Permutation Jensen-Shannon distance: A versatile and fast symbolic tool for complex time-series analysis

Luciano Zunino [1,2,*], Felipe Olivares [3,†], Haroldo V. Ribeiro [4,‡] and Osvaldo A. Rosso [5,§]

[1]*Centro de Investigaciones Ópticas (CONICET La Plata-CIC-UNLP), 1897 Gonnet, La Plata, Argentina*
[2]*Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina*
[3]*Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain*
[4]*Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil*
[5]*Instituto de Física, Universidade Federal de Alagoas, Maceió, Alagoas 57072-970, Brazil*

The main motivation of this paper is to introduce the *permutation Jensen-Shannon distance*, a symbolic tool able to quantify the degree of similarity between two arbitrary time series. This quantifier results from the fusion of two concepts, the Jensen-Shannon divergence and the encoding scheme based on the sequential ordering of the elements in the data series. The versatility and robustness of this ordinal symbolic distance for characterizing and discriminating different dynamics are illustrated through several numerical and experimental applications. Results obtained allow us to be optimistic about its usefulness in the field of complex time-series analysis. Moreover, thanks to its simplicity, low computational cost, wide applicability, and less susceptibility to outliers and artifacts, this ordinal measure can efficiently handle large amounts of data and help to tackle the current big data challenges.

I. INTRODUCTION

At present, investigations related to complex systems are spread along all scientific fields and, typically, time series of measured variables are employed to analyze the dynamical behavior of such systems. Consequently, the identification of possible hidden dynamical structures from time series is crucial to achieve a more reliable comprehension of the mechanisms that govern the system that generates these complex temporal fluctuations. Unveiling the nature of the underlying process from observed data is indispensable for classification, modeling, and forecasting purposes. Different methods have been developed to deal efficiently with this nontrivial task. Without being exhaustive, the interested readers are directed to Refs. [1–10] for a brief survey of some of these methods. It has also been found that symbolic techniques are especially efficient for recognizing distinct features of the phenomenon represented by the sequence of observations of a suitable observable [11–13]. Moreover, this fact is true not only for low-dimensional dynamical systems since symbolic approaches can be successfully applied to time series generated by systems of much higher dimensionality [14]. In particular, during the past few years, the ordinal symbolization introduced by Bandt and Pompe (BP) [15] has proven great potential for extracting useful information about the intrinsic dynamics of complex systems [16–25]. This success can be mainly attributed to the fact that the causal information that stems from the system's dynamical evolution is naturally taken into account in the BP recipe. As stated by Amigó *et al.* [26], "ordinal patterns are not symbols *ad hoc* but they actually encapsulate qualitative information about the temporal structure of the underlying data." Furthermore, this approach does not rely on generating partitions [27,28], resulting in a practical option to symbolize datasets generated by unknown dynamic processes with unknown levels of noise. Due to these useful properties, permutation entropy, the most representative and widely used descriptor from the BP encoding scheme, which estimates the average rate at which new information appears in a time series per observation [29], has been successfully applied in countless applications since its introduction almost 20 years ago (please see Refs. [30,31] for a review of some of these applications).

On the other hand, different methodologies have been previously introduced to measure the degree of similarity between the symbol sequence statistics of two signals. To the best of our knowledge, up to now, there has been no optimal algorithm for quantifying it in practice. Particularly, Yang *et al.* [32] have proposed a method based on rank-order statistics of symbolic sequences and it has been successfully applied to detect temporal structures in human heartbeat time series and to categorize biological signals [33]. However, there exist some discrepancies regarding its properties and practical utility [34,35]. Ouyang *et al.* [36] have defined a dissimilarity measure based on the rank-frequency distribution of ordinal patterns to investigate the dynamical characteristics of epileptic electroencephalogram (EEG) data, while Parlitz

---

*Corresponding author: lucianoz@ciop.unlp.edu.ar
†olivaresfe@gmail.com
‡hvr@dfi.uem.br
§oarosso@gmail.com

045310-1

et al. [37] have identified equivalent dynamics quantifying the similarity of their relative frequencies of occurrence of ordinal patterns through the Hellinger distance. Much more recently, an approach based on the differences between Rényi entropy spectra has been implemented by Xu and Beck [14]. Finally, the *alphabetic Jensen-Shannon divergence*, a relative distance (in a distributional sense) between time series, has been introduced by Mateos *et al.* [38] for detecting dynamical changes. A binary encoding is undertaken as a coarse-grained representation of the signals and then used to estimate the Jensen-Shannon divergence. Discrimination between stochastic and chaotic dynamics and the determination of change points in time series are two successful applications achieved by implementing this scheme.

Following a similar philosophy to that of Refs. [36–38], we shed some light on this issue by introducing the *permutation Jensen-Shannon distance*. The Jensen-Shannon divergence is implemented as a discriminant quantifier between two signals but using the ordinal symbolic representation introduced by BP for estimating it. It is plausible to assume that the distribution of ordinal patterns is a fingerprint of the underlying dynamics and that, in practical contexts, different dynamics are associated with different ordinal patterns probability distributions. Consequently, the Jensen-Shannon divergence between the ordinal mapping of two time series provides an alternative to quantify the degree of similarity between the temporal structures of arbitrary complex systems. Moreover, this tool straightforwardly inherits all the important advantages associated with the BP scheme, namely simplicity, low computational cost, noise robustness, and invariance under scaling of the data. These features make our proposed ordinal distance especially well suited for the analysis of experimental data. Besides, another advantage to highlight is the versatility of the proposed approach since hypothesis tests related to the nature of an arbitrary time series can be easily carried out by estimating its permutation Jensen-Shannon distance to reference time series appropriately generated according to the null model. In the following, we introduce the permutation Jensen-Shannon distance, present numerical analysis and empirical applications, and, finally, summarize the main conclusions achieved in this work. A performance comparison with other approaches has been addressed in the Appendix.

II. PERMUTATION JENSEN-SHANNON DISTANCE

The Jensen-Shannon divergence [39] between the probability distributions $P = \{p_1, \ldots, p_n\}$ and $Q = \{q_1, \ldots, q_n\}$ is a symmetrized version of the Kullback-Leibler divergence $D_{KL}(P, Q) = \sum_{i=1}^{n} p_i \ln(p_i/q_i)$ given by

$$D_{JS}(P, Q) = \frac{1}{2}\left[D_{KL}\left(P, \frac{P+Q}{2}\right) + D_{KL}\left(Q, \frac{P+Q}{2}\right)\right]$$
$$= S\left(\frac{P+Q}{2}\right) - \frac{1}{2}S(P) - \frac{1}{2}S(Q), \quad (1)$$

where S is the Shannon entropy function, i.e., $S(P) = -\sum_{i=1}^{n} p_i \ln p_i$ (as usual, we assume the convention that $0 \ln 0 = 0$ and $\ln(0/0) = 0$). This is a dissimilarity measure, bounded between 0 and $\ln 2$, that can also be interpreted as the entropy of the average distribution minus the average of

the entropies [40]. The minimum value is achieved if and only if the two distributions under comparison are identical, while the maximum value is obtained whenever their supports are disjoint (that is, $p_i q_i = 0$ for $i = 1, \ldots, n$). The entropic definition of the Jensen-Shannon divergence leads to a natural generalization that allows us to weigh differently the compared distributions. Denoting π_P and π_Q ($\pi_P \geqslant 0$, $\pi_Q \geqslant 0$ and $\pi_P + \pi_Q = 1$) the weights associated with the probability distributions P and Q, respectively, this generalization is defined as

$$D_{JS}^{\pi_P, \pi_Q}(P, Q) = S(\pi_P P + \pi_Q Q) - \pi_P S(P) - \pi_Q S(Q). \quad (2)$$

Furthermore, this measure of discernability can also be generalized to quantify the difference between more than two distributions [39]:

$$D_{JS}^{\pi_1, \pi_2, \ldots, \pi_n}(P_1, P_2, \ldots, P_n)$$
$$= S\left(\sum_{i=1}^{n} \pi_i P_i\right) - \sum_{i=1}^{n} \pi_i S(P_i), \quad (3)$$

with P_i the probability distributions and π_i positive weights such that $\sum_{i=1}^{n} \pi_i = 1$. But what is more important for our present purposes is the fact that the square root of the Jensen-Shannon divergence, $[D_{JS}(P, Q)]^{1/2}$, is a true metric for probability distributions. For further mathematical details please see Ref. [41]. In fact, an entire monoparametric family of metrics is obtained for $[D_{JS}(P, Q)]^{\gamma}$ with $\gamma \in (0, 1/2]$ [42]. The Jensen-Shannon divergence has previously allowed us to shed some light on some relevant issues: the segmentation of nonstationary symbolic sequences, such as DNA sequences, into stationary subsequences [43,44]; the definition of a statistical measure of complexity [45]; the quantification of time symmetry breaking [46]; the characterization of network evolution processes [47]; the study of proteins dynamics [48]; and the statistical analysis of language (Ref. [49] and references therein).

The estimation of the Jensen-Shannon distance, i.e., $[D_{JS}(P, Q)]^{1/2}$, as any measure from information theory, requires us first to know the probability distribution of the time series under analysis. We propose to implement this quantifier over the symbolic sequences obtained via the ordinal mapping procedure introduced by BP. This encoding scheme considers the order relations between some equidistant successive values of time series instead of the values themselves. Consequently, it avoids amplitude threshold dependencies that affect other more conventional symbolization recipes based on range partitioning [50]. The probability distribution is then approximated by the histogram of the ordinal patterns and it characterizes the shape of the time-series waveforms in short temporal windows quantitatively. The ordinal approach offers the possibility to find an appropriate symbolic representation from a time series based on the temporal dynamics in a simple and natural way, with only a weak stationary assumption on the underlying process [51].

This procedure can be better illustrated through a simple numerical example. Given the short time series $X = \{4, 1, 6, 5, 10, 7, 2, 8, 9, 3\}$, two parameters, the order of the permutation symbols D ($D \geqslant 2$ with $D \in \mathbb{N}$, length of the ordinal pattern) and the lag τ ($\tau \in \mathbb{N}$, time separation between

elements) should be chosen. Next, the time series is partitioned into subsets of length D with lag τ akin to phase-space reconstruction by means of time-delay embedding. The elements in each new partition (of length D) are replaced by their rank in the subset. This symbolic ordering reflects the visual pattern in this subset. For example, if $D = 3$ and $\tau = 1$, then there are eight three-dimensional vectors of three consecutive data points associated with X. The first one $(x_0, x_1, x_2) = (4, 1, 6)$ is mapped to the ordinal pattern (102). The second three-dimensional vector is $(x_0, x_1, x_2) = (1, 6, 5)$, and (021) will be its related permutation. The procedure continues so on until the last sequence, (8,9,3), is mapped to its corresponding motif, (120). There are a couple of technical issues worth mentioning. First, note that we implement an overlapped scheme for successive ordinal patterns. A comparison between the overlapping versus the nonoverlapping ordinal pattern selection has confirmed that the former approach reduces the variance of motifs frequencies thanks to the larger number of ordinal patterns to count [52]. Second, if two elements in the vector have the same value, then they are ranked according to their temporal order or, alternatively, a small random perturbation is added to break ties. The effect of a significant number of equalities on the BP encoding scheme, and more specifically on the permutation entropy estimated values, has been carefully analyzed in Refs. [53,54].

The probability of each ordinal pattern can then be estimated by simply computing the relative frequencies of the $D!$ possible permutations π_i:

$$p(\pi_i) = \frac{C(\pi_i)}{N - (D-1)\tau}, i = 1, \ldots, D!, \tag{4}$$

with $C(\pi_i)$ the number of occurrences of the ordinal pattern π_i in an arbitrary scalar time series $\{x_t\}_{t=1}^N$ of length N. In such a way, an ordinal pattern probability distribution, $P = \{p(\pi_i), i = 1, \ldots, D!\}$, is obtained. Returning to our numerical example: $p(\pi_1) = p(012) = 1/8$, $p(\pi_2) = p(021) = 1/4$, $p(\pi_3) = p(102) = 3/8$, $p(\pi_4) = p(120) = 1/8$, $p(\pi_5) = p(201) = 0$, and $p(\pi_6) = p(210) = 1/8$. When there is not any temporal dependence between the values in the time series, all possible ordinal patterns appear with the same probability, i.e., $P = \{p(\pi_i) = 1/D!, i = 1, \ldots, D!\}$. Nontrivial dynamics manifest themselves in a nonuniform distribution of the ordinal patterns: Some motifs occur more often than others and this can be interpreted as an underlying dynamical signature. The distribution of ordinal patterns is invariant with respect to order-preserving changes, such as translations (adding a constant), scalings (multiplying by a positive constant), and/or any monotone increasing nonlinear transformations of the original data [55]. Technically speaking, the ordinal pattern probability distribution P is obtained once the order D and the lag τ are fixed. This encoding scheme does not require the optimal reconstruction of the phase space that is necessary for estimating other quantifiers of chaotic signals. Consequently, D and τ are not usually selected following the methodologies often employed in a conventional phase-space reconstruction (e.g., the first zero of the autocorrelation function, the first minimum of the average mutual information, and the false nearest-neighbor algorithm) [56]. Taking into account that there are $D!$ potential permutations for a D-dimensional vector, the condition $N \gg D!$, with N the

length of the time series, must be satisfied in order to obtain a reliable estimation of P [57]. It is clear that more temporal information is incorporated into the ordinal patterns as the order D increases. For practical purposes, BP suggest in their seminal paper to estimate the frequency of ordinal patterns with $3 \leqslant D \leqslant 7$ and lag $\tau = 1$ (consecutive data points). Since the lag τ physically corresponds to multiples of the sampling time of the signal under analysis, a multiscale analysis can be easily accomplished by analyzing the behavior of any statistic of P as a function of this parameter. Particularly, the estimation with lagged data points, i.e., $\tau \geqslant 2$, offers a more complete comprehension of the underlying dynamics in the case of continuous and/or scale-dependent systems [18,19,51,58,59].

The permutation Jensen-Shannon distance (PJSD) is hence obtained by calculating the square root of Eq. (1), with P and Q the ordinal probability distributions associated with the two signals under analysis. It is worth remarking here that the metric property of $[D_{JS}(P, Q)]^{1/2}$ and all the practical advantages of the ordinal coarse-graining are merged into the PJSD. Henceforth, normalized values of the PJSD with respect to its upper bound, i.e., $(\ln 2)^{1/2}$, are estimated. Since it constitutes a measure of distinguishability between two probability distributions, the symbol composition between different sequences can be quantitatively compared through this metric. Obviously, larger values of this quantifier indicate less similarity between the symbolic mapping of the signals and vice versa. We hypothesized that signals coming from the same underlying dynamics would have small estimated values of this ordinal distance (close to but not exactly zero as a consequence of finite-size effects), whereas larger values (significantly different from zero) will be found when the signals have different origins. Actually, in the former instance, PJSD converges asymptotically to zero with the series size, as will be shown shortly.

III. NUMERICAL ANALYSIS

Next, several numerical datasets are analyzed for characterizing the PJSD behavior, as well as for illustrating the versatility and robustness of the proposed symbolic quantifier within controlled frameworks. It is worth advancing that the main purpose in each of these examples is totally different and that these applications seek to demonstrate the potentiality and flexibility of the PJSD for heterogeneous applications.

A. Convergence to zero for signals with the same dynamics

Trying to characterize the behavior of the PJSD when two signals from the same dynamical source are analyzed, we have estimated the ordinal metric as a function of the time-series size N for two independent numerical realizations of a Gaussian white noise. Figure 1 shows the log-log plot of the average PJSD with $D \in \{3, 4, 5, 6\}$ and lag $\tau = 1$ from an ensemble of 100 independent estimations of sizes $N \in \{2^{11}, 2^{12}, \ldots, 2^{20}\}$. From the observed linear decrease in the log-log plot, it can be concluded that the PJSD tends asymptotically to zero following a power-law behavior with N. To be more precise, PJSD $\propto N^{-1/2}$, with a proportionality factor that depends on the order D. Consequently, the deviation from zero for the PJSD estimation in the case of two signals
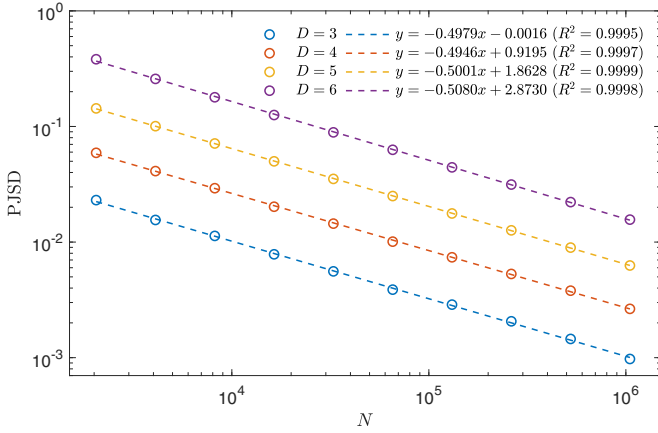
FIG. 1. Log-log plot of the PJSD estimations as a function of the time-series size N for two independent numerical realizations of a Gaussian white noise. Circles indicate the average value from an ensemble of 100 independent estimations with several orders D and lag $\tau = 1$. Dashed lines correspond to the best linear fits to the data. Fitting parameters and R^2 values of these linear regressions are detailed in the legend.
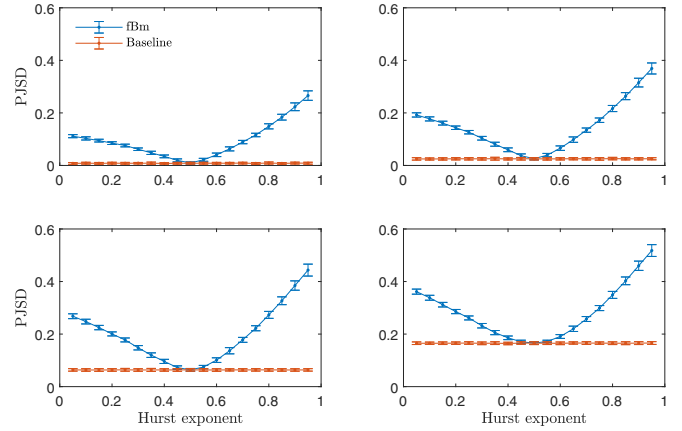

FIG. 2. PJSD estimations with several orders $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and lag $\tau = 1$ are plotted as a function of the Hurst exponent for fBms. Mean and standard deviation (as error bar) from estimations of an ensemble of 100 independent realizations of length $N = 10^4$ data are depicted. PJSD baseline references resulting from the analysis of a pair of shuffled surrogate realizations from each simulation are also included.

with the same dynamics is due to finite-size effects. We have confirmed qualitatively similar behaviors for other stochastic and deterministic dynamics [60].

It has been analytically shown by Grosse *et al.* [44] that, up to a first-order approximation, the expected value of the Jensen-Shannon divergence $D_{JS}(P, Q)$ for two sequences of independent and identically distributed symbols decays inversely proportional to N, i.e., $D_{JS}(P, Q) = C_n/N$, with C_n a constant that depends on n, the number of elements in the probability distributions. Considering that the square root is taken when estimating the PJSD and the fact that $n = D!$ (number of possible permutations) in the BP symbolization scheme, our numerical results are consistent with this analytical approximation. It should also be highlighted here that $D_{JS}(P, Q)$ follows a χ^2 distribution for asymptotically large values of N under the null hypothesis that P and Q are generated from the same probability distribution [41,44]. This result could, in principle, be used for defining a criterion or threshold in order to quantify the statistical significance of estimated values of PJSD.

B. Characterizing long-range correlated time series

It is well known that a wide range of complex systems from diverse fields, including physiology, economy, geophysics, and music, among many others, generates output signals that display different degrees of long-range power-law correlations. Discriminating between these long-memory stochastic dynamics is crucial for several purposes. For example, scaling properties observed for the heart-rate variability are different in healthy and pathological conditions [61]. Fractional Brownian motions (fBms) and their increments, the fractional Gaussian noises (fGns), are ubiquitous models for nonstationary and stationary long-range correlated time series, respectively. These stochastic processes are usually characterized by the Hurst exponent $H \in (0, 1)$ that controls the power-law exponent of their associated spectra. In the

stationary case, i.e., for fGns, the power spectrum is proportional to $1/f^\beta$ with $\beta = 2H - 1$ ($-1 < \beta < 1$), while an analogous generalized power spectrum but with $\beta = 2H + 1$ ($1 < \beta < 3$) can be defined for the fBm nonstationary processes. Antipersistence, absence of correlations, and persistence are obtained for fGns with $H \in (0, 1/2)$, $H = 1/2$, and $H \in (1/2, 1)$, respectively. In fGns with antipersistent behaviors, the series are anticorrelated. That means that positive values are followed by negative values (or vice versa) more frequently than by chance and time series with higher degrees of roughness are observed. On the other hand, long-range correlated series are obtained for persistent fGns where positive values are followed by positive values and negative values are followed by negative values more likely than by chance. Memoryless dynamics are generated if $H = 1/2$: Gaussian white noise for the fGn model and ordinary Brownian motion for the fBm one. Smoother time series with stronger trends are obtained for fBms with larger values of H. Further details about these fractal stochastic models can be found in the seminal paper by Mandelbrot and Van Ness [62].

We have analyzed numerical simulations of fBm and fGn with Hurst exponents $H \in \{0.05, 0.1, \dots, 0.95\}$ generated by implementing the function *wfbm* of MATLAB. One hundred independent realizations of length $N = 10^4$ data were simulated for each Hurst exponent. Mean and standard deviation (included as error bar) of the PJSD, estimated for several orders $D \in \{3, 4, 5, 6\}$ and lag $\tau = 1$, between the simulated records and their shuffled counterparts (white noise and random walk null models for fGn and fBm, respectively [63]) are displayed in Fig. 2 for fBms and in Fig. 3 for fGns. Shuffled or scrambled surrogates are considered as time series of reference since we are trying to quantify the degree of temporal correlations, and these randomized resampled sequences are constrained realizations that satisfy the null hypothesis; i.e., fully random data having exactly the same amplitude distribution as the original time series are generated [64,65].
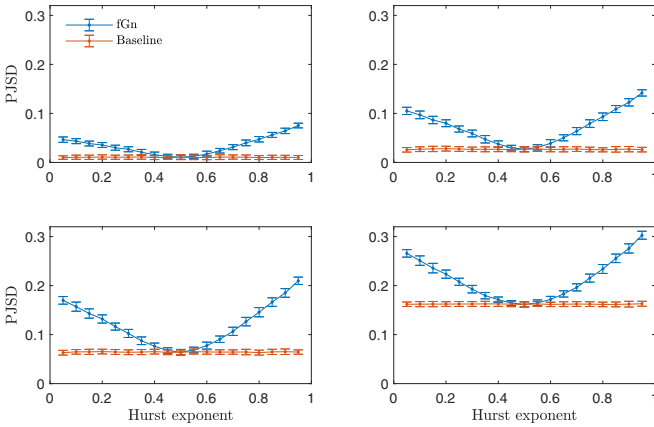
FIG. 3. The same as in Fig. 2 but for fGns.

FIG. 4. Estimated ordinal pattern probabilities for fGns with Hurst exponents $H \in \{0.05, 0.1, \ldots, 0.95\}$ (increasing from blue to yellow). For the sake of a better visualization, only results for $D = 3$ (top plot) and $D = 4$ (bottom plot), and lag $\tau = 1$, are displayed. Mean and standard deviation (as error bar) from estimations of an ensemble of 100 independent realizations of length $N = 10^4$ data are depicted. Ordinal patterns are indexed following the lexicographic order (please see Fig. 2 of Ref. [18]).

Moreover, a baseline is established by estimating the PJSD between two shuffled realizations of each simulation. In such a way, finite-size and amplitude distribution effects are taken into account. The presence of any temporal structure can be robustly concluded only when a significant deviation from this baseline is confirmed. Actually, a clear discrimination of the temporal correlations is observed for both models: The PJSD reaches values larger than the baseline for $H \neq 1/2$. This discrimination is improved for longer simulations with $N = 10^5$ data [60]. The deviations are asymmetric, with larger distances for highly persistent dynamics than those reached for highly antipersistent ones. A similar asymmetric behavior is found when the permutation entropy is implemented for the fGns characterization [60,66]. In relation to the selection of the lag τ, we have confirmed that the correlations discrimination is worse for fGns (especially for $H < 1/2$) and remains qualitatively similar for fBms when larger values ($1 < \tau \leqslant 20$) are used [60]. This latter finding for fBms will be further explored in Sec. III D. Finally, by comparing the results obtained for the two models, it can be concluded that PJSD analysis on the process (fBm) allows us to unveil the presence of the underlying temporal correlations much more precisely. This is in agreement with previous findings [67] and should be taken as a rule of thumb when possible. For a comparison with the results obtained by implementing other dissimilarity measures, please see the Appendix.

C. Distinguishing persistent from antipersistent stochastic dynamics

Results previously obtained for the fGns show that persistent and antipersistent dynamics are indistinguishable when the PJSD to the shuffled surrogate realizations is implemented to characterize the underlying temporal correlations. This ambiguous behavior, also observed when the permutation entropy is used as a discriminative measure [60], could be considered an important weakness, especially for classification tasks. Ordinal probability distributions for fGns with $H > 1/2$ and $H < 1/2$ are clearly different as you can conclude from Fig. 4. However, the problem lies in the fact that their distances to the equiprobable distributions (associated with the uncorrelated dynamics of the shuffled realizations) achieve similar values. Beyond this limitation, it is important
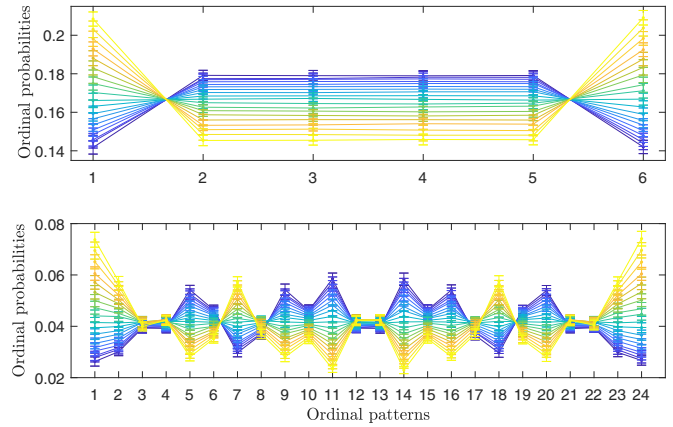
to emphasize that PJSD can discriminate persistence from antipersistence. To demonstrate this, we have estimated the PJSD between two fGns with Hurst exponents H_1 and H_2 in the set $\{0.05, 0.1, \ldots, 0.95\}$, generated by following the same conditions described in the previous section. Figure 5 shows the associated distance matrices with several orders D and $\tau = 1$ for these synthetic fGns of length $N = 10^4$ data. More precisely, mean distance from 100 independent estimations for each pair of Hurst exponents is plotted. As expected, the minimal distances are obtained along the diagonal with values near (but not) zero due to finite-size effects. Furthermore, the distance between two fGns increases as long as their Hurst exponents are more separated from each other. Improved discrimination is achieved for longer time-series sizes [60].
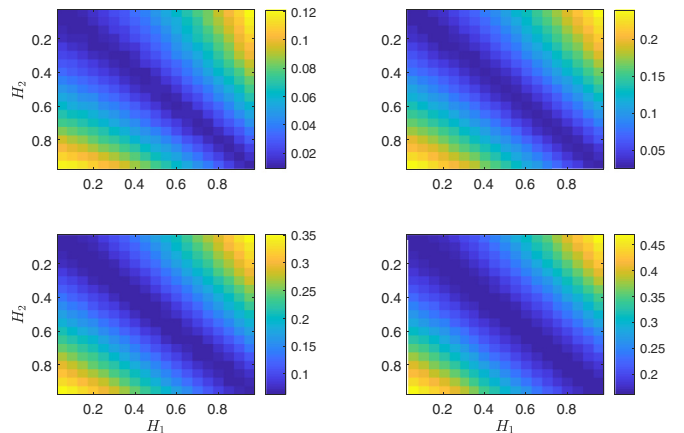

FIG. 5. PJSD matrices for two fGns with Hurst exponents H_1 and H_2 in the set $\{0.05, 0.1, \ldots, 0.95\}$. Mean from estimations of an ensemble of 100 independent realizations of length $N = 10^4$ data for several orders $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and lag $\tau = 1$ are depicted.
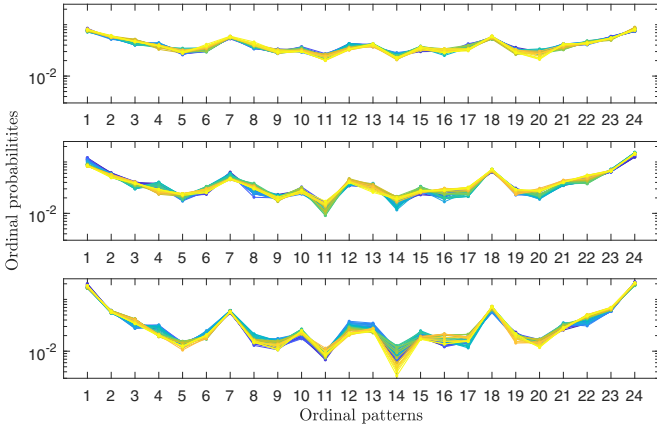
FIG. 6. Estimated ordinal pattern probabilities with $D = 4$ and lags $1 \leqslant \tau \leqslant 40$ (increasing from blue to yellow) for one representative fBm numerical simulation of length $N = 10^4$ data with Hurst exponent $H = 0.25$ (top plot), $H = 0.5$ (center plot), and $H = 0.75$ (bottom plot). Ordinal patterns are indexed following the lexicographic order.

Consequently, we can confirm that the PJSD is actually an efficient tool for distinguishing persistent from antipersistent fGns dynamics and potentially useful for classifying them.

D. Ordinal self-dissimilarity

Our analysis for the fBms developed in Sec. III B shows that results are invariant under changes of the lag τ (at least for small values of this parameter). The self-similar property of these processes analytically justifies this finding: fBms look qualitatively similar independently of the distance from which we look at them (Ref. [68], see in particular page 48 for a more formal definition of self-similarity). Indeed, it has been theoretically shown that the ordinal patterns probabilities do not depend on the lag τ for fBms [69]. Following the definition introduced by Bandt [58], fBms fulfill the *order self-similarity* property. For illustration, we show in Fig. 6 the estimated ordinal pattern probabilities with $D = 4$ and lags τ between 1 and 40 for fBms with three Hurst exponents $H \in \{0.25, 0.5, 0.75\}$. Since numerical realizations of length $N = 10^4$ data are analyzed, the theoretical order self-similarity is only approximately verified and some discrepancies are visually evident when the lag τ increases from 1 (blue curve) to 40 (yellow curve). Keeping this in mind, we propose to estimate the PJSD of an arbitrary time series symbolized with an order D and two different lags τ_1 and τ_2 to quantify the degree of self-dissimilarity of the time series itself at these two temporal scales. By fixing $\tau_1 = 1$ and varying τ_2 from 2 to $\tau_2^{\max}$, this *ordinal self-dissimilarity* seeks to characterize how different, from an ordinal perspective, the time series looks when the scale is increased with respect to the initial scale used to examine it. Smaller estimated values of the ordinal self-dissimilarities, close to the baseline reference, are expected in the case of self-similar dynamics, such as that associated with fBms processes, while larger values will be obtained for systems whose dynamics depend on the observation scale. Moreover, it is reasonable to conjecture that intrinsic characteristic scales of the dynamics that generates
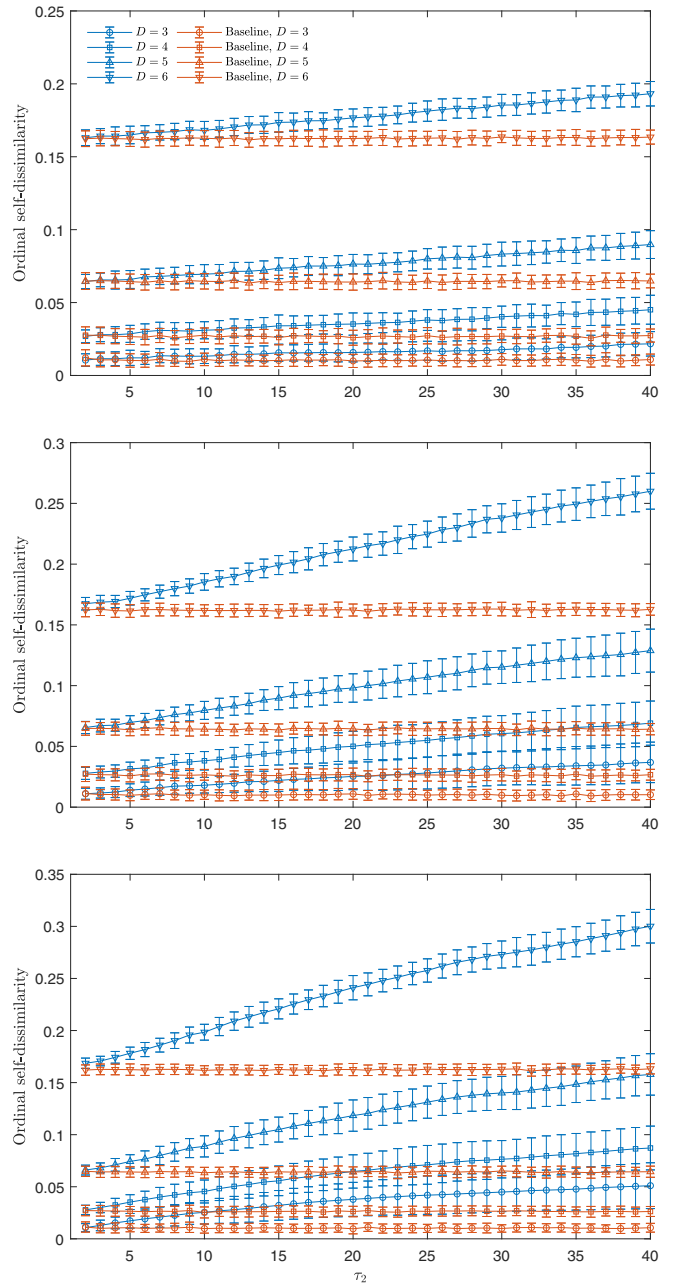

FIG. 7. Ordinal self-dissimilarity as a function of the observation scale τ_2 with $\tau_2^{\max} = 40$ for fBms with Hurst exponents $H = 0.25$ (top plot), $H = 0.5$ (center plot), and $H = 0.75$ (bottom plot). Mean and standard deviation (as error bar) from estimations of an ensemble of 100 independent realizations of length $N = 10^4$ data for several orders $D \in \{3, 4, 5, 6\}$ are displayed. Baseline references are calculated from shuffled surrogates of the original simulations.

the time series, such as temporal delays and/or periodicities, could be revealed through this analysis. With the aim of illustrating this approach, we have tested fBms numerical realizations. Figure 7 shows the ordinal self-dissimilarity with different orders D as a function of the observation scale τ_2 with $\tau_2^{\max} = 40$ for fBm simulations of length $N = 10^4$ data with the same three Hurst exponents previously considered. Once again, average and standard deviation from an ensemble

of 100 independent realizations have been plotted. Baseline references are calculated repeating the procedure for shuffled surrogates of the original simulations. The expected invariance is only verified for small scales. Moreover, the ordinal self-dissimilarity moves further away from the baseline reference for larger Hurst exponent and larger orders D. It is also confirmed that the behavior is qualitatively similar but with smaller values of the distances if longer numerical realizations are considered [60]. Even though ordinal pattern probabilities are *a priori* similar (on average) for different lags τ, the ordinal self-dissimilarity seems to be a sensitive tool able to quantify and magnify their subtle differences.

E. Distances between different complex deterministic dynamics

Within the proposed approach, we conjecture that the PJSD estimated between signals coming from different dynamical origins will achieve values significantly larger than the baseline or reference values obtained when two signals from the same system are compared. The rich dynamical behavior of the logistic map is considered as a testbed to illustrate this fact. Realizations of this one-dimensional quadratic map $x_{t+1} = rx_t(1 - x_t)$ for values of the parameter r between 3.5 and 4 with step $\Delta r = 0.001$ (giving a total of 501 values of r) are mutually compared through the PJSD. It is well known that different deterministic dynamics are generated within this range of the parameter value, from simple periodic to fully chaotic regimes. Average values of the PJSD matrices estimated with several orders D and lag $\tau = 1$ from an ensemble of 20 independent realizations of length $N = 10^4$ data for each pair of values of the parameter r are displayed in Fig. 8. As was expected, minima of the PJSD near zero are observed along the diagonal. It is also concluded that the discrimination between different dynamics improves for larger orders D. This is especially evident in the periodic windows, in agreement with results previously obtained by Parlitz *et al.* [37]. Thus, as a rule of thumb, the largest possible order D that satisfies the condition $N \gg D!$ should be chosen to maximize the accuracy for discriminating between different dynamics. Please see the Appendix to compare with results obtained by implementing other dissimilarity measures.

F. Unveiling noise-induced effects

It is well known that the presence of dynamical noise in nonlinear dynamical systems can lead to several interesting phenomena such as stochastic resonance, noise-induced order, and noise-induced chaos [70]. Within this context, it is critical to appropriately identify the parameters of the nonlinear systems and the noise levels for which these phenomena are realized. To illustrate that the PJSD is useful for this purpose, we have analyzed the paradigmatic noisy logistic map defined by

$$x_{t+1} = rx_t(1 - x_t) + \mu_t, \tag{5}$$

with μ_t a Gaussian random variable with zero mean and standard deviation σ. As it was done in Ref. [70], the σ value is interpreted as the noise level. Realizations of this noisy map for values of the bifurcation parameter r between 3.54 and 3.94 with step $\Delta r = 0.01$ (giving a total of 41

values of r) and noise levels σ between 0.0005 and 0.002 with step $\Delta \sigma = 0.0005$ (giving a total of 40 values of σ) are *ordinally* contrasted against their noise-free counterparts via the PJSD. More precisely, average values of the PJSD with different orders D and lag $\tau = 1$ between 100 independent noisy realizations of length $N = 10^4$ data (for each value of r and σ) and the noise-free realization with the same value of the bifurcation parameter are calculated. Results obtained are displayed in Fig. 9.

It is observed that the PJSD increases abruptly for some particular pair of r and σ values independently of the order D. In other words, noisy and original (noise-free) dynamics are significantly different from an ordinal perspective, and, consequently, noise-induced relevant effects on the original clean dynamics are concluded for such instances. Actually, noise-induced chaos has already been confirmed for some of the bifurcation parameter values our analysis highlights, namely $r = 3.63$, $r = 3.74$, and $r = 3.83$ [70]. Although further research is needed to confirm the chaotic nature of the noisy realizations, we conjecture that transition from periodic to chaotic states due to the presence of dynamical noise is the main source of the observed behavior.

G. Discriminating reversible from irreversible time series

The possibility to discriminate between reversible and irreversible dynamics from the analysis of time series generated by the underlying generating systems is relevant for modeling purposes. The irreversible nature of the time series confirms that Gaussian linear processes and nonlinear static transformations of them should be excluded as potential models, and this allows us to shed some light on the physical mechanisms that govern the system under study. In other words, any difference between the actual (forward) time series and its time-reversed (backward) counterpart is a symptom of the presence of nonlinearities in the process that generated the observed time series. Despite the fact that several methods have been introduced to quantify the degree of time irreversibility in practice, there is not an *a priori* optimal approach (Ref. [71] and references therein). Motivated by this issue, we propose here that the amount of time irreversibility can be measured as the normalized distance (in a distributional sense) between the ordinal probability distributions estimated from the forward and backward series, i.e., $[D_{\mathrm{JS}}(P_{\mathrm{for}}, P_{\mathrm{back}})/\ln 2]^{1/2}$ with P_{for} and P_{back} the time-forward and time-reversed ordinal pattern probability distributions, respectively. It is important to emphasize that the idea of implementing ordinal related tools as a way of assessing the irreversibility of a time series is not new and has been recently explored by Zanin *et al.* [72] and Martínez *et al.* [73] and that we are simply trying to give a step forward in this direction.

In the following, several numerical analyses on theoretically validated reversible and irreversible stochastic and chaotic dynamics have been developed to characterize the performance of the proposed approach. In Fig. 10, we have plotted, in log-log scale, the PJSD between the forward and backward series for simulated fGns with Hurst exponents $H \in \{0.1, 0.2, \dots, 0.9\}$ as a function of the realization length N ($N \in \{2^{10}, 2^{11}, \dots, 2^{20}\}$). Average PJSD estimated with orders $D \in \{3, 4, 5, 6\}$ and lag $\tau = 1$ from an ensemble of 100
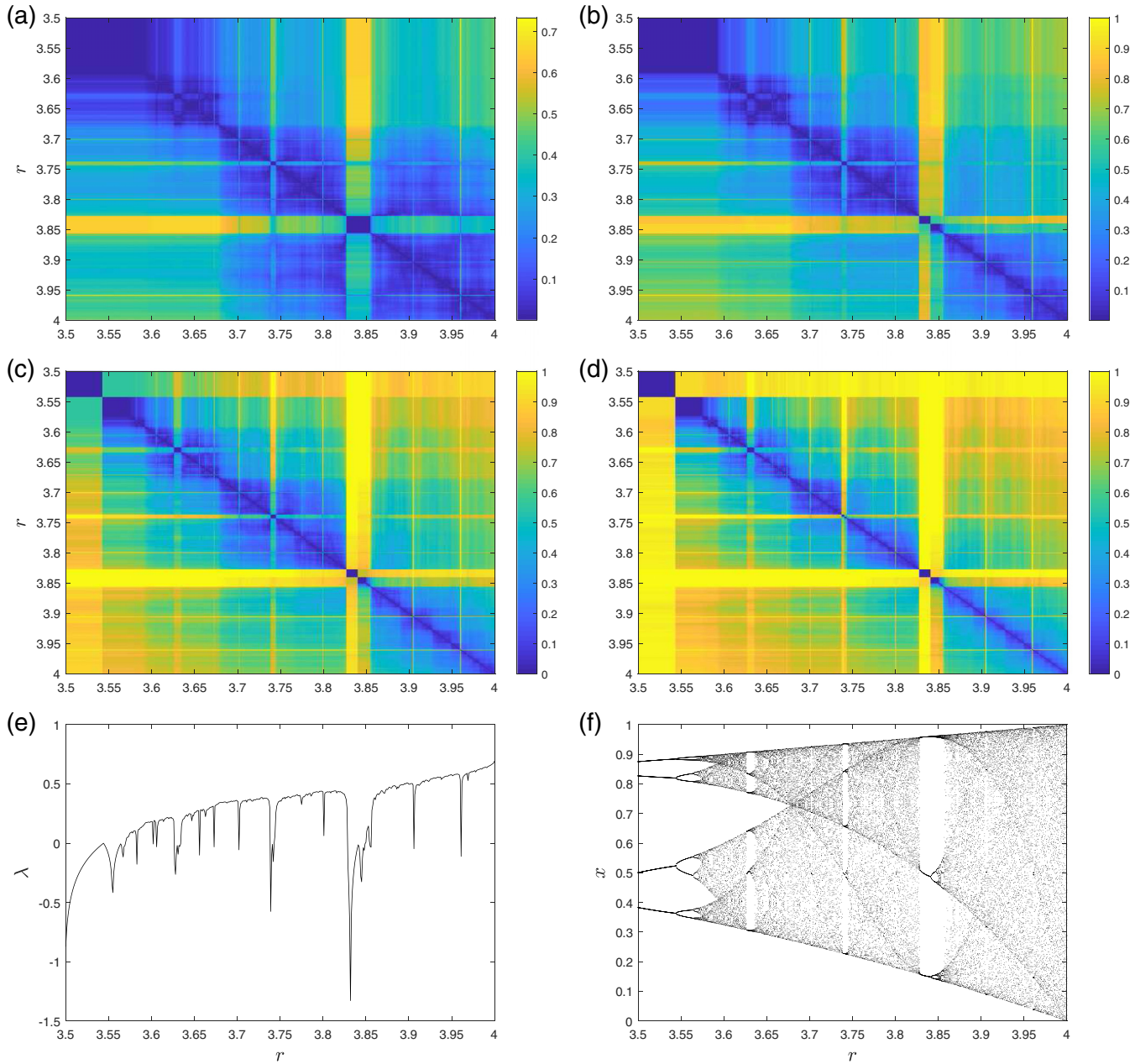
FIG. 8. [(a)–(d)] PJSD matrices estimated with $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and lag $\tau = 1$ for the logistic map with parameter r between 3.5 and 4 and step $\Delta r = 0.001$. Average values from estimations of an ensemble of 20 independent realizations of length $N = 10^4$ data for each pair of values of the parameter r are shown. (e) Lyapunov exponents and (f) bifurcation diagram of the logistic map for this parameter range are also included for comparison purposes.

independent numerical simulations for each value of H are displayed. It is concluded that the PJSD converges asymptotically to zero when N increases with the same power-law behavior PJSD $\propto N^{-1/2}$ observed in Sec. III A. This result is independent of the Hurst exponent H (that increases from the blue to the yellow curves) and with a proportionality factor that depends on the order D. Consequently, the observed deviations from zero are purely related to finite-size effects, and a reversible dynamics is concluded, as expected, for all these linearly correlated stochastic noises. We have found qualitatively equivalent findings for Gaussian and uniform white noises [60].

Next, we consider the irreversibility analysis of two stochastic systems: a static nonlinear transformation of a Gaussian process (commonly known with the acronym STAR) and a linear autoregressive model driven by a non-Gaussian noise (NGRP). More precisely, the STAR is given by $x_t = \tanh^2(y_t)$ where y_t is the first-order autoregressive AR(1) process $y_t = 0.6y_{t-1} + \epsilon_t$ with ϵ_t pseudorandom values from the standard normal distribution while the NGRP is defined by $x_t = 0.3x_{t-1} + \xi_t$ where ξ_t are pseudorandom numbers uniformly distributed in the interval $(-0.5, 0.5)$ [74]. These are examples of stochastic systems with static and dynamic nonlinearities and, consequently, they are time reversible and
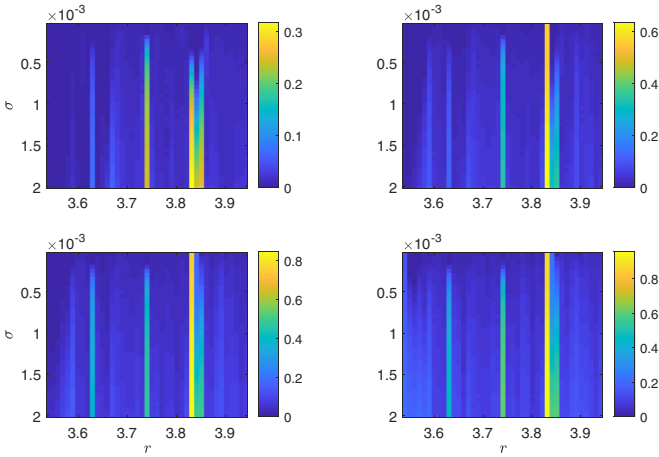
FIG. 9. PJSD matrices estimated with $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and lag $\tau = 1$ for the noisy logistic map with parameter r between 3.54 and 3.94 with step $\Delta r = 0.01$ and noise levels σ between 0.0005 and 0.002 with step $\Delta \sigma = 0.0005$. Average values of PJSD estimations between an ensemble of 100 independent noisy realizations of length $N = 10^4$ data (for each value of r and σ) and their noise-free counterparts are shown. PJSD reaches larger values for all orders D when the bifurcation parameter $r \in \{3.63, 3.74, 3.83, 3.84, 3.85\}$ and the noise level σ increases.

irreversible, respectively [73,74]. We have developed an analysis similar to that previously described for the fGns. The average PJSD between the forward and backward series from an ensemble of 100 independent realizations is depicted in Fig. 11 as a function of the time-series length N in log-log scale. Ordinal distances are estimated with $D \in \{3, 4, 5, 6\}$ and lag $\tau = 1$. Results for the STAR model are displayed in the top plot. The PJSD tends asymptotically to zero, allowing us to conclude a reversible dynamics, as it was expected, for this case. Behaviors of the ordinal approach for the NGRP
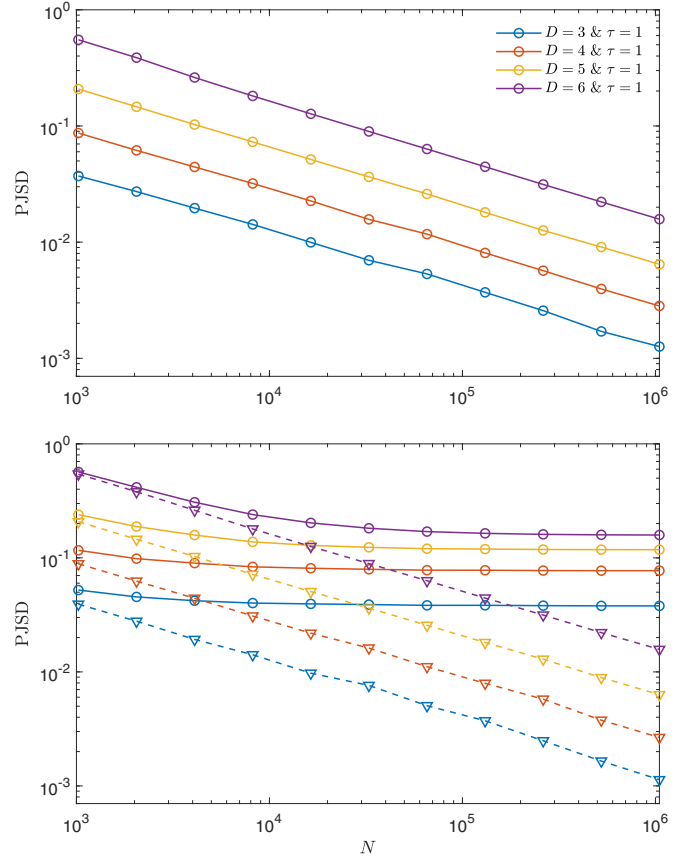

FIG. 11. Irreversibility analysis for the STAR (top plot) and NGRP (bottom plot) models. The PJSD between the forward and backward series estimated with several orders D and $\tau = 1$ is plotted (in log-log scale) as a function of the time-series length N. Average values from 100 independent realizations are displayed. Behaviors displayed by the standard Gaussian AR(1) process (dashed lines with triangle markers) have also been included for comparison purposes in the bottom plot.
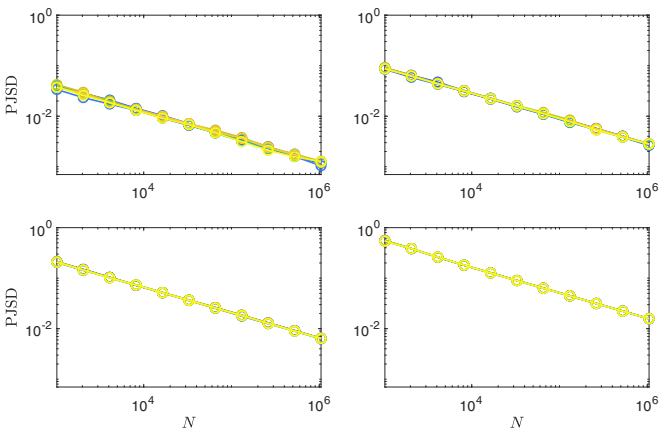

FIG. 10. Irreversibility analysis for fGns with several orders $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and $\tau = 1$. Log-log plots of the average PJSD between the forward and backward series from 100 independent realizations for each Hurst exponent $H \in \{0.1, 0.2, \ldots, 0.9\}$ (increasing from blue to yellow) as a function of the time-series length N. It should be noted that as D increases the dispersion between the nine curves reduces and they are totally overlapped for $D = 5$ and $D = 6$.

model are detailed in the bottom plot. We have also included the results for the standard Gaussian AR(1) counterpart as a reference (please see the dashed lines with triangle markers). An irreversible dynamics is concluded for the NGRP model since the PJSD converges to nonzero values for the different order values. Qualitatively similar results have also been found for another linear system excited by a non-Gaussian noise. It is the third-order autoregressive process $x_t = 0.4x_{t-3} - 0.3x_{t-2} + 0.2x_{t-1} + \zeta_t$ with ζ_t the squaring of a white process with uniform amplitude distribution between -0.5 and 0.5 [74]. For a matter of space, the results for this irreversible system are detailed in the Supplemental Material [60]. We can conclude that the robustness of the proposed ordinal approach to distinguish between reversible and irreversible time series is confirmed in these more challenging stochastic models. Moreover, we have found that static and dynamic nonlinearities can be successfully handled. We have found that other dissimilarity measures have problems to unveil the irreversible nature of the NGRP model. Please see the Appendix for more details.
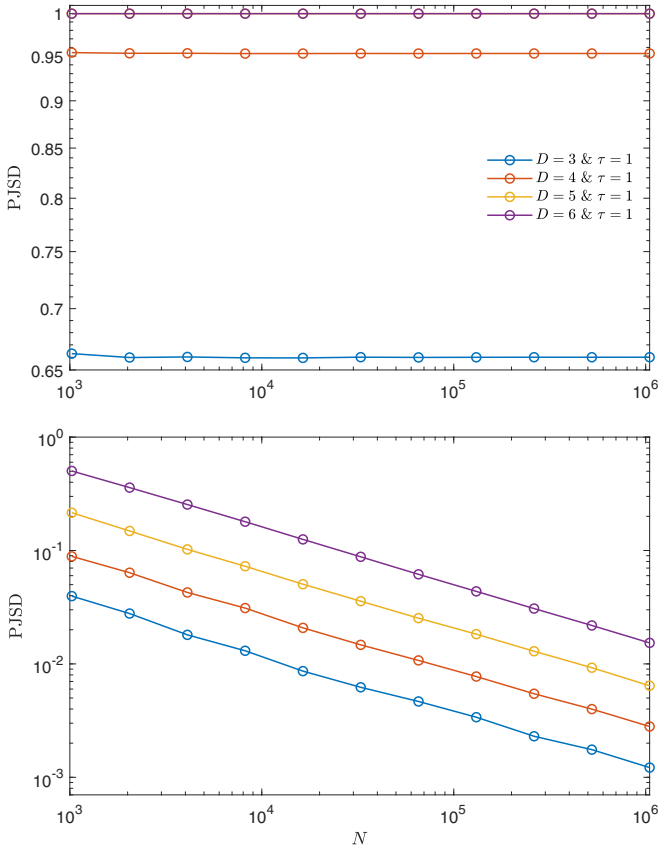
FIG. 12. The same as in Fig. 11 but for the logistic map in the fully chaotic regime ($r = 4$) (top plot) and the x component of the Arnold's cat map with $k = 2$ (bottom plot). The maximal ordinal distance is obtained for the logistic map with $D = 5$ and $D = 6$.

FIG. 13. The same as in Fig. 11 but for the β transform with $\beta = \sqrt{20}$ (top plot), $\beta = \sqrt{200}$ (center plot), and $\beta = \sqrt{2000}$ (bottom plot). The same vertical scale is used in the three plots for easier comparison.

We have also tested the performance of the proposed ordinal approach to discriminate the reversibility or irreversibility of several chaotic maps. More precisely, the logistic map in the fully chaotic regime ($r = 4$), the Hénon map ($x_{t+1} = 1 + y_t - ax_t^2$, $y_{t+1} = bx_t$ with $a = 1.4$ and $b = 0.3$), and the Arnold's cat map [$x_{t+1} = x_t + y_t \mod(1)$, $y_{t+1} = x_t + ky_t \mod(1)$ with $k = 2$] have been analyzed. The temporal evolution of the x component is considered in the two-dimensional maps. Figure 12 shows the results obtained for the logistic (top plot) and the Arnold's cat map (bottom plot). Results for the Hénon map are qualitatively similar to that observed for the logistic map [60]. On the one hand, and taking into account that dissipative chaotic systems are irreversible [71,72], the (practically constant) nonzero estimated values reached by the PJSD for the different orders D in the case of the logistic and Hénon maps confirm irreversibility for them in full agreement with their dissipative nature. On the other hand, since the PJSD asymptotically tends to zero with N following the same power-law behavior previously described for linear systems, a reversible dynamics is concluded for the Arnold's cat map as it was expected for a conservative chaotic system [71]. We have found that the ordinal distance achieves its maximum possible value when the irreversibility of the logistic map is analyzed with the larger orders ($D = 5$ and $D = 6$). This is because the two ordinal pattern probability distributions under comparison have disjoint supports. This means that observed
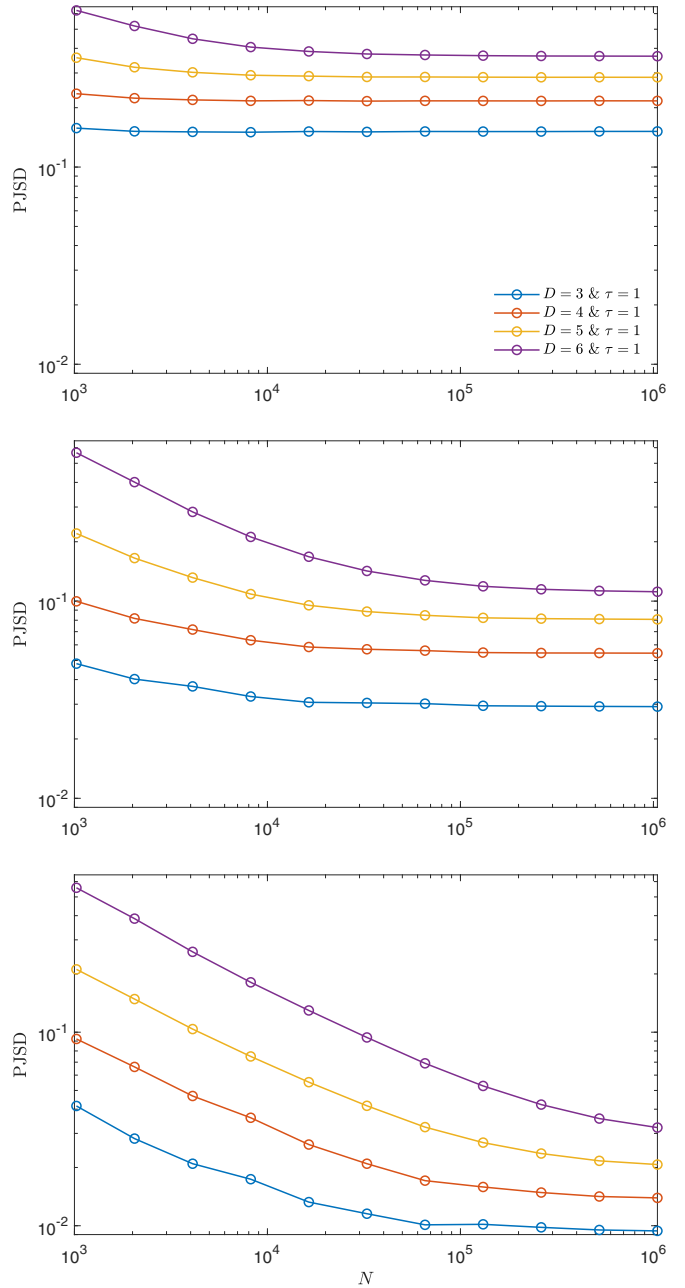
ordinal patterns of the forward time series are unobserved ordinal patterns of the reversed time series and vice versa.

Finally, we have investigated the β transform, defined by the iterated map $x_{t+1} = \beta x_t \mod (1)$, with increasing values of the parameter β. The detection of the irreversibility associated with this map becomes more challenging for larger values of the parameter since the dynamical complexity of this system increases monotonically with β [74]. More precisely, a Kolmogorov-Sinai entropy equal to $\ln \beta$ is theoretically obtained [75]. Figure 13 illustrates the ability of the PJSD approach to unveil the irreversibility of the β transform for

three different values of $\beta \in \{\sqrt{20}, \sqrt{200}, \sqrt{2000}\}$. In the three cases it is observed that the PJSD between the forward and backward series converges to a nonzero value but longer time series are necessary to realize the underlying irreversible nature for larger values of β. Moreover, the ordinal distance tends to lower values as long as β increases. Thus, it is possible to conjecture that the ordinal approach can quantify the irreversibility degree associated with this complex system. Even though further analyses are required to have a rigorous statistical approach, the current results suggest that this strategy can be useful for testing the temporal irreversibility in time series.

IV. EMPIRICAL APPLICATIONS

Some examples of real data applications are detailed below. They are mainly included to demonstrate the significance and applicability of the PJSD into specific practical settings.

A. Ordinal self-dissimilarity for a financial time series

We have estimated the ordinal self-dissimilarity for the historical temporal price evolution of crude oil, one of the most influential commodities in the present world economy. Considering that the ordinary Brownian motion is a widely accepted first approach to model financial time series, we conjecture that the behavior for the oil prices should be analogous to that obtained for the numerical simulations analyzed in Sec. III D. Our main objective is to find some evidence to either support or reject this hypothesis. Daily closing spot prices of the West Texas Intermediate (WTI), often used as a benchmark in oil pricing, from January 2, 1986, to February 16, 2021, have been tested. This time series of $N = 8851$ oil prices, quoted in U.S. dollars per barrel and freely available to download at the U.S. Energy Information Administration (EIA) website, is displayed in Fig. 14 (top plot). At the bottom of this figure, the estimated ordinal self-dissimilarity with different orders D is plotted as a function of the observation scale τ_2 with $\tau_2^{\max} = 40$. Since observed trends are similar to those obtained for the ordinary Brownian motion (please compare with the middle plot of Fig. 7 taking into account that these results correspond to numerical realizations of length $N = 10^4$ data), we can conclude that, from an ordinal self-dissimilarity perspective, this stochastic process seems to be suitable for modeling the financial data under analysis. It is also worth mentioning that our analysis is in accord with recent results obtained by Bandt [58], which showed that the assumption of equality of frequencies of ordinal patterns in the case of the same financial time series is only justified for small values of the lag parameter ($\tau \leqslant 6$).

B. Estimating the Hurst exponent from crude oil prices

Although our previous application demonstrates that the ordinal self-dissimilarity obtained for time series of crude oil prices agrees with that displayed for simulations of an ordinary Brownian motion model, we seek to find further evidence in favor of this random walk behavior. If this was the case, then a memoryless *efficient* dynamics immune to speculative strategies would be associated with the financial time series.
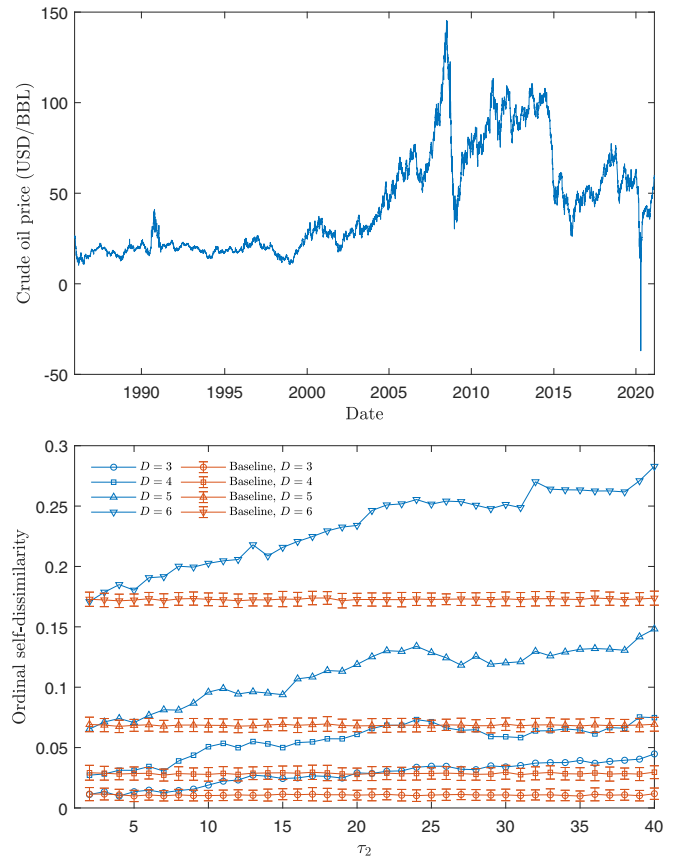

FIG. 14. Time series of the daily closing spot prices of the West Texas Intermediate (top plot) and its estimated ordinal self-dissimilarity as a function of the observation scale τ_2 with $\tau_2^{\max} = 40$ for several orders $D \in \{3, 4, 5, 6\}$ (bottom plot). Baseline references are calculated from shuffled surrogates of the crude oil prices (mean and standard deviation from estimations of an ensemble of 100 independent shuffled realizations are displayed).

With this goal in mind, we propose to quantify the ordinal distance between the WTI daily data and numerical realizations of fBm with the same length ($N = 8851$ data) and different Hurst exponents ($H \in \{0.05, 0.1, \ldots, 0.95\}$). More precisely, the PJSD between the financial data and 100 independent fBm numerical simulations for each of these Hurst exponents is estimated. A minimum of the proposed quantifier is expected for the Hurst exponent which best fits the data within an ordinal framework. Results obtained for order $D = 4$ and different lags $1 \leqslant \tau \leqslant 40$ are shown in Fig. 15. Qualitative similar behaviors are obtained for other orders [60]. Even when certain deviations with the lag τ are observed, we can conclude that within the fBm family, the ordinary Brownian motion, i.e., the fBm with Hurst exponent $H = 1/2$, is the more appropriate model for the ordinal patterns of the WTI daily oil prices. This allows us to suggest the *ordinal efficiency* of the oil prices. We have confirmed that a Hurst exponent $H \approx 1/2$ is also estimated with the widely accepted detrended fluctuation analysis methodology [60,76]. It should be stressed that the success of the methodology described in this section is strongly dependent on the quality of the fBm generator and that spurious results could potentially be achieved if numerical
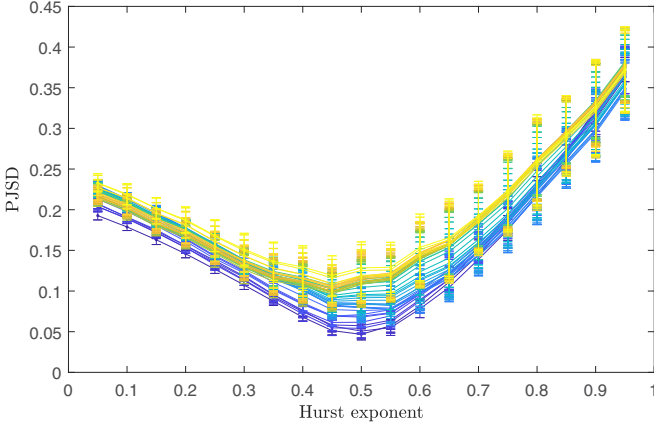
FIG. 15. PJSD between the WTI daily data and fBm numerical simulations of the same length ($N = 8851$ data) with different Hurst exponents ($H \in \{0.05, 0.1, \ldots, 0.95\}$). Mean and standard deviation (as error bar) from estimations with $D = 4$ and different lags $1 \leqslant \tau \leqslant 40$ (increasing from blue to yellow) of an ensemble of 100 independent numerical realizations for each Hurst exponent are shown.

simulations have significant bias. Beyond this limitation, we consider that the proposed ordinal approach can serve as a complementary and validation tool for more traditional Hurst exponent estimators.

C. Parameter adjustment for modeling sunspot number fluctuations

Sunspots are identified as dark spots on the visible surface of the sun, with a mean lifetime of the order of a few days [77]. The number of these spots during a given period of time is a widely accepted proxy for the solar activity and, consequently, modeling its dynamics is of great interest. The best known low-frequency component of the sunspot number time series is its quasiperiodic 11-year cycle. Yet, several attempts have been made to model the interplay between this long-term variation and the presence of high-frequency fluctuations (Ref. [78] and references therein). In particular, Blanter *et al.* [79] have demonstrated that the evolution of the sunspot number can be modelled as a modulated noise, whose long-term component is the well-known 11-year cycle. In a nutshell, the model consists of an autoregressive process of first order, as the high-frequency component, multiplied by a low-frequency function, $x(t) = \eta(t)[\sin(\omega t) + c]$, with $\eta(t + 1) = \alpha \eta(t) + \xi(t)$, where α quantifies the autocorrelation of the noise $\eta(t)$ and $\xi(t)$ is an independent randomly variable uniformly distributed in [0,1]. Nonnegative values are guaranteed by setting $c \geqslant 1$. It is worth mentioning that the autocorrelation parameter α can be related to the mean lifetime τ of sunspots ($\tau = 1/(1 - \alpha)$) [79].

Here we aim to estimate the autocorrelation parameter α for which the model more accurately reproduces the daily evolution of the sunspot numbers. With this goal in mind, we propose to look for the α value that minimizes the PJSD between the modulated-noise model and the daily sampled sunspot number. Following Ref. [79], we consider two different periods of the daily international sunspot number data from the World Data Center SILSO, Royal Observatory of
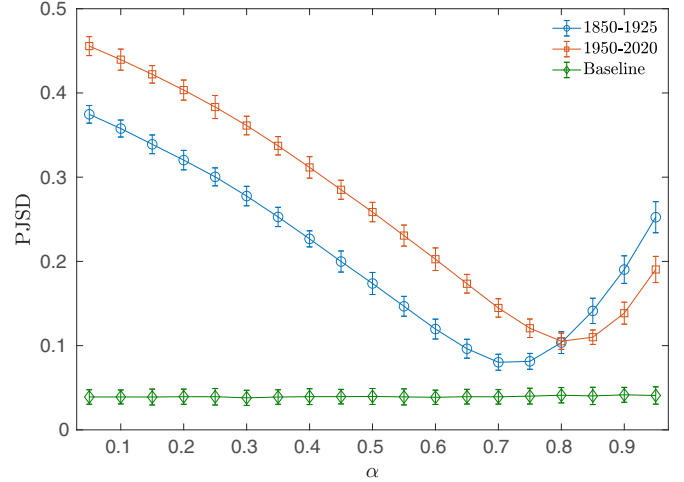

FIG. 16. PJSD with $D = 5$ and $\tau = 1$ between signals from the modulated-noise model (with $c = 1.5$ and $\omega = 2\pi/11$) and the daily sunspot time series, both previously smoothed by an 11-year moving average sliding window, as a function of the autocorrelation parameter α. Mean and three standard deviations (as error bar) calculated from the distance between the empirical sequence and 100 independent realizations of the model are plotted.

Belgium, Brussels [80], spanned between 1850 to 1925 and 1950 to 2020, respectively. We have estimated the PJSD with $D = 5$ and $\tau = 1$ between the sunspot numbers for these two periods and 100 independent realizations of the modulated-noise model with $c = 1.5$ and $\omega = 2\pi/11$ for different values of $\alpha \in \{0.05, 0.1, \ldots, 0.95\}$. Additionally, a baseline was established by estimating the PJSD between two independent realizations of the model. These results are depicted in Fig. 16. A minimum of the PJSD is found for $\alpha = 0.7$ and 0.8 for the first and second periods, respectively. For these values, we can say that the modulated-noise model better fits the real observations within an ordinal framework. Similar results are found for other orders [60]. The change in the autocorrelation of the daily solar activity can be interpreted as an increment on the average lifetime of the sunspots in the period 1925–1950. Indeed, our results are in accordance with those obtained in Ref. [79], where α was estimated from the daily sunspot signal by calculating its Markov radius of correlation for the same two time periods approximately.

D. Multiscale analysis of electroencephalograms from healthy and epileptic patients

Trying to demonstrate the potential of the PJSD in the analysis of neurophysiological data, we have implemented it for discriminating between different brain electrical activity time series. More precisely, five sets of EEG for different groups and recording regions have been considered: surface (scalp) EEG recordings from five healthy volunteers in an awake state with eyes open (Set A) and closed (Set B), intracranial EEG recordings from five epilepsy patients during the seizure free interval from outside (Set C) and from within (Set D) the seizure generating area, and intracranial EEG recordings of epileptic seizures (Set E). One hundred representative EEG segments of length $N = 4097$ data, free
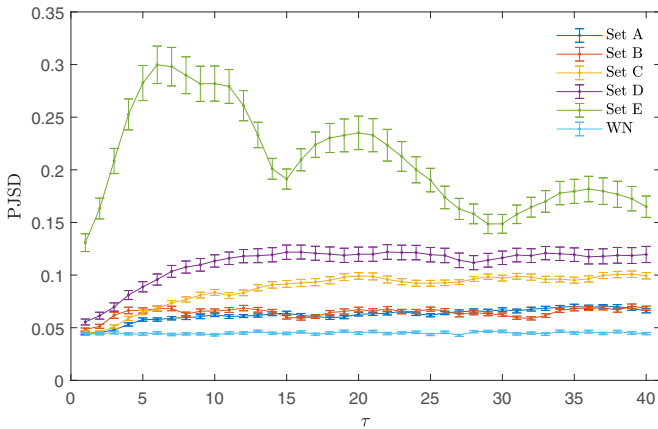
FIG. 17. PJSD with $D = 4$ between the forward and backward EEG records as a function of the lag τ. Mean and standard error (as error bar) calculated from the 100 representative time series for each set are plotted. Results obtained for a set of 100 independent Gaussian white noise realizations of the same length (WN) as linear reference are also displayed. The irreversibility degree and, consequently, the nonlinearity changes as a function of the considered temporal scale for the different EEG sets while the linear reference shows a constant behavior. Actually, it can be visually concluded that the discriminative power between the different datasets increases for larger lag τ.

of artifacts and satisfying a weak stationarity criterion, were appropriately selected for each of these sets. Please see the original paper by Andrzejak *et al.* [81] for further details about this database.

It has been previously shown that the degree of nonlinearity in these sets is different [7,81–83]. Consequently, tools or statistical tests that quantify the presence of nonlinear structures can be particularly suited for the discrimination task. Considering this and also the fact that time irreversibility is a signature of nonlinearity, we propose the PJSD between the forward and backward EEG recordings as discriminatory tool. In this case, due to the limited length of the EEG recordings, we cannot study the behavior of the PJSD with the time-series length N as done in Sec. III G. Instead, we simply compare the estimated ordinal distances for the five EEG groups. Since the discriminative power could depend on the timescale and the optimal temporal scale is *a priori* unknown, a multiscale analysis, i.e., PJSD versus τ, is developed. Moreover, in order to have the reference of a linear process to compare, results obtained for an additional set of 100 independent Gaussian white noise (WN) realizations of the same length ($N = 4097$ data) have also been included. PJSD estimations with $D = 4$ and $1 \leqslant \tau \leqslant 40$ for the six considered sets are displayed in Fig. 17. Qualitatively similar results are obtained for other orders [60].

As expected, coexistence of timescales with different degrees of time irreversibility are concluded for the complex fluctuations of the physiological time series. On the one hand, it is clearly observed that the segments of ictal activity (Set E) have the highest degree of nonlinearity for any lag τ, allowing them to be clearly distinguished from the other brain electrical activities. On the other hand, larger timescales improve the discrimination between the other groups (Sets A–D and WN).

Indeed, for $\tau = 1$, the irreversibility degree estimated from records of Sets A–C and those from the linear WN reference are highly overlapped. The nonlinearity nature of Sets A–D seems to be better unveiled at larger timescales and, hence, a multiscale analysis is crucial to differentiate these EEG time series from the white noise linear realizations. Moreover, higher degrees of irreversibility or nonlinearity are observed for the pathological cases (Sets C and D). As a matter of fact, this increment of time irreversibility with pathology is in perfect agreement with the results previously obtained by implementing different techniques on the same [7,81–83] and other [84,85] epileptic EEG databases. It should also be noted that this finding is totally opposite to the loss of time irreversibility with disease over multiple timescales observed for human heartbeat time series and interpreted as evidence of loss of functionality and adaptability [86]. Contrarily, for the epileptic brain electrical activities analyzed in this section, the time-reversal asymmetry is lower in the healthy condition. It is worth adding that, even when the surface EEG recordings of healthy subjects with eyes open and closed (Sets A and B) achieve the lower indications of nonlinear dynamics, they are clearly distinguished from the WN reference, especially for larger values of the lag τ. Last, EEG data from these two healthy groups show similar time irreversibility behaviors as a function of τ and can be hardly distinguished.

E. Differentiating between real and simulated heartbeat time series

In this last experimental application, we investigated the database from a challenge to generate realistic sequences of interbeat (RR) intervals [87]. More precisely, a collection of 50 (26 physiologic and 24 synthetic) RR interval time series, each between 20 and 24 hours in length, are tested. The real series were obtained from long-term ECG recordings of adults between the ages of 20 and 50 with no evidence of cardiac abnormalities, while each of 12 different generators was implemented to simulate synthetic RR interval time series (two series had been simulated by each generator). The length N of these sequences is variable and it goes from a minimum of 71 953 up to a maximum of 128 507 data. It should also be noted that two of the generators violated the required rule that portions of real (physiologic) RR interval sequences cannot be incorporated in the output. One of them, that generated the series designated as rr36 and rr42, used smoothed averages of real time series while the other one simply time-reversed entire (real) 24-h series (datasets labeled as rr14 and rr16 correspond to this last not-allowed strategy). See Ref. [88] for further information about this open access database.

Our ultimate goal in this analysis is to identify a property of the underlying processes generating the time series that allows us to discriminate between the real and artificial heartbeat signals. According to some previous results, looking for signatures of nonlinear dynamical behavior seems to be an appropriate strategy for achieving robust discrimination of these data [86,89]. Following this hypothesis, once again the degree of time irreversibility is proposed as a potential discriminatory tool. It has also been previously demonstrated the importance of considering the multiple timescales inherent in the healthy cardiac interbeat variability for a more
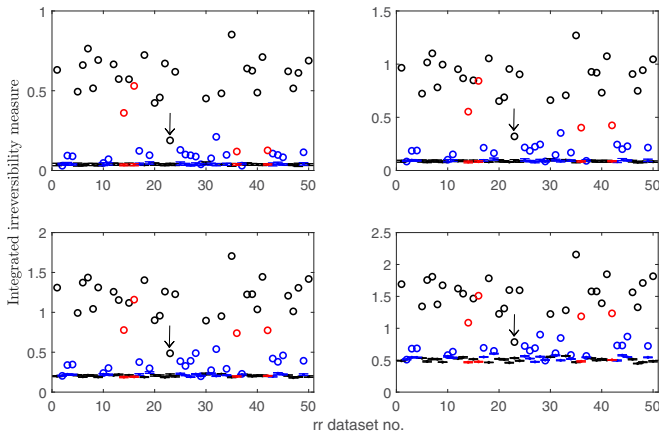
FIG. 18. Integrated irreversibility measure with different orders $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and $\tau_{max} = 10$ for the different sequences of the RR challenge database. Results obtained for the physiologic (black circles), synthetic (blue circles), and the four problematic artificial sequences (red circles) are shown. Mean and standard deviation (as error bar) of estimations for an ensemble of 10 independent shuffled realizations for each RR interval time series are also displayed in order to get a baseline for linear/reversible dynamics with the same length. The arrow indicates the true physiologic dataset rr23 with a low integrated irreversibility degree, similarly to those estimated for simulated datasets.

proper characterization of its intrinsic time-reversal asymmetry [86,89]. Consequently, as in the analysis of the EEG database, a multiscale analysis is recommended, i.e., PJSD between the forward and the backward time heartbeat time series are estimated and contrasted at multiple levels of temporal resolution. After several preliminary analyses, we have further found that the summation of the estimated irreversibility degree, i.e., $\sum_{\tau=1}^{\tau=\tau_{max}} [D_{JS}(P_{for}, P_{back})/\ln 2]^{1/2}$, up to a suitable maximum temporal scale, τ_{max}, for a value of the order D is better for the discrimination purpose. This global quantifier for characterizing the time-reversal asymmetry, which considers a predefined range of scales simultaneously, is in line with the *multiscale asymmetry index* proposed by Costa *et al.* [86,89].

Results obtained for the RR challenge database by estimating this integrated irreversibility measure with several orders D and $\tau_{max} = 10$ are shown in Fig. 18. In order to have a linear reversible reference to compare with, ensembles of 10 independent shuffled realizations for each RR interval time series have also been analyzed. The mean and standard deviation of the 10 shuffled estimations of the proposed global irreversibility quantifier for each tested sequence are also displayed in Fig. 18. The value for τ_{max} has been chosen through visual comparison and it might not be the optimal one in a quantitative context. However, we consider that our qualitative analysis is enough for the illustration of the technique and that the identification of the optimal value for τ_{max} is beyond the scope of the present work. A clear discrimination between real and simulated sequences is concluded from Fig. 18. True physiological records (black circles) have larger estimated values of the global quantifier, confirming the existence of nonlinear temporal structures in their underlying

dynamics. Estimations for the synthetic datasets (blue circles) are significantly lower. Essentially, algorithms included in the challenge cannot reproduce the nonlinearity intrinsically present in the true physiological records. Only the true physiologic dataset rr23 (indicated with an arrow) shows a low irreversibility degree that makes us erroneously assume that it is an artificial record. It is worth noting here that curiously the same misclassification has been observed in the analysis developed in Ref. [89] (see Fig. 4 in this reference). This common misclassification reinforces the hypothesis suggested by the authors of the aforementioned article that unknown factors such as age, level of physical activity, and/or drug effects could be potential reasons for the low level of the integrated reversal asymmetry related to this particular physiologic record. Furthermore, the results obtained for the four problematic synthetic sequences deserve special attention. On the one hand, the two records that are simply time reversal of real data, rr14 and rr16 (the two red circles on the left), are classified as physiologic as it was obviously expected. On the other hand, the other two records, rr36 and rr42 (the two red circles on the right), which are smoothed averages of real time series, show an interesting behavior as a function of the order D. Their integrated irreversibility degree reaches a low value for $D = 3$, similarly to that obtained for synthetic records. However, estimated values for the global quantifier increase as D increases, achieving values similar to that obtained for true RR sequences when $D = 6$. We conjecture that their inherent physiological (irreversible/nonlinear) nature remains partially hidden at the lowest orders but is successfully revealed at the highest one. We consider that all these findings confirm that the proposed multiscale time irreversibility analysis is particularly valuable for testing the quality of algorithms designed to simulate heart rate time series. We are also optimistic about its usefulness for characterizing the potential limitations of generators of other physiological signals.

V. CONCLUSIONS

We have introduced the PJSD, an ordinal metric able to quantify the degree of similarity between two (or more) time series. Through several numerical and experimental analyses, we have confirmed its usefulness for characterizing and classifying time series. Comparisons against other approaches have allowed us to verify the enhanced performance of the proposed tool for some relevant purposes. Taking into account the robustness to noise effects and the invariance under scaling of the data associated with the ordinal symbolization approach, the PJSD seems to be a quantitative metric especially suited for the analysis of real-world signals. We also consider this approach to be versatile because the hypothesis to be tested can be easily set up by conveniently choosing the time series taken as reference. Actually, the diversity of applications included in this paper is explicit evidence of this. Consequently, a wide range of complex phenomena can be scrutinized with this ordinal symbolic quantitative metric. In particular, we conjecture that the PJSD can be potentially useful for quantifying the functional connectivity in complex network analysis; e.g., to reveal specific disease effects from brain networks analysis. Moreover, it can be easily and quickly implemented, paving the way to monitor the behavior of big data series in real time.

It is also possible to hypothesize that multiple scientific fields can be benefited through its implementation. For all these reasons, we consider that the proposed ordinal distance is a useful addition to the repertoire of existing methods for complex signals analysis. Further analytical investigations that support the heuristic results detailed in this work will be undertaken in future research. Exploring possible generalizations using other divergences, such as the Jensen-Tsallis one that allows a more complete characterization in the case of heavy-tailed distributions [49], could be another potential avenue of study. Interested readers are welcome to contact the authors for an implemented code of the PJSD in MATLAB. This method will also be available in ordpy, a Python package for data analysis with ordinal methods [90].

APPENDIX: PERFORMANCE COMPARISON AGAINST OTHER APPROACHES

With the aim of achieving a better characterization regarding the significance of the results obtained with the PJSD, we have developed some comparative analysis with other two quantifiers of similarity or dissimilarity between signals; namely the information-based similarity index (IBSI) [32,33] and the alphabetic Jensen-Shannon divergence (aJSD) [38]. A binary coarse graining depending on the relative amplitudes of successive values is first implemented in both of them and then binary sequences of length m (m-bit words) are obtained. This procedure is known as *alphabetic mapping*. The relative frequencies of the 2^m possible m-bit words are calculated and then used to quantify the dissimilarity between two arbitrary signals. On the one hand, in the case of IBSI, the m-bit words are first sorted and ranked according to their frequencies of occurrence for each signal. Then, an average deviation of the ranks for the common words, with a suitable weighting factor that takes into account the importance (relative frequency) of the word, is proposed for characterizing the dissimilarity between the two symbolic sequences. On the other hand, for the aJSD, the Jensen-Shannon divergence between the probability distributions of the two signals under analysis, estimated following the alphabetic mapping scheme, is proposed as measure of dissimilarity. For further details, see Refs. [32,33] for the former tool and Ref. [38] for the latter one.

As a first numerical test, we compare performances of IBSI, aJSD, and PJSD for discriminating the presence of temporal correlations in long-range correlated time series. The analysis developed is similar to that described in Sec. III B, and, for this reason, the information is not repeated here. Both fractal stochastic models, fBm (nonstationary process)
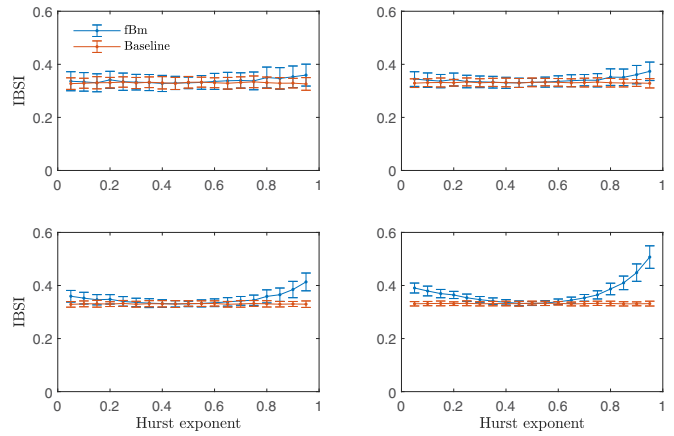

FIG. 19. Estimations of IBSI with several lengths $m \in \{8, 9, 10, 11\}$ (increasing from top left to bottom right) are plotted as a function of the Hurst exponent for fBms. Mean and standard deviation (as error bar) from estimations of an ensemble of 100 independent realizations of length $N = 10^4$ data are depicted. Baseline references resulting from the analysis of a pair of shuffled surrogate realizations from each simulation are also included.

and fGn (stationary noise), with different Hurst exponent H have been analyzed. Word lengths m between 8 and 11 have been chosen since these values were proposed in the original papers. Results obtained for IBSI are detailed in Figs. 19 and 20 while those for aJSD are displayed in Figs. 21 and 22. We have accordingly adjusted the vertical scale of these plots in order to make the comparison easier with the results obtained for PJSD (Figs. 2 and 3). By comparing the behavior and range of the three quantifiers, we conclude that PJSD offers the best characterization for both stochastic models. Performance differences observed in favor of PJSD are clear and evident when comparing against IBSI. Although much less noticeable, PJSD also outperforms aJSD with a larger range of estimated values that points out a higher ability for characterizing long-range memory effects.

Next, we analyze how useful the three measures result for unveiling the presence of slightly correlated dynamics. This is a relevant and well-known issue in several fields, where a fully uncorrelated dynamic is taken as the null hypothesis.
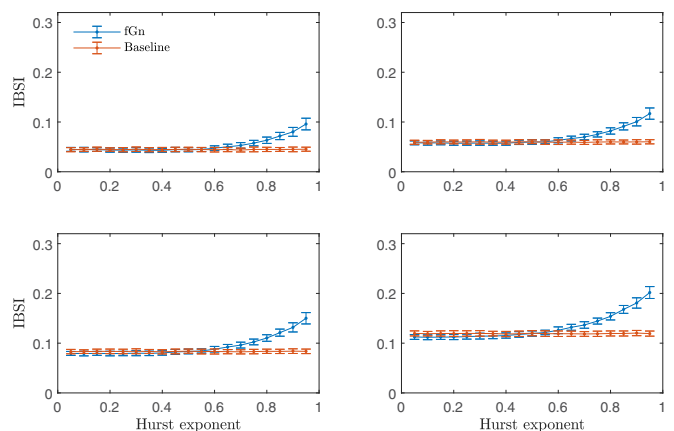

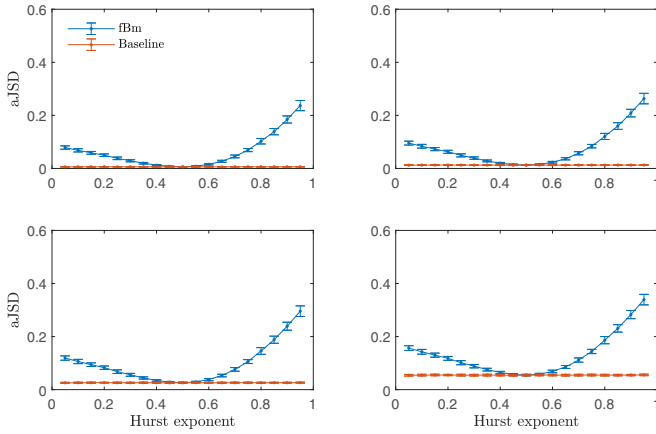FIG. 20. The same as in Fig. 19 but for fGns.

FIG. 21. Estimations of aJSD with several lengths $m \in \{8, 9, 10, 11\}$ (increasing from top left to bottom right) are plotted as a function of the Hurst exponent for fBms. Mean and standard deviation (as error bar) from estimations of an ensemble of 100 independent realizations of length $N = 10^4$ data are depicted. Baseline references resulting from the analysis of a pair of shuffled surrogate realizations from each simulation are also included.

Trying to provide a more rigorous statistical answer to this issue, we have estimated the three quantifiers for ensembles of 10 000 independent realizations of fGns with $H = 0.5$ (fully random dynamics) and $H = 0.6$ (slightly persistent dynamics). Simulations of length $N = 10^4$ data were generated for this numerical test. Taking into consideration that the main objective is to identify the presence of memory effects in the time series, dissimilarities between the original and their shuffled counterparts were calculated. Histograms of estimated values for IBSI, aJSD, and PJSD are depicted in Figs. 23–25, respectively. The two distributions of estimated values are better separated for the PJSD while the overlap increases in the case of IBSI and aJSD. This finding allows us to confirm a higher performance of the PJSD for uncovering the presence of slightly correlated dynamics.

Dissimilarity between the original signal and its shuffled counterpart has been previously proposed as nonrandomness index [32]. Indeed, this is analogous to what we have done when fGns are analyzed. Following this idea, we have found
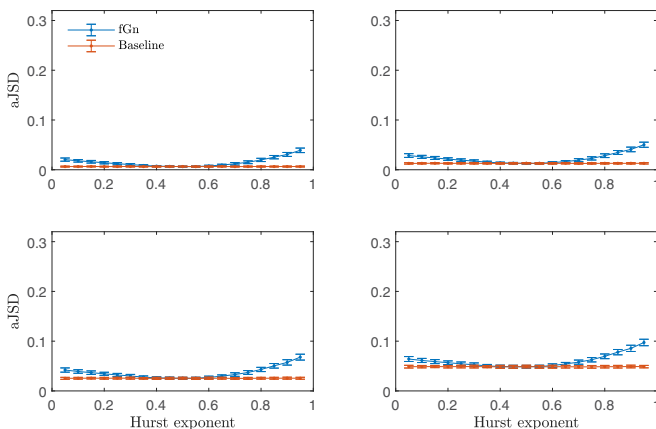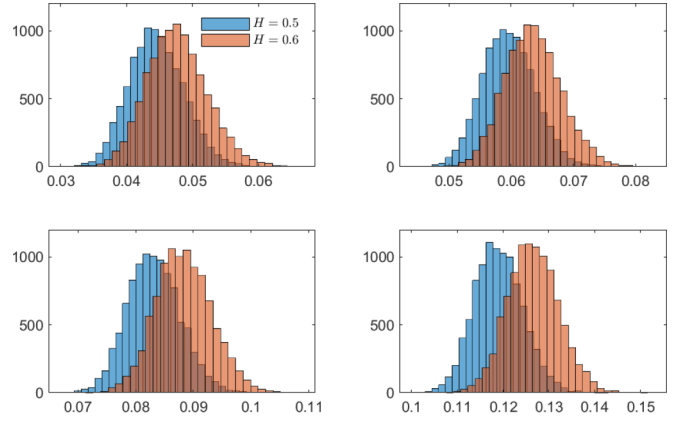

FIG. 23. Distributions of IBSI estimated values for ensembles of 10 000 independent numerical realizations of fGns with $H = 0.5$ (fully uncorrelated dynamics) and $H = 0.6$ (slightly correlated dynamics) with respect to their shuffled counterparts. Results obtained for simulations of length $N = 10^4$ data with word lengths $m \in \{8, 9, 10, 11\}$ (increasing from top left to bottom right) are shown.

that the three dissimilarity measures show monotonic increasing behaviors when they are implemented to quantify the nonrandomness degree of colored noises. Figure 26 shows these behaviors when $1/f^k$ noises with k between 0 and 3 and step $\Delta k = 0.1$ are analyzed. The Fourier filtering method has been implemented for generating the colored noises. We address the reader to Ref. [91] for more details about this algorithm. The mean and standard deviation (displayed as error bars) of the three dissimilarity measures for 100 independent realizations of length $N = 2^{14}$ data for each value of k have been plotted. Results for other values of N ($N \in \{2^{10}, 2^{11}, \ldots, 2^{17}\}$) are qualitatively similar. It is visually clear that IBSI has less accuracy for discriminating colored noises in both extremes of the power-law exponent range ($k \in [0, 1]$ and $k \in [2, 3]$) while aJSD suffers from a similar lack of sensitivity but only when $k \in [0, 1]$. PJSD shows the best discrimination, especially for larger D. Moreover, an ideal quasilinear behavior is observed for $D = 6$ when $k > 0.5$.

Next, the performance of these dissimilarity measures for classifying different colored noises is quantified. For such a purpose, we train a linear support vector machine algorithm

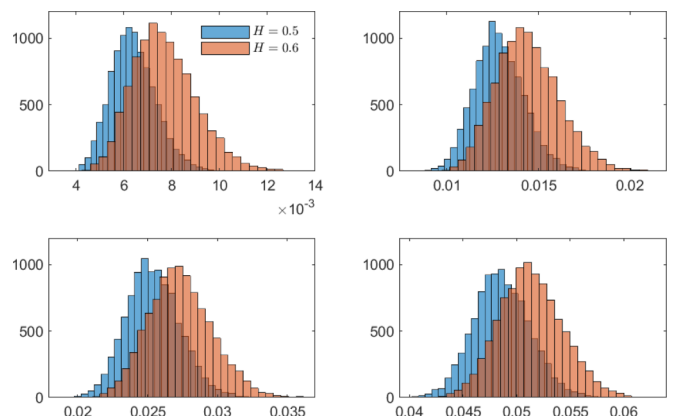FIG. 22. The same as in Fig. 21 but for fGns.

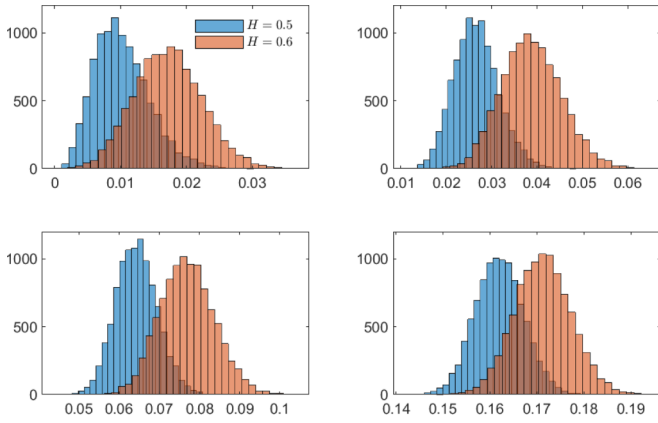FIG. 24. The same as in Fig. 23 but for aJSD.

FIG. 25. The same as in Fig. 23 but for PJSD. In this case, results for orders $D \in \{3, 4, 5, 6\}$ (increasing from top left to bottom right) and lag $\tau = 1$ are displayed.

with a fivefold cross validation considering the 31 values of k ($k \in \{0, 0.1, \ldots, 2.9, 3\}$) as possible classes for the algorithm and the estimated nonrandomness index as single feature. We generate ensembles with 100 simulations for each value of the power-law exponent k and for each length N. Overall accuracies (fractions of correctly classified values of k) as a function of the time-series length N for the three dissimilarity measures are compared in Fig. 27. As it was expected, the overall accuracy for PJSD is enhanced regardless of the time-series length N. Differences observed with IBSI are larger. When comparing with aJSD, differences are lower but always in favor of PJSD. The only exception is when PJSD is estimated with $D = 6$ for $N = 2^{10}$. This can be attributed to finite-size effects since the condition $N \gg D!$ is not satisfied in this case.

The ability to distinguish between different complex deterministic dynamics, reproducing the analysis detailed in Sec. III E, has also been contrasted. We have first ob-
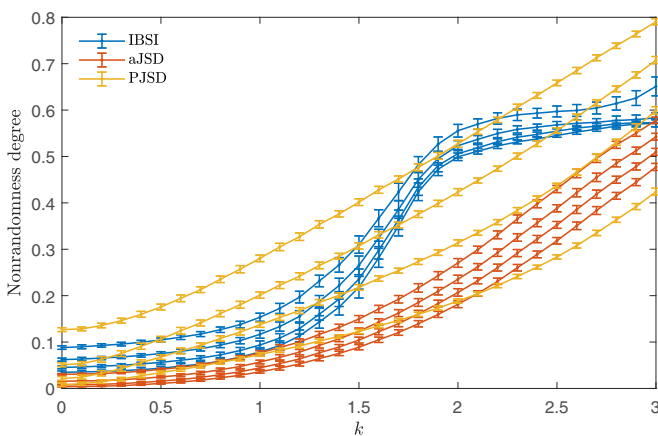

FIG. 26. Nonrandomness degree for colored noises $1/f^k$ with k ranging from 0 to 3 with steps of size 0.1. Mean and standard deviation (as error bars) from estimations of an ensemble of 100 independent realizations of length $N = 2^{14}$ data are displayed. Word lengths $m \in \{8, 9, 10, 11\}$ are considered for the IBSI and aJSD estimations while orders $D \in \{3, 4, 5, 6\}$ with lag $\tau = 1$ are used for the PJSD estimations. The value of m and D increases in the upward direction for each color group.
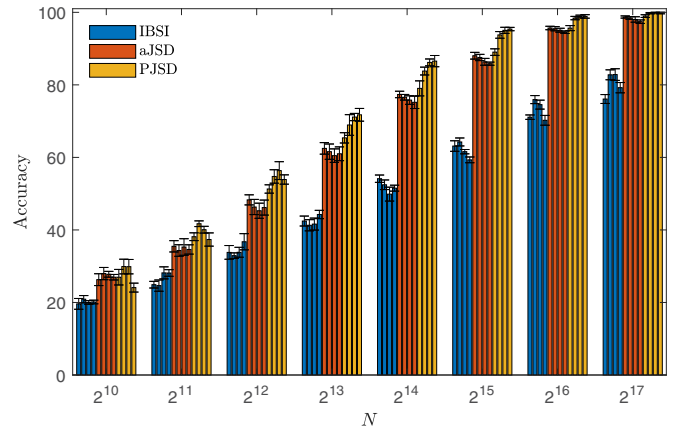

FIG. 27. Overall accuracy as a function of the time-series length N for classifying colored noises. A linear support vector machine algorithm with a fivefold cross validation has been implemented as machine learning classifier. Word lengths $m \in \{8, 9, 10, 11\}$ are considered for the IBSI and aJSD estimations while orders $D \in \{3, 4, 5, 6\}$ with lag $\tau = 1$ are used for the PJSD estimations. The value of m and D increases from left to right for each color group.

served that IBSI has problems to be estimated in several instances in which the dissimilarity between pure periodic and low-dimensional chaotic dynamics is considered. This is essentially due to the fact that the signals associated with these
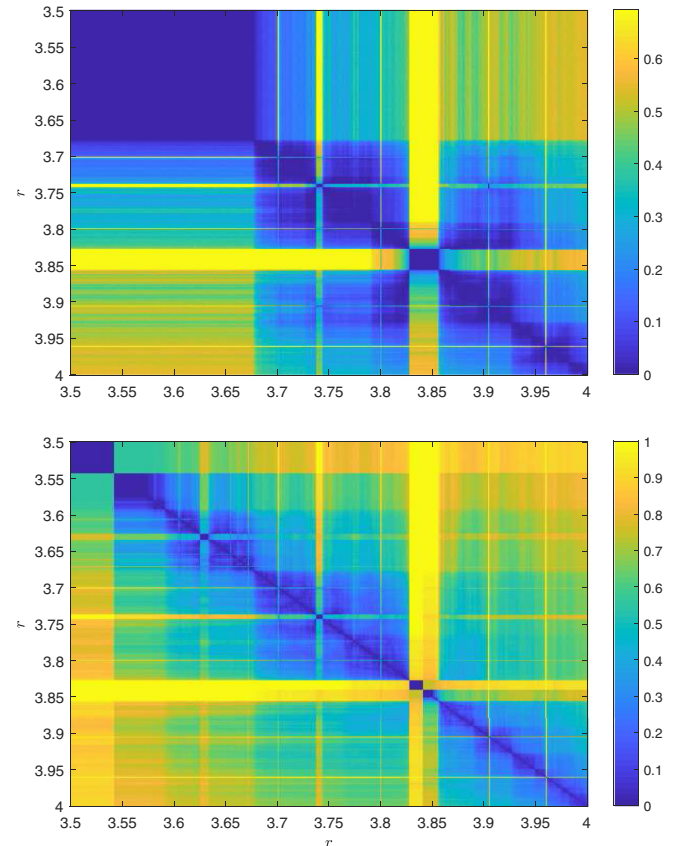

FIG. 28. Matrices distances obtained for the aJSD with $m = 8$ (top plot) and PJSD with $D = 5$ and $\tau = 1$ (bottom plot) for the logistic map with parameter r between 3.5 and 4 and step $\Delta r = 0.001$. Numerical realizations of length $N = 10^4$ data are analyzed.
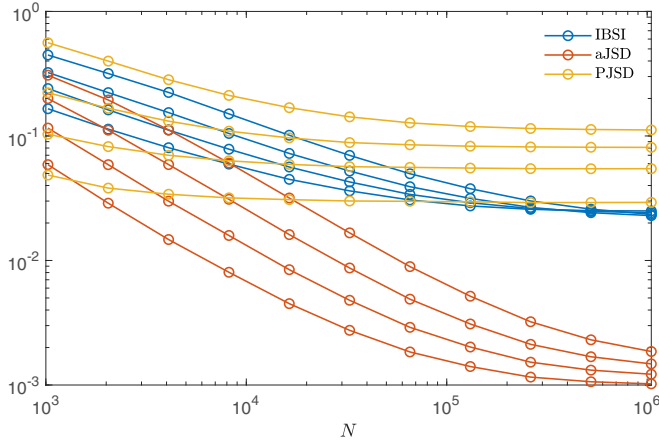
FIG. 29. Irreversibility analysis for the β transform with $\beta = \sqrt{200}$. Dissimilarities between the forward and backward series are plotted (in log-log scale) as a function of the time-series length N. Average values from 100 independent realizations are displayed. Word lengths $m \in \{8, 9, 10, 11\}$ are considered for the IBSI and aJSD estimations while orders $D \in \{3, 4, 5, 6\}$ with lag $\tau = 1$ are used for the PJSD estimations. The value of m and D increases in the upward direction for each color group.

FIG. 30. The same as in Fig. 29 but for the NGRP model.

two regimes have very few common m-bit words, making statistically unreliable the estimation of an average deviation between their associated ranks. Actually, we have found several couples of control parameters of the logistic map for which only one m-bit word is shared between the simulated time series. Consequently, IBSI has been discarded for this particular analysis. The estimated matrices for the aJSD and PJSD in the case of logistic map for 501 equidistant values of the control parameter $r \in [3.5, 4]$, are compared in Fig. 28. For a fairer comparison, we have fixed the values of $m = 8$ and $D = 5$ in order to have a similar number of elements in the probability distributions ($2^8 = 256$ and $5! = 120$). By comparing the two matrices, it is concluded that PJSD achieves a more accurate discrimination of the different deterministic dynamics, especially in the periodic windows. Moreover, expected minima along the diagonal are also better defined in the PJSD matrix distance, evidencing a more robust distinction between nearby values of the control parameters. It is worth mentioning here that we have confirmed that results obtained for aJSD with larger m ($m \in \{9, 10, 11\}$) are very similar to those obtained for $m = 8$ and that they do not show significant
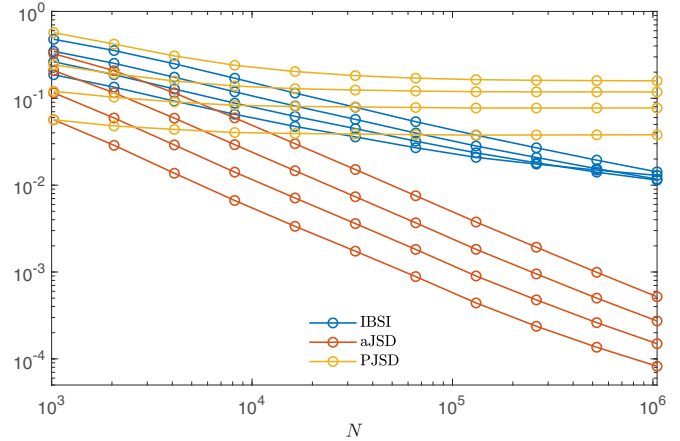
improvements. This behavior is different to that observed for the PJSD (Fig. 8), for which improved discrimination between dynamics is reached for larger orders D.

Finally, we have confirmed that IBSI and aJSD are also useful to discriminate between reversible and irreversible time series following a procedure analogous to the one proposed for the PJSD. That is, the behavior of the dissimilarity between the forward and backward series as a function of the time-series length is analyzed to conclude in favor of a reversible (asymptotic convergence to zero) or irreversible (convergence to a nonzero value) dynamics. We have identified, however, a couple of cases in which both IBSI and aJSD have clear difficulties within this framework. The first one is the β transform with a large value of the parameter β. Figure 29 illustrates the irreversibility analysis for IBSI, aJSD, and PJSD when $\beta = \sqrt{200}$. It is concluded that the convergence to a nonzero value is realized for IBSI and aJSD but for larger values of N. Consequently, these two quantifiers need longer time series to identify the irreversible character of the dynamics. With short time series both of them may be unable to conclude the irreversible nature of the original dynamics. The second case that deserves special consideration is the NGRP model. According to the results displayed in Fig. 30, IBSI and aJSD are not able to unveil the underlying irreversible nature. In fact, on the contrary, reversibility could be erroneously concluded from the aJSD analysis. Undoubtedly, PJSD offers an enhanced characterization of this stochastic system with dynamic nonlinearities.

[1] M. Costa, A. L. Goldberger, and C.-K. Peng, Multiscale Entropy Analysis of Complex Physiologic Time Series, Phys. Rev. Lett. **89**, 068102 (2002).

[2] J. B. Gao, J. Hu, W. W. Tung, and Y. H. Cao, Distinguishing chaos from noise by scale-dependent Lyapunov exponent, Phys. Rev. E **74**, 066204 (2006).

[3] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, Recurrence plots for the analysis of complex systems, Phys. Rep. **438**, 237 (2007).

[4] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, From time series to complex networks: The visibility graph, Proc. Natl. Acad. Sci. USA **105**, 4972 (2008).

[5] B. Podobnik and H. E. Stanley, Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series, Phys. Rev. Lett. **100**, 084102 (2008).

[6] M. De Domenico and V. Latora, Fast detection of nonlinearity and nonstationarity in short and noisy time series, Europhys. Lett. **91**, 30005 (2010).

[7] J. F. Donges, R. V. Donner, and J. Kurths, Testing time series irreversibility using complex network methods, Europhys. Lett. **102**, 10004 (2013).

[8] Q. Li, Z. Fu, and N. Yuan, Beyond Benford's law: Distinguishing noise from chaos, PLoS One **10**, e0129161 (2015).

[9] Z.-K. Gao, M. Small, and J. Kurths, Complex network analysis of time series, Europhys. Lett. **116**, 50001 (2016).

[10] P. A. Bernaola-Galván, M. Gómez-Extremera, A. R. Romance, and P. Carpena, Correlations in magnitude series to assess nonlinearities: Application to multifractal models and heartbeat fluctuations, Phys. Rev. E **96**, 032218 (2017).

[11] C. S. Daw, C. E. A. Finney, and E. R. Tracy, A review of symbolic analysis of experimental data, Rev. Sci. Instrum. **74**, 915 (2003).

[12] A. Porta, M. Baumert, D. Cysarz, and N. Wessel, Enhancing dynamical signatures of complex systems through symbolic computation, Philos. Trans. R. Soc. A **373**, 20140099 (2015).

[13] T. Miyano and H. Gotoda, Estimation of the degree of dynamical instability from the information entropy of symbolic dynamics, Phys. Rev. E **96**, 042203 (2017).

[14] D. Xu and C. Beck, Symbolic dynamics techniques for complex systems: Application to share price dynamics, Europhys. Lett. **118**, 30001 (2017).

[15] C. Bandt and B. Pompe, Permutation Entropy: A Natural Complexity Measure for Time Series, Phys. Rev. Lett. **88**, 174102 (2002).

[16] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán, True and false forbidden patterns in deterministic and random dynamics, Europhys. Lett. **79**, 50001 (2007).

[17] O. A. Rosso, H. A. Larrondo, M. T. Martín, A. Plastino, and M. A. Fuentes, Distinguishing Noise from Chaos, Phys. Rev. Lett. **99**, 154102 (2007).

[18] U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, and N. Wessel, Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics, Comput. Biol. Med. **42**, 319 (2012).

[19] L. Zunino, M. C. Soriano, and O. A. Rosso, Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach, Phys. Rev. E **86**, 046210 (2012).

[20] X. Zhao, P. Shang, and J. Huang, Permutation complexity and dependence measures of time series, Europhys. Lett. **102**, 40005 (2013).

[21] Q. Li and Z. Fu, Permutation entropy and statistical complexity quantifier of nonstationarity effect in the vertical velocity records, Phys. Rev. E **89**, 012905 (2014).

[22] L. Zunino, F. Olivares, and O. A. Rosso, Permutation min-entropy: An improved quantifier for unveiling subtle temporal correlations, Europhys. Lett. **109**, 10005 (2015).

[23] A. Aragoneses, L. Carpi, N. Tarasov, D. V. Churkin, M. C. Torrent, C. Masoller, and S. K. Turitsyn, Unveiling Temporal Correlations Characteristic of a Phase Transition in the Output Intensity of a Fiber Laser, Phys. Rev. Lett. **116**, 033902 (2016).

[24] A. Politi, Quantifying the Dynamical Complexity of Chaotic Time Series, Phys. Rev. Lett. **118**, 144101 (2017).

[25] H. V. Ribeiro, M. Jauregui, L. Zunino, and E. K. Lenzi, Characterizing time series via complexity-entropy curves, Phys. Rev. E **95**, 062106 (2017).

[26] J. M. Amigó, K. Keller, and V. A. Unakafova, Ordinal symbolic analysis and its application to biomedical recordings, Philos. Trans. R. Soc. A **373**, 20140091 (2015).

[27] J. Garland, R. James, and E. Bradley, Model-free quantification of time-series predictability, Phys. Rev. E **90**, 052910 (2014).

[28] E. Bradley and H. Kantz, Nonlinear time-series analysis revisited, Chaos **25**, 097610 (2015).

[29] M. Neuder, E. Bradley, E. Dlugokencky, J. W. C. White, and J. Garland, Detection of local mixing in time-series data using permutation entropy, Phys. Rev. E **103**, 022217 (2021).

[30] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, Permutation entropy and its main biomedical and econophysics applications: A review, Entropy **14**, 1553 (2012).

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

[32] Albert C.-C. Yang, S.-S. Hseu, H.-W. Yien, A. L. Goldberger, and C.-K. Peng, Linguistic Analysis of the Human Heartbeat Using Frequency and Rank Order Statistics, Phys. Rev. Lett. **90**, 108103 (2003).

[33] C.-K. Peng, A. C.-C. Yang, and A. L. Goldberger, Statistical physics approach to categorize biologic signals: From heart rate dynamics to DNA sequences, Chaos **17**, 015115 (2007).

[34] A. Kraskov, W. Nadler, H. Stögbauer, and P. Grassberger, Comment on "Linguistic Analysis of the Human Heartbeat Using Frequency and Rank Order Statistics," Phys. Rev. Lett. **92**, 109801 (2004).

[35] Albert C.-C. Yang, S.-S. Hseu, H.-W. Yien, A. L. Goldberger, C.-K. Peng, Yang *et al.*, Reply, Phys. Rev. Lett. **92**, 109802 (2004).

[36] G. Ouyang, C. Dang, D. A. Richards, and X. Li, Ordinal pattern based similarity analysis for EEG recordings, Clin. Neurophysiol. **121**, 694 (2010).

[37] U. Parlitz, H. Suetani, and S. Luther, Identification of equivalent dynamics using ordinal pattern distributions, Eur. Phys. J.: Spec. Top. **222**, 553 (2013).

[38] D. M. Mateos, L. E. Riveaud, and P. W. Lamberti, Detecting dynamical changes in time series by using the Jensen Shannon divergence, Chaos **27**, 083118 (2017).

[39] J. Lin, Divergence measures based on the Shannon entropy, IEEE Trans. Inf. Theory **37**, 145 (1991).

[40] F. Nielsen, On a generalization of the Jensen-Shannon divergence and the Jensen-Shannon centroid, Entropy **22**, 221 (2020).

[41] D. M. Endres and J. E. Schindelin, A new metric for probability distributions, IEEE Trans. Inf. Theory **49**, 1858 (2003).

[42] T. M. Osán, D. G. Bussandri, and P. W. Lamberti, Monoparametric family of metrics derived from classical Jensen-Shannon divergence, Physica A **495**, 336 (2018).

[43] P. Bernaola-Galván, R. Román-Roldán, and J. L. Oliver, Compositional segmentation and long-range fractal correlations in DNA sequences, Phys. Rev. E **53**, 5181 (1996).

[44] I. Grosse, P. Bernaola-Galván, P. Carpena, R. Román-Roldán, J. Oliver, and H. E. Stanley, Analysis of symbolic sequences using the Jensen-Shannon divergence, Phys. Rev. E **65**, 041905 (2002).

[45] P. W. Lamberti, M. T. Martin, A. Plastino, and O. A. Rosso, Intensive entropic non-triviality measure, Physica A **334**, 119 (2004).

[46] E. H. Feng and G. E. Crooks, Length of Time's Arrow, Phys. Rev. Lett. **101**, 090602 (2008).

[47] L. C. Carpi, O. A. Rosso, P. M. Saco, and M. Gómez Ravetti, Analyzing complex networks evolution through Information Theory quantifiers, Phys. Lett. A **375**, 801 (2011).

[48] J. Sparacino, P. W. Lamberti, and C. M. Arizmendi, Shock detection in dynamics of single-headed motor proteins KIF1A via Jensen-Shannon divergence, Phys. Rev. E **84**, 041907 (2011).

[49] M. Gerlach, F. Font-Clos, and E. G. Altmann, Similarity of Symbol Frequency Distributions with Heavy Tails, Phys. Rev. X **6**, 021009 (2016).

[50] E. M. Bollt, T. Stanford, Y.-C. Lai, and K. Życzkowski, What symbolic dynamics do we get with a misplaced partition?: On the validity of threshold crossings analysis of chaotic time-series, Physica D **154**, 259 (2001).

[51] C. Bandt, Small order patterns in big time series: A practical guide, Entropy **21**, 613 (2019).

[52] D. J. Little and D. M. Kane, Variance of permutation entropy and the influence of ordinal pattern selection, Phys. Rev. E **95**, 052126 (2017).

[53] L. Zunino, F. Olivares, F. Scholkmann, and O. A. Rosso, Permutation entropy based time series analysis: Equalities in the input signal can lead to false conclusions, Phys. Lett. A **381**, 1883 (2017).

[54] D. Cuesta-Frau, M. Varela-Entrecanales, A. Molina-Picó, and B. Vargas, Patterns with equal values in permutation entropy: Do they really matter for biosignal classification?, Complexity **2018**, 1324696 (2018).

[55] A. M. Unakafov and K. Keller, Change-point detection using the conditional entropy of ordinal patterns, Entropy **20**, 709 (2018).

[56] M. Riedl, A. Müller, and N. Wessel, Practical considerations of permutation entropy—A tutorial review, Eur. Phys. J.: Spec. Top. **222**, 249 (2013).

[57] M. Staniek and K. Lehnertz, Parameter selection for permutation entropy measurements, Int. J. Bifurcat. Chaos **17**, 3729 (2007).

[58] C. Bandt, Order patterns, their variation and change points in financial time series and Brownian motion, Stat. Pap. **61**, 1565 (2020).

[59] F. Olivares and L. Zunino, Multiscale dynamics under the lens of permutation entropy, Physica A **559**, 125081 (2020).

[60] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.105.045310 for several complementary analyses.

[61] A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. C. Ivanov, C.-K. Peng, and H. E. Stanley, Fractal dynamics in physiology: Alterations with disease and aging, Proc. Natl. Acad. Sci. USA **99**, 2466 (2002).

[62] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. **10**, 422 (1968).

[63] It should be noted that the random walk references for the fBm processes are generated by integrating the randomly resampled normalized fluctuations associated with each simulation.

[64] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer, Testing for nonlinearity in time series: the method of surrogate data, Physica D **58**, 77 (1992).

[65] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska, Surrogate data for hypothesis testing of physical systems, Phys. Rep. **748**, 1 (2018).

[66] L. Zunino, D. G. Pérez, M. T. Martín, M. Garavaglia, A. Plastino, and O. A. Rosso, Permutation entropy of fractional Brownian motion and fractional Gaussian noise, Phys. Lett. A **372**, 4768 (2008).

[67] D. DeFord and K. Moore, Random walk null models for time series data, Entropy **19**, 615 (2017).

[68] J. Beran, *Statistics for Long-memory Processes* (Chapman & Hall, London, 1994)

[69] C. Bandt and F. Shiha, Order patterns in time series, J. Time Ser. Anal. **28**, 646 (2007).

[70] J. B. Gao, S. K. Hwang, and J. M. Liu, When Can Noise Induce Chaos? Phys. Rev. Lett. **82**, 1132 (1999).

[71] L. Lacasa, A. Nuñez, É. Roldán, J. M. R. Parrondo, and B. Luque, Time series irreversibility: a visibility graph approach, Eur. Phys. J. B **85**, 217 (2012).

[72] M. Zanin, A. Rodríguez-González, E. Menasalvas Ruiz, and D. Papo, Assessing time series reversibility through permutation patterns, Entropy **20**, 665 (2018).

[73] J. H. Martínez, J. L. Herrera-Diestra, and M. Chavez, Detection of time reversibility in time series by ordinal patterns analysis, Chaos **28**, 123111 (2018).

[74] C. Diks, J. C. van Houwelingen, F. Takens, and J. DeGoede, Reversibility as a criterion for discriminating time series, Phys. Lett. A **201**, 221 (1995).

[75] K. Keller, A. M. Unakafov, and V. A. Unakafova, Ordinal patterns, entropy, and EEG, Entropy **16**, 6212 (2014).

[76] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, Detecting long-range correlations with detrended fluctuation analysis, Physica A **295**, 441 (2001).

[77] E. M. Blanter, J.-L. Le Mouël, F. Perrier, and M. G. Shnirman, Short-term correlation of solar activity and sunspot: Evidence of lifetime increase, Solar Phys. **237**, 329 (2006).

[78] A. Shapoval, J.-L. Le Mouël, M. Shnirman, and V. Courtillot, Dynamics of sunspot series on time scales from days to years: Correlation of sunspot births, variable lifetimes, and evolution of the high-frequency spectral component, J. Geophys. Res.: Space Phys. **122**, 11,874 (2017).

[79] E. M. Blanter, M. G. Shnirman, and J.-L. Le Mouël, Solar variability: Evolution of correlation properties, J. Atmos. Sol. Terr. Phys. **67**, 521 (2005).

[80] SILSO, World Data Center—Sunspot Number and Long-term Solar Observations, Royal Observatory of Belgium, on-line Sunspot Number catalogue [http://www.sidc.be/SILSO/].

[81] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys. Rev. E **64**, 061907 (2001).

[82] L. Zunino and C. W. Kulp, Detecting nonlinearity in short and noisy time series using the permutation entropy, Phys. Lett. A **381**, 3627 (2017).

[83] C. W. Kulp, L. Zunino, T. Osborne, and B. Zawadzki, Using missing ordinal patterns to detect nonlinearity in time series data, Phys. Rev. E **96**, 022218 (2017).

[84] M. J. van der Heyden, C. Diks, J. P. M. Pijn, and D. N. Velis, Time reversibility of intracranial human EEG recordings in mesial temporal lobe epilepsy, Phys. Lett. A **216**, 283 (1996).

[85] K. Schindler, C. Rummel, R. G. Andrzejak, M. Goodfellow, F. Zubler, E. Abela, R. Wiest, C. Pollo, A. Steimer, and H. Gast, Ictal time-irreversible intracranial EEG signals as markers of the epileptogenic zone, Clin. Neurophysiol. **127**, 3051 (2016).

[86] M. Costa, A. L. Goldberger, and C.-K. Peng, Broken Asymmetry of the Human Heartbeat: Loss of Time Irreversibility in Aging and Disease, Phys. Rev. Lett. **95**, 198102 (2005).

[87] RR Interval Time Series Modeling—The PhysioNet/Computing in Cardiology Challenge 2002, https://physionet.org/content/challenge-2002/1.0.0/.

[88] G. B. Moody, RR interval time series modeling: The PhysioNet/Computers in Cardiology Challenge 2002, Comput. Cardiol. **29**, 125 (2002).

[89] M. D. Costa, C.-K. Peng, and A. L. Goldberger, Multiscale analysis of heart rate dynamics: Entropy and time irreversibility measures, Cardiovasc. Eng. **8**, 88 (2008).

[90] A. A. B. Pessa and H. V. Ribeiro, ordpy: A Python package for data analysis with permutation entropy and ordinal network methods, Chaos **31**, 063110 (2021).

[91] H. A. Makse, S. Havlin, M. Schwartz, and H. E. Stanley, Method for generating long-range correlations for large systems, Phys. Rev. E **53**, 5445 (1996).