DIFFERENTIATING PATIENTS WITH OBSTRUCTIVE SLEEP APNEA FROM HEALTHY CONTROLS BASED ON HEART RATE – BLOOD PRESSURE COUPLING QUANTIFIED BY ENTROPY-BASED INDICES

PAWEŁ PILARCZYK

Faculty of Applied Physics and Mathematics & Digital Technologies Center, Gdańsk University of Technology, 80-233 Gdańsk, Poland.
Orcid: 0000-0003-0597-697X

GRZEGORZ GRAFF

Faculty of Applied Physics and Mathematics & BioTechMed Center, Gdańsk University of Technology, 80-233 Gdańsk, Poland.
Orcid: 0000-0001-5670-5729

JOSÉ M. AMIGÓ

Centro de Investigación Operativa (CIO), Universidad Miguel Hernández, 03202 Elche, Elche, Spain.
Orcid: 0000-0002-1642-1171

KATARZYNA TESSMER

Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, 80-233 Gdańsk, Poland.
Orcid: 0000-0003-4487-8037

KRZYSZTOF NARKIEWICZ

Department of Hypertension and Diabetology, Medical University of Gdańsk, 80-210 Gdańsk, Poland.

E-mail addresses: pawel.pilarczyk@pg.edu.pl, grzegorz.graff@pg.edu.pl (corresponding author), jm.amigo@umh.es, katarzyna.tessmer@pg.edu.pl, krzysztof.narkiewicz@gumed.edu.pl, beata.graff@gumed.edu.pl.

2

Orcid: 0000-0002-9181-9884

BEATA GRAFF

Department of Hypertension and Diabetology, Medical University of Gdańsk, 80-210 Gdańsk, Poland.
Orcid: 0000-0002-9181-9884

ABSTRACT. We introduce an entropy-based classification method for pairs of sequences (ECPS) for quantifying mutual dependencies in heart rate and beat-to-beat blood pressure recordings. The purpose of the method is to build a classifier for data in which each item consists of the two intertwined data series taken for each subject. The method is based on ordinal patterns, and uses entropy-like indices. Machine learning is used to select a subset of indices most suitable for our classification problem in order to build an optimal yet simple model for distinguishing between patients suffering from obstructive sleep apnea and a control group.

Heart rate (HR) and blood pressure (BP) are two important physiological variables that are tightly interrelated. Monitoring their mutual influence can provide important information about a person's cardiovascular health. In many cases, the altered effect of HR on BP or vice versa is a consequence of a serious disease. Therefore, indices that reflect mutual relation between these two quantities may support healthcare professionals in the development of more effective diagnostic and treatment approaches. In particular, indices based on ordinal patterns and entropy can serve the purpose of quantitatively measuring mutual relation between two time series, as well as the complexity of each of them. Using these measures to quantify mutual dependencies between HR and BP sequences, we introduce a new method that involves machine learning techniques for differentiating between healthy subjects and persons suffering from obstructive sleep apnea.

1. INTRODUCTION

In recent years, the interest in investigating the complex regulatory mechanisms of the human organism has notably increased. One of the most important directions of research in this area concerns studying interaction between various components of the cardiovascular system. This has created the need for more effective mathematical tools that would be useful in this kind of analysis.

The research goal of this paper is to develop an entropy-based classification method for the analysis of pairs of sequences (ECPS for short): heart rate (HR) and beat-to-beat blood pressure (BP), in order to differentiate between the group of healthy subjects and patients suffering from obstructive sleep apnea (OSA).

The problem of mutual relations between HR and BP is of growing interest in the community of exact sciences, and medical sciences as well. One of the main challenges in this kind of analysis is, according to [27], *extracting meaningful parameters usable*

for diagnostics and risk stratification, and we aim at addressing this point with respect to the patients with OSA.

OSA is a chronic condition characterized by pauses in breathing and/or decreases in airflow caused by repeated total or partial obstruction to the airway occurring during sleep. OSA is highly prevalent especially in patients with cardiovascular diseases. However, in clinical practice many times it is not recognized and not treated [46]. Although the apnea-hypopnea index (AHI), calculated on the basis of a polysomnographic recording (PSG), remains the main tool for the diagnosis and planning of a treatment strategy in OSA patients, more detailed personal approach ("Precision sleep medicine") is currently being advocated [20].

There are several studies of OSA in which classical linear measures are applied [37], nonlinear methods have also been used in the last years on a wider scale [44].

It was shown that OSA leads to alterations in baroreflex which is one of the most important homeostatic mechanisms maintaining blood pressure. Moreover, previous findings suggested that baroreflex sensitivity is reduced before the onset of cardiovascular complications in OSA patients [21].

We propose a comprehensive collection of indices based on ordinal patterns and entropy (OPE indices for short) that are known to be successful in the evaluation of mutual relations between sets of data. We describe these indices in detail in Appendix A.

In real applications, many of the OPE indices may be correlated with each other. Moreover, the variability of some others may not be relevant to the actual differences between the data classes upon consideration. Therefore, in our ECPS method, we propose to use machine learning in order to select a small number of the most relevant indices, and to construct a simple nonlinear classifier that provides optimal results. In this way we develop a model that would distinguish patients that suffer from OSA from healthy subjects, based solely on the measurement of their HR and BP taken during an outpatient exam conducted during the day. An overview of our approach is illustrated in Figure 1.

In [13], symbolic dynamics methods for the HR and BP sequences were successfully applied to study the cardiovascular regulation during different sleep stages in OSA syndrome. Our approach, also based on the analysis of ordinal patterns, follows this promising line of research. However, the novelty of our approach is threefold.

First, we analyze the HR and BP sequences taken during wakefulness, that is, at the time when the sleep apnea problem does not appear; however, long-term cardiovascular consequences of that might be detected.

Second, we compute a large number of entropy-based indices. Last but not least, we use machine learning to select the most relevant indices which yield an optimal classifier.

Finally, we demonstrate the advantage of the selected OPE indices over models built upon traditional measures of HR and BP variability that are typically used in medical practice.
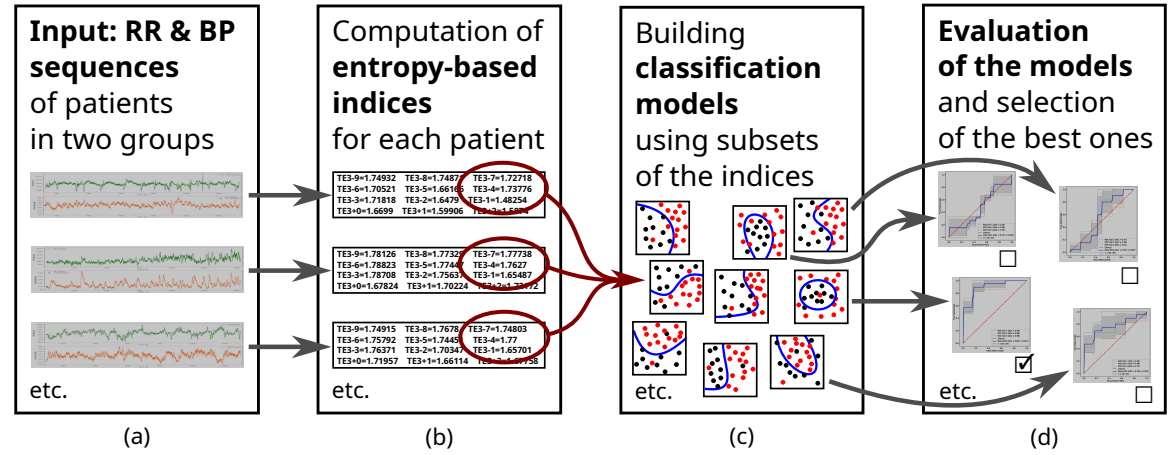
FIGURE 1. Overview of the ECPS method for constructing an optimal classifier of patients: (a) The input to the method consists of pairs of time series (HR and BP sequences) defined for each patient; the patients are split into two groups (CON and OSA). (b) For each patient separately, a collection of entropy-based indices are computed (see Appendix A). (c) Several classification SVM models are built using triplets of indices. (d) All the constructed models are evaluated using ROC curves, and the best ones are selected as optimal classifiers found using the ECPS method.

2. MATERIALS AND METHODS

2.1. **Patients.** There were 19 patients diagnosed with obstructive sleep apnea (OSA) and 19 healthy volunteers (CON) included in the study. The participants were recruited in 2013–2018 from the outpatient hypertension clinic of the University Clinical Centre in Gdańsk and with the help of local advertisements. The study complied with the Declaration of Helsinki; a written informed consent was obtained from each studied person. All of the subjects were males. There were no significant differences with respect to age between OSA and CON groups (48.84 ± 6.48 vs 51.11 ± 8.80, respectively). Except for one OSA person, the patients were treated for hypertension but they did not differ significantly according to office blood pressure compared to the subjects in the CON group.

2.2. **Obtaining raw RR and BP sequences.** For each study participant, a 20-minute recording of ECG (with the PowerLab system and Lab Chart software, sampling rate 1000 Hz) and of non-invasive beat-to-beat blood pressure (with FINOMETER device) were taken in the supine position (except for one patient in the OSA group for whom a recording of 18 minutes is available only due to technical reasons). All the subjects were asked to relax but not to fall asleep. We denote the two time series obtained for each patient p as follows:

- $\{r_{p,i}\}_{i=0}^{k_p}$ – a sequence of RR interval lengths, further called "the RR sequence" for short,
- $\{b_{p,i}\}_{i=0}^{k_p}$ – a sequence of systolic blood pressure (BP) readouts, further called "the BP sequence" for short.
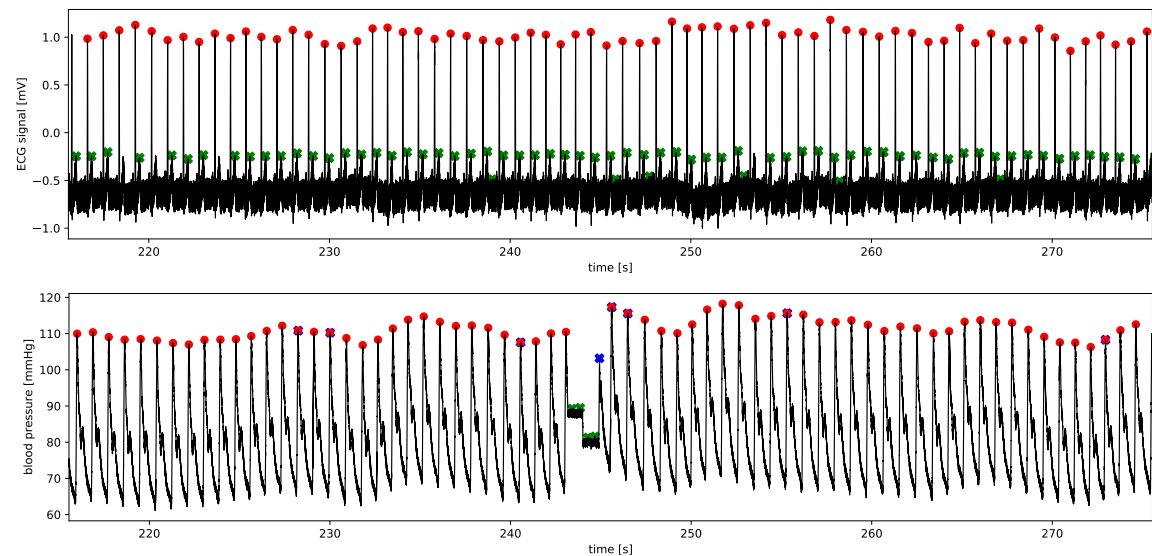
FIGURE 2. Samples of ECG (top figure) and blood pressure (bottom figure) signals with detected peaks marked with red dots. The green and blue crosses indicate other local maxima found in the signal that were excluded from the list of R's and systolic blood pressure readouts. An artifact caused by the blood pressure measuring device calibration procedure can be seen in the middle of the bottom figure.

The length of an RR interval is the distance in time between consecutive R peaks in the ECG signal recording. The associated *time of its occurrence* is defined as the moment of occurrence of the second R in the ECG signal (the right end-point of the RR interval), and is denoted as $\text{time}(r_{p,i})$. The time of occurrence of each element in the BP sequence is denoted as $\text{time}(b_{p,i})$.

Raw RR and BP sequences were automatically generated from raw ECG and blood pressure signals using heuristic algorithms; see Figure 2 for a real sample.

2.3. **Cleaning and matching the RR and BP sequences.** The automatically detected R and BP peaks were visually inspected and manually corrected where necessary. Outliers in both sequences were marked as *artifacts*. These sequences may also contain other points marked as artifacts or there might be some missing data points, which is due to signal disturbances or other kinds of problems, such as ectopic heartbeats (some detected automatically, some marked manually) in the RR sequence, or short periods of calibration of the blood pressure measuring device (detected automatically; shown in Figure 2).

These kinds of problems are unavoidable in the real clinical data. Therefore, in order to match the sequences in a reliable way for cause-and-effect analysis, their subsequences $\{r_{p,\kappa(i)}\}$ and $\{b_{p,\theta(i)}\}$ were selected in such a way that the chosen RR and BP readouts were not artifacts, and were interleaving in time: each RR interval occurrence time must be followed by a BP occurrence, which must in turn be followed by another RR interval occurrence, and so on. Specifically,

$$\text{time}(r_{p,\kappa(i)}) < \text{time}(b_{p,\theta(i)}) < \text{time}(r_{p,\kappa(i+1)}) \qquad \text{for all } i.$$
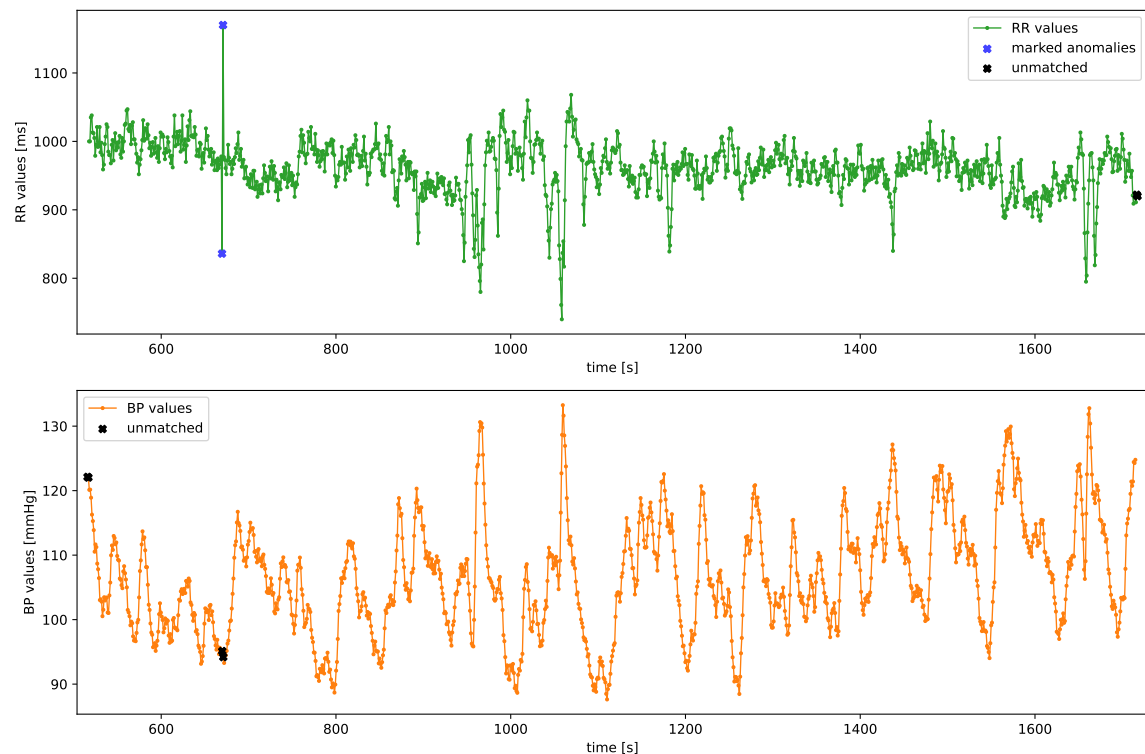
FIGURE 3. Matching an RR sequence with a BP sequence. Two ectopic beats in the RR sequence (at time 670 s) were marked as artifacts (blue crosses), which left the corresponding BP values unmatched (marked with black crosses). There are also a few unmatched BP values at the beginning of the sequence, preceding the first RR detected, and some unmatched RR values at the end of the sequence; this kind of lack of matching often happens at the beginning or at the end of the recording and is normal.

All the elements of the RR and BP sequences that were not included in these subsequences were marked as *artifacts* by the software. See Figure 3 for an example.

We use the two subsequences for further computations. Whenever a tuple of elements of any of the sequences is encountered whose indices are not consecutive (that is, there is some artifact between them), such a tuple is not taken into consideration.

2.4. **Ensuring proper definition of ordinal patterns.** It is required in (A.1) that each subsequence that defines an ordinal pattern must consist of pairwise different terms. However, we cannot expect that real time series that comes from measurements made with finite resolution would satisfy this assumption. One solution to this problem might be to put terms with lower indices first, but then certain patterns would be favored, which we would like to avoid. Therefore, we propose the following procedure (cf. [7] and the discussion about equal values in ordinal patterns in the recent paper [48]).

Before computing ordinal patterns from time series, we add pairwise different small pseudo-random numbers to all the elements in all the sequences. By "small" we mean numbers of 3 orders of magnitude smaller than the granularity of the specific

sequence. The "granularity" is the smallest positive distance between any two terms in the sequence, not necessarily consecutive. For example, if we are given a sequence of integers then we add pseudo-random numbers from the open interval $(0, 0.001)$.

Justification for this procedure is the following. In case of equal results of measurements, it may turn out that increasing the precision would reveal that one of the numbers is actually greater. However, we do not have this information, so we make a "random choice" in order to compensate for the indeterminacy.

2.5. **Computation of the entropy-based indices.** We chose two different lengths L of ordinal patterns obtained from the RR and BP series: 3 and 4, and the shift Λ varying from 0 to 9. More precisely, we considered the following collection of quantities; see Appendix A for the definitions and explanations of the entropy-based indices listed here.

- permutation entropy for both RR and BP series, as defined by Equation (A.4): `RR_PE3`, `RR_PE4`, `BP_PE3` and `BP_PE4` (note that all permutations are bijective);
- statistical complexity for both RR and BP series, as defined by Equation (A.11): `RR_SC3`, `RR_SC4`, `BP_SC3` and `BP_SC4`;
- transcript entropy between the RR and BP series shifted by $\Lambda \in \{0, 1, \ldots, 9\}$, as defined by Equation (A.13):
  `TE3+0`, ..., `TE3+9` and `TE4+0`, ..., `TE4+9`,
  or between the BP and RR series shifted by $\Lambda \in \{1, 2, \ldots, 9\}$:
  `TE3-1`, ..., `TE3-9` and `TE4-1`, ..., `TE4-9`;
- self-transcript entropy of the RR and BP series with the shift $\Lambda \in \{1, 2, \ldots, 9\}$, as defined by Equation (A.14):
  `RR_STE3+1`, ..., `RR_STE3+9`, and `RR_STE4+1`, ..., `RR_STE4+9`,
  as well as `BP_STE3+1`, ..., `BP_STE3+9` and `BP_STE4+1`, ..., `BP_STE4+9`;
- mutual information between RR and BP series shifted by $\Lambda \in \{-9, \ldots, 9\}$, as defined by Equation (A.15):
  `MI3-9`, ..., `MI3+9` and `MI4-9`, ..., `MI4+9`;
- transcript mutual information from RR to BP with the shift $\Lambda \in \{1, 2, \ldots, 9\}$, as defined by Equation (A.16); see also [4, formula (19)]:
  `TMI3+1`, ..., `TMI3+9` and `TMI4+1`, ..., `TMI4+9`,
  and also from BP to RR with the same range of shifts:
  `TMI3-1`, ..., `TMI3-9` and `TMI4-1`, ..., `TMI4-9`.

With the above selection of entropy-based indices, we expect to capture most characteristics of the signals to be analyzed below. Indeed, entropy is a measure of randomness, achieving its maximum value for equiprobable ordinal patterns and transcripts, e.g., white noise. On the other hand, statistical complexity focuses rather on the "structure" of the signal, taking the largest values for signals with intermediate normalized permutation entropies and vanishing for both white noise and zero-entropy signals [22]. As for the mutual information, it is a nonlinear measure of dependence between two random variables [9], which is symmetric in the conventional case and asymmetric if transcripts are used; as said in Section A.3.2, transcript mutual information is a causality indicator. In addition, all these indices are model-free, although they depend on one or two parameters: L (the length of the ordinal patterns) and Λ (coupling delay). The values of L and Λ were chosen based on our experience

[4, 5, 6, 14] and, as shown by the results, they turned out to be satisfactory for our purposes.

2.6. **Computation of the classical HRV and BPV indices.** For comparison, several classical heart rate variability (HRV) parameters were also computed for the RR series using the "hrv-analysis" Python module.[1] They were computed for the RR series after exclusion of artifacts (such as outliers or ectopic beats, as explained in Section 2.3). We analyzed the following parameters:

- time domain features: `mean_nni`, `sdnn`, `sdsd`, `pnni_50`, `pnni_20`, `rmssd`, `median_nni`, `range_nni`, `cvsd`, `cvnni`, `mean_hr`, `max_hr`, `min_hr`, `std_hr`;
- geometrical features: `triangular_index`, `tinn`;
- CSI CVI features: `csi`, `cvi`, `Modified_csi`;
- Poincaré plot features: `sd1`, `sd2`, `ratio_sd2_sd1`;
- frequency domain features: `lf`, `hf`, `lf_hf_ratio`, `lfnu`, `hfnu`, `total_power`, `vlf`.

Note that we excluded sample entropy from the list of the classical HRV parameters, because this already is an entropy-based index, and our intention was to compare our entropy-based indices with indices of different type.

We also computed the following quantities for determining the degree of blood pressure variability (BPV):

- time domain features: `mean`, `sd`;
- frequency domain features: `hf`, `lf`, `vlf`;

see [29, Figure 2 and Table 2] for justification of our selection of the features in addition to the most obvious choice of the mean and standard deviation. Similarly to the HRV indices, we did not include sample entropy in our comparison. The results of the computation of all these indices are available in [30].

2.7. **Choosing an optimal set of indices.** The OPE indices listed in Section 2.5 give rise to 156 variables that can be taken as a basis for statistical analysis and machine learning. In our ECPS method, we are specifically interested in building an effective classifier.

Since it happens that variables irrelevant to the classification may degrade the quality of the classifier, we propose to choose a small subset of the indices that would yield an optimal classifier. Since it is not obvious which lengths L, which coupling delays Λ, and which specific indices are optimal for a specific application, we propose to test all the possible triplets of indices, and choose the triplet that provides optimal results. These results can be assessed using a chosen method for the assessment of the classification model of the type that we use.

Taking this into consideration, we used receiver operating characteristic curves (ROC curves) and we computed average area under the curve (AUC) for the ROC curves obtained in 5 repetitions of a 3-fold cross-validation test for the non-linear support vector machine (SVM) binary classifier with the radial basis function (RBF) kernel (see e.g. [16]) for all the possible triplets of OPE indices. Although it may seem a computationally demanding task, in fact all these computations were completed by a Python script on a modern PC within less than 30 minutes. After the model has

---

[1]Heart Rate Variability analysis, `https://github.com/Aura-healthcare/hrv-analysis/`.

been constructed, applying it to new data (e.g., diagnosing another patient) is very fast.

More specifically, we generated the collection C of all the 620,620 subsets of cardinality 3 of the set of all the 156 features (such as `MI4+3`) computed for the ordinal patterns, listed in Section 2.5. For each set $c \in C$ we conducted the following test. We split the set of subjects into 3 subsets of approximately equal size with approximately the same proportion of subjects in both groups (which is known under the term "with stratification"). We then built a classifier using the standard nonlinear SVM model with the RBF kernel on two of these three sets, and tested this model on the third set. The quality of the classifier can be assessed visually on the basis of the ROC curves, and numerically by computing the area under the curve (AUC). The higher the AUC value, the better the classifier: the value of about 0.5 indicates a classifier that is as good as random choice, while the value of 1.0 indicates the perfect classifier that makes no mistakes. We computed the mean of the three AUCs. We repeated this procedure of training the SVM model and testing it for the three possibilities of choosing the sets for building the model. This procedure of splitting the set of subjects, building the classifier, and testing it was repeated 5 times in order to ensure that the good or bad result is not due to a more or less lucky subdivision of the set of subjects into the three subsets. In this way, for each triplet of the indices under consideration, we computed the mean AUC, the standard deviation of all the 15 AUCs obtained in the 5 repetitions, as well as the minimum and the maximum values of the five mean AUCs. The computed values were gathered in a table, which we then sorted by mean AUC.

3. Results

Recall that whenever any of the ordinal patterns under analysis contained an artifact in the RR or BP series, its appearance was not counted. The percentage of this kind of situations was recorded. In our dataset, there were no ordinal patterns omitted in the sequences for 21 subjects. The number of omitted patterns was below 1% in the data of further 11 subjects, below 2% for the next 5 subjects, and reached the maximum of 2.6% for one subject. In total, less than 0.305% of all the patterns were omitted in all the analyses. The results of the computation of all these indices are available in [30].

Before proceeding with machine learning, we applied the U Mann-Whitney test to verify whether any single index could be used alone to differentiate between the two groups of patients (with p-value below 0.05). The result was positive for two classical HRV indices: `HRV_mean_nni` (p-value 0.0308) and `HRV_median_nni` (p-value 0.0329), and four OPE indices: `TMI4-9`, `TMI4+5`, `TMI4+6`, `TMI4+7` (p-values between 0.0102 and 0.0307).

We also computed Pearson linear correlations (with p-values) between all the OPE and classical HRV and BPV indices on the set of all the subjects together and also on the CON and OSA groups. The results are gathered in CSV files and in illustrations available in [30]. We found several strong correlations between some HRV indices, mainly due to the way they were defined, and other correlations that were well known already. The classical BPV indices did not exhibit such strong correlations with any of the classical HRV or BPV indices. In the case of the OPE indices that we computed,

however, we found a considerable number of very strong correlations, in some cases even above 0.98 in absolute value. Finally, we verified correlations between the OPE indices and the classical HRV and BPV indices, but we did not find very strong correlations there.

The best triplets of the indices derived from ordinal patterns following the procedure described in Section 2.7 are listed in Table 1, ordered by mean AUC. The table also contains the standard deviation of the AUCs computed in the five attempts, as well as the minimum and the maximum mean AUC encountered in the five attempts.

Mean AUC ± SD	Min AUC	Max AUC	the triplet of indices
0.857 ± 0.086	0.800	0.897	RR_STE3+6,MI3+4,MI4+9
0.853 ± 0.070	0.811	0.885	RR_STE3+6,MI3+4,MI4+8
0.846 ± 0.072	0.800	0.875	RR_STE3+6,MI3+5,MI4+8
0.846 ± 0.076	0.794	0.888	RR_STE3+6,MI3+4,MI4+7
0.838 ± 0.068	0.790	0.879	MI3+4,RR_STE4+5,MI4+8
0.833 ± 0.104	0.737	0.870	RR_STE3+6,MI3+5,MI4+7
0.831 ± 0.113	0.741	0.876	RR_STE3+6,MI4+7,MI4+8
0.831 ± 0.088	0.776	0.895	MI3+4,RR_STE4+5,MI4+9
0.822 ± 0.094	0.779	0.838	RR_STE3+6,MI3+5,MI4+9
0.819 ± 0.078	0.780	0.857	RR_STE3+4,RR_STE3+6,MI4+8

TABLE 1. The triplets of OPE indices that yield the best non-linear SVM models based on the OPE indices.

In order to obtain a reference point to benchmark the performance of the OPE indices, an analogous computation was conducted using the 33 classical HRV and BPV features described in Section 2.6 (5,456 triplets), and the best results are shown in Table 2.

Mean AUC ± SD	Min AUC	Max AUC	the triplet of indices
0.694 ± 0.137	0.614	0.763	HRV_range_nni,HRV_min_hr,HRV_triangular_index
0.691 ± 0.161	0.532	0.779	HRV_min_hr,HRV_triangular_index,BPV_sd
0.689 ± 0.120	0.583	0.762	HRV_pnni_20,HRV_range_nni,HRV_triangular_index
0.669 ± 0.144	0.548	0.771	HRV_median_nni,HRV_range_nni,HRV_triangular_index
0.668 ± 0.138	0.599	0.730	HRV_cvnni,HRV_min_hr,HRV_Modified_csi
0.667 ± 0.107	0.562	0.751	HRV_pnni_20,HRV_range_nni,HRV_cvnni
0.665 ± 0.100	0.592	0.728	HRV_pnni_20,HRV_cvnni,HRV_lf
0.664 ± 0.140	0.572	0.746	HRV_range_nni,HRV_mean_hr,HRV_triangular_index
0.661 ± 0.151	0.566	0.755	HRV_cvnni,HRV_min_hr,HRV_lf
0.660 ± 0.139	0.544	0.755	HRV_mean_nni,HRV_range_nni,HRV_triangular_index

TABLE 2. The triplets of indices that yield the best non-linear SVM models built upon the classical HRV and BPV indices.

The results of the computation of all these triplets of indices are available in [30]. One CSV file contains all the OPE triplets, and another CSV file contains all the classical HRV and BPV triplets, together with the data indicated in Tables 1 and 2, sorted by the mean AUC value in the descending order.

The ROC curves and AUCs computed for the best of the 5 cases of the 3-fold cross validation for the best triplets of indices are shown in Figure 4 for the OPE indices, and in Figure 5 for the others.
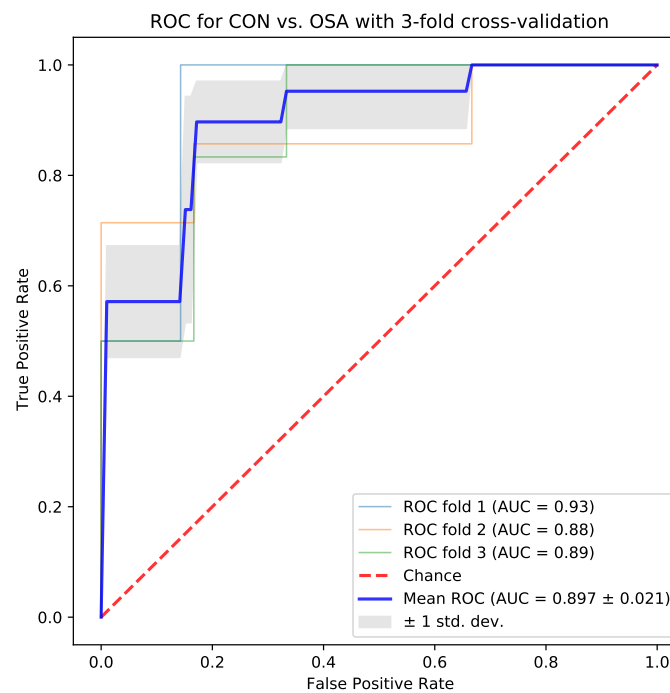

FIGURE 4. Results for one of the nonlinear SVM classification models built upon the top-ranked triplet of OPE indices: RR_STE3+6, MI3+4, MI4+9.

For the purpose of measuring the overall performance of the specific indices that participated in all the tested triplets, we computed the average value of 1% of the best mean AUCs in which each of the features was included. The best ten OPE indices turned out to be RR_STE3+6, TMI4+1, RR_STE4+5, MI4+9, MI4+8, RR_STE3+3, BP_-STE3+3, TE3+1, MI4−5, and MI4+7 (in this order), and the best ten classical HRV and BPV indices were: HRV_pnni_20, HRV_triangular_index, BPV_sd, HRV_range_nni, HRV_cvi, HRV_Modified_csi, BPV_vlf, HRV_min_hr, HRV_cvnni, and HRV_lf,] (in this order). It is interesting to see that the indices successful in statistical tests are not present in these lists. Indeed, as far as the TMI4-based indices are concerned, their values are statistically different in the two groups, but the ranges overlap considerably, which apparently does not help in building a successful classifier.

The quality of the indices found in the best triplets can be illustrated by the fact that one can see relatively good separation of the data groups already in a 2-dimensional plot in two selected coordinates; see Figure 6 for an example. We would like to remark that it is a common practice to apply PCA in order to visualize the separation of data groups in the plane of greatest variability of the variables. We made such an attempt using all the OPE indices together, but – although we do not show the corresponding figures here – we did not see any separation in the first few components. This proves that variability of some of the indices was in fact useless for distinguishing the two groups. Our machine learning approach helped us ignore such
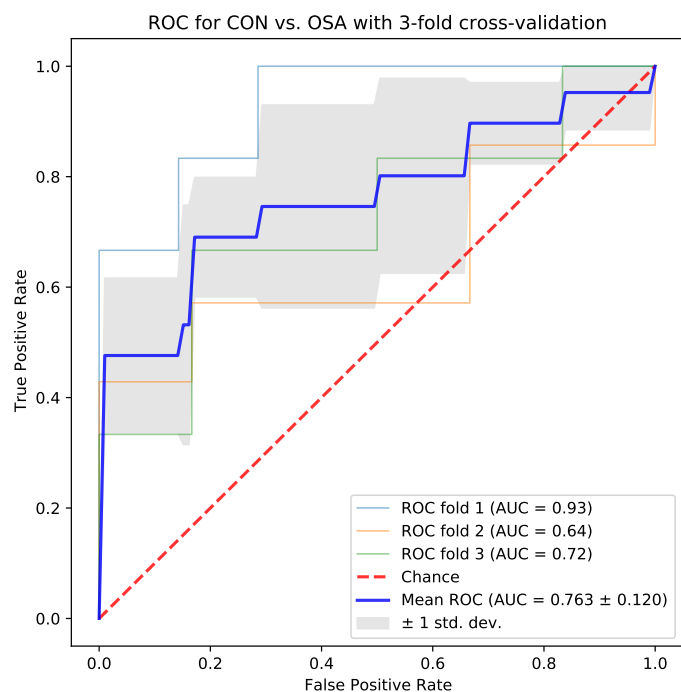
FIGURE 5. Results for one of the nonlinear SVM classification models built upon the top-ranked triplet of HRV and BPV indices: `HRV_range_nni`, `HRV_min_hr`, `HRV_triangular_index`.

indices and extract the useful ones, so that we obtained satisfactory effect by simply selecting two of the optimal variables found.
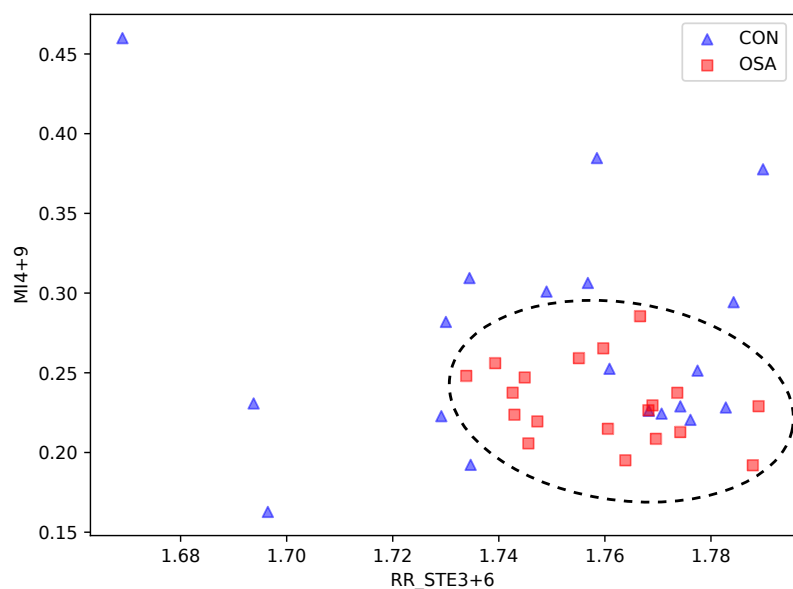

FIGURE 6. Data points projected onto the plane spanned by two OPE indices `RR_STE3+6` and `MI4+9` that participate in the best triplet found. Note the relatively good non-linear separation of the groups.

4. Disscussion

Our results, including the application of the ECPS method (Section 2.7) to the two groups of subjects, confirm the importance of conducting the analysis of mutual relation between heart rate variability (HRV) and blood pressure variability (BPV), particularly if one aims at the development of machine learning methods that could support one in diagnosing certain diseases. With the help of several indices based on ordinal patterns and entropy (OPE indices), we developed a satisfactory classifier capable of distinguishing patients that suffered from obstructive sleep apnea (OSA) from a control group (CON). A similarly built classifier that was based on the classical analysis of HRV and BPV sequences without delving into their mutual relationship (and without considering the entropy) achieved considerably worse results.

We would like to point out the fact that in order to measure the quality of the classifier in Section 2.7, we did not only look at mean AUC, but also at the consistency between results obtained for different train–test splits of the data, reflected by their standard deviation that is shown in the tables. Indeed, as one can see in Tables 1 and 2, some results with relatively high AUCs in fact corresponded to models whose reliability could be questioned, because in spite of high AUC for some subdivisions of the set of subjects, in some other cases the results were considerably worse. This is very well reflected in Figures 4 and 5, where one can see the wide gray belt and one poor classification result (AUC=0.64) for the standard HRV and BPV indices, as opposed to constantly valuable results achieved for the OPE indices.

Correlations between the OPE indices and the classical HRV and BPV indices might provide some hints for the interpretation of the actual physiological features that were reflected by the OPE indices. In particular, stronger correlations in general could be noticed in the CON group than in the OSA group, which might indicate impaired regulation between HR and BP in the latter patients.

At this point, we would like to make a comment on why we chose to investigate triplets of indices, and not pairs or quadruplets (or larger tuplets). In fact, we did conduct analogous computations for pairs and quadruplets of indices. However, the results for pairs were clearly worse than those for the triplets, while the computations for quadruplets yielded essentially as good results as for the triplets, but were considerably more costly in terms of computation time, because we had to run each individual machine learning test for each of the 23,738,715 quadruplets, as opposed to the 620,620 triplets. We suspect that the fact that the results were not significantly improved could be attributed to strong correlation (negative or positive) between many OPE indices, which one can easily see in the illustrations provided in [30].

It might also be tempting to build a classification model based on all the indices together. Indeed, we conducted such an experiment. We used all the 156 OPE indices together for one model, and all the 33 classical HRV and BPV indices together for the other model. We tried both linear SVM and the nonlinear SVM with the RBF kernel. The results of 3-fold cross-validation in all the tested cases were discouraging – the quality of the SVM classifier was essentially equivalent to the "random choice" classifier, as the resulting mean AUCs were rarely above 0.6. This shows why it is beneficial to choose a selection of indices in the ECPS method, as described in Section 2.7.

We would like to emphasize the fact that the ECPS method may be enriched by adding other types of indices, and then using the wider selection for choosing an optimal subset. Indeed, the interaction between time series that represent HR and BP has been observed and analyzed using a variety of methods in many recent studies, such as [10, 11, 12, 17, 23, 32, 41, 45]; see also the review papers [27, 36] and the references therein. For example, momentary information transfer was introduced in [31] for multivariate analysis of coupling between two time series. In order to verify whether adding this index to our analysis could yield better results, we additionally computed this index with delays in $\{-9, \ldots, 9\}$, as defined by [31, Formula (9)]; we shall further denote these indices as `MIT-9`, ..., `MIT+9`. Then we applied the analysis described in Section 2.7 to the resulting set of 175 features (instead of the original 156 ones). Some of the new indices indeed appeared in the tested triplets at relatively high positions, e.g., `MIT-7` in a triplet yielding AUC 0.835 (position 6 in the ranking) and AUC 0.831 (position 8), and `MIT-4` in a triplet yielding AUC 0.824 (position 10). This indicates the usefulness of the `MIT` indices, although adding them did not improve the best triplet found.

Another question is whether our method is specific to the OSA patients, or maybe it might be applied to other pairs of time series, too. In order to verify this, we tested our method on an artificial set of 20 RR and BP sequences generated by bidirectionally coupled logistic maps, as defined in [31, Formula (25) and Figure 2(a)]. Already using basic statistical tests, a multitude of the OPE indices, and also all the `MIT` indices allowed to significantly differentiate between the group of these pairs of sequences and a set of 20 other pairs of sequences generated by uncoupled logistic maps. This shows that with machine learning support, the ECPS method is capable of finding much more subtle differences in pairs of time series than those discussed in [31].

The quantitative definitions and qualitative meanings of the OPE indices were given in Appendix A and Section 2.5, respectively. It is remarkable that the best four OPE indices in statistical tests (Section 3) were based on transcripts; namely, this was the transcript mutual information with different coupling delays. Similarly, among the best ten OPE indices that participated in all the tested triplets (Section 3), six of them were also based on transcripts; namely, self-transcript entropy appeared four times, and transcript mutual information and transcript entropy appeared once. We conclude that transcripts have a noteworthy potential for discrimination. Note that the SC indices did not participate in high ranked classification models, even though they are often taken into consideration in the literature. It is also a noticeable fact that among all the classical HRV and BPV indices, the only statistically significant differences between the groups were exhibited by the mean and median length of time intervals between heartbeats.

The results illustrated in Figure 6 showed that, when considering simultaneously both indices `RR_STE3+6` and `MI4+9`, patients with OSA have lower values of `MI4+9` and the same or greater values of `RR_STE3+6`. Recall that `MI4+9`, the mutual information, is a symmetric measure of dependence. Its lower values indicate weaker coupling that reflects impaired information flow between RR and BP series. The greater values of `RR_STE3+6` are related to higher predictability of the RR series of OSA patients, which may be interpreted as a loss of complexity in heart rate in comparison of healthy subjects, a typical syndrome in pathology or aging. Let us remark that we observed

the same regularities in case of data points projected onto the plane spanned by two other OPE indices: RR_STE3+6 and MI3+4 that are present in the best triplet found.

It might be tempting to create a "profile" of a healthy patient with respect to the indices that we computed, or to indicate ranges of values of concern with respect to OSA. However, the separation of the two groups of patients is not linear, as shown in Figure 6, which makes description of such a profile cumbersome. Moreover, the number of patients tested in our research is too small to make a reliable profile; a more involved research would be necessary. Nevertheless, based on our results one might make an attempt, for example, to claim that RR_STE3+6 values between 1.74 and 1.79 might be an indicator of possible OSA problems, provided that MI4+9 is within 0.17 and 0.27 and MI3+4 is between 0.03 and 0.08.

The main conclusion from the specific application of our ECPS method to the RR and BP time series is that the OPE methods might show altered cardiovascular regulation in awake OSA patients. Although there is a number of studies on HRV and BPV in OSA patients, they differ with the time of assessment (mostly night-time), parameters tested (usually HR only) and indices used for the parameters' variability analysis (typically standard, linear methods) [33, 43].

Non-linear parameters have been recently proven to have a great potential for showing specific features of HRV in a large group of OSA patients. With the increasing severity of the disease, lower complexity of heart rate during sleep [19] was found in OSA subjects, and, what is more, also during wakefulness [34].

Studying the relation between heart rate and blood pressure in continuous recordings might be challenging. Most of the studies in OSA subjects in this area were based on the assessment of baroreflex sensitivity (BRS). The question remains, however, whether the standard BRS assessment is able to catch the whole range of autonomic disturbances in OSA patients [21].

Another approach has been proposed recently that allows the assessment of non-linear nature of cardiovascular regulation based on information transfer quantification [11, 12, 17, 41]. To our best knowledge, this is the first study which successfully combines the above kind of approach with the use of advanced symbolic methods and machine learning for the elucidation of changes in awake OSA patients. However, the interpretation of physiological meaning of OPE indices is challenging and needs further investigation in a larger group of patients. For example, in our study the best discrimination power was achieved by the MI and TMI indices for larger lags, which cannot be directly explained by baroreflex or Frank-Starling mechanisms. We hypothesized that in group of OSA patients altered respiration has to be considered first looking for the explanation.

In our recent pilot study, patients with apnea episodes and desaturations characterized with the increased respiratory variability in short recordings during wakefulness which could impact cardiovascular coupling [15]. There is also an emerging area of growing interest considering various vascular or blood pressure response to sympathetic outflow (i.e., sympathetic transduction), depending for example on age, sex or obesity status, which might explain the delay of 6–9 heartbeats. It was also hypothesized that altered sympathetic transduction might be a compensatory mechanism in subjects with baroreflex dysfunction; in fact, it is known that baroreflex impairment is present in OSA patients [47]. Summarizing, further studies are needed to assess

the actual mechanisms underlying our findings related to the role of long delays as well as the effectiveness of the ordinal pattern parameters in differentiating the two groups of subjects that we considered.

Data Availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Acknowledgments

References

[1] J. M. Amigó, R. Monetti, T. Aschenbrenner and W. Bunk, "Transcripts: An algebraic approach to coupled time series," *Chaos* 22 013105, (2012).

[2] J. M. Amigó and K. Keller, "Permutation entropy: One concept, two approaches," *European Physical Journal Special Topics* 222, 263–273 (2013).

[3] J. M. Amigó, T. Aschenbrenner, W. Bunk and R. Monetti, "Dimensional reduction of conditional algebraic multi-information via transcripts," *Information Sciences* 278, 298–310 (2014).

[4] J. M. Amigó, R Monetti, N. Tort-Colet and M. V. Sanchez-Vives, "Infragranular layers lead information flow during slow oscillations according to information directionality indicators," *Journal of Computational Neuroscience* 39, 53–62 (2015).

[5] J. M. Amigó, R. Monetti, B. Graff and G. Graff, "Computing algebraic transfer entropy and coupling directions via transcripts," *Chaos* 26 113115, (2016).

[6] J. M. Amigó, B. Graff, G. Graff, R. Monetti and K. Tessmer, "Detecting coupling directions with transcript mutual information: A comparative study," *Discrete and Continuous Dynamical Systems B* 24, 4079–4097 (2019).

[7] C. Bandt and B. Pompe, "Permutation entropy – a natural complexity measure for time series," *Physical Review Letters* 88, 174102 (2002).

[8] C. Bandt, G. Keller and B. Pompe, "Entropy of interval maps via permutations," *Nonlinearity* 15, 1595–1602 (2002).

[9] T. M. Cover and J. A. Thomas, "Elements of Information Theory," *Wiley-Interscience*, (2006).

[10] L. Faes, G. Nollo and A. Porta, "Mechanisms of causal interaction between short-term RR interval and systolic arterial pressure oscillations during orthostatic challenge," *Appl. Physiol.* 114, 1657–1667 (2013).

[11] L. Faes, G. Nollo and A. Porta, "Non-uniform multivariate embedding to assess the information transfer in cardiovascular and cardiorespiratory variability series," *Comput Biol Med.* 42(3), 290–7 (2012).

[12] L. Faes, G. Nollo and A. Porta, "Information domain approach to the investigation of cardiovascular, cardio-pulmonary and vasculo-pulmonary causal couplings," *Front Physiol.* 2:80, (2011).

[13] A. Gapelyuk, M. Riedl, A. Suhrbier, J. F. Kraemer, G. Bretthauer, H. Malberg, J. Kurths, T. Penzel and N. Wessel, "Cardiovascular regulation in different sleep stages in the obstructive sleep apnea syndrome," *Biomed Tech (Berl).* 56(4), 207-13 (2011).

[14] G. Graff, B. Graff, A. Kaczkowska, D. Makowiec, J. M. Amigó, J. Piskorski, K. Narkiewicz and P. Guzik, "Ordinal pattern statistics for the assessment of heart rate variability," *European Physical Journal Special Topics* 222, 525–534 (2013).

[15] B. Graff, G. Graff, P. Pilarczyk, M. Szymańska and K. Narkiewicz, "Assessment of respiratory pattern variability during wakefulness in subjects with sleep apnea," 12th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO), Vysoké Tatry, Štrbské Pleso, Slovakia, 2022, pp. 1–2.

[16] J. Han, M. Kamber and J. Pei, "Data Mining: Concepts and Techniques," in *The Morgan Kaufmann Series in Data Management Systems*, (2011).

[17] M. Javorka, J. Krohova, B. Czippelova, Z. Turianikova, Z. Lazarova, K. Javorka and L. Faes, "Basic cardiovascular variability signals: mutual directed interactions explored in the information domain," *Physiol Meas.* 38(5), 877–894 (2017).

[18] H. Kantz and T. Schreiber, "Nonlinear Time Series Analysis," *Cambridge University Press* (1997).

[19] D. Liang, S. Wu, L. Tang, K. Feng, G. Liu, "Short-Term HRV Analysis Using Nonparametric Sample Entropy for Obstructive Sleep Apnea," *Entropy* Feb 24;23(3):267. (2021).

[20] D. C. Lim, D. R. Mazzotti, K. Sutherland, J. W. Mindel, J. Kim, P. A. Cistulli, U. J. Magalang, A. I. Pack, P. de Chazal and T. Penzel, "Reinventing polysomnography in the age of precision medicine," *Sleep Med Rev.* 52:101313, (2020).

[21] C. Lombardi, M. F. Pengo and G. Parati, "Obstructive sleep apnea syndrome and autonomic dysfunction," *Auton Neurosci.* 221:102563, (2019).

[22] P. Martínez Coq, A. Rey, O. A. Rosso, R. Armentano and W. Legnani, "Detection of cardiac arrhythmia patterns in ECG through $H \times C$ plane," *Chaos* 32 123118, (2022).

[23] N. Marwan, Y. Zou, N. Wessel, M. Riedl and J. Kurths, "Estimating coupling directions in the cardiorespiratory system using recurrence properties," *Philos Trans A Math Phys Eng Sci.* 371(1997):20110624, (2013).

[24] R. Monetti, W. Bunk, T. Aschenbrenner and F. Jamitzky, "Characterizing Synchronization in Time Series using Information Measures Extracted from Symbolic Representations," *Phys. Rev. E* 79 046207, (2009).

[25] R. Monetti, J. M. Amigó, T. Aschenbrenner and W. Bunk, "Permutation complexity of interacting dynamical systems," *European Physical Journal Special Topics* 222, 421–436 (2013).

[26] R. Monetti, W. Bunk, T. Aschenbrenner, S. Springer and J. M. Amigó, "Information directionality in coupled time series using transcripts," *Phys. Rev. E* 88 022911, (2013).

[27] A. Müller, J. F. Kraemer, T. Penzel, H. Bonnemeier, J. Kurths and N. Wessel, "Causality in physiological signals," *Physiol Meas.* 37(5):R46-72, (2016).

[28] A. Myers and F. A. Khasawneh, "On the automatic parameter selection for permutation entropy," *Chaos* 30 033130, (2020).

[29] G. Parati et al., "Blood pressure variability: methodological aspects, clinical relevance and practical indications for management — a European Society of Hypertension position paper," *J Hypertens.* 41(4):527-544, (2023).

[30] P. Pilarczyk, "ECG BP Transcripts. Results of Computations," (2023); see `https://www.pawelpilarczyk.com/ecg2bp/` (accessed July 25, 2023).

[31] B. Pompe, J. Runge, "Momentary information transfer as a coupling measure of time series," *Phys Rev E* 83, 051122 (2011).

[32] A. Porta, A. M. Catai, A. C. Takahashi, V. Magagnin, T. Bassani, E. Tobaldini, P. van de Borne and N. Montano, "Causal relationships between heart period and systolic arterial pressure during graded head-up tilt," *Am J Physiol Regul Integr Comp Physiol.* 300(2):R378-86, (2011).

[33] H. Qin, N. Steenbergen, M. Glos, N. Wessel, J. F. Kraemer, F. Vaquerizo-Villar and T. Penzel, "The Different Facets of Heart Rate Variability in Obstructive Sleep Apnea," *Front Psychiatry* 12:642333, (2021).

[34] H. Qin, B. T. Keenan, D. R. Mazzotti, F. Vaquerizo-Villar, J. F. Kraemer, N. Wessel, S. Tufik, L. Bittencourt, P. A. Cistulli, P. de Chazal, K. Sutherland, B. Singh, A. I. Pack, N. H. Chen, I. Fietze, T. Gislason, S. Holfinger, U. J. Magalang and T. Penzel, "Heart rate variability during wakefulness as a marker of obstructive sleep apnea severity," *Sleep* 44(5):zsab018, (2021).

[35] T. Sauer, J. A. Yorke and M. Casdagli, "Embedology," *J. Stat. Phys.* 65, 579–616 (1991).

[36] S. Schulz, F. C. Adochiei, I. R. Edu, R. Schroeder, H. Costin, K. J. Bar and A. Voss, "Cardiovascular and cardiorespiratory coupling analyses: a review," *Phil. Trans. R. Soc. A* 371(1994):20120191, (2013).

[37] V. C. C. Sequeira, P. M. Bandeira and J. C. M. Azevedo, "Heart rate variability in adults with obstructive sleep apnea: a systematic review," *Sleep Sci.* 12(3):214-221, (2019).

[38] J. C. Sprott, "Chaos and Time-Series Analysis," *Oxford University Press*, (2003).

[39] M. Staniek and K. Lehnertz, "Symbolic transfer entropy," *Physical Review Letters* 100(15):158101, (2008).

[40] J. Stark, "Delay embeddings for forced systems. I. Deterministic forcing," *J. Nonlinear Sci.* 9, 255–332 (1999).

[41] J. Svačinová, M. Javorka, Z. Nováková, E. Závodná, B. Czippelová and N. Honzíková, "Development of causal interactions between systolic blood pressure and inter-beat intervals in adolescents," *Physiol Res.* 64(6):821-9, (2015).

[42] F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence, Warwick 1980*, 366–381 (1981).

[43] S. Ucak, H.U. Dissanayake, K. Sutherland, P. de Chazal, P.A. Cistulli, "Heart rate variability and obstructive sleep apnea: Current perspectives and novel technologies," *J Sleep Res.* 30(4):e13274 (2021).

[44] C. Varon and S. Van Huffel, "Complexity and Nonlinearities in Cardiorespiratory Signals in Sleep and Sleep Apnea," *Complexity and Nonlinearity in Cardiovascular Signals*, 503–537 (2017).

[45] N. Wessel, A. Gapelyuk, J. F. Kraemer, K. Berg and J. Kurths, "Spontaneous baroreflex sensitivity: sequence method at rest does not quantify causal interactions but rather determines the heart rate to blood pressure variability ratio," *Physiol Meas.* 41(3):03LT01, (2020).

[46] Y. Yeghiazarians, H. Jneid, J. R. Tietjens, S. Redline, D. L. Brown, N. El-Sherif, R. Mehra, B. Bozkurt, C. E. Ndumele and V. K. Somers, "Obstructive Sleep Apnea and Cardiovascular Disease: A Scientific Statement From the American Heart Association," *Circulation* 144(3):e56-e67, (2021).

[47] B.E. Young, J.L. Greaney, D.M. Keller, P.J. Fadel, "Sympathetic transduction in humans: recent advances and methodological considerations," Am J Physiol Heart Circ Physiol. 1;320(3): H942-H953, (2021).

[48] M. Zanin and F. Olivares, "Ordinal patterns-based methodologies for distinguishing chaos from noise in discrete time series," *Commun Phys* 4(1):190, (2021).

Appendix A. Mathematical tools: Ordinal patterns and entropy-based indices (OPE indices)

Let $(\Omega, \mathcal{B}, \mu)$ be a probability space where, for the sake of this paper, Ω is a compact manifold, $\mathcal{B}$ is the Borel sigma-algebra of Ω, and μ is a probability over the measurable space $(\Omega, \mathcal{B})$. If, furthermore, the map $f \colon \Omega \to \Omega$ is μ-invariant, i.e., f is measurable and $\mu(f^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}$, then $(\Omega, \mathcal{B}, \mu, f)$ is called a (discrete-time, measure-preserving) *dynamical system*. The working hypothesis of nonlinear time series analysis (NTSA) is that real world (univariate) time series are scalar observations of a generally higher dimensional dynamical system $(\Omega, \mathcal{B}, \mu, f)$ which, moreover, is either unknown, only partially known, or too complex for its knowledge to be of any practical use. This being the case, it is assumed hereafter that any time series $(x_n)_{n \geq 0}$, $x_n \in \mathbb{R}$, is not, in general, an orbit of an autonomous

dynamic f but an orbit of a non-autonomous dynamic $\varphi \circ f$, where $\varphi \colon \Omega \to \mathbb{R}$ is a so-called *observation function*, typically a projection $\mathbf{x} \mapsto x$ from Ω onto one of its local coordinates. From the point of view of time series analysis, this results in that the observations $(x_n)_{n \geq 0}$ can be viewed as random, unless $\Omega \subset \mathbb{R}$ and φ is invertible (e.g., the identity). For this reason, we will write

$$(x_n)_{n \geq 0} = (\varphi \circ f^n(\mathbf{x}_0))_{n \geq 0} = (\varphi(\mathbf{x}_n))_{n \geq 0} =: \ \mathbf{X}(\mathbf{x}_0)$$

and say that $(x_n)_{n \geq 0}$ is a trajectory or realization of the random process $\mathbf{X}$. Here f^n is the nth iterate of f for $n \geq 1$, f^0 is the identity, and $\mathbf{x}_n = f^n(\mathbf{x}_0) \in \Omega$. Formally, $\mathbf{X}(\mathbf{x}_0)$ is a deterministic orbit with dynamical noise.

Because of the applications of NTSA to real world observations, specially to irregular time series, the underlying dynamical system $(\Omega, \mathcal{B}, \mu, f)$ is generally supposed to be chaotic, and Ω to be an attractor. Some techniques of NTSA aim to reconstruct and characterize the attractors of such systems from the observations $(x_n)_{n \geq 0}$ using time delay coordinates [42, 35, 40]. Such characterization include Lyapunov exponents, several dimensions and dynamical entropies such as the Kolmogorov-Sinai entropy [18, 38]. However, since our goal in this paper is to classify patients into two groups based on their heart rate and blood pressure, it will suffice to characterize the structure and complexity of those recordings. To this end, there are two main options: linear (statistical) and nonlinear tools. As the former have been well investigated already, we focus here on the latter. Specifically, we will calculate various kinds of entropy and mutual information using ordinal patterns, which are the building blocks of the so-called ordinal methodology. This is a convenient way to convert continuous-valued time series into finite-valued (or symbolic-valued) time series for the reasons that we mention in the following section.

A.1. **Ordinal patterns and transcripts.** To characterize the complexity of the time series $\mathbf{X}(\mathbf{x}_0)$ we will use the ordinal methodology, which trades off blocks $x_n^L := x_n, x_{n+1}, ..., x_{n+L-1}$ of length $L \geq 2$ for rank vectors called ordinal patterns or simply permutations of length L [7]. Ordinal patterns are conceptually simple and their calculation can be done in real time because knowledge of the data range is not required in advance. In addition, since ordinal patterns can be identified with permutations, one can take advantage of their algebraic structure, e.g., through transcripts, as we will do in Section A.3. Last but not least, the Shannon entropy rate of the ordinal patterns as $L \to \infty$ equals the Kolmogorov-Sinai entropy of the underlying dynamics in the case of one-dimensional interval maps under very mild assumptions [8], which shows that ordinal patterns can capture the complexity of time series. For generalizations of the ordinal methodology to higher dimensional maps, the interested reader is referred to [2].

Next we introduce the conceptual and notational setting of the ordinal methodology. Let $(r_0, r_1, ..., r_{L-1})$ be a permutation of $\{0, 1, ..., L-1\}$, $L \geq 2$. We say that the block x_n^L of $\mathbf{X}(\mathbf{x}_0)$ defines the *ordinal pattern of length L* (or simply ordinal L-pattern) $\mathbf{r} = (r_0, r_1, ..., r_{L-1})$ if

$$(A.1) \qquad\qquad x_{n+r_0} < x_{n+r_1} < \ldots < x_{n+r_{L-1}}$$

(other rules can be found in the literature). We use the notation $\mathbf{r} = \mathrm{rank}(x_n^L)$ or $\mathbf{r}_n = \mathrm{rank}(x_n^L)$ to associate ordinal patterns and blocks. Since $\mathbf{r}$ is a permutation of

$\{0, 1, ..., L-1\}$, sometimes one also speaks of permutations instead of ordinal patterns and writes $\mathbf{r} \in \mathcal{S}_L$, where $\mathcal{S}_L$ denotes the set (actually a group) of such permutations. Therefore, the number of ordinal L-patterns is $|\mathcal{S}_L| = L!$.

For example, if $L = 5$ and

$$(A.2) \qquad x_n = 0.4, \ x_{n+1} = -0.5, \ x_{n+2} = 1.5, \ x_{n+3} = -0.8, \ x_{n+4} = 1.4,$$

then $\text{rank}(x_n^5) = (3, 1, 0, 4, 2)$. In case of ties, several conventions can be adopted. If the number of ties in $(x_n)_{n \geq 0}$ is small (as we will suppose here), it can be agreed that the earlier entry is the smaller, or a small amplitude noise can be added to the signal to eliminate the ties [28].

More generally, the *ordinal representation* $\mathbf{R}(L) = (\mathbf{r}_n)_{n \geq 0}$ of a time series $\mathbf{X}(\mathbf{x}_0) = (x_n)_{n \geq 0}$ is obtained by rank-ordering a sliding window of size L along $\mathbf{X}(\mathbf{x}_0)$, i.e., $\mathbf{r}_n = \text{rank}(x_n^L)$. Therefore, ordinal representations are finite-state random processes because they are obtained by discretizing continuous-valued times series. In practice, $\mathbf{X}(\mathbf{x}_0)$ is also randomized by observational noise.

Also in practice, time series have a finite length N. In this case, the ordinal representation of $(x_n)_{0 \leq n \leq N-1}$ of parameter $L \leq N$ is $(\mathbf{r}_n)_{0 \leq n \leq N-L}$, a symbolic representation of length $N - L + 1$. Furthermore, it may be convenient not to consider the complete sequence $(x_n)_{0 \leq n \leq N-1}$ but only an equally spaced subsequence $(x_{m\tau})_{0 \leq m \leq M-1}$, where $\tau \geq 1$ is called a *delay time* (or lag) and $M = \left\lfloor \frac{N-1}{\tau} + 1 \right\rfloor$. Therefore, an ordinal representation has in general two parameters: the length $L \geq 2$ of the ordinal patterns (sometimes called the *embedding dimension* of the representation) and the delay time τ ($= 1$ unless otherwise stated). Delay times greater than 1 are used, for example, when the sampling frequency of an analog signal is high compared to the inverse of the typical time constant of the sampled signal. For methods to select the parameters L and τ and the subtleties involved, see e.g. [28].

A.2. **Entropies and entropy-like quantities based on ordinal patterns.** From the toolbox of Information Theory, we will resort to the perhaps simplest, uni- and two-variate quantities, namely, entropy and mutual information.

A.2.1. *Permutation entropy.* Suppose that $\mathbf{X}(\mathbf{x}_0) = (x_n)_{n \geq 0}$ are scalar observations of a chaotic dynamical system $(\Omega, \mathcal{B}, \mu, f)$, where Ω is a minimal attractor and, hence, f is ergodic on Ω. Then the probabilities $p(\mathbf{r})$ of the ordinal L-patterns generated by $\mathbf{X}$ (i.e., the probability that $\text{rank}(x_n^L) = \mathbf{r}$ for any $n \geq 0$) can be estimated by the empirical probabilities $\hat{p}_N(\mathbf{r})$, that is, the normalized count of sliding windows x_n^L with ordinal pattern $\mathbf{r} \in \mathcal{S}_L$. For a time series of length $N \gg L$ we have

$$(A.3) \qquad \hat{p}_N(\mathbf{r}) = \frac{\#\{n : \text{rank}(x_n^L) = \mathbf{r}, \ 0 \leq n \leq N - L\}}{N - L + 1}.$$

The *permutation entropy of order L* of a process $\mathbf{X}$, whether random or deterministic, is the Shannon entropy of the probability distribution of ordinal L-patterns, $\{p(\mathbf{r}) : \mathbf{r} \in \mathcal{S}_L\}$, generated by $\mathbf{X}$, i.e.,

$$(A.4) \qquad \texttt{X\_PE(L)} = -\sum_{\mathbf{r} \in \mathcal{S}_L} p(\mathbf{r}) \ln p(\mathbf{r}).$$

By convention, $0 \cdot \ln 0 := \lim_{x \to 0+} x \ln x = 0$. We use natural logarithms (entropy in nats) throughout.

The maximum of the Shannon entropy is achieved for flat distributions [9]. Therefore, the maximum of the permutation entropy of order L occurs when $p(\mathbf{r}) = 1/L!$ for all $\mathbf{r} \in \mathcal{S}_L$, which implies $\texttt{X\_PE(L)} \leq \ln L!$. This bound prompts to define the *normalized permutation entropy* of order L,

$$(A.5) \qquad \texttt{X\_NPE(L)} = \frac{\texttt{X\_PE(L)}}{\ln L!} = -\frac{1}{\ln L!} \sum_{\mathbf{r} \in \mathcal{S}_L} p(\mathbf{r}) \ln p(\mathbf{r}),$$

so that

$$(A.6) \qquad 0 \leq \texttt{X\_NPE(L)} \leq 1.$$

A.2.2. *Statistical complexity.* Among the entropy-like quantities that we will use in Section 2 for classification is the statistical (permutation) complexity of order L of a continuous-valued process $\mathbf{X}$. Its definition, adapted to our setting and notation, comprises two ingredients: the normalized permutation entropy $\texttt{X\_NPE(L)}$, Equation (A.5), and the *desequilibrium* $Q(L)$, which we define next.

Let

$$(A.7) \qquad D_{JS}(L) = \texttt{U\_PE(L)} - \frac{1}{2}\texttt{X\_PE(L)} - \frac{1}{2}\ln L!$$

be the *Jensen-Shannon divergence* from the probability distribution $\{p(\mathbf{r}) : \mathbf{r} \in \mathcal{S}_L\}$ of the ordinal L-patterns of the process $\mathbf{X}$ to the probability distribution

$$(A.8) \qquad \{p(\mathbf{u}) : \mathbf{u} \in \mathcal{S}_L\} = \left\{ \frac{1}{2}\left(p(\mathbf{r}) + \frac{1}{L!} \right) : \mathbf{r} \in \mathcal{S}_L \right\}$$

of the ordinal L-patterns of an auxiliary process $\mathbf{U}$. $D_{JS}(L)$ is a symmetrized version of the Kullback-Leibler divergence [9]; its square root is a distance, in our case between the probability distributions $\{p(\mathbf{r}) : \mathbf{r} \in \mathcal{S}_L\}$ and $\{p(\mathbf{u}) : \mathbf{u} \in \mathcal{S}_L\}$. Then [22]

$$(A.9) \qquad Q(L) = Q_0 D_{JS}(L)$$

is the *desequilibrium of order L* of $\mathbf{X}$, where Q_0 is the normalization constant

$$(A.10) \qquad Q_0 = -2\left[\frac{L!+1}{L!}\ln(L!+1) - 2\ln(2L!) + \ln L! \right]^{-1}.$$

Finally, the *statistical complexity of order L* of $\mathbf{X}$ is defined as

$$(A.11) \qquad \texttt{X\_SC(L)} = Q(L) \cdot (\texttt{X\_NPE(L)}).$$

A.3. **Entropies and entropy-like quantities based on transcripts.** As mentioned in Section A.1, ordinal representations have an additional property that can also be exploited in time series analysis, namely: their symbols (permutations) belong to an algebraic group. Indeed, $\mathcal{S}_L$ is a group under the composition (or "product") of permutations, called the symmetric group of degree L. Precisely, transcripts is perhaps the simplest way to capitalize on the algebraic structure of $\mathcal{S}_L$.

Thus, consider two ordinal representations $(\mathbf{r}_n)_{n \geq 0}$ and $(\mathbf{s}_n)_{n \geq 0}$ of the time series $\mathbf{X}(\mathbf{x}_0) = (x_n)_{n \geq 0}$ and $\mathbf{Y}(\mathbf{y}_0) = (y_n)_{n \geq 0}$, respectively, with the same embedding dimension $L \geq 2$. We say that the symbolic sequence $(\mathbf{t}_n^\Lambda)_{n \geq 0}$, where $\Lambda \in \mathbb{Z}$ and $\mathbf{t}_n^\Lambda \in \mathcal{S}_L$ for all $n \geq 0$, is a transcript representation from the process $\mathbf{X}(\mathbf{x}_0)$ to the process $\mathbf{Y}(\mathbf{y}_0)$ with *coupling delay* Λ if

$$(A.12) \qquad \mathbf{t}_n^\Lambda = \mathbf{s}_{n+\Lambda} \circ \mathbf{r}_n^{-1},$$

where $\mathbf{r}_n^{-1}$ is the inverse of $\mathbf{r}_n$ in the group $\mathcal{S}_L$ [24, 1]. Note that $\mathcal{S}_L$ is non-commutative for $L \geq 3$, so the order of the factors in (A.12) matters. To distinguish the particular case $\mathbf{X}(\mathbf{x}_0) = \mathbf{Y}(\mathbf{y}_0)$ from the general case $\mathbf{X}(\mathbf{x}_0) \neq \mathbf{Y}(\mathbf{y}_0)$, one speaks of self-transcripts and cross-transcripts, respectively.

The algebraic properties of transcripts and their generalization to more than two time series have been studied in [25, 3].

A.3.1. *Transcript entropy.* The *transcript entropy of order L and coupling delay Λ* from the process $\mathbf{X}$ to the process $\mathbf{Y}$ is the Shannon entropy of the time series $(\mathbf{t}_n^{\Lambda})_{n \geq 0}$ (see Equation (A.12)), i.e.,

$$(A.13) \qquad \texttt{XY\_TE(L,}\Lambda\texttt{)} = -\sum_{\mathbf{t} \in \mathcal{S}_L} p(\mathbf{t}^{\Lambda}) \ln p(\mathbf{t}^{\Lambda}) = \texttt{XY}^{\Lambda}\texttt{\_TE(L)},$$

where $\mathbf{Y}^{\Lambda} = (y_{n+\Lambda})_{n \geq [-\Lambda]_+}$ (i.e., $n \geq 0$ if $\Lambda \geq 0$ and $n \geq -\Lambda$ if $\Lambda < 0$). Here the probabilities $p(\mathbf{t}^{\Lambda})$ are estimated by the frequencies

$$\hat{p}_N(\mathbf{t}^{\Lambda}) = \frac{\#\{n : \mathrm{rank}(x_n^L) = \mathbf{r} \ \& \ \mathrm{rank}(y_{n+\Lambda}^L) = \mathbf{t}^{\Lambda} \circ \mathbf{r}, \ [-\Lambda]_+ \leq n \leq N - L\}}{N - L + 1}$$

for any $\mathbf{r} \in \mathcal{S}_L$, since then and only then $\mathrm{rank}(y_{n+\Lambda}^L) \circ (\mathrm{rank}(x_n^L))^{-1} = \mathbf{t}^{\Lambda} \circ \mathbf{r} \circ \mathbf{r}^{-1} = \mathbf{t}^{\Lambda}$. Moreover, $N \gg L$ and $[-\Lambda]_+ \ll N - L$ for statistical significance.

If $\mathbf{Y}(\mathbf{y}_0) = \mathbf{X}(\mathbf{x}_0)$ and $\Lambda \neq 0$, then we speak of self-transcript entropy,

$$(A.14) \qquad \qquad \texttt{X\_STE(L,}\Lambda\texttt{)},$$

to distinguish this special case from the general "cross-transcript" entropy (A.13) with $\mathbf{Y}(\mathbf{y}_0) \neq \mathbf{X}(\mathbf{x}_0)$. If the data $\mathbf{X}(\mathbf{x}_0)$ and $\mathbf{Y}(\mathbf{y}_0)$ are clear from the context, we shorten the notations (A.13) and (A.14) to $\texttt{TE(L,}\Lambda\texttt{)}$ and $\texttt{STE(L,}\Lambda\texttt{)}$, respectively. If $\Lambda = 0$, the parameter Λ is dropped.

A.3.2. *Permutation and transcript mutual information.* The *permutation mutual information* of order L and coupling delay Λ between the processes $\mathbf{X}$ and $\mathbf{Y}$ is defined as

$$(A.15) \qquad \texttt{XY\_MI(L,}\Lambda\texttt{)} = \texttt{X\_PE(L)} + \texttt{Y\_PE(L)} - \texttt{XY}^{\Lambda}\texttt{\_PE(L)} \geq 0,$$

where $\texttt{XY}^{\Lambda}\texttt{\_PE(L)}$ is the permutation entropy of order L of the joint probability distribution $\{p(\mathbf{r}, \mathbf{s}^{\Lambda}) : \mathbf{r}, \mathbf{s}^{\Lambda} \in \mathcal{S}_L\}$, where $p(\mathbf{r}, \mathbf{s}^{\Lambda})$ is the probability that $\mathrm{rank}(x_n^L) = \mathbf{r}$ and $\mathrm{rank}(y_{n+\Lambda}^L) = \mathbf{s}^{\Lambda}$ for any $n \geq 0$.

Finally, we will also use a particular sort of transcript-based mutual information that was proposed in [26] as a causality indicator. We define the *transcript mutual information of order L and coupling delay Λ* from $\mathbf{X}$ to $\mathbf{Y}$ as

$$(A.16) \qquad \texttt{XY\_TMI(L,}\Lambda\texttt{)} = \texttt{XY\_TE(L)} + \texttt{Y}^{\Lambda}\texttt{Y\_TE(L)} - \texttt{XY\_Y}^{\Lambda}\texttt{Y\_TE(L)},$$

where (see Equation (A.13)) $\texttt{XY\_TE(L)}$, $\texttt{Y}^{\Lambda}\texttt{Y\_TE(L)}$ and $\texttt{XY\_Y}^{\Lambda}\texttt{Y\_TE(L)}$ are the Shannon entropies of the transcript probability distributions $\{p(\mathbf{s} \circ \mathbf{r}^{-1}) : \mathbf{r}, \mathbf{s} \in \mathcal{S}_L\}$, $\{p(\mathbf{s} \circ (\mathbf{s}^{\Lambda})^{-1}) : \mathbf{r}, \mathbf{s}^{\Lambda} \in \mathcal{S}_L\}$, and the joint transcript probability distribution $\{p(\mathbf{s} \circ \mathbf{r}^{-1}, \mathbf{s} \circ (\mathbf{s}^{\Lambda})^{-1}) : \mathbf{r}, \mathbf{s}, \mathbf{s}^{\Lambda} \in \mathcal{S}_L\}$, respectively.

While the permutation mutual information $\texttt{XY\_MI(L,}\Lambda\texttt{)}$ is symmetric under the exchange of $\mathbf{X}$ and $\mathbf{Y}$ (hence, useless for detecting the information direction in coupled processes), for the transcript mutual information $\texttt{XY\_TMI(L,}\Lambda\texttt{)} \neq \texttt{YX\_TMI(L,}\Lambda\texttt{)}$. If the processes $\mathbf{X}$ and $\mathbf{Y}$ are clear from the context, we shorten the previous notation

to $\mathrm{MI(L,\Lambda)}$ and $\mathrm{TMI(L,\Lambda)}$. The transcript mutual information is a two-dimensional approximation of the three dimensional symbolic transfer entropy [39] and it was shown in [4, 5, 6] to be an efficient indicator of information directionality.