, p_1)-*Bernoulli shift system*.

Interpretation: Each binary sequence x_0^∞ is a possible outcome of the random experiment.

1. Shift systems

Generalization. (*Dice rolling, etc.*) If X_n are i.i.d. random variables, $\mathcal{X} = \{0, ..., k-1\}$, and

$$\Pr\{X_n = i\} = p_i,$$

the shift space model is called the $(p_0, ..., p_{k-1})$-*Bernoulli shift system*.

They exhibit all properties of low-dimensional chaos:

- Sensitivity to the initial condition (*butterfly effect*)
- Ergodicity
- Transitivity (existence of a dense orbit = Boltzmann's *Ergodenhypothese*)

Next we address Question 2.

Definition. A (finite) *partition* of Ω is a finite family of subsets $\alpha = \{A_0, ..., A_{k-1}\}$ s.t.

(1) $A_i \cap A_j = \emptyset$ for $i \neq j$,

(2) $A_0 \cup A_1 \cup ... \cup A_{k-1} = \Omega$.

Example. *Partition of a 1D interval* $\Omega = [a, b]$ *into* k *bins* (binning): Let

$$\Delta = \frac{b-a}{k},$$

then

$$A_0 = [a, a+\Delta), A_1 = [a+\Delta, a+2\Delta), ..., A_{k-1} = [a+(k-1)\Delta, a+k\Delta].$$

The process of partitioning a state space is called "*coarse-graining*" or "*quantification*".

2. Symbolic dynamics

Definition. Given

- a measure-preserving dynamical system (Ω, f, μ), and
- a partition $\alpha = \{A_0, ..., A_{k-1}\}$ of Ω,

we associate to each $x \in \Omega$ its *itinerary* wrt α, i.e.

$$x \mapsto i_0, i_1, ..., i_n, ... \text{ with } i_n = j \text{ if } f^n(x) \in A_j.$$

Example. Let $\Omega = [0,1]$, $f(x) = 4x(1-x)$, and

$$\alpha = \{A_0, A_1\},\ A_0 = [0, 1/2),\ A_0 = [1/2, 1],$$

and $x_0 = 0.1$. Then

$$\begin{aligned} \text{orbit of } x_0 &= 0.1, 0.36, 0.921\,6, 0.289\,01, 0.821\,94, 0.585\,42, ... \\ \text{itinerary of } x_0 &= 0, 0, 1, 0, 1, 1, ... \end{aligned}$$

2. Symbolic dynamics

Let $i_0, i_1, ..., i_n, ...$ be the itinerary of x wrt $\alpha = \{A_0, ..., A_{k-1}\}$. Set

$$\mathbf{X}^{\alpha}(x) = i_0, i_1, ..., i_n, ... \equiv \{X_n^{\alpha}(x)\}_{n\geq 0}$$

Fact. $\mathbf{X}^{\alpha}$ is a *stationary*, *finite-alphabet* RP, $\mathcal{X} = \{0, ..., k-1\}$, with

$$\Pr\{X_0^{\alpha} = i_0, X_1^{\alpha} = i_1, ..., X_n^{\alpha} = i_n\} = \mu\left(A_{i_0} \cap f^{-1} A_{i_1} \cap ... \cap f^{-n} A_{i_n}\right).$$

Definition. $\mathbf{X}^{\alpha}$ is called the *symbolic dynamics of f wrt α*.

- *If f is invertible, the itineraries and symbolic dynamics are two-sided.*

3. Kolmogorov-Sinai entropy

Let $\mathbf{X}^\alpha = \{X_n^\alpha\}$ be the symbolic dynamics of f wrt to $\alpha = \{A_0, ..., A_{k-1}\}$.

Definition.

- The entropy of f wrt α is

$$h_\mu(f,\alpha) = h(\mathbf{X}^\alpha)$$

- The *metric* (or *Kolmogorov-Sinai*) *entropy* of f is

$$h_\mu(f) = \sup_\alpha h_\mu(f,\alpha)$$

Fact. If $(\mathcal{X}^\infty, \sigma, m)$ is the shift space model of a random process $\mathbf{X}$, then

$$h_m(\sigma) = h(\mathbf{X}).$$

3. Kolmogorov-Sinai entropy

A partition γ of Ω is called a *generating partition* or a *generator* of f if

$$h_\mu(f) = h_\mu(f, \gamma).$$

The computation of $h_\mu(f)$ is in general difficult. **Exceptions**:

1. A generator of f is known (seldom, but there are numerical methods).
2. If the invariant measure is smooth (i.e., $\mu(dx) = \rho(x)dx$ with ρ differentiable), the KS entropy is the sum of the positive Lyapunov exponents (*Pesin's formula*).
3. A closed formula is known for some maps (Bernoulli shifts, etc.)

Otherwise. Calculate $h_\mu(f, \alpha)$ for ever finer box partitions $\alpha_1, \alpha_2, ..., \alpha_n, ...$

$$\lim_{n\to\infty} h_\mu(f, \alpha_n) = h_\mu(f).$$

4. Generating partitions

Example. Let $(\mathcal{X}^\infty, \sigma, m)$ be a $(p_0, ..., p_{k-1})$-Bernoulli (one-sided) shift space. The partition $\gamma = \{C_0, ..., C_{k-1}\}$,

$$C_0 = \{x_0^\infty : x_0 = 0\},\ C_1 = \{x_0^\infty : x_0 = 1\}, ...,\ C_{k-1} = \{x_0^\infty : x_0 = k-1\}$$

can be proved to be a generator of the shift transformation, so

$$h_m(\sigma) = h_m(\sigma, \gamma) = -\sum_{i=0}^{k-1} m(C_i) \log m(C_i).$$

4. Generating partitions

Let (Ω, f, μ) be a measure-preserving dynamical system. There exist generators of f under quite general conditions.

Fact. Let γ be a generator of f. Then

the shift space model of $\mathbf{X}^{\gamma}$ is an "isomorphic copy" of (Ω, f, μ)

Consequences.

itinerary $\mathbf{X}^{\gamma}(x_0)$	$\leftrightarrow$	initial condition x_0
KS entropy $h_m(\sigma) = h(\mathbf{X}^{\gamma})$	$=$	KS entropy $h_\mu(f)$

4. Generating partitions

If two DS are isomorphic, their generators correspond.

Example. The logistic and tent maps are isomorphic to the $(\frac{1}{2}, \frac{1}{2})$-Bernoulli (one-sided) shift via measure-preserving transformations[1] that map the generator

$$\gamma = \{C_0, C_1\}, \text{ where } C_0 = \{x_0^\infty : x_0 = 0\}, C_1 = \{x_0^\infty : x_0 = 1\},$$

of the Bernoulli shift into the partition

$$\alpha = \{A_0, A_1\}, \text{ where } A_0 = [0, \tfrac{1}{2}),\ A_1 = [\tfrac{1}{2}, 1],$$

of $\Omega = [0, 1]$.

[1] J.M.Amigó, *Permutation Complexity in Dynamical Systems*, Springer Verlag, 2010.

Application. Numerical generation of random binary sequences.

1. Take a number $x_0 \in [0,1]$.
2. Let f be the logistic or (better) the tent map. Set

$$b_n = \begin{cases} 0 & \text{if } f^n(x_0) < 0.5 \\ 1 & \text{if } f^n(x_0) \geq 0.5 \end{cases}$$

Then the binary sequence $\{b_n\}_{n\geq 0}$ is i.i.d. with

$$\begin{aligned} \Pr\{b_n = 0\} &= \mu\{[0,\tfrac{1}{2})\} = \tfrac{1}{2}, \\ \Pr\{b_n = 1\} &= \mu\{[\tfrac{1}{2},1]\} = \tfrac{1}{2}. \end{aligned}$$

Warning. Computers are finite-state machines!

Summary:

RANDOM PROCESSES		DYNAMICAL SYSTEMS
Stationary random process	$\longrightarrow$	Shift space model
Symbolic dynamics wrt α	$\longleftarrow$	DS + Partition α
Symbolic dynamics wrt γ	$=$	DS + Generator γ
Shannon entropy	$\longleftrightarrow$	Kolmogorov-Sinai entropy

Ordinal patterns provide a natural way to define a symbolic dynamics.

Definition. The "*ordinal L-pattern*", "*rank vector*" or "*type*" of L points $x_0, x_1, ..., x_{L-1}$ in a linearly ordered set Ω is the *permutation*

$$\{0, 1, ..., L-1\} \longrightarrow \{\pi_0, \pi_1, ..., \pi_{L-1}\}$$

such that

$$x_{\pi_0} < x_{\pi_1} < ... < x_{\pi_{L-1}}.$$

Notation.

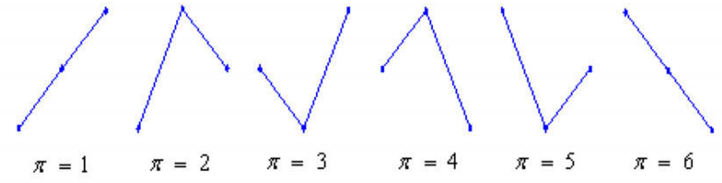
- $\pi = \langle \pi_0, \pi_1, ..., \pi_{L-1} \rangle$
- $\{\text{ordinal } L\text{-patterns}\} = \mathcal{S}_L \quad (\#\mathcal{S}_L = L!)$

Convention. If $x_i = x_j$ then we set $x_i < x_j$ if $i < j$.

Examples.

1. $\Omega = \mathbb{R}$,
$$x_0 = \sqrt{3},\ x_1 = e,\ x_2 = 2,\ x_3 = -1.7,$$
then
$$\pi = \langle 3, 0, 2, 1 \rangle .$$

2. Ordinal patterns of length $L = 3$.



5. Ordinal symbolic dynamics

If $x_0, x_1, ..., x_{L-1} = x_0, f(x_0), ..., f^{L-1}(x_0)$ has type π, then we say that x_0 *has type* π.

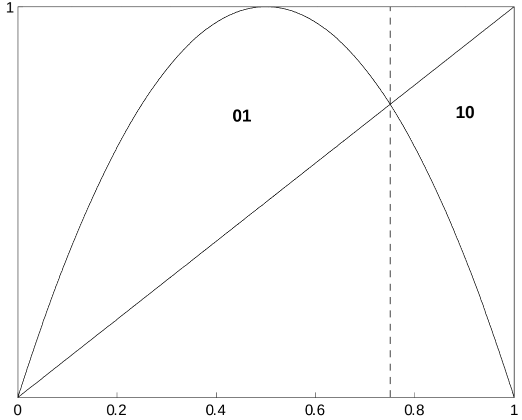
Example. $I = [0, 1]$, $f(x) = 4x(1-x)$, then

$$(f^n(0.6416))_{n\geq 0} = 0.6416,\ 0.9198,\ 0.2951,\ 0.8320, 0.5590,\ 0.9861, \ldots$$

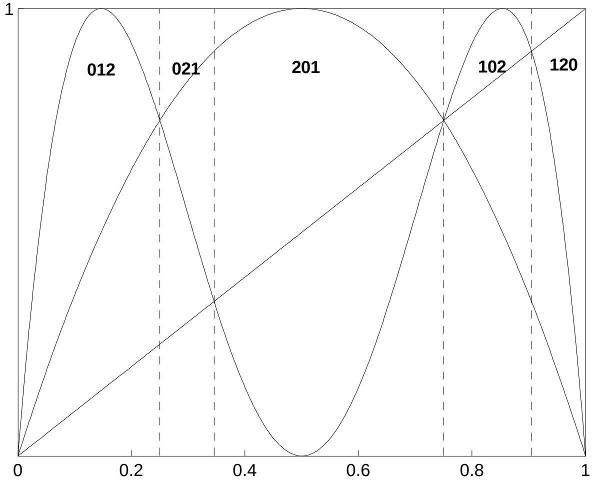
Hence $x = 0.6416$ has the types

$$\langle 0,1 \rangle, \langle 2,0,1 \rangle, \langle 2,0,3,1 \rangle, \langle 2,4,0,3,1 \rangle, \langle 2,4,0,3,1,5 \rangle, \ldots$$

Example (cont'd). *Visualization of ordinal 2-patterns*



Example (cont'd). *Visualization of ordinal 3-patterns*



5. Ordinal symbolic dynamics

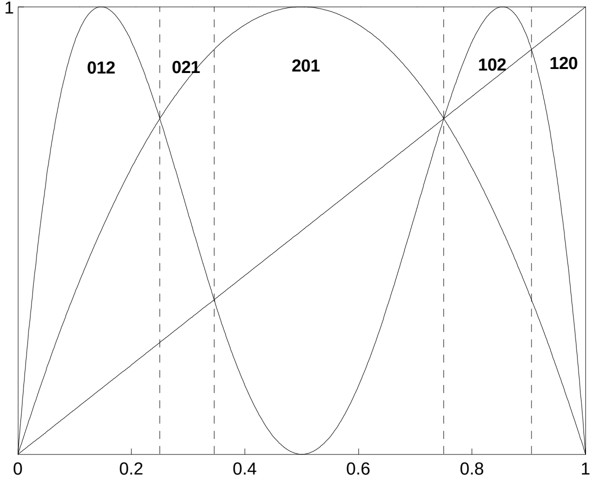
- *Ordinal symbolic dynamics* is the symbolic dynamics which symbols are ordinal patterns of a fixed length L.
- The state space Ω gets divided in $L!$ disjoint subsets P_π, $\pi \in \mathcal{S}_L$, namely
$$P_\pi = \{x \in \Omega : x \text{ has type } \pi \in \mathcal{S}_L\}.$$
- The partition
$$\mathcal{P}_L = \{P_\pi \neq \emptyset : \pi \in S_L\}$$
is called the *ordinal partition* of Ω of length L.
- Use $3 \leq L \leq 7$ in applications.

Definition. An ordinal L-pattern π is said to be *forbidden* for f if $P_\pi = \emptyset$, i.e., there is no $x \in \Omega$ of type π. Otherwise they are called *admissible*.

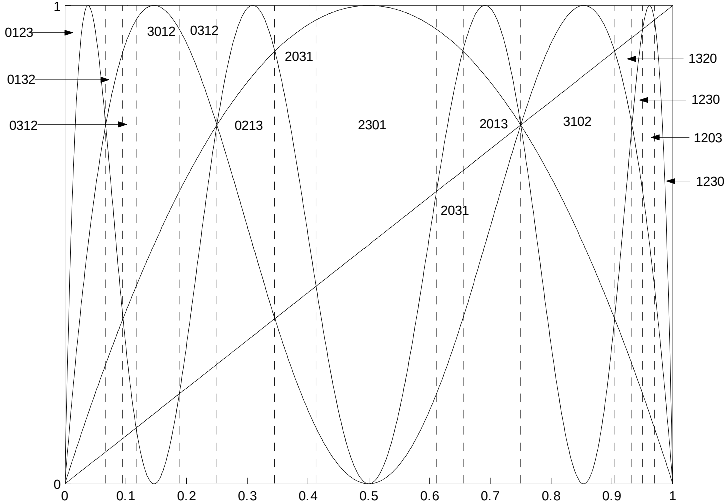
If Ω is an interval of $\mathbb{R}$, $f : \Omega \to \Omega$ is called *piecewise monotone* if there is a *finite* partition of Ω into intervals, such that f is continuous and monotone on each of those intervals.

Fact. *If f is a piecewise strictly monotone interval map, then it has forbidden pattern of sufficiently large length.*

Example. The logistic map has 1 forbidden 3-pattern (**210**)



Example. The logistic map has 12 forbidden 4-patterns.



A) Permutation entropy of a random process

- $\mathbf{X} = \{X_n\}_{n\geq 0}$ a random process
- $p(\pi)$ the probability that $X_0, ..., X_{L-1}$ has type $\pi \in \mathcal{S}_L$

Then, the *permutation entropy of order* L of $\mathbf{X}$ is

$$h^*(X_1, ...X_L) = -\frac{1}{L-1} \sum_{\pi \in \mathcal{S}_L} p(\pi) \log p(\pi),$$

and the *permutation entropy of* $\mathbf{X}$ is

$$h^*(\mathbf{X}) = \lim_{L\to\infty} h^*(X_1, ...X_L) = -\lim_{L\to\infty} \frac{1}{L-1} \sum_{\pi \in \mathcal{S}_L} p(\pi) \log p(\pi).$$

Fact[2]**.** If $\mathbf{X}$ is finite-alphabet and stationary, then $h^*(\mathbf{X}) = h(\mathbf{X})$.

[2] J.M.Amigó, Physica D 241 (2012) 789.

B) Permutation entropy of a dynamical system

- (Ω, f, μ) a measure-preserving DS
- $\mathcal{P}_L = \{P_\pi \neq \emptyset : \pi \in S_L\}$ the ordinal partition

Then, the *metric permutation entropy* of order L of f is

$$h_\mu^*(f;\mathcal{P}_L) = -\frac{1}{L-1}\sum_{\pi\in\mathcal{S}_L}\mu(P_\pi)\log\mu(P_\pi),$$

and the *permutation entropy of* f is

$$h_\mu^*(f) = -\lim_{L\to\infty} h_\mu^*(f;\mathcal{P}_L) = -\lim_{L\to\infty}\frac{1}{L-1}\sum_{\pi\in\mathcal{S}_L}\mu(P_\pi)\log\mu(P_\pi),$$

Fact[3]**.** If Ω is a 1D interval and f is piecewise-monotone,

$$h_\mu(f) = h_\mu^*(f) = \lim_{L\to\infty} h_\mu^*(f;\mathcal{P}_L).$$

$\Longrightarrow$ *The ordinal partitions* $\mathcal{P}_2, \mathcal{P}_3, ..., \mathcal{P}_L, ...$ *build a generating sequence.*

---

[3] C. Bandt, G. Keller, and B. Pompe, Nonlinearity 15 (2002) 646.

Detection of determinism in noisy signals is an application of ordinal symbolic dynamics.

Consider a finite, noisy time series

$$\xi_n = f^n(x_0) + w_n$$

$(0 \leq n \leq N-1)$ where w_n is white noise.

Facts.

- Deterministic signals have forbidden patterns (but they are 'destroyed' by the noise)
- Random signals have no forbidden patterns (but finite signals may have missing ordinal patterns)

6. Detection of determinism

Null hypothesis:

$$\boxed{H_0\text{: the } \xi_n \text{ are i.i.d.}}$$

Detection method 1: *Count and shuffle.*

1. Count the number of missing pattern is a sliding window of size L
2. Randomize the sequence
3. Repeat step 1 an compare.

If the counts in steps 1 and 3 are very different, reject H_0.

Null hypothesis:

$$\boxed{H_0\text{: the } \xi_n \text{ are i.i.d.}}$$

Detection method 2: *Chi-square test*.

1. Take a sliding window of size L and compute

$$\chi^2(L) = \sum_{\pi \in \mathcal{S}_L} \frac{(\nu_\pi - K/L!)^2}{K/L!} = \frac{L!}{K} \sum_{\pi \in \mathcal{S}_L:\,\text{visible}} \nu_\pi^2 - K,$$

   where ν_π is the number of windows of type $\pi \in \mathcal{S}_L$.

2. Reject H_0 with confidence level α 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.

Numerical simulation. The *Lorenz map*

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

has an attractor with $D_1 = 2$.

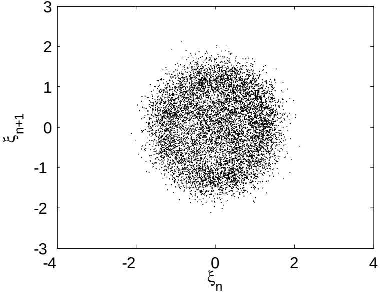


Figure. Return map of the x-component contaminated with Gaussian white noise ($\sigma = 0.25$)

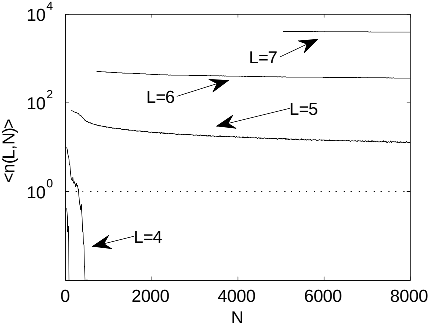


Figure. Average number of missing L-patterns for the x-component of noisy Lorenz time series ξ_1^N ($\sigma = 0.25$).

6. Detection of determinism

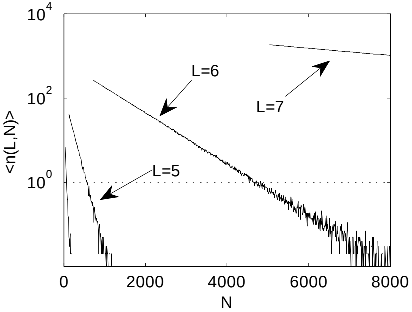


Figure. Average number of missing L-patterns for Gaussian white noise w_1^N ($\sigma = 0.25$).

6. Detection of determinism

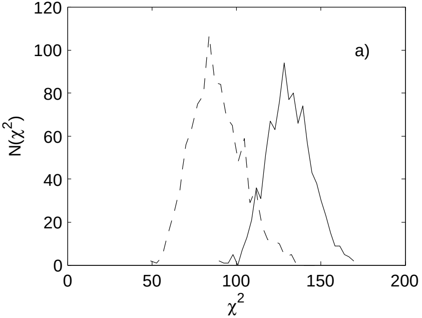


Figure. Distribution of $\chi^2(L = 4)$ for noisy Lorenz time series ζ_1^{1000} with $\sigma = 0.25$ (continuous line) and $\sigma = 0.50$ (dashed line). Rejection threshold: $\chi^2_{23,0.95} = 35.17$.

6. Detection of determinism

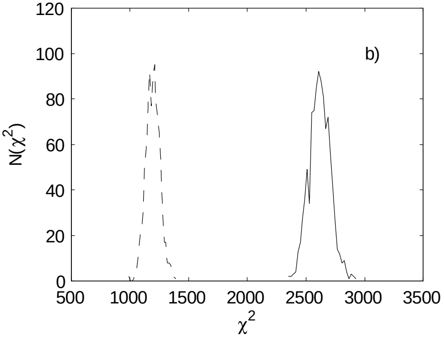


Figure. Distribution of $\chi^2(L = 5)$ for noisy Lorenz time series ζ_1^{8000} with $\sigma = 0.25$ (continuous line) and $\sigma = 0.50$ (dashed line). Rejection threshold: $\chi^2_{119,0.95} = 145.46$.

Comparison with the BDS test of independence.

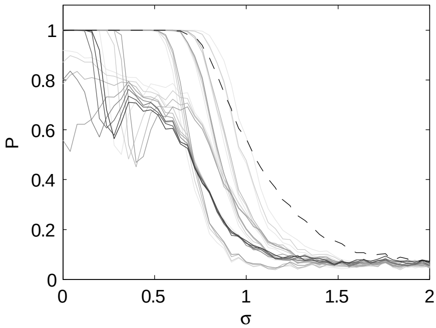


Figure. Rejection probability for a noisy Lorenz time series using the BDS test with different parameters (continuous lines) and forbidden patterns (dashed line).

References

1. J.M. Amigó, *Permutation complexity in dynamical systems*. Springer Verlag, 2010.
2. G.H. Choe, *Computational ergodic theory*, Springer Verlag, 2005.
3. P. Walters, *An introduction to ergodic theory*, Springer Verlag, 2000.