

This is the Author's Pre-print version of the following article: *Nazul Merino-Negrete, Cesar Maldonado, Raúl Salgado-García, Sorting ECGs by lag irreversibility, Physica D: Nonlinear Phenomena, Volume 459, 2024, 134022*, which has been published in final form at: <https://doi.org/10.1016/j.physd.2023.134022>

© 2023 This manuscript version is made available under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Sorting ECGs by lag irreversibility

Nazul Merino Negrete^a, Cesar Maldonado^a, Raúl Salgado-García^b

^a*IPICYT/División de Control y Sistemas Dinámicos, Camino a la Presa San José 2055.
Lomas 4a. sección. San Luis Potosí S.L.P. México*

^b*Centro de Investigación en Ciencias-IICBA, Universidad Autónoma del Estado de
Morelos. Avenida Universidad 1001, colonia Chamilpa, 62209, Cuernavaca, México.*

Abstract

In this work we introduce the lag irreversibility function as a method to assess time-irreversibility in discrete time series. It quantifies the degree of time-asymmetry for the joint probability function of the state variable under study and the state variable lagged in time. We test its performance in a time-irreversible Markov chain model for which theoretical results are known. Moreover, we use our approach to analyze electrocardiographic recordings of four groups of subjects: healthy young individuals, healthy elderly individuals, and persons with two different disease conditions, namely, congestive heart failure and atrial fibrillation. We find that by studying jointly the variability of the amplitudes of the different waves in the electrocardiographic signals, one can obtain an efficient method to discriminate between the groups already mentioned. Finally, we test the accuracy of our method using the ROC analysis.

Keywords: Lag irreversibility function Time Irreversibility estimators
Electrocardiograms Heart Rate Variability ROC Analysis .

1. Introduction

In recent years, assessing the time-reversal asymmetry has become an important tool for the analysis of time series [1]. This is because the irreversibility of a (stochastic or deterministic) process is a feature that provides information about the nature of the underlying dynamics [2, 3, 4]. For instance, the time-irreversibility is closely related to the entropy production in physical systems [5], a signature of out-of-equilibrium processes [6, 5, 7]. Also, the time-reversal asymmetry is a property related to the presence of non-linear correlations as well as the presence of non-Gaussian fluctuations, among other interesting properties [8].

Several approaches has been proposed to determine the irreversibility of a time series [1, 9, 10]. For example, a natural way to assess irreversibility is comparing directly the joint distribution of the process forward and backward in time [2, 11]. This is usually done by encoding the time series into a symbolic sequence and then computing the occurrence of words and their respective time-reversed versions [8]. Another well-established method to quantify

the degree of irreversibility in symbolic sequences is to compute the Kullback-Leibler divergence between the process forward in time and its time-reversal version [7, 12, 13]. Many other approaches that is worth to mention are: visibility graph approach [14], ordinal patterns analysis [15], recurrence-time statistics [16, 17, 18] and recurrence plots [19], just to mention a few examples.

Since time irreversibility is a common feature of the out-of-equilibrium phenomena, it has been widely studied in physics, biology, mathematics, chemistry, economic sciences, among other disciplines [1, 5, 6, 20, 21]. For instance, time-irreversibility has been implemented in the analysis of financial time series in [20, 22, 23]. Using a permutation patterns approach, in [10], M. Zanin studied the irreversibility properties of different chaotic dynamical systems. In [12], Porporato *et al.*, estimated the irreversibility by means of the Kullback-Leibler divergence of a stochastic process defined by a combination of an asymmetric jumps Poisson process and an Ornstein-Uhlenbeck process. They later applied that same methodology to quantify the asymmetry of the time series of the discharge measurements for the Po River in Italy. With respect to the physics literature there is a huge amount of studies concerning the relation with the fluctuation-dissipation theorem, as well as for applications in out-of-equilibrium systems. For a comprehensive study on irreversibility in statistical mechanics, see [6] and references therein. The irreversibility of biological and physiological signals has recently arisen interest. For instance, there are studies in electroencephalographic time series [24], in Markovian models of Spike Trains [25], and also it has been shown that the irreversibility analysis can be useful for discriminating between coding and non-coding DNA sequences [26, 17].

Here we focus in the analysis of irreversibility of time series obtained from some properties (derived signals) of electrocardiograms (ECG's). In that context, for instance, the temporal asymmetry of heart rate variability signals has been studied in [27] by Cammarota *et al.*, using a ternary symbolization of the RR-intervals, which was later extended in [28]. Using indices of irreversibility such as Poincaré plots and related methods, it has been found that the heart rate variability signals display a temporal asymmetry [29, 30, 31]. In [32], Costa *et al.*, suggested that loss of irreversibility might be related to aging and disease of individuals. They analyzed the signal using different scales of the time series and estimated the irreversibility of the probability density function of the system and its time reversal, which considers the energy and heat fluxes in the process. Related methods, such as estimation of the change of entropy in (a so-called) natural time are given by [33] and [34]. Recently, in [35], under an assumption of Markovianity of the signals, the authors estimated the entropy production of the symbolized time series for both the electrocardiographic signals and the heart rate variability of three groups of patients; healthy ones, patients with Atrial Fibrillation and individuals with Congestive Heart Failure. There, it is shown that it is possible to discriminate the healthy ones from the diseased groups by means of the entropy production. Other studies that analyze time-irreversibility in order to determine the health condition of individuals from the electrocardiographic time series can be found in [36, 37, 38, 39, 40, 41].

In the present work we define an index to quantify the time irreversibility, which we call *lag irreversibility function*. This quantity can be applied not only to single-signal processes, but to processes possessing more than one signal, i.e., the lag irreversibility function can be evaluated for multivariate time series. This is important in applications, for instance, we apply our irreversibility index to ECG signals, since it takes into account information on different stages (or features) of the ECG through their corresponding heart rate variability signals. This may contribute to get a better estimate of the time-reversal asymmetry using not so long recordings. So, by computing the lag irreversibility function we are able to discriminate signals for four groups of individuals with different health conditions, including healthy elderly patients. We test the accuracy of our method using the ROC analysis.

The article is organized as follows. In Section 2 we define the lag irreversibility function and we introduce the necessary concepts that we use for the rest of the paper. Section 3 is devoted to describe the methodology employed to obtain the signals, the filtering process, the symbolization of the time series and the methodology to estimate the lag irreversibility function. We give our results in Section 4, as well as we test them using the ROC analysis. Finally we give some concluding remarks in Section 5.

2. Lag irreversibility

2.1. Irreversibility and entropy production

Let us start by stating some preliminary concepts about irreversibility of stochastic processes. Let $\mathcal{X} = \{X_t : t \in \mathbb{N}\}$ be a discrete-time stochastic process where the state variables take values on a finite state space \mathcal{S} . Also, we denote by \mathbf{X}_t^{t+n} a finite random path or trajectory of the process starting at time t and ending a time $t+n$, i.e.,

$$\mathbf{X}_t^{t+n} = (X_t, X_{t+1}, \dots, X_{t+n}). \quad (1)$$

The realizations of the process or the finite trajectories will be denoted by bold lower case letters, i.e., a realization of the random path \mathbf{X}_t^{t+n} will be denoted simply by \mathbf{a} , where \mathbf{a} belongs to the set \mathcal{S}^{n+1} . If necessary, we will use subscripts and superscripts to emphasize the finite character of the realizations, \mathbf{a}_0^n .

We know that the stochastic process \mathcal{X} is reversible if for all $n \in \mathbb{N}$, all $t \in \mathbb{N}_0$ and any finite sequence $\mathbf{a} \in \mathcal{S}^{n+1}$ we have that

$$\mathbb{P}(\mathbf{X}_t^{t+n} = \mathbf{a}) = \mathbb{P}(\mathbf{X}_t^{t+n} = \bar{\mathbf{a}}), \quad (2)$$

where the overline in \mathbf{a} denotes the reversed realization of the process, i.e., if $\mathbf{a} = (a_0, a_1, \dots, a_{n-1}, a_n)$, then,

$$\bar{\mathbf{a}} = (a_n, a_{n-1}, \dots, a_1, a_0). \quad (3)$$

If the identity (2) does not hold it is said that the process is irreversible. In that case the probability of observing a given trajectory \mathbf{a} does not coincide with the probability of observing the reverse trajectory $\bar{\mathbf{a}}$.

The degree of irreversibility of a process \mathcal{X} is commonly measured by means of the Kullback-Leibler divergence between path distribution $\mathbb{P}(\mathbf{X}_t^{t+n} = \mathbf{a})$ and the reversed path distribution $\mathbb{P}(\mathbf{X}_t^{t+n} = \bar{\mathbf{a}})$:

$$D_{\text{KL}}(\mathbb{P}(\mathbf{X}_t^{t+n} = \mathbf{a}) \parallel \mathbb{P}(\mathbf{X}_t^{t+n} = \bar{\mathbf{a}})) := \sum_{\mathbf{a} \in \mathcal{S}} \mathbb{P}(\mathbf{X}_t^{t+n} = \mathbf{a}) \log \left(\frac{\mathbb{P}(\mathbf{X}_t^{t+n} = \mathbf{a})}{\mathbb{P}(\mathbf{X}_t^{t+n} = \bar{\mathbf{a}})} \right). \quad (4)$$

The previous quantity is related to the entropy production rate, which is an important concept. From the physical point of view, it is not only a measure of the time irreversibility, but it provides information about the nature of the corresponding physical system. For example, it is known that if a given time series comes from a thermodynamic system at equilibrium, then the entropy production rate of such a process will be zero, and thus, being time-reversible. Otherwise, if such a time series comes from a system out from equilibrium, the series will be clearly irreversible. Despite its importance, from the practical point of view, the Kullback-Leibler divergence between the path and reversed path distributions has its drawbacks while being estimated from a sample trajectory (or finite realization). This is mainly due to the fact that it is not only necessary to take a limit for $n \rightarrow \infty$, but also it is necessary to obtain estimations for the n -dimensional marginals $\mathbb{P}(\mathbf{X}_t^{t+n} = \mathbf{a})$, for which their fluctuations might be large when n increases, since the size of the sample is finite. This fact, among others, has motivated the introduction of several techniques to assess the time-irreversibility, something that has been achieved with different success in every case [1].

2.2. Lag irreversibility

As we mentioned above, the main problem in evaluating the degree of irreversibility from a time series is the finiteness of the sample. Some estimators like those based on recurrence times, such as hitting time or waiting time, require a large sample trajectory. Other estimators might be more efficient and a comparative study can be found in [1]. The main advantage of the method we introduce here for determining the irreversibility of given time series, is that the number of parameters to be estimated is relatively small compared to other methods. This method will be referred to as *lag irreversibility* and the basic idea of the proposed irreversibility index goes as follows. First consider a discrete-time stochastic process $\mathcal{X} = \{X_t : t \in \mathbb{N}\}$ and let $\tau \in \mathbb{N}$ be a non-negative integer. Assume that such a stochastic process is stationary and consider the random path $\mathbf{X}_t^{t+\tau} = (X_t, X_{t+1}, \dots, X_{t+\tau})$. As we mentioned above, if the process \mathcal{X} is irreversible, then

$$\mathbb{P}(\mathbf{X}_t^{t+\tau} = \mathbf{a}_0^\tau) \neq \mathbb{P}(\mathbf{X}_t^{t+\tau} = \bar{\mathbf{a}}_0^\tau). \quad (5)$$

If we sum over $a_1, a_2, \dots, a_{\tau-1}$ it is clear that

$$\sum_{a_1} \sum_{a_2} \cdots \sum_{a_{\tau-1}} \mathbb{P}(\mathbf{X}_t^{t+\tau} = \mathbf{a}_0^\tau) = \mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau), \quad (6)$$

where $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ is the joint probability function for the state variables X_t and $X_{t+\tau}$. Equivalently we have that,

$$\sum_{a_1} \sum_{a_2} \cdots \sum_{a_{\tau-1}} \mathbb{P}(\mathbf{X}_t^{t+\tau} = \bar{\mathbf{a}}_0) = \mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0), \quad (7)$$

where $\mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)$ is the joint probability function for the state variables X_t and $X_{t+\tau}$, but reversed in time. The above equations, together with the condition of irreversibility given by Eq. (5) implies that if the process \mathcal{X} is irreversible, then the joint distribution for the state variables X_t and $X_{t+\tau}$ will have the following asymmetry property,

$$\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau) \neq \mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0). \quad (8)$$

The latter motivates our definition of *pairwise reversibility*.

Definition 2.1. *We say that a stochastic process $\mathcal{X} = \{X_t : t \in \mathbb{N}\}$ is pairwise reversible if for all $\tau \in \mathbb{N}$, we have that*

$$\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau) = \mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0). \quad (9)$$

Otherwise, we say that the process is pairwise irreversible.

To check pairwise irreversibility we need to estimate the joint probability function $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ for every τ , a task that is computationally less expensive than estimating the joint probability function of the whole trajectory. A natural way to evaluate the pairwise irreversibility is to compute the Kullback-Leibler divergence of the joint probability $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ with respect to $\mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)$, which leads to the following definition of the lag irreversibility function.

Definition 2.2. *We define the lag irreversibility (LI) function, $L(\tau)$, as the Kullback-Leibler divergence of the joint probability $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ with respect to $\mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)$, i.e.,*

$$\begin{aligned} L(\tau) &:= D_{KL}(\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau) \parallel \mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)) \\ &= \sum_{a_0} \sum_{a_\tau} \mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau) \log \left(\frac{\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)}{\mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)} \right). \end{aligned} \quad (10)$$

It is important to remark that we named *lag irreversibility function* to the Kullback-Leibler divergence $L(\tau)$ because, in some sense, we are measuring the asymmetry of the joint probability function of the state variable X_t and the state variable lagged a time τ , i.e., $X_{t+\tau}$. This asymmetry, as we argued above, gives information about the irreversibility of the process throughout the estimation of $L(\tau)$ for every τ .

At this point, it is instructive to compute analytically the lag irreversibility function $L(\tau)$ in a specific Markov chain model.

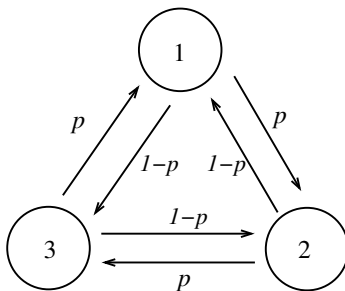


Figure 1: The three-states Markov chain model. This Markov chain is known to be reversible if $p = 1/2$. For parameter values lower or larger than $p = 1/2$ the Markov chain is irreversible, and the irreversibility get larger as p gets away the value $p = 1/2$.

Example 2.1. Let $\mathcal{X} = \{X_t : t \in \mathbb{N}\}$ be a three-states stationary Markov chain with state space $\mathcal{S} = \{1, 2, 3\}$ and stochastic matrix \mathbf{Q} given by

$$\mathbf{Q} = \begin{bmatrix} 0 & p & 1-p \\ 1-p & 0 & p \\ p & 1-p & 0 \end{bmatrix}. \quad (11)$$

It is known that for a Markov chain, the joint probability function $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ can be written as

$$\mathbb{P}(X_t = a; X_{t+\tau} = b) = \pi_a (\mathbf{Q}^\tau)_{a,b}, \quad (12)$$

where $a, b \in \mathcal{S}$ and π_a is the a -th element of the stationary probability vector $\boldsymbol{\pi}$, i.e., the probability vector satisfying the stationary equation $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{Q}$.

It is not hard to see that $L(\tau)$ can be written in this case as,

$$L(\tau) = \sum_{a \in \mathcal{S}} \sum_{b \in \mathcal{S}} \pi_a (\mathbf{Q}^\tau)_{a,b} \log \left(\frac{(\mathbf{Q}^\tau)_{a,b}}{(\mathbf{Q}^\tau)_{b,a}} \right). \quad (13)$$

Eq. (13) is actually the expression for the lag irreversibility function for any one-step Markov chain. This expression can be straightforwardly evaluated for the three-states Markov chain introduced above. In Fig. 2 we can appreciate the behavior of the LI function for several values of the parameter p . All the parameter values we use to compute the lag irreversibility function for this Markov model correspond to the case in which the process is irreversible. Actually the only parameter value for which the three-states Markov chain is reversible corresponds to $p = 1/2$. It is not hard to see that the LI function, $L(\tau)$, is zero for all $\tau \in \mathbb{N}$ in the case $p = 1/2$. This is because in this case the stochastic matrix is symmetric, implying that

$$(\mathbf{Q}^\tau)_{a,b} = (\mathbf{Q}^\tau)_{b,a},$$

making $L(\tau) = 0$ for all τ . For other values of the parameter p , the LI function is positive and increases as p gets away from $p = 1/2$. In Fig. 2 we show the

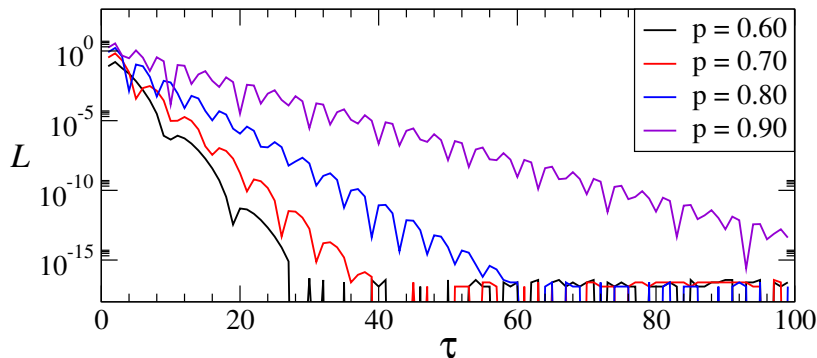


Figure 2: Lag-irreversibility function for the three-states Markov chain. We show the LI functions for $p = 0.60$ (solid black line), $p = 0.70$ (solid red line), $p = 0.80$ (solid blue line), and $p = 0.90$ (solid violet line). Observe that the LI function gets larger as p increases. Also is important to notice that all the LI functions decay exponentially fast and the larger p the slower decaying rate.

LI functions for $p = 0.60$, $p = 0.70$, $p = 0.80$, and $p = 0.90$. As we see, the LI function gets larger as the parameter p increases. Also, as p increases, the entropy production also increases, which is consistent with the fact that the LI function can be used as a measure of the irreversibility of the process. We should notice that the larger entropy production rate, the lower decaying rate in the LI function, which allows to conclude that the decay rate of the LI function might be a measure (or at least an indicator) of the irreversibility as well.

As we saw in the preceding example, we have that the LI function $L(\tau)$ decays as τ increases. This is actually a consequence of the mixing property of the Markov chain, i.e., as τ gets larger, the random variables X_t and $X_{t+\tau}$ become independent, that is known as the *decay of correlations*. If the random variables X_t and $X_{t+\tau}$ become independent, then the joint probability function $\mathbb{P}(X_t = a; X_{t+\tau} = b)$ becomes symmetric, i.e.,

$$|\mathbb{P}(X_t = a; X_{t+\tau} = b) - \mathbb{P}(X_t = b; X_{t+\tau} = a)| \rightarrow 0, \quad (14)$$

as $\tau \rightarrow \infty$. This implies that, if a given process comply with the mixing property, we will have that

$$\lim_{\tau \rightarrow \infty} L(\tau) = 0.$$

2.3. Lag irreversibility estimator

The empirical lag irreversibility estimator is defined by the empirical joint probability of the state variables X_t and $X_{t+\tau}$, both in the original sequence and in the time-reversed sequence. We assume that these variables are produced by an unknown processes, but that they can be estimated directly from the trajectories. Let $x = x_1, x_2, \dots, x_n$ be a realization of the stationary process $\mathcal{X} = \{X_t : t \in \mathbb{N}\}$, with state space $\mathcal{S} = \{a_1, a_2, \dots, a_N\}$. We say that

process \mathcal{X} at time i is in state a_j if $x_i = a_j$. The joint probability functions $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ and $\mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)$, can be estimated by means of

$$\widehat{\mathbb{P}}(X_t = a_0; X_{t+\tau} = a_\tau) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} \chi_{a_0}(x_t) \chi_{a_\tau}(x_{t+\tau}), \quad (15a)$$

and

$$\widehat{\mathbb{P}}(X_t = a_\tau; X_{t+\tau} = a_0) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} \chi_{a_\tau}(x_t) \chi_{a_0}(x_{t+\tau}), \quad (15b)$$

respectively, where $\chi_a(\cdot)$ is the indicator function for the state a . Equations (15a) and (15b) give us the empirical joint probability that we use to estimate $\widehat{L}(\tau)$, by directly substitute them in Equation (10).

3. Methodology

As we mentioned above, the main purpose of this work is to use the time irreversibility as a property for discriminating among four different groups of patients, namely, Healthy Young (HY), Healthy Elderly (HE), Congestive Heart Failure (CHF) and Atrial Fibrillation (AF). However, quantifying the degree of irreversibility from ECG recordings is neither a direct nor a trivial task, so we propose a methodology consisting of six stages. (1) retrieving and selecting data (ECG recordings) from databases, (2) cleaning ECG recordings and computing variability signals from ECGs (see definitions below), (3) symbolic encoding of variability signals, (4) designing estimation method, (5) assessing irreversibility for classification and (6) evaluating classification through ROC analysis. Next, we give a brief description of stages (1)-(4) which are the preliminary steps before presenting of our main results, stages (5) and (6), which are fully described in Section 4 below.

3.1. Retrieving and selecting data

Datasets were obtained from the *PhysioBank* database [42], by means of the open-source WFDB Software Package [43]. Healthy Young (HY) and Healthy Elderly (HE) groups were selected from *FANTASIA* database [44], while Congestive Heart Failure (CHF) and Atrial Fibrillation (AF) groups were selected from the *BIDM Congestive Heart Failure* [45] and *MIT-BIH Atrial Fibrillation* [46] databases, respectively.

According to the database, Healthy Young group, as well as Healthy Elderly group, are comprised of 20 electrocardiograms (ECGs) each, which were acquired from patients at supine rest. The sampling frequency is 250 Hz and every ECG record is 120 minutes long. It is also important to mention that the HY patient group, from which these ECGs were obtained, is comprised of 10 men and 10 women aged between 21 and 34. On the other hand, HE patient group is also comprised of 10 men and 10 women, but aged between 68 and 85.

For the CHF group, the databases provide 15 ECG recordings from which, 11 corresponds to men aged from 22 to 71 and 4 corresponds to women aged

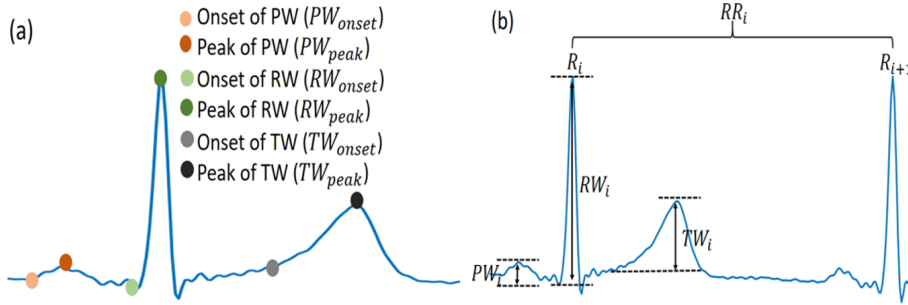


Figure 3: Components of the ECG signal. The amplitudes of waves P, R and T, at the i -th heartbeat, are obtained by means of $PW_i = PW_{peak_i} - PW_{onset_i}$, $RW_i = RW_{peak_i} - RW_{onset_i}$ and $TW_i = TW_{peak_i} - TW_{onset_i}$, respectively.

from 54 to 63. Every ECG record in this group is 20 hours long with a sampling frequency of 250 Hz. For the AF group the databases provides 25 ECG records of adult subjects¹ and each recording is 10 hours long with a sampling frequency of 250 Hz. ECG recordings for AF and CHF groups were obtained by means of ambulatory electrocardiography.

At this point, it is important to stress that, although HY, HE and AF databases have 20 or more electrocardiographic recordings, CHF database contains only 15 ECG samples. Therefore, we only considered 15 ECG records for each group. Also, in order to have homogeneous samples, we consider 5000 heartbeats from ECGs to carry out all the estimates. This choice is due to the fact that this is the least common number of heartbeats within these selected databases.

3.2. Cleaning ECG records and variability signals

The ECG recordings retrieved from databases are not completely clean, and might possess some artifacts that could not allow the appropriate identification of P, Q, R, S and T waves in the ECG. To deal with this issue, we made use of *NeuroKit2* package [47], which is The Python Toolbox for Neurophysiological Signal Processing. This tool allows to clean ECG signals and detect the components, i.e., the P, Q, R, S, and T waves. In particular, we can automatically clean the ECG and locate the onsets and peaks of the P, R and T waves (see Figure 3(a)), using the `ecg_process`, `ecg_peaks` and `ecg_delineate` commands.

Next, once we obtained the onsets and peaks of P, R and T waves, we compute the wave amplitude by means of the difference between the voltage at its onset point and its peak. In other words, the wave amplitude is defined as $W_{peak} - W_{onset}$, where W_{onset} is the voltage at the starting point of the wave and W_{peak} is the voltage that reaches the peak of the wave. For example, for the i -th heartbeat, the P-wave amplitude is given by $PW_i = PW_{peak_i} - PW_{onset_i}$;

¹No mention about the patient age is found in the database.

we repeat the procedure analogously for the R and T waves (see Figure 3(b)). Regarding the RR intervals, whose length is called *heart rate variability* (HRV), it is possible to obtain them using the `hrv` function of *Neurokit*.

This definition of wave amplitude gives us four time series, namely, PW , RW , TW and the heart rate variability RR . These quantities are those that we refer to as *variability signals*, which are discrete time series of the form $x = (x_1, x_2, \dots, x_m)$, where $m = 5000$, is the number of heartbeats contained in data. To eliminate spurious non-stationary characteristics, we consider the differences of consecutive entries of each time series, i.e., we use the vector $y = (y_1, y_2, \dots, y_{m-1})$, where $y_i = x_{i+1} - x_i$, for $i = 1, 2, \dots, m-1$. Specifically, for the RR intervals, P-wave, R-wave, and T-wave, their respective difference time series are given by $\Delta RR = \{\Delta RR_i : \Delta RR_i = RR_{i+1} - RR_i, i = 1, 2, \dots, m-1\}$, $\Delta PW = \{\Delta PW_i : \Delta PW_i = PW_{i+1} - PW_i, i = 1, 2, \dots, m-1\}$, $\Delta RW = \{\Delta RW_i : \Delta RW_i = RW_{i+1} - RW_i, i = 1, 2, \dots, m-1\}$ and $\Delta TW = \{\Delta TW_i : \Delta TW_i = TW_{i+1} - TW_i, i = 1, 2, \dots, m-1\}$.

3.3. Symbolic encoding

Once we have the time series of the ΔRR , ΔPW , ΔRW and ΔTW signals, we proceed to encode them by partitioning the state space of the time series. An intuitive encoding technique was proposed in [8], which considers a uniform partition of the state space. However, this uniform partition does not take into account the typical fluctuations of the signal. In order to capture these fluctuations inherent to physiological processes, in [35], the authors proposed different partitions composed by elements with non-uniform sizes. There, it is considered a “center cell” around the mean μ value of the time series, and the size of the element of the partition is defined using the standard deviation σ of the data and fitted by means of a parameter γ . After that, the authors make an exploratory study in order to select a suitable value for γ in the sense that the obtained results permit a better discrimination among the groups of patients.

Here, we simplify the method for defining a useful partition from [35]. We consider the simplest partition of the state space that allows a good estimation of the irreversibility properties, which is a ternary partition. And in order to draw information from the time series, we perform a *single signal encoding* and a *joint signal encoding*, considering non-uniform sizes of the partitions. This enables us, on the one hand, to simplify the method for obtaining the partition; and, on the other hand, to obtain more information about the irreversibility properties of the data sequences. For the *single signal encoding* (see Figure 4 for a schematic representation), given a time series of differences $y = (y_1, y_2, \dots, y_{m-1})$, with mean μ and standard deviation σ , its corresponding symbolic sequence can be obtained by using the rule:

$$s_i := \begin{cases} 1 & \text{if } y_i \leq \mu - \gamma\sigma, \\ 2 & \text{if } \mu - \gamma\sigma < y_i < \mu + \gamma\sigma, \\ 3 & \text{if } y_i \geq \mu + \gamma\sigma. \end{cases}$$

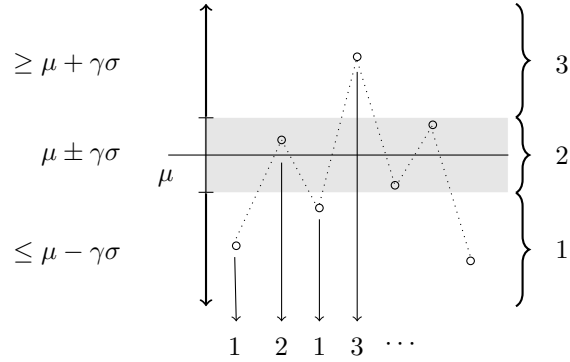


Figure 4: Schematic diagram of the encoding method for a single signal.

For the case of *joint signal encoding*, given two time series of differences $y^1 = (y_1^1, y_2^1, \dots, y_{m-1}^1)$ (with parameters μ_1 and σ_1) and $y^2 = (y_1^2, y_2^2, \dots, y_{m-1}^2)$ (with parameters μ_2 and σ_2), their joint symbolic sequence can be obtained through the following rule:

$$s_i = \begin{cases} 1 & \text{if } y_i^1 \geq \mu_1 + \gamma\sigma_1 & \wedge & y_i^2 \geq \mu_2 + \gamma\sigma_2, \\ 2 & \text{if } y_i^1 \geq \mu_1 + \gamma\sigma_1 & \wedge & \mu_2 - \gamma\sigma_2 < y_i^2 < \mu_2 + \gamma\sigma_2, \\ 3 & \text{if } y_i^1 \geq \mu_1 + \gamma\sigma_1 & \wedge & y_i^2 \leq \mu_2 - \gamma\sigma_2, \\ 4 & \text{if } \mu_1 - \gamma\sigma_1 < y_i^1 < \mu_1 + \gamma\sigma_1 & \wedge & y_i^2 \geq \mu_2 + \gamma\sigma_2, \\ 5 & \text{if } \mu_1 - \gamma\sigma_1 < y_i^1 < \mu_1 + \gamma\sigma_1 & \wedge & \mu_2 - \gamma\sigma_2 < y_i^2 < \mu_2 + \gamma\sigma_2, \\ 6 & \text{if } \mu_1 - \gamma\sigma_1 < y_i^1 < \mu_1 + \gamma\sigma_1 & \wedge & y_i^2 \leq \mu_2 - \gamma\sigma_2, \\ 7 & \text{if } y_i^1 \leq \mu_1 - \gamma\sigma_1 & \wedge & y_i^2 \geq \mu_2 + \gamma\sigma_2, \\ 8 & \text{if } y_i^1 \leq \mu_1 - \gamma\sigma_1 & \wedge & \mu_2 - \gamma\sigma_2 < y_i^2 < \mu_2 + \gamma\sigma_2, \\ 9 & \text{if } y_i^1 \leq \mu_1 - \gamma\sigma_1 & \wedge & y_i^2 \leq \mu_2 - \gamma\sigma_2. \end{cases}$$

3.4. Estimation method

Once we have the symbolic sequences of the data, we proceed to build the estimator. First, we estimate the joint probability functions $\mathbb{P}(X_t = a_0; X_{t+\tau} = a_\tau)$ and $\mathbb{P}(X_t = a_\tau; X_{t+\tau} = a_0)$, for $\tau = 1, \dots, 20$, using equations (15a) and (15b), respectively. Next, we estimate the *lag irreversibility function*, $\hat{L}(\tau)$, by directly plugging the estimated joint probability functions into equation (10).

4. Results

In this section we estimate the lag irreversibility function for single and joint variability signals. We do that by means of an empirical estimation of the joint probabilities, as described in Section 3.4. For the symbolization process, we consider a ternary partition of the state space, making an exploratory study for finding the best parameter value γ .

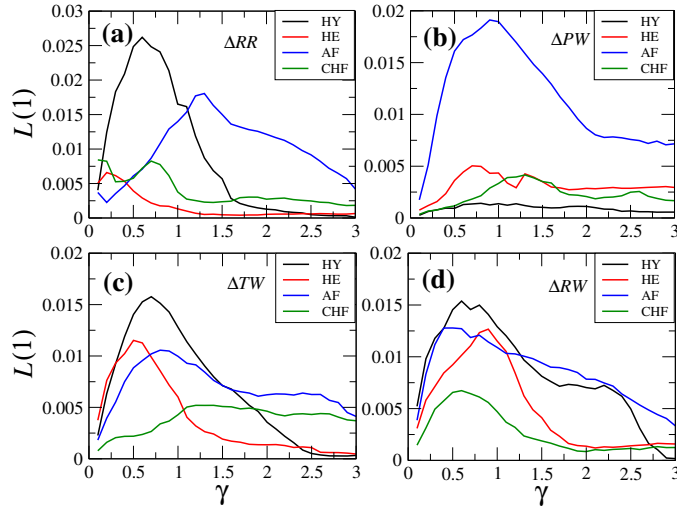


Figure 5: Lag-irreversibility function of single variability signals, for different values of γ and for $\tau = 1$.

4.1. Fixing the value of γ

In Figure 5 we show the values of the LI function for single variability signals, for different values of γ and $\tau = 1$. In this figure we can see that the heart variability signal (ΔRR) exhibits a higher value of LI for the HY group and it decreases with aging and disease. Furthermore, we see from the latter that a suitable choice for the parameter is $\gamma = 3/10$, in the sense that it allows a better discrimination between groups, i.e., for healthy young subjects the LI value (≈ 0.01832) is more than three times higher than the LI value for CHF (≈ 0.00526) and about five times higher than the LI value for AF group (≈ 0.00354). For healthy elderly subjects its LI value (≈ 0.00614) is approximately 16% higher than for CHF (≈ 0.00526) and 73% higher than for AF (≈ 0.00354). Finally, the LI value in AF is roughly 48% higher respect the CHF group. From these results we suggest that $\gamma = 3/10$ is a suitable choice to estimate $\hat{L}(\tau)$ as a function of τ (see Section 4.2 below).

The main advantage of the lag irreversibility function is found in the case of the joint variability signals. We carry out the same exploratory study, that is, we estimate LI function for $\tau = 1$ and different values of γ , which we show in Figure 6. We can see that joint coding cases involving heart rate variability ($(\Delta RR, \Delta PW)$, $(\Delta RR, \Delta RW)$, $(\Delta RR, \Delta TW)$) allow better and consistent discrimination between groups of healthy patients from those with some adverse health condition. For example, for the case $(\Delta RR, \Delta PW)$, the LI value of the healthy young group is ≈ 0.0421 , which is four times greater than that for the CHF group (≈ 0.0106) and two times greater than the value for the AF group (≈ 0.0207), while the LI value of the group HE (≈ 0.0356) is three times greater than that for the CHF group and two times greater than the value for the AF

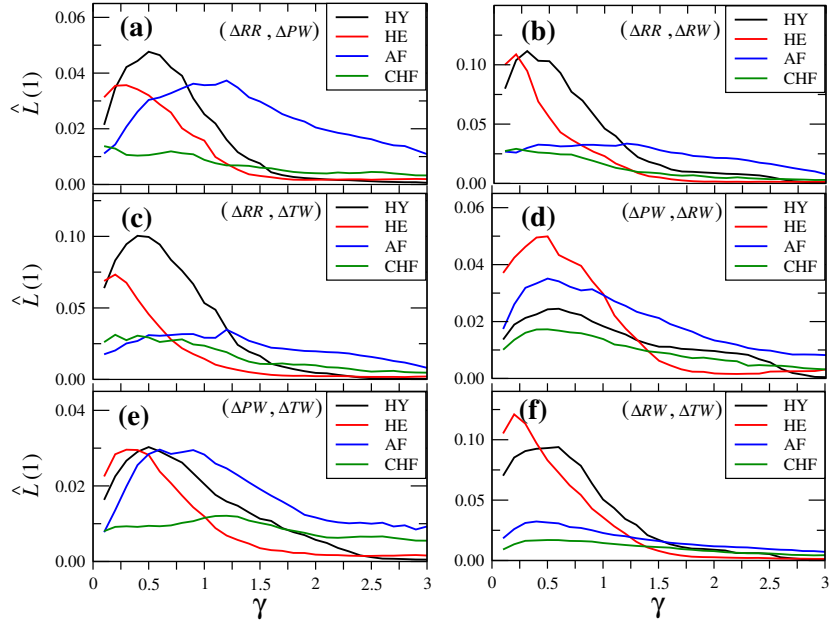


Figure 6: Lag-irreversibility function of joint variability signals, for different values of γ and for $\tau = 1$.

group. Similar results are obtained for $(\Delta RR, \Delta RW)$, $(\Delta RR, \Delta TW)$. All these results were obtained using $\gamma = 3/10$ for the three cases (see Table 1 below). Accordingly, we will use that value to estimate $\hat{L}(\tau)$ as a function of τ , in Section 4.3.

	$\Delta RR, \Delta PW$	$\Delta RR, \Delta RW$	$\Delta RR, \Delta TW$
HY	0.04217	0.11155	0.09455
HE	0.03566	0.09478	0.06762
CHF	0.01063	0.02748	0.02745
AF	0.02075	0.02993	0.02519
$\hat{L}(\tau = 1)$ and $\gamma = 3/10$			

Table 1: $\hat{L}(1)$ values for joint variables $(\Delta RR, \Delta PW)$, $(\Delta RR, \Delta RW)$ and $(\Delta RR, \Delta TW)$ using parameter $\gamma = 3/10$.

4.2. Lag-irreversibility of single variability signals

In Fig. 7 we can see the LI function of all the single variability signals as a function of τ . First of all, we can see in Fig. 7(a) that the LI function for ΔRR signal for the HY group is larger than for the other groups. The latter means

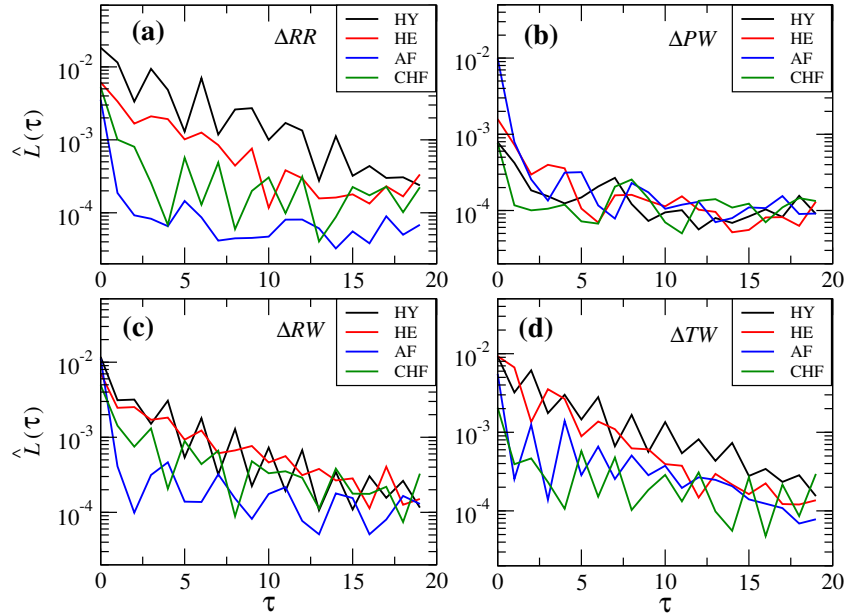


Figure 7: Lag-irreversibility function for single variability signals. We show the estimated LI functions, $\hat{L}(\tau)$ for (a) the heart rate variability, HRV, (b) the P-wave variability, ΔPV , (c) the R-wave variability, ΔRV , and (d) the T-wave variability, ΔTV .

that LI function for the heart rate variability signal might be used as a tool for discriminating between healthy young individuals from the other individuals under study. On the other hand, in Fig. 7 (b), (c) and (d), we observe that the LI functions for ΔPW , ΔRW and ΔTW do not allow us to reach the same conclusion. This is because the LI functions look very similar to each other among the different health conditions of individuals.

4.3. Lag-irreversibility of joint variability signals

Provided that the LI function for a single variability signal gives concluding results only for the ΔRR signal, we go further and consider a joint variability signal study. We expect that the data sequences contain physiological information that can be drawn from the different waves and variability signals. Moreover it is expected that these variability signals are not necessarily independent. This is because all the variability signals are, in some sense, different stages of the heartbeat of the same patient. Thus, the joint distribution carry some information of the dynamic process that cannot be obtained by solely looking at a single variability signal, such as the heart rate variability or a given wave amplitude variability.

In Fig. 8 we can see the resulting LI functions estimated from the six cases of pair variability signals: (a) $(\Delta RR, \Delta PW)$, (b) $(\Delta RR, \Delta RW)$, (c) $(\Delta RR, \Delta TW)$, (d) $(\Delta PW, \Delta RW)$, (e) $(\Delta PW, \Delta TW)$ and (f) $(\Delta RW, \Delta TW)$. In order

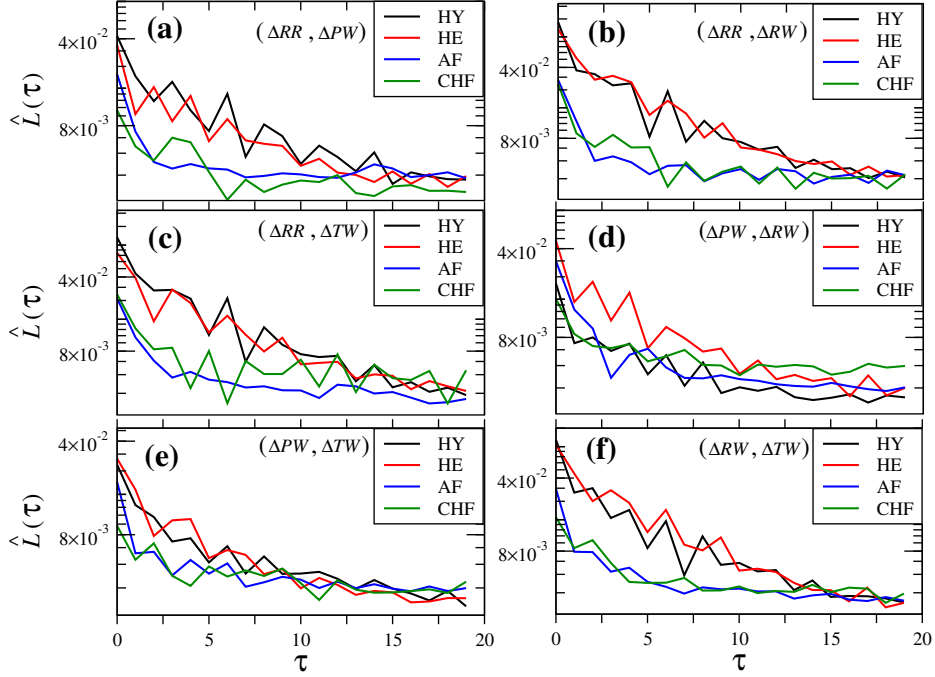


Figure 8: Lag-irreversibility functions for joint variability signals. We show the estimated LI function for joint signal, i.e., we consider two signals obtained from an ECG to obtain the LI function. There are only six possible pair joint signals: (a) $(\Delta RR, \Delta PW)$, (b) $(\Delta RR, \Delta RW)$, (c) $(\Delta RR, \Delta TW)$, (d) $(\Delta PW, \Delta RW)$, (e) $(\Delta PW, \Delta TW)$ and (f) $(\Delta RW, \Delta TW)$. For each case we estimate the LI function using the same encoding scheme as the one used in the single signal analysis, but jointly to take into account two signals. This encoding procedure results in a time series with a states space of nine symbols (see Section 3).

to estimate the LI function, we first performed the symbolic encoding scheme described in Section 3.3. For every pair of variability signals we obtained a symbolic time series made up of nine symbols, which we used to estimate the LI function using the method described in Section 3.4. In every panel of Fig. 8 we show the average LI function for every case of joint encoding. We can qualitatively observe in that figure that almost all cases allow us to distinguish between healthy individuals (young and elderly) from those individuals with an adverse health condition. A clear exception is observed in Fig. 8 (d) where the analysis of the pair $(\Delta PW, \Delta RW)$ fail to achieve a satisfactory discrimination between health conditions.

In order to quantitatively evaluate how accurate is this method for discriminating signals according to the health condition, we carried out an analysis through the so-called “Receiver operating characteristics graphs”, a technique that is referred to as *ROC analysis* [48].

4.4. ROC analysis

In this section we perform the ROC analysis, which is a binary method that allows to evaluate classifiers by means of a graph whose x - and y -axes represent the false-positive rate (denoted by g) and the true-positive rate (denoted by h), respectively [48]. To obtain the ROC curve, one chooses a threshold value of $\widehat{L}(\tau)$, above which the patient is predicted as a negative case (healthy, for instance) and below that threshold is considered to be as a positive case (unhealthy); that is, each threshold value would provide a proportion of true positive and true negative cases and their respective error proportions (false positive or false negative). Here we use this methodology to discriminate between eight pairs of patient groups, namely, H vs AF, H vs CHF, HY vs UH, HY vs HE, HE vs AF, HE vs CHF, HE vs UH and AF vs CHF; the group UH (unhealthy group), is made up of patients exhibiting any of the adverse health conditions, AF or CHF. With this analysis, we can evaluate the performance of the method by means of *area under the ROC curve* (AUC) [48, 49]. The AUC quantifies the area under the curve derived by plotting h vs g , and it takes values in the range $[1/2, 1]$; the closer the values of AUC are to 1, the greater the efficiency of the method to discriminate between the two groups. On the other hand, for practical reasons, we are interested in having the highest value of AUC, using the shortest signal time as possible. Consequently, in Figure 9, we show the AUC values as a function of time (heartbeats), for the eight pairs of patient groups and the six coding cases.

In Figure 9, we can see that with cases $(\Delta RR, \Delta PW)$, $(\Delta RR, \Delta RW)$ and $(\Delta RR, \Delta TW)$, which involve heart rate variability (ΔRR), it is possible to distinguish more clearly between the two groups of healthy patients (HY and HE) and those with medical conditions. Specifically, the two pairs of groups that show the greatest area under the curve ($AUC > 0.80$) are HY vs AF, HY vs CHF and HY vs UH, i.e., the methodology makes it possible to better distinguish between the group of healthy young patients from those with any of the medical conditions. The two pairs of groups that yield the lowest AUC value (≈ 5.5) are HY vs HE and AF vs CHF. In cases $(\Delta RR, \Delta RW)$ and $(\Delta RR, \Delta TW)$, the maximum AUC values (> 0.8) can be obtained using approximately 3 000 heartbeats. Regarding cases $(\Delta PW, \Delta RW)$, $(\Delta PW, \Delta TW)$ and $(\Delta RW, \Delta TW)$, they yield AUC values similar to the previous cases, but to distinguish between the unhealthy groups. Explicitly, the case $(\Delta PW, \Delta TW)$ allows reaching AUC values close to 0.8, to distinguish between the two groups with medical conditions.

The results described above are shown quantitatively in Tables 2 and 3, which show the AUC values obtained when using the full signal (5 000 heartbeats). We can see that values close to 0.80 or greater are reached with pairs HY vs CHF, HY vs AF and HY vs UH, when joint coding includes heart rate variability ($(\Delta RR, \Delta PW)$, $(\Delta RR, \Delta RW)$, $(\Delta RR, \Delta TW)$). In other words, we can say that, within the accuracy of our statistical analysis, this methodology is good enough for distinguishing between these groups of patients (up to 80% of the analyzed cases). An interesting case not yet reported in the literature is

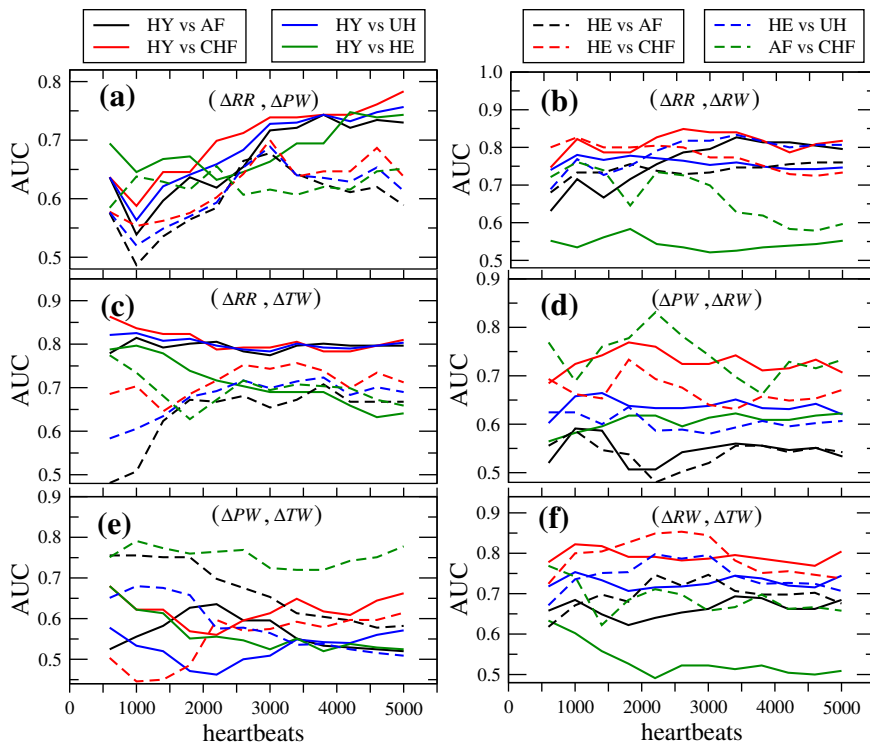


Figure 9: ROC analysis for the joint variability signal analysis.

the group of HE patients. We see that our methodology makes it possible to discriminate this group from healthy young or individuals with AF or CHF up to 70% of the analyzed cases ($AUC > 0.70$). As mentioned above, these maximum AUC values can be achieved using approximately 3000 heartbeats, which, in real time, is approximately 30 minutes of electrocardiographic signal.

5. Conclusions

The main purpose of this work is twofold. Firstly, we have introduced a method for assessing time-irreversibility from time series. This method is based on the concept of pairwise reversibility, which is essentially a symmetry property of the joint distribution of a lagged pair of state variables. Based on this property we define an “index” to quantify irreversibility, which we called lag-irreversibility (LI) function, which is the cornerstone of our analysis. We showed through a simple example of three-states Markov chain, how the LI function measures the degree of irreversibility of the process. In particular we argued that whenever the process is reversible the LI function, $L(\tau)$, is zero for all $\tau \in \mathbb{N}$, and that it must be positive if the process is irreversible.

	AUC		
	$\Delta RR, \Delta PW$	$\Delta RR, \Delta RW$	$\Delta RR, \Delta TW$
HY vs AF	0.7300	0.8266	0.7966
HY vs CHF	0.7833	0.8488	0.8300
HY vs UH	0.7566	0.7466	0.8033
HY vs HE	0.7433	0.5722	0.6611
HE vs AF	0.6200	0.7600	0.6677
HE vs CHF	0.6866	0.7733	0.7344
HE vs UH	0.6133	0.8066	0.7011
AF vs CHF	0.6511	0.6188	0.6988

Table 2: Area under the ROC curve for joint variability signals with ΔRR , using $\tau = 1$.

	AUC		
	$\Delta PW, \Delta RW$	$\Delta PW, \Delta TW$	$\Delta RW, \Delta TW$
HY vs AF	0.5333	0.5200	0.6844
HY vs CHF	0.7333	0.6622	0.8044
HY vs UH	0.6422	0.5711	0.7444
HY vs HE	0.6222	0.5244	0.5088
HE vs AF	0.5555	0.5822	0.7022
HE vs CHF	0.6711	0.6144	0.7555
HE vs UH	0.6066	0.5377	0.7244
AF vs CHF	0.7333	0.7777	0.6666

Table 3: Area under the ROC curve for joint variability signals, using $\tau = 1$.

Secondly, we proceeded to apply our method. That is, we evaluated the time-irreversibility of time series obtained from electrocardiograms of individuals grouped into four different categories depending on their health condition: *i*) healthy young, *ii*) healthy elderly, *iii*) atrial fibrillation and *iv*) congestive heart failure. Every ECG of each health condition was processed to obtain four variability signals: the heart rate, the P-wave, the R-wave, and the T-wave amplitude variabilities. Thus the ECG of each individual gave us four time series, which were used to perform the irreversibility analysis by means of the LI function. We showed that the lag-irreversibility function of every single variability signal was not efficient enough to discriminate among the different groups. Then we turn to the joint analysis of the variability signals, where it relies the main advantage of our method. For this purpose, we consider pairs of variability signals in order to evaluate the irreversibility as a joint process. We used a symbolization scheme of three states for every signal, which resulted in a symbolic time series of nine symbols (considering all the possible combinations). This time series was used to estimate the LI function for every group. We thus showed that the lag irreversibility function was able to distinguish be-

tween healthy individuals (groups *i* and *ii*) from individuals with an adverse health condition (groups *iii* and *iv*). The latter was made through a ROC analysis, which allows to determine the efficiency and sensibility of our test by a statistical analysis of false positives and negatives.

Acknowledgements

NM and CM thank CInC-UAEM for the warm hospitality during a short visit where part of the work was done. CM was financially supported by CONAHCYT grant Ciencia Básica (Ciencia de Frontera) No. A1-S-15528. RSG thanks IPICYT for its hospitality during academic visits and thanks CONAHCYT by partial financial support through grant number CF-2019-1327701.

References

- [1] M. Zanin, D. Papo, Algorithmic approaches for assessing irreversibility in time series: Review and comparison, *Entropy* 23 (11) (2021). doi:10.3390/e23111474.
URL <https://www.mdpi.com/1099-4300/23/11/1474>
- [2] A. Porporato, J. R. Rigby, E. Daly, Irreversibility and fluctuation theorem in stationary time series, *Phys. Rev. Lett.* 98 (2007) 094101. doi:10.1103/PhysRevLett.98.094101.
URL <https://link.aps.org/doi/10.1103/PhysRevLett.98.094101>
- [3] P. Gaspard, Time-reversed dynamical entropy and irreversibility in markovian random processes, *J. Stat. Phys.* 117 (3-4) (2004) 599–615.
- [4] V. Latora, M. Baranger, Kolmogorov-sinai entropy rate versus physical entropy, *Phys. Rev. Lett.* 82 (3) (1999) 520.
- [5] G. T. Landi, M. Paternostro, Irreversible entropy production: From classical to quantum, *Rev. Mod. Phys.* 93 (2021) 035008. doi:10.1103/RevModPhys.93.035008.
URL <https://link.aps.org/doi/10.1103/RevModPhys.93.035008>
- [6] P. Gaspard, *The Statistical Mechanics of Irreversible Phenomena*, Cambridge University Press, 2022. doi:10.1017/9781108563055.
- [7] E. Roldán, J. M. R. Parrondo, Entropy production and kullback-leibler divergence between stationary trajectories of discrete systems, *Phys. Rev. E* 85 (2012) 031129. doi:10.1103/PhysRevE.85.031129.
URL <https://link.aps.org/doi/10.1103/PhysRevE.85.031129>
- [8] C. Daw, C. Finney, M. Kennel, Symbolic approach for measuring temporal “irreversibility”, *Physical Review E* 62 (2) (2000) 1912.
- [9] A. Kathpalia, N. Nagaraj, Time-reversibility, causality and compression-complexity, *Entropy* 23 (3) (2021) 327.

- [10] M. Zanin, Assessing time series irreversibility through micro-scale trends, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 31 (10) (2021) 103118. doi:10.1063/5.0067342.
URL <https://doi.org/10.1063/5.0067342>
- [11] C. Diks, J. Van Houwelingen, F. Takens, J. DeGoede, Reversibility as a criterion for discriminating time series, *Phys. Lett. A* 201 (2-3) (1995) 221–228.
- [12] A. Porporato, J. Rigby, E. Daly, Irreversibility and fluctuation theorem in stationary time series, *Physical review letters* 98 (9) (2007) 094101.
- [13] E. Roldán, J. M. R. Parrondo, Estimating dissipation from single stationary trajectories, *Phys. Rev. Lett.* 105 (2010) 150607. doi:10.1103/PhysRevLett.105.150607.
URL <https://link.aps.org/doi/10.1103/PhysRevLett.105.150607>
- [14] L. Lacasa, A. Nunez, É. Roldán, J. M. Parrondo, B. Luque, Time series irreversibility: a visibility graph approach, *Eur. Phys. J. B* 85 (6) (2012) 1–11.
- [15] J. H. Martínez, J. L. Herrera-Diestra, M. Chavez, Detection of time reversibility in time series by ordinal patterns analysis, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28 (12) (2018) 123111. doi:10.1063/1.5055855.
- [16] J.-R. Chazottes, F. Redig, Testing the irreversibility of a Gibbsian process via hitting and return times, *Nonlinearity* 18 (6) (2005) 2477–2489. doi:10.1088/0951-7715/18/6/004.
URL <http://dx.doi.org/10.1088/0951-7715/18/6/004>
- [17] R. Salgado-García, C. Maldonado, Estimating entropy rate from censored symbolic time series: A test for time-irreversibility, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 31 (1) (2021) 013131. arXiv:<https://doi.org/10.1063/5.0032515>, doi:10.1063/5.0032515.
URL <https://doi.org/10.1063/5.0032515>
- [18] Y. Gao, I. Kontoyiannis, E. Bienenstock, Estimating the entropy of binary time series: Methodology, some theory and a simulation study, *Entropy* 10 (2) (2008) 71–99. doi:10.3390/entropy-e10020071.
URL <https://www.mdpi.com/1099-4300/10/2/71>
- [19] P. Guzik, J. Piskorski, T. Krauze, A. Wykretowicz, H. Wysocki, Heart rate asymmetry by poincare plots of rr intervals: 51 (4) (2006) 272–275. doi:10.1515/BMT.2006.054.
URL <https://doi.org/10.1515/BMT.2006.054>
- [20] R. Flanagan, L. Lacasa, Irreversibility of financial time series: A graph-theoretical approach, *Physics Letters A* 380 (20) (2016) 1689–1697.

doi:<https://doi.org/10.1016/j.physleta.2016.03.011>.
URL <https://www.sciencedirect.com/science/article/pii/S0375960116002401>

- [21] D.-Q. Jiang, D. Jiang, *Mathematical theory of nonequilibrium steady states: on the frontier of probability and dynamical systems*, Springer Science & Business Media, 2004.
- [22] J. Li, P. Shang, Time irreversibility of financial time series based on higher moments and multiscale kullback–leibler divergence, *Physica A: Statistical Mechanics and its Applications* 502 (2018) 248–255. doi:<https://doi.org/10.1016/j.physa.2018.02.099>.
URL <https://www.sciencedirect.com/science/article/pii/S0378437118301869>
- [23] J. Xia, P. Shang, J. Wang, W. Shi, Classifying of financial time series based on multiscale entropy and multiscale time irreversibility, *Physica A: Statistical Mechanics and Its Applications* 400 (2014) 151–158.
- [24] M. Zanin, B. Güntekin, T. Aktürk, L. Hanoglu, D. Papo, Time irreversibility of resting-state activity in the healthy brain and pathology, *Frontiers in physiology* 10 (2020) 1619.
- [25] R. Cofré, C. Maldonado, Information entropy production of maximum entropy markov chains from spike trains, *Entropy* 20 (1) (2018). doi: [10.3390/e20010034](https://doi.org/10.3390/e20010034).
URL <https://www.mdpi.com/1099-4300/20/1/34>
- [26] R. Salgado-García, Time-irreversibility test for random-length time series: The matching-time approach applied to dna, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 31 (12) (2021) 123126. arXiv:<https://doi.org/10.1063/5.0062805>, doi: [10.1063/5.0062805](https://doi.org/10.1063/5.0062805).
URL <https://doi.org/10.1063/5.0062805>
- [27] C. Cammarota, E. Rogora, Time reversal, symbolic series and irreversibility of human heartbeat, *Chaos, Solitons & Fractals* 32 (5) (2007) 1649–1654.
- [28] F. Hou, X. Huang, Y. Chen, C. Huo, H. Liu, X. Ning, Combination of equiprobable symbolization and time reversal asymmetry for heartbeat interval series analysis, *Physical Review E* 87 (1) (2013) 012908.
- [29] J. Piskorski, P. Guzik, Geometry of the poincaré plot of rr intervals and its asymmetry in healthy adults, *Physiological Measurement* 28 (3) (2007) 287. doi:[10.1088/0967-3334/28/3/005](https://doi.org/10.1088/0967-3334/28/3/005).
- [30] C. K. Karmakar, A. Khandoker, J. Gubbi, M. Palaniswami, Defining asymmetry in heart rate variability signals using a poincaré plot, *Physiological measurement* 30 (11) (2009) 1227.

- [31] C. Yan, P. Li, L. Ji, K. Yao, C. Karmakar, C. Liu, Area asymmetry of heart rate variability signal, *BioMedical Engineering OnLine* 16 (2017) 112. doi:10.1186/s12938-017-0402-3.
- [32] M. Costa, A. L. Goldberger, C.-K. Peng, Broken asymmetry of the human heartbeat: loss of time irreversibility in aging and disease, *Phys. Rev. Lett.* 95 (19) (2005) 198102.
- [33] P. A. Varotsos, N. V. Sarlis, E. S. Skordas, M. S. Lazaridou, Identifying sudden cardiac death risk and specifying its occurrence time by analyzing electrocardiograms in natural time, *Applied Physics Letters* 91 (6) (2007) 064106. arXiv:<https://doi.org/10.1063/1.2768928>, doi:10.1063/1.2768928. URL <https://doi.org/10.1063/1.2768928>
- [34] N. Sarlis, S.-R. Christopoulos, M. Bemplidaki, Change δs of the entropy in natural time under time reversal: Complexity measures upon change of scale, *EPL (Europhysics Letters)* 109 (1) (2015) 18002.
- [35] C. Maldonado, N. Merino Negrete, Irreversibility indices as discriminators of heart conditions from electrocardiographic signals, Preprint (2023). doi: <https://doi.org/10.1016/j.physleta.2016.03.011>.
- [36] A. Porta, K. R. Casali, A. G. Casali, T. Gneccchi-Ruscione, E. Tobaldini, N. Montano, S. Lange, D. Geue, D. Cysarz, P. Van Leeuwen, Temporal asymmetries of short-term heart period variability are linked to autonomic regulation, *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 295 (2) (2008) R550–R557, pMID: 18495836. doi: 10.1152/ajpregu.00129.2008. URL <https://doi.org/10.1152/ajpregu.00129.2008>
- [37] K. R. Casali, A. G. Casali, N. Montano, M. C. Irigoyen, F. Macagnan, S. Guzzetti, A. Porta, Multiple testing strategy for the detection of temporal irreversibility in stationary time series, *Phys. Rev. E* 77 (6) (2008) 066204.
- [38] A. Porta, G. D’addio, T. Bassani, R. Maestri, G. D. Pinna, Assessment of cardiovascular regulation through irreversibility analysis of heart period variability: a 24 hours holter study in healthy and chronic heart failure populations, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367 (1892) (2009) 1359–1375. arXiv:<https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2008.0265>, doi:10.1098/rsta.2008.0265. URL <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2008.0265>
- [39] F. Hou, J. Zhuang, C. Bian, T. Tong, Y. Chen, J. Yin, X. Qiu, X. Ning, Analysis of heartbeat asymmetry based on multi-scale time irreversibility

- test, *Physica A: Statistical Mechanics and its Applications* 389 (4) (2010) 754–760.
- [40] J. Piskorski, P. Guzik, Asymmetric properties of long-term and total heart rate variability, *Medical and biological engineering and computing* 49 (2011) 1289–97.
- [41] C. K. Karmakar, A. Khandoker, M. Palaniswami, Phase asymmetry of heart rate variability signal, *Physiological Measurement* 36 (2) (2015) 303. doi:10.1088/0967-3334/36/2/303. URL <https://dx.doi.org/10.1088/0967-3334/36/2/303>
- [42] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, H. E. Stanley, Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, *Circulation* 101 (23) (2000) e215–e220.
- [43] I. Silva, G. B. Moody, An open-source toolbox for analysing and processing physionet databases in matlab and octave, *Journal of open research software* 2 (1) (2014).
- [44] N. Iyengar, C. Peng, R. Morin, A. L. Goldberger, L. A. Lipsitz, Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics, *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 271 (4) (1996) R1078–R1084.
- [45] D. S. Baim, W. S. Colucci, E. S. Monrad, H. S. Smith, R. F. Wright, A. Lanoue, D. F. Gauthier, B. J. Ransil, W. Grossman, E. Braunwald, Survival of patients with severe congestive heart failure treated with oral milrinone, *Journal of the American College of Cardiology* 7 (3) (1986) 661–670.
- [46] G. Moody, A new method for detecting atrial fibrillation using rr intervals, *Computers in Cardiology* (1983) 227–230.
- [47] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, S. H. A. Chen, NeuroKit2: A python toolbox for neurophysiological signal processing, *Behavior Research Methods* 53 (4) (2021) 1689–1696. doi:10.3758/s13428-020-01516-y.
- [48] T. Fawcett, An introduction to roc analysis, *Pattern recognition letters* 27 (8) (2006) 861–874.
- [49] S. J. Mason, N. E. Graham, Areas beneath the relative operating characteristics (roc) and relative operating levels (rol) curves: Statistical significance and interpretation, *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 128 (584) (2002) 2145–2166.