Chapter 2
First Applications

In this chapter we present four applications of permutation entropy and ordinal patterns: entropy estimation, complexity analysis, recovery of parameters from itineraries, and synchronization analysis of time series. The scope is to give the reader a multifaceted picture of ordinal analysis in action. Two more applications (to determinism detection and to space–time chaos) will be discussed at length in Chaps. 9 and 10, respectively.

2.1 Entropy Estimation

Real or numerical time series, say \((x_n)_{n \in \mathbb{N}_0}\) with \(x_n \in \mathbb{R}\), can be produced in principle by discrete-time or continuous-time dynamical systems, which for convenience we think as including also the corresponding stochastic systems. In the continuous-time case, \(x_n\) can be thought as readouts of an analogue signal at discrete times, as it actually happens in practice. Formally, continuous-time dynamical systems are constructed from the solutions of ordinary differential equations and are called flows [98]. When solving differential equations numerically, the time variable is discretized anyway [173].

Permutation entropy made its first appearance in the analysis of univariate time series, i.e., sequences of real numbers—the only ones we will consider in this section. Given a finite time series\(^1\) \(x^{N-1}_0 = x_0, x_1, \ldots, x_{N-1}\), take a sliding window of size \(2 \leq L \ll N\) along the time series (each window comprising a symbol block \(x^{n+L-1}_n = x_n, \ldots, x_{n+L-1}, 0 \leq n \leq N-L\)) and count the number of blocks realizing a particular ordinal pattern \(\pi \in \mathcal{S}_L\). The relative frequency of each \(\pi \in \mathcal{S}_L\) in the sequence \(x^{N-1}_0\) is then

\[
\hat{p}(\pi) = \frac{\left| \{ n : 0 \leq n \leq N - L, x^{n+L-1}_n \text{ is of type } \pi \} \right|}{N - L + 1}.
\]  

\(^1\)For notational simplicity, we assume that one symbol is output per time unit. In this way, a time series can be labeled as a sequence.
This estimator of the probability of \( \pi \) converges with probability 1 to the true value in the limit of infinitely long time series, under the proviso that the underlying stochastic process is stationary or, at least, that the probability for \( x_n < x_{n+k}, \)

\( 1 \leq k \leq L - 1 \), does not depend on \( n \) [28]. Let us mention in passing that the ordinal pattern probability distributions have been calculated for some random processes and pattern lengths, like Gaussian, fractional Brownian, and autoregressive moving-average (ARMA) processes for \( L \leq 4 \) [30, 213]; see also [190].

The permutation entropy per symbol of order \( L \) of \( x_0^{N-1} \) is then defined as

\[
h^*_L(x_0^{N-1}) = -\frac{1}{L} \sum_{\pi \in \mathbb{S}_L} \hat{p}(\pi) \log \hat{p}(\pi). \tag{2.2}
\]

In the case of infinitely long sequences, one defines the permutation entropy of a sequence \( x_0^\infty \) as

\[
h^*(x_0^\infty) = \lim_{L \to \infty} h^*_L(x_0^\infty), \tag{2.3}
\]

provided the limit exists.

The general procedure followed so far is well known to the practitioners of non-linear time analysis: \( L \) is the embedding dimension and the delay time \( T \) is here 1 (since we take consecutive entries). As the window of size \( L \) slides along the time series \( x_0^\infty \), the vectors \( \mathbf{x}_n = x_{n+L-1} \in \mathbb{R}^L \) describe the so-called reconstructed trajectory in the \( L \)-dimensional embedding space [1, 112, 166, 197]. The changes to be done when the sequences

\[
x_n, x_{n+T}, \ldots, x_{n+(L-2)T}, x_{n+(L-1)T},
\]

have a delay time \( T > 1 \), are merely a matter of form but not of concept. Note that for deterministic sequences \( x_n = f^n(x_0), n \geq 0 \), subsequence (2.4) is an orbit segment of \( f^T \).

In general, \( h^*_L \) and \( h^* \) are defined for arbitrary-alphabet sequences whose symbols can be linearly ordered, while Shannon entropy applies to finite-alphabet sequences.\(^2\) In practice all alphabets are finite because of the finite precision of the observation device and/or the finite real number representation of the computers. Such being the case, let \( \mathbf{X} = \{X_n\}_{n \in \mathbb{N}_0} \) be the actual data source of the sequences \( x_0^\infty \), where now \( x_i \) are “discretized” values drawn from a finite alphabet \( \mathbb{S} \), and

\[
h(\mathbf{X}) = -\lim_{n \to \infty} \frac{1}{L} \sum_{x_0, \ldots, x_{L-1} \in \mathbb{S}} p(x_0, \ldots, x_{L-1}) \log p(x_0, \ldots, x_{L-1}),
\]

\(^2\)Real-valued data sources call for the concept of differential entropy [59].
its Shannon entropy. Usually, \( h(X) \) is estimated by means of the so-called plug-in, maximum likelihood, or naive estimator

\[
\hat{h}_L(x_0^{N-1}) = -\frac{1}{L} \sum \hat{p}(a_0 \ldots a_{L-1}) \log \hat{p}(a_0 \ldots a_{L-1}),
\]

where the summation is over all blocks \( a_0^{L-1} = a_0 \ldots a_{L-1} \in S^L \), and

\[
\hat{p}(a_0 \ldots a_{L-1}) = \frac{|\{n: 0 \leq n \leq N - L, x_n^{n+L-1} = a_0^{L-1}\}|}{N - L + 1}
\]

is the relative frequency of \( a_0^{L-1} \) in \( x_0^{N-1} \).

Important for us is that if the process \( X \) is stationary and ergodic, then \( h^*(x_0^{\infty}) = h(X) \) for a “typical” sequence (Chap. 6, Theorem 8). Therefore, in such cases \( h^*_L(x_0^{N-1}) \), with \( L \ll N \), can be used as an estimator of \( h(X) \) instead of (2.5).

The numerical estimation of entropy via ordinal patterns will be discussed with more detail in Sect. 6.4, once the theoretical underpinnings of metrical permutation entropy of maps have been elucidated. At this point it suffices to advance that the computation is fast but the convergence is in general slow.

The slow convergence of \( h^*_L \) to the Shannon entropy seems to require great values of \( L \) for an accurate estimation. On the other hand, the superexponential growth of \( |S_L| = L! \) makes exhaustive sampling computationally unfeasible for, say, \( L \gtrsim 12 \), even if there would be enough data at our disposal. In Chap. 7 we shall learn sampling techniques that work pretty well in these cases. In practice, the estimation of both Shannon entropy and permutation entropy (or, for that matter, of any quantity involving the limit \( L \to \infty \)) suffers from undersampling when \( L \) becomes sufficiently large as compared to the length \( N \) of the sequence. Undersampling means that the observed relative frequencies (of blocks or ordinal patterns) are no longer good estimators of the corresponding probabilities, simply because the samples are too small to be statistically significant. The following first-order correction due to finite sample effects was proposed by Herzel [93]:

\[
\hat{h}_L(x_0^{N-1}) \leftarrow \hat{h}_L(x_0^{N-1}) - \frac{M_1}{2M_2},
\]

where \( M_1 \) is the number of words \( a_0^{L-1} \) with positive probabilities and \( M_2 \) is the number of samples (\( M_2 = N - L + 1 \) when the sequence is sampled by means of overlapping sliding windows, see (2.6)). In principle, the samples should be independent, but as stated in [94], the results are also satisfactory when the words overlap. Other corrections have been discussed by Grassberger [88] (who generalizes (2.7)) and Schmitt et al. [181] (who exploit Shannon–McMillan–Breiman’s theorem of asymptotic equidistribution). Sometimes extrapolation techniques perform fine when undersampling occurs. One of them [195, 6] calls for plotting the partial entropies \( h^*_L \) against \( 1/L \); if the graph exhibits a distinctive linear part (showing...
that $h^*_L/L$ has already converged), then one extrapolates with a straight line this linear part till it intercepts the vertical axis ($1/L \rightarrow 0$), Fig. 2.1. See [127] for other methods to estimate the Shannon entropy and [167] for a review on entropy estimation.

Summing up, permutation entropy ("counting ordinal patterns") provides a conceptually simple and computationally fast method to estimate Shannon entropy. When compared to the usual block-based estimators ("counting blocks"), there is a difference that can be important in applications: the number of ordinal $L$-patterns does not depend on the alphabet. Specifically, the maximal number of length-$L$ blocks (Shannon entropy) and length-$L$ ordinal patterns (permutation entropy) grows with $L$ as

$$|S|^L = e^{L \ln |S|} \quad \text{and} \quad L! \sim e^{L \ln L},$$

respectively, where $S$ is the alphabet. It follows that if $|S|$ is very large, undersampling might set in earlier for block-based estimation than for ordinal pattern-based estimation. This occurs precisely with real-world or computer-generated data. Such an advantage has been reported in the literature, also in the computation of the Rényi entropy

$$h_{R\alpha}(X) = \lim_{L \to \infty} \frac{1}{L} \frac{1}{1 - \alpha} \log \left( \sum_{x_0, \ldots, x_{L-1} \in S} p(x_0, \ldots, x_{L-1})^\alpha \right), \quad (2.8)$$

where $\alpha \geq 0, \alpha \neq 1$ ($\lim_{\alpha \to 1} h_{R\alpha}(X) = h(X)$), and the Tsallis entropy
2.2 Permutation Complexity

\[ h_{T_q}(X) = \lim_{L \to \infty} \frac{1}{L} \frac{1}{q-1} \sum_{x_0, \ldots, x_{L-1} \in S} (p(x_0, \ldots, x_{L-1}) - p(x_0, \ldots, x_{L-1}y^L)), \]  

(2.9)

where \( q \in \mathbb{R}, q \neq 1 \) [213]. When \( p(x_0, \ldots, x_{L-1}) \) is replaced in (2.8) and (2.9) by \( p(\pi), \pi \in S_L \) (or estimated by the relative frequency \( \hat{p}(\pi) \)), one speaks of the Rényi permutation entropy and the Tsallis permutation entropy, respectively. To complete the picture, let us add that the situation reverses when the alphabet comprises few symbols. But in this case, Lempel–Ziv complexity (specifically, LZ-76) can be a better choice than block counting [6]; see [82] for the entropy estimation in binary sequences.

Although less used than the “Shannon permutation entropy” \( h^* \), one can also define the topological permutation entropy or permutation capacity,

\[ h^*_0(x^\infty_0) = -\lim_{L \to \infty} h^*_{0,L}(x^\infty_0), \]  

(2.10)

provided the limit exists, where the rate of finite order is given as

\[ h^*_{0,L}(x^\infty_0) = -\frac{1}{L} \log N(L), \]  

(2.11)

\( N(L) \) being the number of distinct ordinal patterns defined by sliding windows \( x_n^{n+L-1} \) of size \( L \). That is, we just count now how many different \( L \)-patterns are realized, instead of computing the relative frequency of those \( L \)-patterns. It follows that \( h^*_0 \) is an upper bound of \( h^* \). When the sequences \( x^\infty_0 \) are seen as outputs of an information source \( X \), then \( N(L) \) stands for the number of admissible \( L \)-patterns in the messages that \( X \) can emit, and one speaks of the permutation capacity or topological permutation entropy of \( X \) (Chap. 7).

The ordinal pattern-based approach to Shannon entropy can also be extended to the metric and topological entropy of maps; see Chaps. 7 and 8. The situation is specially simple for one-dimensional, piecewise monotone interval maps \( f:I \to I \).

In this case, we only need to numerically estimate the probabilities \( \mu(P_\pi) \) of the admissible \( L \)-patterns (\( P_\pi \neq \emptyset \)), or just the number of distinct admissible patterns, to get an estimate of the metric or topological entropy of \( f \), respectively (see (1.30), (1.31), and (1.32)). Thus, the estimation of \( h^*_\mu(f) \) and \( h^*_{\text{top}}(f) \) boils down again to counting ordinal \( L \)-patterns. The computation of \( h^*_{\text{top}}(f) \) is also simpler than for its standard counterpart. The higher dimensional case will also be considered in Chaps. 6 and 7.

2.2 Permutation Complexity

Although complexity, (pseudo-)randomness, disorder, irregularity, typicality, etc., are terms that have been introduced eventually in different settings to mean more or less the same dynamical behavior, complexity is the preferred one when there is no
measure (or probability) involved. In fact, Bandt and Pompe introduced permutation
entropy in [28] via (2.1), (2.2), and (2.3) as a “natural complexity measure for time
series.” The time series can be the output of a random process or an orbit of a
dynamical system. By analyzing the complexity of a signal (if no other informa-
tion available), we are inquiring into the complexity of the source. An axiomatic
characterization of complexity was proposed in [163].

The measurement of complexity and its eventual time variation is an issue of
utmost important in the analysis of biomedical, economic, physical, and technical
time series. Think of the forecasting of transitions to abnormal health conditions,
financial crashes, severe weather, earthquakes, etc. Over the years, a battery of
methods has been proposed and developed with this purpose or adapted from other
fields like information theory and networks. Let us mention some of these methods
(see also the references therein):

- Cross-correlation sum analysis [111]
- Lempel–Ziv complexity [208, 196, 90, 6, 78]
- Mutual information [90]
- Nonlinear cross-prediction analysis [183]
- Recurrence plots [73, 144, 200] and recurrence quantification analysis [81]
- Relative entropy [180]
- Statistical complexity [56, 143] (statistical complexity was introduced by
  Crutchfield and collaborators within a theory called computational mechanics
  [60, 185, 24])
- Statistical tests in the reconstructed phase space [120]
- Topological methods [209]

Permutation entropy and other related quantities are specially well suited to mea-
sure the complexity of random and deterministic dynamical systems for several
reasons.

First of all, permutation entropy in its different variants involves counting ordinal
patterns. With the exception of a few cases, the number of ordinal $L$-patterns realized
by a map $f$ increases with $L$. Therefore, the (logarithm of the) rate of this increasing
is a natural measure (as stated by Bandt and Pompe) for quantifying the complexity
of a deterministic time series or, more generally, of a dynamical system. In the met-
ric variant, each admissible $L$-pattern contributes to the entropy a term containing
its relative frequency or probability, respectively. In the topological variant, all such
patterns make the same contribution to the entropy; formally, they are assigned the
same probability. Since random, unconstrained processes have no forbidden patterns
with probability 1 (hence, they have a superexponential growth of admissible ordinal
patterns with length), their complexity, as measured by the permutation entropy, is
infinite. At the other end, a periodic or quasiperiodic dynamic has vanishing or neg-
ligible permutation entropy. Complex systems lie between order and randomness.
From a practical point of view, we can characterize them as having a positive, finite
permutation entropy. Both metric and topological permutation entropies increase as
the sequence “looks” more random.
Second, unlike other proposals for complexity measures, permutation entropy applies in principle both to finite-alphabet and arbitrary-alphabet sequences, albeit it is more interesting in the second case.

Technically we are assuming that the limits involved in the corresponding definitions (like (2.3) and (2.10)) converge. In practice, limits have to be estimated using a finite number of terms—real sequences are finite anyway. What we mean is that the actual tools of permutation complexity are going to be the permutation entropy rates of finite order, like \( h^*_{L}(\alpha_0^{N-1}) \) and \( h^*_{0,L}(\alpha_0^{N-1}) \), and other related quantities based on finite-length ordinal patterns, like probability distributions, information-theoretical tools (relative entropy, mutual information, etc.), complexity functionals. Moreover, since the maximal value of \( h^*_{L}(\alpha_0^{N-1}) \) and \( h^*_{0,L}(\alpha_0^{N-1}) \) is \( \log L! \), we can eventually divide both entropy rates by \( \log L! \) to obtain dimensionless quantities ranging between the two non-complex extremes: 0 (order) and 1 (randomness).

Finally, permutation entropy rates of finite order are computationally fast for the pattern lengths used in practice (3 \( \leq L \leq 7 \))—also for the Rényi (2.8) and Tsallis (2.9) permutation entropies. This allows calculation in real time, which is a significant advantage in applications. We come back to this point in the next chapters.

Application of ordinal patterns and permutation entropy to complexity analysis of data has been reported in different fields. For instance

- biomedical series [116, 45, 118]
- financial series [146, 147]
- physical series [28]
- statistical series [30, 146, 147, 212]

Let us underline at this point that the application by Keller [116] of ordinal patterns to electroencephalogram (EEG) data from children with epileptic disorders dates from about the same time as permutation entropy was formulated [28].

Similarly, one of the first applications of permutation entropy was the detection of dynamical changes in time series and, in particular, epileptic seizure detection from EEGs by Cao et al. [45]. Regarding the second application, the authors analyzed continuous EEG measurements recorded intracranially (also called depth EEG) with typically 28 electrodes. Figure 2.2 shows the normalized permutation entropy rate of order \( L = 5 \) for three different patients. Each signal is more than 5 h long, with a sample frequency of 200 Hz and time delay 3 (i.e., only every third entry in the EEG signal is taken into account, what amounts to sampling the signal with frequency \( 200/3 \) Hz). According to [45], the change of permutation complexity in all these cases indicates that the dynamics of the brain first becomes more regular right after the seizure, then its irregularity increases as it approaches the normal state.

Since these and other pioneering works, ordinal analysis of time series has remain a popular technique. In some cases, ordinal analysis has been incorporated into more general schemes, such as the method of recurrence plots, introduced by Eckmann et al. [73] to visualize the recurrences of dynamical systems. This method, which is being used to analyze virtually any natural data [144], is based on the recurrence matrix of a scalar or vectorial trajectory \( (x_i)_{i=0}^{N-1} \) of a system in its state space \( S \), defined as
Fig. 2.2 [Reproduced with permission from [45].] Variation of the normalized $h^*_t$ with time for EEG signals of (a) patient 1, channel 1, (b) patient 2, channel 1, and (c)–(e) patient 3, channels 1–3

$$R_{i,j}(\varepsilon) = H(\varepsilon - \|x_i - x_j\|), \quad i,j = 0, \ldots, N - 1,$$

(2.12)

where $\varepsilon$ is a threshold distance, $H(\cdot)$ is the Heaviside function ($H(x) = 0$ if $x < 0$ and $H(x) = 1$ otherwise), and $\|\cdot\|$ is a norm in $S$. Instead of using spatial closeness
as in (2.12), *ordinal patterns recurrence plots* are based on the ordinal patterns \( \pi(i) \) realized by the sequences \( x_i^{i+L-1}, 0 \leq i \leq N - L \). If \( \delta(\pi, \pi') = 1 \) for \( \pi = \pi' \in S_L \), and \( \delta(\pi, \pi') = 0 \) otherwise, set

\[
R_{i,j}(L) = \delta(\pi(i), \pi(j)),
\]

(2.13)

\( \pi(i), \pi(j) \in S_L, 0 \leq i,j \leq N - L \). According to [144], the main advantage of (2.13) is its robustness against non-stationary data.

To distinguish the kind of complexity captured by the tools of ordinal analysis—ordinal patterns, permutation entropy, permutation entropy rates of finite order, and other quantities based on order relations—we propose to call it *permutation complexity*. Therefore, permutation complexity has to do with the ordinal structure of data obtained from deterministic or random dynamical systems. These also include spatially extended systems, like the ones we shall consider in Chap. 10.

### 2.3 Estimation of Control Parameters from Symbolic Sequences

The basis of permutation complexity is the relation between order and dynamics. This relation is specially strong on one-dimensional intervals, where order and metric are intertwined, leading to such interesting results as Sarkovskii’s theorem [179, 150]. It is therefore not surprising that the study of the ordinal structure of time series provides valuable information on the underlying dynamical system. In this section we learn how to recover the “control” parameter of a unimodal map from itineraries. The relationship between the itineraries of parametric unimodal maps and the value of the parameter that controls a particular dynamics was shown in [153, 203, 5].

Let \( \mathcal{U} \) be the class of unimodal maps on an interval \( I = [a, b] \subset \mathbb{R} \). A map \( f: I \rightarrow I \) is unimodal if it is continuous, has a single turning point (called hereafter the critical point) \( x_c \) in \( I \), and is monotone increasing on the left of \( x_c \) and decreasing on the right. The class \( \mathcal{U} \) includes maps defined in a parametric way, say, \( f_v(x) = \varphi(v, x) \), where \( x \in I, v \in J \subset \mathbb{R} \) will be called the control parameter, and \( \varphi \) is a map on \( I \times J \).

The class \( \mathcal{U} \) includes the *logistic family* \( g_v: [0, 1] \rightarrow [0, 1] \),

\[
g_v(x) = vx(1 - x),
\]

(2.14)

where \( 0 \leq v \leq 4 \), and the *tent family* \( \Lambda_v: [0, 1] \rightarrow [0, 1] \),

\[
\Lambda_v(x) = \begin{cases} 
  x/v & \text{if } 0 \leq x \leq v, \\
  (1 - x)/(1 - v) & \text{if } v \leq x \leq 1,
\end{cases}
\]

(2.15)

where \( 0 < v < 1 \); see Fig. 2.3. In particular, \( g_4 \) is the logistic map (1.19) and \( \Lambda_{1/2} \) the symmetric tent map (1.17). The critical point of \( g_v \) does not depend on \( v \): \( x_c = \frac{1}{2} \)
for all \( v \). On the opposite side, the critical point of \( \Lambda_v \) coincides with the parameter value: \( x_c = v \). As usual in the literature, we will also refer to \( g_v \) and \( \Lambda_v \) just as the logistic and tent maps, respectively, when the parameter \( v \) is thought to be fixed.

Note that \( \Lambda_v \) preserves the Lebesgue measure for all \( v \in (0, 1) \).

For \( f \in \mathcal{U} \), let \( \Phi_1(x) \) be the itinerary of \( x \in [a, b] \) with respect to the partition \( \{A_0, A_1\} \), with \( A_0 = [a, x_c) \) and \( A_1 = [x_c, b] \). Specifically,

\[
\Phi_1(x) = \Phi_1^0(x), \Phi_1^1(x), \ldots, \Phi_1^n(x), \ldots = (\Phi_1^i(x))_{i=0}^\infty,
\]

where

\[
\Phi_1^n(x) = \begin{cases} 
0 & \text{if } f^n(x) < x_c, \\
1 & \text{if } f^n(x) \geq x_c.
\end{cases}
\]

As a result, any orbit \( O_f(x) \) can be encoded into a binary sequence. Whenever convenient, we will write \( \Phi_1(f, x) \) instead of \( \Phi_1(x) \) to make clear which unimodal map is generating the itinerary of \( x \).

An interesting aspect of the binary sequences \( \Phi(x) \) is that they can be endowed with a signed lexicographical order (sometimes called Gray ordering) \( \leq \) that is equivalent to the order in \( [a, b] \) in the following weakened sense:

(E1) If \( x < y \), then \( \Phi(x) \leq \Phi(y) \).

(E2) If \( \Phi(x) < \Phi(y) \), then \( x < y \).

A sufficient condition for \( x < y \) if and only if \( \Phi(x) \leq \Phi(y) \) is given in [57, Theorem II.5.4]. The order between binary sequences is defined as follows. Given \( \Phi(x) \neq \Phi(y) \), let \( i_{\text{min}} \) be the first index such that \( \Phi_i(x) \neq \Phi_i(y), i \geq 0 \). Depending on \( i_{\text{min}} \), three cases can occur:

(O1) If \( i_{\text{min}} = 0 \), then \( \Phi(x) < \Phi(y) \) iff \( \Phi_0(x) < \Phi_0(y) \).
\(O2\) If \(i_{\text{min}} > 0\) and \(\{\Phi_i(x): 0 \leq i < i_{\text{min}}\}\) contains an even number of 1’s, then \(\Phi(x) < \Phi(y)\) iff \(\Phi_{i_{\text{min}}}(x) < \Phi_{i_{\text{min}}}(y)\).

\(O3\) If \(i_{\text{min}} > 0\) and \(\{\Phi_i(x): 0 \leq i < i_{\text{min}}\}\) contains an odd number of 1’s, then \(\Phi(x) < \Phi(y)\) iff \(\Phi_{i_{\text{min}}}(x) > \Phi_{i_{\text{min}}}(y)\).

Given \(x, f_v(x), \ldots, f_v^{L-1}(x)\), suppose that their corresponding itineraries, namely,

\[
(\Phi_i(x))_{i=0}^\infty, (\Phi_i(x))_{i=1}^\infty, \ldots, (\Phi_i(x))_{i=L-1}^\infty,
\]

are all different. Then, according to (E1)–(E2),

\[
f^{\pi_0}(x) < \ldots < f^{\pi_{L-1}}(x) \iff (\Phi_i(x))_{i=\pi_0}^\infty < \ldots < (\Phi_i(x))_{i=\pi_{L-1}}^\infty. \tag{2.17}
\]

Before proceeding further, let us point out that this setting can be extended to l-modal maps, i.e., continuous and piecewise strictly monotone self-maps of compact intervals with \(l\) local maxima, which map endpoints to endpoints. For the applications we will discuss, it is sufficient to consider only unimodal maps \((l = 1)\).

In some applications, one is confronted with the following task: given the “sharp” orbit \(O_{f_v}(x_0)\) of \(x_0 \in [a, b]\) under \(f_v \in \mathcal{U}\), find the value of the parameter \(v\). In practice, the exact values of \(O_{f_v}(x_0)\) are seldom known because of the finite precision of real number computation, so one has only access to a (finite segment of a) “coarse-grained” orbit \((\hat{x}_i)_{i=0}^\infty\), where \(\hat{x}_i\) is an approximation to \(x_i = f_v(x_0)\). In some chaos-based cryptosystems, the situation is even worse: the plaintext (i.e., the message to be encrypted prior to its transmission or storage) is encoded via the symbolic sequences (2.16) of a chaotic map \(f_v \in \mathcal{U}\), the value of \(v\) being part of the secret key of the cipher (see, e.g., [131]). Therefore, the cryptanalyst has eventually only access to the binary code \(\Phi(f_v, x)\) (via a so-called chosen-text attack) to recover the control parameter \(v\). D. Arroyo has shown how to recover \(v\) with the aid of the ordinal patterns of \(f_v\) and their itineraries \(\Phi(f_v, x)\), if \(f_v\) is ergodic with respect to its natural measure \(\mu_v\) for all values of \(v\) [21].

For simplicity, the estimation of \(v\) from the symbolic sequences \(\Phi(f_v, x)\) will be illustrated using the tent map \(\Lambda_v\), which is chaotic for all \(v \in (0, 1)\). Since the natural invariant measure of \(\Lambda_v\) is the Lebesgue measure, the probability that \(x\) is of type \(\pi \in S_L\) when drawn uniformly from \([0, 1]\) equals the length of \(P_\pi = \{x \in [0, 1]: x \text{ defines } \pi\}\) (as in (1.27)). By ergodicity, the relative frequency of \(\pi\) in an orbit of \(\Lambda_v\) coincides with the length of \(P_\pi\), except possibly for a set of initial conditions with length zero. For the tent map, the length of the sets \(P_\pi\) can be determined analytically. The simplest case corresponds to the \(L\)-pattern \(\langle 0, 1, \ldots, L-1 \rangle\):

\[
P_{\langle 0,1,\ldots,L-1 \rangle} = (0, \phi_L(v)),
\]

where \(\phi_L(v)\) is the leftmost intersection of \(\Lambda_v^{L-1}\) and \(\Lambda_v^{L-2}\). Therefore, the length of \(P_{\langle 0,1,\ldots,L-1 \rangle}\), hence the probability that \(x\) is of type \(\langle 0, 1, \ldots, L-1 \rangle\) when drawn uniformly from the interval \([0, 1]\) is \(\phi_L(v)\).
In order to calculate $\phi_L(v)$, use

$$\Lambda_n^L(x) = \begin{cases} \frac{x}{v^n} & \text{if } 0 \leq x < v^n, \\ \frac{(v^n - x)}{(v^n - 1)}(1 - v) & \text{if } v^n \leq x \leq v^{n-1}. \end{cases}$$

Equating $\Lambda_{v-1}^L$ and $\Lambda_{v-2}^L$, it follows

$$\phi_L(v) = \frac{v^{L-2}}{2 - v}.$$  \hfill (2.18)

Note that this function is 1-to-1 in the interval $0 \leq v \leq 1$ for $L \geq 2$, with $\phi_L(0) = \frac{1}{2}$, $\phi_L(0) = 0$ for $L \geq 3$, and $\phi_L(1) = 1$ for $L \geq 2$. This fact allows to determine $v$ from $\phi_L(v)$. Furthermore, from the equation

$$\frac{d}{dv} \phi_L(v) = \frac{v^{L-3}}{(2-v)^2} \left[ 2(L-2) - (L-3)v \right] = \begin{cases} 0 & \text{if } v = 0, \\ L-1 & \text{if } v = 1, \end{cases}$$  \hfill (2.19)

it follows that $\phi_L(v)$ is a $\cup$-convex map on $0 \leq v \leq 1$ for $L \geq 2$ that converges to 0 on $0 \leq v < 1$ (i.e., it “flattens”) as $L \to \infty$. As a result, the higher the $L$ the lower the precision with which $v$ can be numerically read off from $\phi_L(v)$. Consequently, $L = 3, 4$ are the best choices for a quality estimation of $v$.

In more general terms, suppose that each $f_v \in \mathcal{U}$ is ergodic for $v \in J$ with the same invariant measure $\mu$. Furthermore assume for the time being that $f_v(a) = a$ and $f_v(x) > x$ on a non-empty vicinity of $a$. Let $(a, c)$ be the maximal interval in $(a, \infty)$ such that $f_v(x) > x$. We claim that the interval

$$I_v^L = (a, c) \cap f_v^{-1}(a, c) \cap \cdots \cap f_v^{-(L-1)}(a, c)$$

coincides with $P_{(0,1,\ldots,L-1)}$. Indeed, if $x \in I_v^L$, then $f_v^i(x) \in (a, c)$ for $0 \leq i \leq L-1$, and

$$x < f_v(x) \Rightarrow f_v(x) < f_v^2(x) \Rightarrow \cdots \Rightarrow f_v^{L-2}(x) < f_v^{L-1}(x).$$

Hence, $I_v^L \subset P_{(0,1,\ldots,L-1)}$. Conversely, if $x \in P_{(0,1,\ldots,L-1)}$, i.e.,

$$x < f_v(x) < f_v^2(x) < \cdots < f_v^{L-1}(x),$$

then $f_v^i(x) \in (a, c)$ for $0 \leq i \leq L-1$. Thus, $P_{(0,1,\ldots,L-1)} \subset I_v^L$. This proves

$$I_v^L = P_{(0,1,\ldots,L-1)},$$  \hfill (2.20)

If otherwise $f_v(a) = a$ but $f_v(x) < x$ on a non-empty vicinity of $a$, then let $(a, c)$ be the maximal interval in $(a, \infty)$ such that $f_v(x) < x$. In this case, a similar reasoning (reversing the inequalities) shows that

$$I_v^L = P_{(L-1,L-2,\ldots,1,0)}.$$  \hfill (2.21)
Since the tent map, our workhorse in this section, complies with (2.20), we restrict attention to this case (similar arguments apply mutatis mutandis to case (2.21)). Because of ergodicity, the relative frequency at which a typical trajectory visits $I_L^{i,v}$ is $\mu(I_L^{i,v})$. If $\mu(I_L^{i,v})$ happens to be different for each $v$, then $\mu(I_L^{i,v})$ can be used to determine or estimate the control parameter $v$. In this case, the relative frequency of the ordinal pattern $\{0, 1, \ldots, L - 1\}$ in an orbit $O_{f_0}(x)$ is just the number of times that $f_i^{j+1}(x) \in (a, c)$ for $i \in \mathbb{N}_0$ and $j = 0, 1, \ldots, L - 1$.

Figure 2.4 shows the relative frequencies of the ordinal patterns (a) $\langle 0, 1, 2, 3 \rangle$, (b) $\langle 0, 1, 3, 2 \rangle$, (c) $\langle 0, 3, 1, 2 \rangle$, and (d) $\langle 3, 0, 1, 2 \rangle$ found in a numerical simulation with the tent map. As expected, curve (a) approximates the function

$$\phi_4(v) = \frac{v^2}{2 - v}$$

with great precision. Observe that a 1-to-2 functional relation between frequency and $v$, as it occurs in Fig. 2.4 (b)–(d), can also be acceptable, e.g., for cryptographic applications since it implies a reduction of the secret key space.

So far we have shown the possibility of recovering the control parameter $v$ from the relative frequency of the pattern $\pi = \langle 0, 1, \ldots, L - 1 \rangle$ (most conveniently for $L = 3, 4$), in a statistically significant sample of orbits of $\Lambda_v$. The ergodicity of $\Lambda_v$ with respect to the Lebesgue measure on $[0, 1]$ and the 1-to-1 relation between $v$ and

![Fig. 2.4 Ordinal pattern frequencies for the tent map family. Here $L = 4$, and (a) $\pi = \langle 0, 1, 2, 3 \rangle$, (b) $\pi = \langle 0, 1, 3, 2 \rangle$, (c) $\pi = \langle 0, 3, 1, 2 \rangle$, and (d) $\pi = \langle 3, 0, 1, 2 \rangle$](image_url)
the probability $\phi_L(v)$ of observing $\pi$ were instrumental to achieve that goal. What about if we have access only to coarse-grained orbits $\Phi(L,v,x)$?

Let $b_0^{M-1} = b_0 b_1 \ldots b_{M-1}$, $b_i \in \{0, 1\}$, be the initial segment of length $M$ of the symbolic sequence $\Phi(f_v,x)$. Take a sliding window of size $W < M$ along $b_0^{M-1}$. The result is $M-W+1$ consecutive blocks of length $W$:

$$b_0^{W-1} = b_0 \ldots b_{W-1}, \ldots, b_i^{i+W-1} = b_i \ldots b_{i+W-1}, \ldots, b_M^{M-W} = b_{M-W} \ldots b_{M-1}.$$ 

The blocks $b_i^{i+W-1}$, $i = 0, 1, \ldots, M - W - L + 1$, define $M - W - L + 2$ ordinal patterns of length $L$. That is, if

$$b_i^{i+\pi_0+W-1} < b_i^{i+\pi_1+W-1} < \ldots < b_i^{i+\pi_{L-1}+W-1},$$

then $b_i^{i+W-1}$ is of type $\pi = (\pi_0, \pi_1, \ldots, \pi_{L-1})$. The order for finite sequences in (2.22) is defined the same way as for infinite sequences in (O1)-(O3).

Each block $b_i^{i+W-1} = b_i \ldots b_{i+W-1}$ locates $f_v^i(x)$ up to an uncertainty interval whose length goes to zero when $W,M \to \infty$:

$$f_v^i(x) \in A_{b_i} \cap f_v^{-1} A_{b_{i+1}} \cap \ldots \cap f_v^{-(W-1)} A_{b_{i+W-1}}.$$ 

This being the case, the ordinal patterns defined by, say, $x, f_v(x), f_v^2(x)$, and $b_0^{W-1}, b_1^W, b_2^{W+1}$ may be different as soon as two of the latter blocks overlap. Otherwise, those ordinal patterns will be the same because of (2.17). In sum, the relative frequencies of an ordinal $L$-pattern in the finite orbits $(f_v^i(x))_{i=0}^M$ and $(b_i^{i+W-1})_{i=0}^{M-W-L+1}$ will converge to each other in the limit $M \to \infty$, $W \to \infty$ ($W < M$). In practice, we expect the latter to be a good approximation of the former, at least for $L = 3, 4$, and $W$ large enough, so that a good estimation of the control parameter is feasible.

Figure 2.5 shows the relative frequencies of the same 4-patterns as in Fig. 2.4 for the itineraries of the tent map family. Here $M = 10, 104$ and $W = 100$. Except for $v \simeq 0$ (an uninteresting region for cryptographic applications), the approximation is excellent. Some caveats related to the finite precision of the numerical simulations are discussed in [21]. In practical cases, the error in the estimation of the control parameter ranges between $10^{-3}$ and $10^{-4}$. From the viewpoint of cryptographic applications, this amounts to a strong reduction of the key space, which compromises the security of the cipher.

The tent map family is a specimen of a more general family: unimodal, piecewise linear expanding Markov transformations (Annex A, Definition 9). Each topologically transitive transformation in this family (i.e., some power of its transition matrix is strictly positive) has a unique ergodic invariant measure, which furthermore is absolutely continuous with respect to the Lebesgue measure [134]. This measure can be calculated or numerically estimated by a variety of methods (Perron–Frobenius operator, Ulam’s method, or just computation of long time averages) [105]. For the purpose envisaged in this chapter, an exact knowledge of the invariant measures is
2.4 Characterizing Synchronization

As a last application, we are going to summarize the work of R. Monetti et al. [159] on characterizing synchronization in time series using ordinal patterns (therein called “symbols”) and some related probability distributions.

Remember that $S_L$ is a group with respect to the product of ordinal patterns (1.25). This being the case, given $\pi, \pi' \in S_L$ there always exists a unique $\tau = \tau(\pi, \pi') \in S_L$, called transcription from the source pattern $\pi$ to the target pattern $\pi'$, such that

$$\tau \circ \pi = \pi', \quad (2.23)$$

where (see (1.25))
\( \tau \circ \pi = \{ \pi_{t_0}, \pi_{t_1}, \ldots, \pi_{t_{L-1}} \} \).

It follows that \( \tau \) is a transcription from \( \pi \) to \( \pi' \) if and only if \( \tau^{-1} \) is a transcription from \( \pi' \) to \( \pi \).

As the source pattern \( \pi \) and the target pattern \( \pi' \) vary over \( S_L \), their transcription varies according to \( \tau(\pi, \pi') = \pi' \circ \pi^{-1} \). Note that different pairs \((\pi, \pi')\) can share the same transcription. As an example in \( S_3 \), \( \tau(\pi, \pi') = (0, 2, 1) \) for

\[
(\pi, \pi') = (012, 021), (021, 012), (120, 102), (102, 120), (201, 210), (210, 210)
\]

(angular parentheses omitted for brevity). More generally, given \( \tau \in S_L \), there exist \( L! \) pairs \((\pi, \pi') \in S_L \times S_L \) such that \( \tau \) is the transcript from \( \pi \) to \( \pi' \).

Another concept we need is that of order of an element. We say that the order of \( \pi \in S_L \) is \( \text{ord}(\pi) \in \mathbb{N} \) if \( \text{ord}(\pi) \) is the minimal number of times we have to multiply \( \pi \) by itself to obtain the identity permutation \( (0, 1, \ldots, L - 1) \) (this is the only permutation whose order is 1).

The group \( S_L \) can be partitioned into non-overlapping sets of transcriptions according to their order. In mathematical notation, \( S_L = \bigcup_{1 \leq i \leq L!} C_i \), where

\[
C_i = C_i(L) = \{ \tau \in S_L : \text{ord}(\tau) = i \}.
\]

For obvious reasons, the sets \( C_i \) are called order classes. From \( \text{ord}(\tau^{-1}) = \text{ord}(\tau) \), it follows that \( \tau \in C_i \) if and only if \( \tau^{-1} \in C_i \). Note that \( C_1(1) = \{(0, 1, \ldots, L - 1)\} \).

The authors of [159] propose to measure the complexity of a transcription between a source and a target pattern by its order.

A permutation of the form \( i_1 \mapsto i_2 \mapsto \cdots \mapsto i_n \mapsto i_1 \) is called a cycle (or cyclic permutation) of length \( n \) and denoted by \((i_1, i_2, \ldots, i_n)\). The order of a cycle of length \( n \) is trivially \( n \). It is also trivial that any permutation of \( (0, 1, \ldots, L - 1) \) can be written as the product of disjoint cyclic permutations. It follows that the order of any transcription (or any permutation for that matter) is the least common multiple (lcm) of the lengths of its decomposition into cycles. In particular, given \( L \) there are ordinal patterns \( \tau \in C_i(1) \) of orders \( 1 \leq i \leq L \) (just take \( \tau = (0, \ldots, i - 1)(i)(i + 1) \cdots (L - 1) \)). For \( L + 1 \leq i \leq L! \), a hypothetical decomposition \( \tau = (i_1, \ldots, i_{n_1})(j_1, \ldots, j_{n_2}) \cdots (k_1, \ldots, k_{n_p}), \tau \in C_i(1) \), has to fulfill the constraints

(i) \( n_1 + n_2 + \cdots + n_p = L \) and (ii) \( \text{lcm} \{n_1, n_2, \ldots, n_p\} = i \), which will not be the case in general. For example, for \( L = 7 \) and \( i = 10 \) or 12, we can choose \( n_1 = 2 \) and \( n_2 = 5 \), or \( n_1 = 3 \) and \( n_2 = 4 \), respectively. But for \( L = 7 \) and \( i = 8, 9 \), or 11, conditions (i) and (ii) cannot be simultaneously satisfied.

Let us next turn attention to the probability density of transcriptions. Consider source and target ordinal patterns generated by the time series of a coupled dynamics. Due to the symmetry property between source and target patterns pointed out above, it is irrelevant which one refers to which subsystem, any of the two possible assignments being fine. Let \( S^s_L \) and \( S^t_L \) be the state spaces comprising the corresponding admissible source and target patterns of length \( L \), respectively, and let \( \Omega_L(\tau) \) be the set of all pairs \((\pi_s, \pi_t) \in S^s_L \times S^t_L \) such that \( \tau \in S_L \) is a transcription from \( \pi_s \) to \( \pi_t \), i.e.,
\[ \Omega_L(\tau) = \{ (\pi_s, \pi_t) \in S^i_L \times S^j_L : \tau \circ \pi_s = \pi_t \}. \]

The probability density of transcriptions \( P_L(\tau), \tau \in S_L \), can be written as

\[ P_L(\tau) = \sum_{(\pi_s, \pi_t) \in \Omega_L(\tau)} P^J(\pi_s, \pi_t), \]

where \( P^J(\pi_s, \pi_t) \) is the joint probability density. Furthermore, let \( P^s(\pi_s) \) and \( P^t(\pi_t) \) be the marginal probability densities of the patterns \( \pi_s \in S_s \) and \( \pi_t \in S_t \), respectively. The matrix \( M(\pi_s, \pi_t) = P^s(\pi_s)P^t(\pi_t) \) is the probability density matrix of transcriptions for two independent sequences of lengths \( L \). In this case, the corresponding probability density of transcriptions \( P^\text{ind}_L(\tau) \) can be evaluated as follows:

\[ P^\text{ind}_L(\tau) = \sum_{(\pi_s, \pi_t) \in \Omega_L(\tau)} M(\pi_s, \pi_t). \]

A natural measure to assess how much \( P_L(\tau) \) deviates from \( P^\text{ind}_L(\tau) \) is provided by the relative entropy or Kullback–Leibler distance (see Annex B, (B.3))

\[ D(P_L \parallel P^\text{ind}_L) = \sum_{\tau \in S_L} P_L(\tau) \log \frac{P_L(\tau)}{P^\text{ind}_L(\tau)}. \]

To circumvent the asymmetry of the relative entropy with respect to its arguments, one can take the harmonic mean of \( D(P_L \parallel P^\text{ind}_L) \) and \( D(P^\text{ind}_L \parallel P_L) \),

\[ S_{KL}(L) = \frac{D(P_L \parallel P^\text{ind}_L)D(P^\text{ind}_L \parallel P_L)}{D(P_L \parallel P^\text{ind}_L) + D(P^\text{ind}_L \parallel P_L)}. \]

In contrast to the symmetrization via the arithmetic mean, the bound

\[ S_{KL}(L) \leq \min\{D(P_L \parallel P^\text{ind}_L), D(P^\text{ind}_L \parallel P_L)\} \]

furnishes more general conditions for the symmetrized Kullback–Leibler distance to be finite. Moreover we shall write \( S_{KL}^C(L) \) when the Kullback–Leibler distance is calculated using the probability densities \( P_C \) of the order classes (see Fig. 2.7). Finally, if \( P_L(\tau) \) and \( P^\text{ind}_L(\tau) \) are obtained using only transcriptions from an order class \( C_i(L) \), then the notation will be \( S_{KL}^i(L) \). The point in doing so is that the dynamics of coupled systems may lead to the extinction of order classes, a feature referred to as saturation in [159].

Let us apply the method to a bidirectionally coupled Rössler–Rössler system [175] defined by the following set of equations:
First Applications

\[
\begin{align*}
\dot{x}_{1,2} &= -w_1 y_{1,2} - z_{1,2} + k(x_{2,1} - x_{1,2}), \\
\dot{y}_{1,2} &= w_1 x_{1,2} + 0.165y_{1,2}, \\
\dot{z}_{1,2} &= 0.2 + z_{1,2}(x_{1,2} - 10).
\end{align*}
\]

Here \(w_1 = 0.99\) and \(w_2 = 0.95\) are the mismatch parameters and \(k\) is the coupling constant. All the time series were generated using a fourth-order Runge–Kutta method with time step \(\Delta t = 10^{-3}\) and initial conditions: \(x_1(0) = -0.4, y_1(0) = 0.6, z_1(0) = 5.8, x_2(0) = 0.8, y_2(0) = -2,\) and \(z_2(0) = -4.\) This chaotic system exhibits a rich synchronization behavior that ranges from phase (\(k \approx 0.036\)) to lag (\(k \approx 0.14\)) and finally to complete synchronization as \(k\) increases [175]. In [159] the authors only study the \(x\)-components of the Rössler subsystems. Specifically, time series of length \(2^{19}\) (about 775 orbits) were sampled with delay \(T = 150\) and dimension \(L\), to obtain delay vectors

\[(x(n\Delta t), x((n+T)\Delta t), \ldots, x((n+(L-1)T)\Delta t))\]

from either subsystem. Following [80] the delay was chosen so as to minimize the mutual information (Annex B, (B.6)) of the coordinates \(x_1(t)\) and \(x_1(t + T\Delta t)\) for the uncoupled system (\(k = 0\)).

Figure 2.6(a) shows the symmetrized Kullback–Leibler distance \(S^C_{KL}(L)\) obtained using the probability density \(P_C\) of order classes for \(L = 6\) and \(L = 7\). Figure 2.6(b)–(d) shows \(S_{KL}(L)\) obtained with the probability density of transcriptions in all non-empty order classes for \(L = 6\) (\(C_2-C_6\) in subfigure (b)) and \(L = 7\) (\(C_7, C_{10},\) and \(C_{12}\) in subfigure (c) and \(C_2-C_6\) in subfigure (d)). We comment first the salient features of \(S^C_{KL}(6)\) and \(S^C_{KL}(7)\).

The increase of \(S^C_{KL}\) at \(k \approx 0.036\) is due to the transition from (almost) uncoupled dynamics to phase synchronization. For stronger coupling \(k\), \(S^C_{KL}\) increases rather monotonically until \(k \approx 0.11\). For \(k \in [0.11, 0.145]\), \(S^C_{KL}\) displays strong fluctuations revealing the presence of “intermittent-lag synchronization.” This particular synchronization regime is characterized by synchronization periods interrupted by bursts of non-synchronized activity [175, 34]. The strong fluctuations sharply vanish at the onset of lag synchronization (\(k \approx 0.145\)). Lag synchronization is defined by the condition \(x_1(t + \delta t) = x_2(t)\), i.e., the coincidence of the time series when shifted in time by a constant time lag \(\delta t\). Both curves, \(S^C_{KL}(6)\) and \(S^C_{KL}(7)\), increase monotonically in the interval \(k \in [0.145, 0.30]\) reflecting stronger synchronization. This trend is only interrupted within the range \(k \in [0.232, 0.256]\), where a period-5 window occurs.

The periodic windows are better observed in Fig. 2.6(b)–(d). In fact, all curves exhibit a peak at \(k \approx 0.061\) that corresponds to a period-3 window [175]. \(S^C_{KL}(6)\) and \(S^C_{KL}(7)\) indicate a period-6 window at \(k \approx 0.11\). It seems that this window was not reported before [159], probably due to its extremely small size (\(k \in [0.1094, 0.1096]\)). All curves show clear signatures of periodic behavior in the range \(k \in [0.232, 0.256]\). Intermittent-lag synchronization is particularly reflected by the strong fluctuations observed in Fig. 2.6(b) and (c) for \(S^C_{KL}(6)\) and \(S^C_{KL}(7)\), which
Fig. 2.6 [Reproduced with permission from [159].] (a) $S_{KL}^C$ obtained using the probability density of the order classes for $L = 6$ (lower curve) and $L = 7$ (upper curve). (b)–(d) $S_{KL}$ calculated with the probability density of transcriptions and sequence lengths shown in the plots. Vertical full lines from left to right locate the transitions to phase synchronization ($k \approx 0.036$), intermittent-lag synchronization ($k \approx 0.11$), and lag synchronization, respectively. The vertical dashed lines at $k \approx 0.061$ and $k \approx 0.11$ as well as the hatched area ($k \in [0.232, 0.256]$) indicate periodic windows abruptly disappear at $k = 0.145$. Observe that different order classes provide complementary information of the coupled system. For instance, $S_{KL}^C(6)$ characterizes the intermittent-lag synchronization and the onset of lag synchronization better than $S_{KL}^C(7)$. In any case, these partial pieces of information add altogether to a global picture of the various synchronization stages.

Figure 2.6 also reveals that the Kullback–Leibler distance of some higher order classes saturates when the value of the coupling constant $k$ increases. Indeed, Fig. 2.6(b) and (c) shows that the coupled dynamics lead to the extinction of order classes $C_5(6)$, $C_7(7)$, and $C_{10}(7)$ at $k \approx 0.145$, $k \approx 0.09$, and $k \approx 0.145$, respectively.

Figure 2.7(a) and (b) shows the probability density $P_C$ of the order classes for $L = 6$ and $L = 7$, respectively. Note that Fig. 2.6(a) displays the contrast between probability densities as in Fig. 2.7 and those of the independent processes. In particular, a vanishing contrast as for $k \approx 0.005$ indicates that the corresponding probability density $P_C$ (which is clearly non-uniform) is similar to the probability density of transcriptions generated by two independent Rössler systems. In the vicinity of the transition to phase synchronization, $k \approx 0.039$, $P_C$ deviates from the probability density of independent processes (Fig. 2.6(a)), and higher order classes dominate the coupled dynamics (Fig. 2.7). This trend is reversed when increasing $k$, and already at $k \approx 0.062$ (resp. $k \approx 0.074$), the class of order 2, $C_2(6)$ (resp. $C_2(7)$), is prevalent (except at $k = 0.299$ for $L = 6$, in which case $C_1(6)$ prevails).
If following [159] we agree to measure the complexity of a transcription by its order, then the probability density of order classes indicates how complex the relationship between the time series is. Figure 2.7 demonstrates that higher order transcriptions play an important role in the description of complex synchronization states such as phase synchronization ($k \geq 0.036$)—a regimen in which amplitudes remain chaotic and uncorrelated. As $k$ increases, the probability densities of higher order classes decrease and some of them vanish, like $C_5(6)$, $C_7(7)$, and $C_{10}(7)$. In fact, simpler synchronization states such as intermittent-lag and lag synchronizations ($k > 0.11$) are predominantly described by lower order classes ($C_2(L)$ and $C_1(L)$). Clearly, the simplest synchronization state, namely complete synchronization, will only be described by the identity ($C_1(L)$).
Permutation Complexity in Dynamical Systems
Ordinal Patterns, Permutation Entropy and All That
Amigó, J.
2010, X, 280 p. 13 illus. in color., Hardcover
ISBN: 978-3-642-04083-2