

# Multiscale Information Storage of Linear Long-Range Correlated Stochastic Processes\*

Luca Faes<sup>†</sup>      Margarida Almeida Pereira<sup>‡</sup>

Maria Eduarda Silva<sup>§</sup>      Riccardo Pernice<sup>¶</sup>

Alessandro Busacca<sup>||</sup>      Michal Javorka<sup>\*\*</sup>      Ana Paula Rocha<sup>††</sup>

## Abstract

Information storage, reflecting the capability of a dynamical system to keep predictable information during its evolution over time, is a key element of intrinsic distributed computation, useful for the description of the dynamical complexity of several physical and biological processes. Here we introduce a parametric approach which allows to compute information storage across multiple time scales in stochastic processes displaying both short-term dynamics and long-range correlations (LRC). Our analysis is performed in the popular framework of multiscale entropy, whereby a time series is first 'coarse grained' at the chosen time scale through lowpass filtering and downsampling, and then its complexity is evaluated in terms of conditional entropy. Within this framework, our approach makes use of linear fractionally integrated autoregressive (ARFI) models to derive analytical expressions for the information storage computed at multiple time scales. Specifically, we exploit state space models to provide the representation of lowpass filtered and downsampled ARFI processes, from which information storage is computed at any given time scale relating the process variance to the prediction error variance. This enhances the practical

---

\*Accepted author's manuscript (AAM) to be published in [Physical Review E].  
[DOI:10.1103/PhysRevE.99.032115]

<sup>†</sup>luca.faes@unipa.it, Department of Energy, Information engineering and Mathematical models (DEIM), University of Palermo, Viale delle Scienze, Bldg. 9, 90128 Palermo, Italy

<sup>‡</sup>Faculdade de Ciências, Universidade do Porto, Rua Campo Alegre, Porto, Portugal, Centro de Matemática da Universidade do Porto (CMUP)

<sup>§</sup>Faculdade de Economia, Universidade do Porto, Rua Dr Roberto Frias, s/n, 4200-464 Porto, Portugal & CIDMA

<sup>¶</sup>Department of Energy, Information engineering and Mathematical models (DEIM), University of Palermo, Viale delle Scienze, Bldg. 9, 90128 Palermo, Italy

<sup>||</sup>Department of Energy, Information engineering and Mathematical models (DEIM), University of Palermo, Viale delle Scienze, Bldg. 9, 90128 Palermo, Italy

<sup>\*\*</sup>Department of Physiology, Comenius University in Bratislava, Jessenius Faculty of Medicine, Mala Hora 4C, 03601 Martin, Slovakia

<sup>††</sup>Faculdade de Ciências, Universidade do Porto, Rua Campo Alegre, Porto, Portugal, Centro de Matemática da Universidade do Porto (CMUP)

usability of multiscale information storage, as it enables a computationally reliable quantification of a complexity measure which incorporates the effects of LRC together with that of short-term dynamics. The proposed measure is first assessed in simulated ARFI processes reproducing different types of autoregressive (AR) dynamics and different degrees of LRC, studying both the theoretical values and the finite sample performance. We find that LRC alter substantially the complexity of ARFI processes even at short time scales, and that reliable estimation of complexity can be achieved at longer time scales only when LRC are properly modeled. Then, we assess multiscale information storage in physiological time series measured in humans during resting state and postural stress, revealing unprecedented responses to stress of the complexity of heart period and systolic arterial pressure variability, which are related to the different role played by LRC in the two conditions.

## 1 Introduction

Several physical and biological systems, such as climatic systems, econometric systems, the brain or the cardiovascular system, exhibit a rich dynamical activity that stems from the coexistence of self-sustained oscillators, interacting subsystems and feedback loops reacting to internal and external inputs Donges et al. (2009); Rosser (1999); Chialvo (2010); Kaplan et al. (1991). This multifaceted organization results in a complex system evolution over time, which is often revealed by the time course of a systemic variable like the global temperature, the stock market, the brain wave amplitude, or the heart period. In the recent past, several techniques have been proposed which aim at quantifying the richness of a dynamic process, usually indicated as “dynamical complexity” Pincus (1991); Porta et al. (1998); Richman and Moorman (2000); Valente et al. (2018); Porta et al. (2012); Costa et al. (2002). These methods have potentially important applications regarding both the characterization of the system state and the extraction of diagnostic parameters; for instance, a reduction of dynamical complexity may be associated with a reduced capability of subsystems to interact and, in physiological systems, has been proposed as a feature of pathologic behaviors Pincus (1994); Goldberger et al. (2002).

A common approach to assess the complexity of a time series intended as a realization of a dynamic process is that of quantifying the degree of irregularity, or unpredictability, of patterns extracted from the series. This approach has been pursued by studies proposing measures derived from linear or nonlinear prediction Porta et al. (2007b); Erla et al. (2011); Faes et al. (2015), or based on the concept of conditional entropy Pincus (1991); Porta et al. (1998); Richman and Moorman (2000), to quantify the dynamical complexity of a process. On the other hand a closely related complementary measure, which has been taking place recently in the frame of information theory, is the amount of information stored in a dynamic system. The so-called information storage is defined as the information contained in the past history of a stochastic process that can be used to predict its future Lizier et al. (2012). This measure has a

straightforward information-theoretic formulation as it quantifies the information shared between the current state of a process and its past states. Moreover, besides reflecting the regularity of a dynamic process intended as a complementary measure of its complexity, this quantity is recognized as one of the three key component processes constituting every act of information processing in a network of interacting systems (i.e., information storage, transfer, and modification) Wibral et al. (2014). As such, information storage is considered as a crucial aspect of the dynamics of several processes ranging from human brain networks Kitzbichler et al. (2009) to artificial networks Boedecker et al. (2009) and robot motion Ay et al. (2008), and has been successfully proposed to describe the regularity of brain Wibral et al. (2014), cerebrovascular Faes et al. (2013), cardiorespiratory Faes et al. (2015) and cardiovascular Faes et al. (2016) dynamics.

Another fundamental question about the analysis of complexity or regularity is the evaluation of these properties for stochastic processes displaying multiscale dynamical structures. It is indeed well known that a large variety of complex systems exhibit peculiar oscillatory activities which span several different temporal scales Ivanov et al. (1999); Costa et al. (2002); Wang et al. (2013); Faes et al. (2017c). Since measures of complexity/regularity based on conditional entropy/information storage Pincus (1991); Porta et al. (1998); Richman and Moorman (2000); Lizier et al. (2012) are not scale-specific, a common way to explore multiscale properties is to apply them after “rescaling” the observed process to focus on a specific range of temporal scales. This approach has been first proposed with the definition of the concept of multiscale entropy Costa et al. (2002), which consists in computing the conditional entropy of the observed process (through the Sample Entropy measure Richman and Moorman (2000)) after eliminating the fast temporal scales through of a lowpass filter followed by downsampling. Even though this approach lacks of solid theoretical grounds (e.g., it does not achieve a scale decomposition of variance, entropy, or conditional entropy), the resulting measures are of great practical relevance. Indeed, multiscale entropy is extremely popular and has been used, either in its original formulation or in refined versions Valencia et al. (2009), to assess multiscale complexity in a big amount of works within several different fields of science Costa et al. (2003, 2005); Ahmed and Mandic (2011); Humeau-Heurtier (2015); Wang et al. (2013); Faes et al. (2017a); Valencia et al. (2009); Faes et al. (2017c,b); Escudero et al. (2006); Catarino et al. (2011); Martina et al. (2011).

Nevertheless, the approach underlying the computation of multiscale entropy and related measures suffers also from practical shortcomings, the main being its inapplicability to short time series which stems from the difficulty of estimating entropies for multidimensional variables, exacerbated by issues related to the filtering and downsampling steps necessary for its calculation. In the present study, we overcome this limitation exploiting an approach based on linear autoregressive (AR) models Faes et al. (2017c). This approach computes complexity from the parameters of a state space (SS) model which represents the rescaled version of the original AR process obtained through the filtering and downsampling steps. The approach is implemented here for the computa-

tion of storage across multiple temporal scales. Importantly, we also extend the SS formulation to account for the effect of long range correlations. In fact, many natural phenomena exhibit slowly decaying serial correlations, or equivalently a spectrum with an hyperbolic behaviour at the origin,  $f(\lambda \sim |\lambda|^{-2d})$ . This behavior has been named long-memory or long range correlation Beran et al. (2012) and is related to self-similar processes, also known as fractals, and the so called  $1/f$  noise. Long range correlations have an effect upon the scaling properties displayed across time scales by a broad class of dynamic processes Ivanov et al. (1999); Bernaola-Galván et al. (2001); Chen et al. (2005); Xiong et al. (2017), and are thus a fundamental aspect of multiscale processes. Moreover, long range correlations are manifested also at short time scales and within the short time windows typically used for the computation of complexity measures, thus co-existing with short-term dynamics and having an impact on the assessment of their complexity Xiong et al. (2017). In spite of this, methods are lacking which are able to describe quantitatively the multiscale complexity or regularity of stochastic processes in the presence of long range correlations. Here, we fill this gap by providing theoretical formulations and practical estimation of multiscale information storage for stochastic processes with long range correlations, which are suitably described by fractionally integrated autoregressive (ARFI) models.

The framework for multiscale analysis of information storage developed in this work is implemented as a part of the Linear Multiscale Entropy (LMSE) Matlab toolbox. The toolbox is freely available for download from <http://www.lucafaes.net/LMSE.html>.

## 2 Methods

### 2.1 Linear Stochastic Processes with Long Range Correlations

We start recalling the classic parametric approach to the description of linear Gaussian stochastic processes exhibiting both short-term dynamics and long-range correlations, which is based on representing a discrete-time, zero-mean stochastic process  $X_n$ ,  $-\infty < n < \infty$ , as a fractionally integrated autoregressive (ARFI) process fed by uncorrelated Gaussian innovations  $E_n$ . The ARFI process takes the form:

$$A(L)(1-L)^d X_n = E_n \quad (1)$$

where  $L$  is the back-shift operator ( $L^i X_n = X_{n-i}$ ),  $A_L = 1 - \sum_{i=1}^p A_i L^i$  is an autoregressive (AR) polynomial of order  $p$  and  $(1-L)^d$  is the fractional differencing operator defined by Beran et al. (2012):

$$(1-L)^d = \sum_{k=0}^{\infty} G_k L^k \quad , \quad G_k = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)}, \quad (2)$$

with  $\Gamma(\cdot)$  denoting the gamma (generalized factorial) function. The parameter  $d$  in (1) determines the long-term behavior of the process, while the coefficients

of the polynomial  $A(L)$ , i.e.  $A_i, i = 1, \dots, p$ , allow description of the short-term dynamics. The ARFI model is stationary for  $-0.5 < d < 0.5$ , while it is nonstationary but mean reverting for  $0.5 \leq d < 1$ . The ARFI model may be extended to nonstationary settings by allowing  $d$  to be written as  $d = d_{LM} + D > 1$  with  $-0.5 < d_{LM} < 0.5$  and  $D \in \{1, 2, \dots\}$ . The most usual case occurs with  $D = 1$ , when the process is said to have a unit root: the ARFI( $p, d$ ) formulation is, then, used to model the increments of the series, that is the differences between consecutive observations. Note that the process defined in (1) is a particular case of the broader class of ARFIMA( $p, d, l$ ) processes, which also contains the class of autoregressive processes AR( $p$ ); here we restrict our analysis to the description of the ARFIMA( $p, d, 0$ ) process, which we denote as an ARFI( $p, d$ ) process.

The parameters of the ARFI model (1), namely the coefficients of  $A(L)$  and the variance of the innovations  $\Sigma_E$ , are obtained from process realizations of finite length first estimating the differencing parameter  $d$  by means of the Whittle semi-parametric local estimator Beran et al. (2012), then defining the filtered data  $X_n^{(f)} = (1 - L)^d X_n$ , and finally estimating the AR parameters from the filtered data  $X_n^{(f)}$  using the ordinary least squares method to solve the AR model  $A(L)X_n^{(f)} = E_n$ , with model order  $p$  assessed through the Bayesian information criterion Faes et al. (2012).

## 2.2 Measures of Information Storage

The information storage of a dynamical system that produces entropy with a non-zero rate is a quantity related to how much the system is able to share information during its evolution across time. Considering a system  $\mathcal{X}$  whose activity is defined by the stochastic process  $X$ , let us define as  $X_n$  the random variable describing the present state of the system, and as  $X_n^- = [X_{n-1} X_{n-2} \dots]$  the infinite-dimensional vector variable describing its past states. Then, the information stored in the system is defined as

$$S_X = I(X_n; X_n^-) = E \left[ \log \frac{p(x_n, x_{n-1}, \dots)}{p(x_n)p(x_{n-1}, x_{n-2}, \dots)} \right], \quad (3)$$

where  $I(\cdot; \cdot)$  denotes mutual information,  $p(\cdot)$  denotes probability density function, and the expectation is taken over several realizations  $(x_n, x_{n-1}, \dots)$  of the random variables  $(X_n, X_{n-1}, \dots)$ .

Even though information storage has been long recognized as an important aspect of the dynamics of complex systems, it has been formalized only recently as in Eq. (3) as the amount of information shared between the present and the past states of a dynamic process Lizier et al. (2012). From the point of view of the dynamic update of the state of a time-evolving system, the information storage is complementary to a well-known measure of system complexity quantified in terms of entropy rate, i.e. the conditional entropy of the present state of the system given its past states,  $C_X = H(X_n | X_n^-)$  which indeed can be related to

$S_X$  by simple information-theoretic rules:

$$S_X = H(X_n) - H(X_n|X_n^-) = H_X - C_X, \quad (4)$$

where  $H_X = H(X_n)$  is the entropy of the present system state. The conditional entropy is typically computed, from short realizations of the studied process, limiting the past history  $X_n^-$  to a small number of lags, and performing non-parametric estimation of the associated probability density functions Xiong et al. (2017); even with this simplification, multiscale computation of the conditional entropy cannot be reliably performed at long time scales. Here, to overcome these limitations we restrict the analysis to Gaussian processes for which exact expressions for conditional entropy and information storage are derived (see the next subsection for the multiscale formulation). Specifically, we exploit the relation stated in (4) and particularize it to the case of linear systems which can be fully described using an ARFI dynamic process in the form of Eq. (1). Specifically we note that, given the ARFI representation, the entropy of the present state of the process and the conditional entropy of the present given the past can be expressed analytically in terms of the variance of the process  $X_n$ ,  $\Sigma_X$ , and the variance of the innovations  $E_n$ ,  $\Sigma_E$ , as Faes et al. (2015); Cover et al. (1994); Barnett and Seth (2015):

$$H(X_n) = \frac{1}{2} \ln 2\pi e \Sigma_X, \quad (5a)$$

$$H(X_n|X_n^-) = \frac{1}{2} \ln 2\pi e \Sigma_E, \quad (5b)$$

from which the analytical formulation of the information storage follows immediately:

$$S_X = \frac{1}{2} \ln \frac{\Sigma_X}{\Sigma_E}. \quad (6)$$

To compute the information storage according to Eq. (6), we need to find an expression for the process variance  $\Sigma_X$  starting from the ARFI parameters  $d$  and  $A(L)$  obtained as described in Sect. II.A (which allows computation also of the innovation variance  $\Sigma_E$ ). To do this, first we approximate the ARFI process (1) with a finite order AR process by truncating the fractional integration part at a finite lag  $q$ :

$$(1 - L)^d \approx G(L) = \sum_{k=0}^q G_k L^k, \quad (7)$$

so that the ARFI( $p, d$ ) process can be rewritten as an AR( $p + q$ ) process:

$$B(L)X_n = E_n, \quad (8a)$$

$$B(L) = A(L)G(L) = \left(1 - \sum_{i=1}^p A_i L^i\right) \sum_{k=0}^q G_k L^k. \quad (8b)$$

The coefficients of the AR polynomial  $B(L) = 1 - \sum_{k=0}^{p+q} B_k L^k$  result from the multiplication of the two polynomials in (8b), which in the case  $q \geq p$  yields:

$$B_0 = 1 \quad ,$$

$$B_k = \begin{cases} -G_k + \sum_{i=1}^k G_{k-i} A_i, & k = 1, \dots, p \\ -G_k + \sum_{i=1}^p G_{k-i} A_i, & k = p+1, \dots, q \\ \sum_{i=0}^{p+q-k} G_{q-i} A_{i+k-q}, & k = q+1, \dots, q+p \end{cases} . \quad (9)$$

Once the ARFI process with parameters  $d$  and  $p$  is approximated by an AR process of order  $m = p + q$ , we derive the expression for the process variance using the theory of state space (SS) models Barnett and Seth (2015). The SS formulation of the AR( $m$ ) process of Eq. (8) is given by

$$Z_{n+1} = \mathbf{B}Z_n + KE_n \quad (10a)$$

$$X_n = CZ_n + E_n \quad (10b)$$

where  $Z_n = [X_{n-1} \cdots X_{n-m+1} X_{n-m}]^T$  is the  $m$ -dimensional state (unobserved) process and the vectors  $K$  and  $C$  and the matrix  $\mathbf{B}$  are defined as:

$$\mathbf{B} = \begin{bmatrix} B_1 & \dots & B_{m-1} & B_m \\ 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix}, K = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (11)$$

$$C = [B_1 \quad \dots \quad B_{m-1} \quad B_m]$$

The quantities in Eq. (11) are finally exploited to compute analytically the process variance  $\Sigma_X$  from the solution of the following discrete-time Lyapunov equation Barnett and Seth (2015):

$$\begin{aligned} \mathbf{\Omega} &= \mathbf{B}\mathbf{\Omega}\mathbf{B}^T + \Sigma_E K^T K \\ \Sigma_X &= C\mathbf{\Omega}C^T + \Sigma_E, \end{aligned} \quad (12)$$

from which the information storage is computed using Eq. (6).

### 2.3 Multiscale Information Storage

In this section we extend to multiple temporal scales the computation of information storage for stochastic processes which have an ARFI representation. To obtain the rescaled version of a stochastic process at the temporal scale defined by the scale factor  $\tau$ , the approach originally designed in Costa et al. (2002) corresponds to simply take the average of the process over  $\tau$  consecutive samples; this procedure has been refined later on Valencia et al. (2009); Faes et al.

(2017a) by recognizing that it actually entails the two subsequent steps of filtering the process with a lowpass filter with cutoff frequency  $f_\tau = 1/(2\tau)$ , and then downsampling the filtered process using a decimation factor  $\tau$ . According to this refined method, we first apply a linear finite impulse response (FIR) filter to the original process  $X_n$  obtaining the following filtered process:

$$X_n^{(r)} = D(L)X_n \quad , \quad (13)$$

where  $r$  denotes the filter order and the filter coefficients  $D(L) = \sum_{k=0}^r D_k L^k$  are chosen to set up a lowpass FIR configuration with cutoff frequency  $1/2\tau$ . The filtering step transforms the AR( $m$ ) process (which approximates the original ARFI( $p, d$ ) process) into an ARMA( $m, r$ ) process with moving average (MA) part determined by the filter coefficients:

$$B(L)X_n^{(r)} = D(L)B(L)X_n = D(L)E_n \quad . \quad (14)$$

Then, we exploit the connection between ARMA processes and state space (SS) processes Aoki and Havenner (1991) to evidence that the ARMA process (14) can be expressed in SS form as:

$$Z_{n+1}^{(r)} = \mathbf{B}^{(r)} Z_n^{(r)} + K^{(r)} E_n^{(r)} \quad (15a)$$

$$X_n^{(r)} = C^{(r)} Z_n^{(r)} + E_n^{(r)} \quad (15b)$$

where  $Z_n^{(r)} = [X_{n-1}^{(r)} \cdots X_{n-m}^{(r)} E_{n-1} \cdots E_{n-r}]^T$  is the  $(m+r)$ -dimensional state process,  $E_n^{(r)} = D_0 E_n$  is the SS innovation process, and the vectors  $K^{(r)}$  and  $C^{(r)}$  and the matrix  $\mathbf{B}^{(r)}$  are given by:

$$\begin{aligned} C^{(r)} &= [B_1 \quad \cdots \quad B_m D_1 \quad \cdots \quad D_r] \\ K^{(r)} &= [1 \quad \mathbf{0}_{1 \times (m-1)} \quad D_0^{-1} \mathbf{0}_{1 \times (r-1)}] \\ B^{(r)} &= \begin{bmatrix} C^{(r)} & & & & \\ \mathbf{I}_{m-1} & \mathbf{0}_{(m-1) \times (r+1)} & & & \\ \mathbf{0}_{1 \times (m+r)} & & & & \\ \mathbf{0}_{(r-1) \times m} & \mathbf{I}_{r-1} & \mathbf{0}_{1 \times (m+r)} & \mathbf{0}_{(r-1) \times 1} & \end{bmatrix} \end{aligned} \quad (16)$$

with  $\mathbf{I}_a$  and  $\mathbf{0}_{a \times b}$  indicating the  $a$ -dimensional identity matrix and the null matrix of dimension  $a \times b$ . The parameters of the SS process are the three quantities defined in (16) and the variance of the innovations  $\Sigma_{E^{(r)}} = D_0^2 \Sigma_E$ , whereby Eq. (15) defines an SS( $\mathbf{B}^{(r)}, C^{(r)}, K^{(r)}, \Sigma_{E^{(r)}}$ ) process.

The second step of the rescaling procedure is to downsample the filtered process in order to complete the multiscale representation. To do this, we make use of recent theoretical findings showing that the downsampled version of an SS process has itself an SS representation Solo (2016); Barnett and Seth (2015). Here, downsampling the SS process (15) with a factor  $\tau$  yields the process  $X_n^{(\tau)} = X_{n\tau}^{(r)}$ , which has the following SS representation:

$$Y_{n+1} = \mathbf{B}^{(\tau)} Y_n + W_n \quad (17a)$$

$$X_n^{(\tau)} = C^{(\tau)} Y_n + V_n \quad (17b)$$



where  $V_n$  and  $W_n$  are different white noise processes with variances  $\Sigma_W$  and  $\Sigma_V$  and covariance  $\Sigma_{VW}$ , respectively serving as innovations for the downsampled process  $X_n^{(\tau)}$  and for the state process  $Y_n$ . Thus, the process (17) is an SS( $\mathbf{B}^{(\tau)}, C^{(\tau)}, \Sigma_W, \Sigma_V, \Sigma_{VW}$ ) process whose parameters can be obtained as Solo (2016); Barnett and Seth (2015):

$$\begin{aligned}
\mathbf{B}^{(\tau)} &= \left(\mathbf{B}^{(r)}\right)^T \\
C^{(\tau)} &= C^{(r)} \\
\Sigma_V &= \Sigma_{E^{(r)}} \\
\Sigma_{VW} &= \left(\mathbf{B}^{(r)}\right)^{\tau-1} K^{(r)} \Sigma_{E^{(r)}} \\
\Sigma_W(1) &= \Sigma_{E^{(r)}} \left(K^{(r)}\right)^T K^{(r)}, \quad \tau = 1 \\
\Sigma_W(\tau) &= \mathbf{B}^{(r)} \Sigma_W(\tau-1) \left(\mathbf{B}^{(r)}\right)^T \\
&\quad + \Sigma_{E^{(r)}} \left(K^{(r)}\right)^T K^{(r)}, \quad \tau \geq 2
\end{aligned} \tag{18}$$

Then, the SS( $\mathbf{B}^{(\tau)}, C^{(\tau)}, \Sigma_W, \Sigma_V, \Sigma_{VW}$ ) process can be converted in a form similar to that of Eq. (15) which evidences the innovations, yielding the SS ( $\mathbf{B}^{(\tau)}, C^{(\tau)}, K^{(\tau)}, \Sigma_{E^{(\tau)}}$ ) process:

$$Z_{n+1}^{(\tau)} = \mathbf{B}^{(\tau)} Z_n^{(\tau)} + K^{(\tau)} E_n^{(\tau)} \tag{19a}$$

$$X_n^{(\tau)} = C^{(\tau)} Z_n^{(\tau)} + E_n^{(\tau)}, \tag{19b}$$

which provides the SS form of the filtered and downsampled version of the original ARFI( $p, d$ ) process. To move from (17) to (19) it is necessary to consider a discrete algebraic Ricatti equation Solo (2016); Barnett and Seth (2015):

$$\begin{aligned}
\mathbf{P} &= \mathbf{B}^{(\tau)} \mathbf{P} \left(\mathbf{B}^{(\tau)}\right)^T + \Sigma_W - \left(\mathbf{B}^{(\tau)} \mathbf{P} C^{(\tau)} + \Sigma_{VW}\right) \\
&\quad \cdot \left(C^{(\tau)} \mathbf{P} \left(C^{(\tau)}\right)^T + \Sigma_V\right)^{-1} \left(C^{(\tau)} \mathbf{P} \left(\mathbf{B}^{(\tau)}\right)^T + \right. \\
&\quad \left. + \left(\Sigma_{VW}\right)^T\right),
\end{aligned} \tag{20}$$

which leads, after solving for  $\mathbf{P}$ , to the derivation of the two remaining parameters in (19):

$$\Sigma_{E^{(\tau)}} = C^{(\tau)} \mathbf{P} \left(C^{(\tau)}\right)^T + \Sigma_V \tag{21a}$$

$$K^{(\tau)} = \frac{\mathbf{B}^{(\tau)} \mathbf{P} \left(C^{(\tau)}\right)^T + \Sigma_{VW}}{\Sigma_V}. \tag{21b}$$

Finally, the variance of the downsampled process can be computed analytically solving a discrete-time Lyapunov equation similar to that of Eq. (12)

$$\mathbf{\Omega} = \mathbf{B}^{(\tau)} \mathbf{\Omega} \left(\mathbf{B}^{(\tau)}\right)^T + \Sigma_{E^{(\tau)}} \left(K^{(\tau)}\right)^T K^{(\tau)} \tag{22a}$$

$$\Sigma_{X^{(\tau)}} = C^{(\tau)} \mathbf{\Omega} \left(C^{(\tau)}\right)^T + \Sigma_{E^{(\tau)}} \tag{22b}$$

The derivations above lead to compute analytically the parameters of the SS process of Eq. (19), which constitutes a rescaled version –derived through filtering (Eq. (15)) followed by downsampling (Eq. (17))– of the AR approximation (Eq. (8)) of the original ARFI process (Eq.(1)). Among the SS parameters, the ones relevant for the computation of the information storage are the variance of the downsampled process,  $\Sigma_{X(\tau)}$ , and the variance of the corresponding innovations,  $\Sigma_{E(\tau)}$ . These variances can be combined in a similar way to that of Eq. (6) to yield the expression of the information stored in the original process  $X_n$  when it is observed at scale  $\tau$ :

$$S_X(\tau) = \frac{1}{2} \ln \frac{\Sigma_{X(\tau)}}{\Sigma_{E(\tau)}}. \quad (23)$$

### 3 Numerical study

This section is devoted to assess the behavior of the proposed multiscale measure of information storage in stochastic processes with known dynamics. First the behavior is assessed by computing the value of  $S_X(\tau)$  for predetermined values of the parameters influencing short term dynamics and long range correlations of linear stochastic processes. Then the performance of the adopted estimators of  $S_X(\tau)$  is studied from finite length realizations of such processes in a simulation study.

#### 3.1 Sensitivity Analysis

Here we investigate the properties of multiscale information storage by varying the parameters which determine the dynamics of ARFI processes. These parameters are the differencing parameter  $d$  and the AR coefficients composing the polynomial  $A(L)$  in Eq. (1), which are related to long range correlations and short-term dynamics, respectively. Here, the strength of long-range correlations was varied changing the parameter  $d$  in the set  $\{0, 0.05, 0.4, 0.7\}$  so as to move from absent ( $d = 0$ ) to long lasting mean reverting ( $d = 0.7$ ) memory effects. Moreover the AR coefficients were set in order to generate stochastic oscillations with assigned frequency and spectral radius. This was achieved setting pairs of complex conjugate poles in the complex plane as the roots of the AR polynomial, where the modulus ( $\rho$ ) or the phase ( $\phi = 2\pi f$ , where  $f$  is the frequency) of the pole was changed to reproduce varying strength and frequency of the stochastic oscillations. Two configurations were considered: (a) an AR polynomial of order  $p = 2$ , with two poles having fixed frequency  $f = 0.1$  Hz and varying modulus  $\rho \in \{0, 0.5, 0.8, 0.9\}$ ; (b) an AR polynomial of order  $p = 4$  with two pairs of poles, the first with fixed modulus  $\rho_1 = 0.8$  and frequency  $f_1 = 0.1$  Hz, and the second with fixed modulus  $\rho_2 = 0.8$  and varying frequency  $f_2 \in \{0.15, 0.2, 0.25, 0.3\}$  Hz.

In each simulation, starting from the imposed theoretical values of the ARFI parameters, the analysis was performed according to the procedures described in Sect. II.B and II.C; the free parameters were set in accordance with the

literature, indicating  $q = 50$  as an appropriate value for truncating the ARFI process to a finite order Bardet et al. (2003), and  $r = 48$  as a viable choice for the order of the lowpass filter used to implement the change of scale Faes et al. (2017c). The results, obtained for time scales increasing from 1 to 50 ( $f_\tau$  decreasing from 0.5 to 0.01 Hz), are reported in Figs. 1,2. In general, long range correlations tend to bring information storage into the dynamic process, to an extent proportional to the long memory of the process: indeed, for an assigned time scale  $\tau$ ,  $S_X(\tau)$  tends to increase with the parameter  $d$ . This occurs both at the original time scale ( $f_\tau = 0.5$  Hz) and at longer time scales, regardless of the type of the AR process (Fig.1 and Fig. 2); the only exception is the presence of a strong stochastic oscillation ( $\rho = 0.9$  in Fig. 1), where an increase of  $d$  corresponds to a decrease of  $S_X$  at the original time scale. This finding has an implication for the evaluation of entropy measures on dynamic processes in which short term dynamics coexist with long range correlations Granger and Joyeux (1980); Xiong et al. (2017). In these situations one should remember that, since long memory properties have an important effect on the dynamics, such properties should be accounted for to make a proper evaluation of the complexity of the observed process; if one is interested in short-term complexity only Porta et al. (2012), long-range correlations should be removed prior to entropy analysis Xiong et al. (2017).

Another general result is that the information storage tends to decrease at decreasing  $f_\tau$ , as a result of the fact that lengthening the time scale corresponds to removing regular oscillatory components and making the process more complex (i.e., less predictable); at very long time scales the process is left with no predictable dynamics and  $S_X$  decays to zero. While the decrease of  $S_X$  with the time scale is monotonic in the absence of long range correlations (see the curves with  $d = 0$  in Figs. 1,2), the simultaneous presence of short and long memory effects may complicate the multiscale behavior of information storage. In fact, the ARFI process tends to store more information at intermediate time scales ( $f_\tau \approx 0.05$  Hz) than at lower time scales when long range correlations occur simultaneously with an appreciable stochastic oscillation ( $d = 0.4, 0.7$  and  $\rho = 0.8, 0.9$ , Fig. 1), or with a mismatch between the frequencies of two stochastic oscillations with the same amplitude (Fig. 2). These patterns were not revealed by the utilization of multiscale complexity measures not accounting for long range correlations Faes et al. (2017c). Therefore, it seems that the multiscale evaluation of complexity may benefit from the use of an approach able to model dynamical effects occurring at different temporal scales such as short and long range correlations.

We investigated also the effects of the approximation of the ARFI process with a finite order AR process, obtained setting a fixed value for the parameter  $q$  (see Eq. (7)). Fig. 3 reports the curves of multiscale information storage obtained in four representative AR parameter settings when assessed by the typical value  $q = 50$  Bardet et al. (2003) (solid lines) and with the reduced value  $q = 10$  (dashed lines). Overall, we note that excessive truncation leads to an underestimation of the information storage and to smooth of non-monotonic trends of the storage with the time scale. The bias is more evident for higher

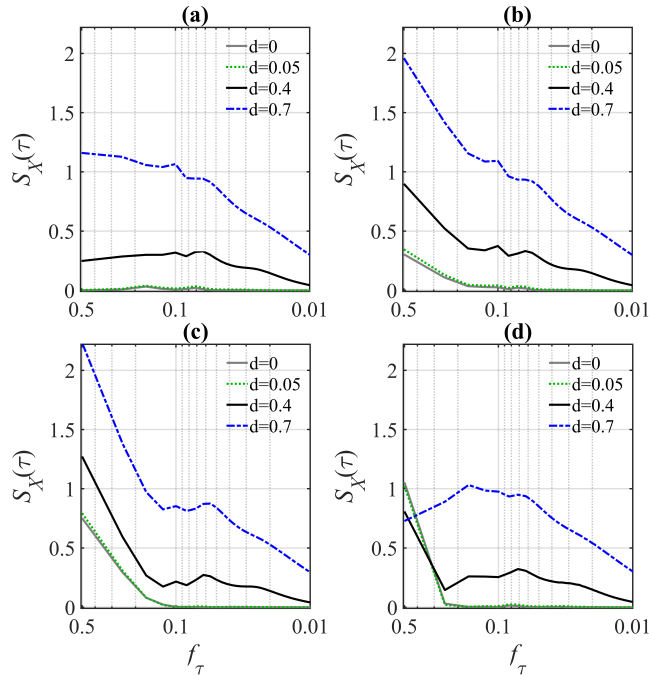


Figure 1: Theoretical profiles of multiscale information storage for simulated ARFI processes with varying amplitude of stochastic oscillations. Plots depict the information storage  $S_X$  computed as a function of the cutoff frequency  $f_\tau$  of the lowpass filter used to change the time scale for an ARFI process characterized by two complex conjugate poles with fixed phase  $\phi = 2\pi 0.1$  and variable modulus  $\rho = 0$  (a),  $\rho = 0.5$  (b),  $\rho = 0.8$  (c), and  $\rho = 0.9$  (d), and variable differencing parameter  $d = 0, 0.05, 0.4, 0.7$ .

values of the differencing parameter  $d$  at long time scales (lower  $f_\tau$ ). Therefore, high values of the parameter  $q$  are recommended to obtain a good approximation of the long memory properties of the observed process, so to limit the negative bias of information storage and to preserve the ability to discern multiscale patterns.

### 3.2 Finite Sample Performance

Here we describe the practical estimation of multiscale information storage computed for the processes simulated as in Sect. III.A. The focus of this analysis was on assessing the computational reliability of the proposed estimator in comparison with two alternative approaches for the multiscale assessment of dynamical complexity: (i) linear multiscale analysis, based on performing pure AR identification without the modeling of long range correlations Faes et al. (2017c), and (ii) refined multiscale entropy analysis, based on computing entropy measures

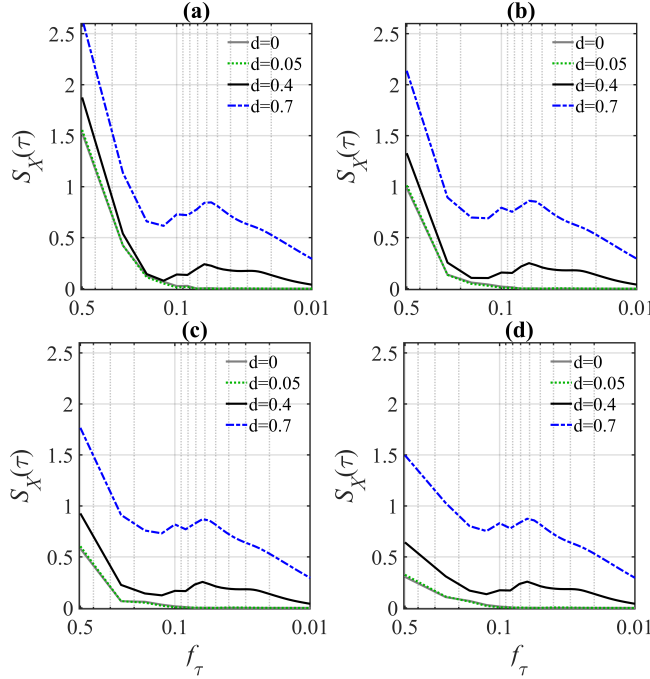


Figure 2: Theoretical profiles of multiscale information storage for simulated ARFI processes with varying frequency of stochastic oscillations. Plots depict the information storage  $S_X$  computed as a function of the cutoff frequency  $f_\tau$  of the lowpass filter used to change the time scale for an ARFI process characterized by a pair of complex conjugate poles with fixed modulus and phase ( $\rho_1 = 0.8, \phi_1 = 2\pi 0.1$ ), and another pair of complex conjugate poles with fixed modulus  $\rho_2 = 0.8$  and variable phase  $\phi_2 = 2\pi f_2$ , where  $f_2 = 0.15$  (a),  $f_2 = 0.2$  (b),  $f_2 = 0.25$  (c), and  $f_2 = 0.3$  Hz (d), as well as variable differencing parameter  $d = 0, 0.05, 0.4, 0.7$ .

after the practical implementation of rescaling executed through the application of a lowpass filter followed by downsampling Valencia et al. (2009). The analysis was performed for a representative example of the simulation described above, where the AR polynomial was obtained from a pair of complex conjugate poles with modulus  $\rho = 0.8$  and frequency  $f = 0.1$  Hz, and for values of the fractional differencing parameter  $d \in \{0, 0.4, 0.7\}$ . In each analyzed case, 100 realizations of the simulation were generated deriving the polynomial  $G(L)$  according to Eq. (2), truncating it to  $q = 50$  terms (Eq. (7)), and feeding the model of Eq. (8) with independent samples drawn from the standard normal distribution. Then, for each realization, multiscale information storage analysis was performed for time scales  $\tau = 1, \dots, 50$  ( $f_\tau = 0.5, \dots, 0.01$  Hz): (i) according to the procedure described in Sect. II (ARFI identification), applying a FIR lowpass filter of order  $r = 48$ ; (ii) according to linear multiscale complexity analysis, i.e. following the

procedure of Sect. II but after forcing  $d = 0$  in Eq. (6) (AR identification); (iii) according to refined multiscale complexity analysis, i.e. filtering the time series with a 6<sup>th</sup> order Butterworth lowpass filter, resampling the filtered series with a downsampling factor equal to  $\tau$ , and computing Sample Entropy Richman and Moorman (2000) on the downsampled time series with standard parameter setting (embedding dimension  $m = 2$ , similarity tolerance  $r = 0.2\Sigma_{E(\tau)}$ ). To allow comparison, the complexity measures  $C_X(\tau)$  derived from the complexity analyses (ii) and (iii) were converted into values of information storage exploiting the equivalence  $S_X(\tau) = 0.5 \ln 2\pi e - C_X(\tau)$  that holds for Gaussian processes. All estimates were compared also with the exact patterns of multiscale information storage obtained from the true values of the parameters.

First, we compared the distribution of  $S_X(\tau)$  estimated from realizations of  $N = 300$  samples with its theoretical values for the three estimation approaches. The results depicted in Fig. 4 show that all approaches return a biased estimate of the information storage, with the bias generally increasing with the differencing parameter  $d$  and with the time scale  $\tau$ . The bias is limited with the proposed approach based on ARFI models, as the true values of  $S_X$  is contained within the dispersion interval of 10<sup>th</sup> – 90<sup>th</sup> percentiles of the estimates (Fig. 4a,d,g). The linear multiscale method based on pure AR identification is highly biased, in the presence of long range correlations (Fig. 4e,h), at intermediate time scales ( $f_\tau \leq 0.1$  Hz) and becomes unreliable at longer time scales ( $f_\tau \leq 0.05$  Hz), returning very low values of information storage. Nevertheless, this method is highly reliable in the absence of long-range correlations (Fig. 4b), performing even better than the ARFI estimator which shows a certain bias (Fig. 4a); this small bias can be related to the variability in the estimation of the parameter  $d$ , which in turn shows non-negligible bias and variance for these estimates obtained with  $N = 300$  (the estimation improves for longer time series, results not shown). On the other hand, the traditional approach based on multiscale complexity analysis is highly unreliable at increasing time scales, as the estimates of  $S_X$  are strongly biased, display a variance that grows dramatically with the time scale, and could not even be computed for  $f_\tau \leq 0.1$  Hz. These results document the necessity of the proposed approach based on ARFI models to capture the dynamical complexity of processes showing both stochastic oscillations and long memory properties.

Next, we studied the dependence of the estimates of information storage on the sample size, repeating the analyses described above for time series of different length in the range  $N \in \{300, 512, 1024, 2048, 4096\}$ . As shown in Fig. 5, a general expected result is that the bias of  $S_X$  decreases at increasing the time series length. The improvement is such that the measure based on ARFI models becomes progressively more accurate at all time scales (Fig. 5a,d,g), while it does not help to obtain a good approximation of  $S_X$  at long time scales for the measure based on AR models (Fig. 5e,h). As to the measure based on multiscale complexity, the improvement brought by analyzing longer time series is only slight and not always clear (Fig. 5c,f,i), confirming the unsuitability of model-free approaches to assess dynamical complexity at long time scales.

## 4 Application to physiological processes

This section reports the application of multiscale information storage in the field of cardiovascular and cardiorespiratory physiology. In this field, it is well known that the dynamics of the cardiac, vascular and respiratory systems, typically assessed from the variability series of the heart period (HP), systolic arterial pressure (SAP) and lung volume (LV), reflect the activity of physiological mechanisms operating across multiple temporal scales. In particular, the assessment of HP and SAP dynamics over temporal scales ranging from seconds to a few minutes allows the detection of short-term cardiovascular regulation, and is typically accomplished through complexity measures like approximate entropy and Sample Entropy Pincus (1991); Richman and Moorman (2000), or even using the information storage Faes et al. (2015); Widjaja et al. (2015). In particular, two main physiological rhythms, associated to specific temporal scales, are commonly detected in short-term heart rate variability Camm et al. (1996); Shaffer and Ginsberg (2017): a low frequency oscillation (LF,  $\sim 0.1$  Hz), which is originated mainly by the sympathetic nervous system but may incorporate also parasympathetic effects and blood pressure regulation via the baroreflex Akselrod et al. (1981); Goldstein et al. (2011); and a high-frequency oscillation (typically occurring within the range 0.15-0.4 Hz), which reflects parasympathetic activity and occurs in synchrony with respiration according to the mechanism known as respiratory sinus arrhythmia Eckberg and Eckberg (1982); Grossman and Taylor (2007). On the other hand, it is also known that cardiovascular oscillations exhibit long-range correlations properties that are manifested in scaling behavior and power law correlations which are commonly assessed using fractal techniques Ivanov et al. (1999); Bernaola-Galván et al. (2001). These long memory properties are physiologically associated to the so-called very low frequency cardiovascular oscillations (VLF,  $< 0.04$  Hz), which are originated by very slow-acting physiological processes driven by the renin-angiotensin system Taylor et al. (1998) and possibly including physical activity as well as thermoregulatory and endothelial influences on the heart Shaffer and Ginsberg (2017); Kleiger et al. (2005).

Given this coexistence of short-term dynamics and long-range correlations, the evaluation of the dynamical complexity of cardiovascular and respiratory processes remains a challenge that can be thoroughly faced only employing multiscale approaches. Here we investigate how the dynamical complexity of HP, SAP and LV, assessed with our measure of information storage quantified from ARFI processes, varies across multiple time scales reflecting separate but simultaneously active physiological mechanisms. Moreover we address the issue of quantifying the impact of long-range correlations, typically manifested in short cardiovascular time series in terms of slow trends superimposed to the short-term dynamics, on the values of information storage computed from short cardiovascular recordings.

## 4.1 Experimental Protocol and Measurement of Physiological Time series

We consider the time series of HP, SAP and LV, interpreted as realizations of the stochastic processes descriptive of the cardiac, vascular and respiratory dynamics, measured in a group of 61 healthy subjects ( $17.5 \pm 2.4$  years old, 37 females) monitored in the resting supine position (SU) and in the upright position (UP) reached through passive head-up tilt Javorka et al. (2017). The acquired signals were the surface electrocardiogram (ECG), the finger arterial blood pressure recorded noninvasively by the photoplethysmographic method, and the respiration signal recorded through respiratory inductive plethysmography. For each subject and experimental condition, the values of HP, SAP and LV were measured on a beat-to-beat basis respectively as the sequences of the temporal distances between consecutive R peaks of the ECG, the maximum values of the arterial pressure waveform taken within the consecutively detected heart periods, and the values of the respiratory signal sampled at the onset of the consecutively detected heart periods. A detailed description of experimental protocol and signal measurement is reported in Ref. Javorka et al. (2017).

The analysis was performed on segments of  $N = 300$  consecutive points, free of artifacts and deemed as weak-sense stationary through visual inspection, extracted from the time series for each subject and condition. Three different approaches were followed to compute multiscale information storage: (i) the “eAR” approach, based on pure AR model identification, i.e. performing the whole procedure described in Sect. II after forcing  $d=0$  in Eq. (6); (ii) the “eARD” approach, based on pure AR identification as in (i), but applied to the filtered data  $X_n^{(f)} = (1 - L)^d X_n$ , after estimating the parameter  $d$  from the original time series; (iii) the “eARFI” approach, based on complete ARFI model identification, i.e., following the whole procedure described in Sect. II with  $d$  estimated from the original time series and considered in the computations. Pursuing these approaches we compare, respectively, (i) the traditional complexity analysis where long-range correlations are neither removed nor modeled, (ii) the analysis performed only on the short-term dynamics after removing long-range correlations, and (iii) the complexity analysis performed by modeling the long range correlations and considering them together with the short-term dynamics. Such a comparison is meant to infer the role of long-range correlations vs. that of short-term dynamics in contributing to the information storage and to its variation between conditions.

The ARFI model fitting each individual time series was identified first estimating the fractional differencing parameter  $d$  using the Whittle estimator, then filtering the time series with the fractional integration polynomial truncated at a lag  $q = 50$ , and finally estimating the parameters of the polynomial relevant to the short-term dynamics through least squares AR identification. The AR model order  $p$  was selected as the minimum of the BIC figure of merit Stoica and Selen (2004) in the range 2-16. Then, multiscale information storage was computed implementing a FIR lowpass filter of order  $r = 48$ , for time scales  $\tau$  in the range  $(1, \dots, 400)$ , which corresponds to lowpass cutoff frequencies



$f_\tau = (0.5, \dots, 0.00125)$  Hz.

Here the effects of SU and UP conditions on the information storage profiles are assessed at any assigned time scale by means of paired comparisons.

## 4.2 Results and Discussion

The results of the multiscale computation of information storage for the HP, SAP and LV time series are depicted respectively in Figs. 6, 7, and 8, reporting the distribution across subjects of the index  $S_X$  (left column) computed following the eAR (first row), eARd (second row) and eARFI (third row) estimation approaches and evaluated as a function of the time scale in the two analyzed physiological states (SU and UP). In each figure, results of the statistical analysis are also visualized (right column) reporting the mean and 95% confidence intervals of the paired difference between the values of  $S_X$  computed in the UP and SU conditions; a statistically significant variation from SU to UP is detectable at a given time scale if the confidence intervals do not encompass the zero line.

Fig. 6 reports the results of multiscale information storage analysis for the HP time series. Using the eAR method whereby long range correlations are not modeled (Fig. 6a), at scale 1 ( $f_\tau = 0.5$ ) the information stored in the HP process is significantly higher in the UP condition compared with SU. This reflects a widely known behavior of heart rate variability, whose complexity is known to decrease with head-up tilt due to an activation of the sympathetic nervous system which has a regularizing effect on the cardiac dynamics Porta et al. (2017, 2007a). Higher values of  $S_X$  during orthostatic stress are observed also at increasing the time scale, and are detectable up to  $f_\tau \sim 0.1$  Hz (Fig. 6b), reflecting the results obtained in Ref. Faes et al. (2017c) where a lower multiscale entropy is detected in the same data. Here we see also that the tilt-induced increase of the information storage is present also when the slow trends affecting HP, likely due to long-range correlations, are filtered out using the eARd method (Fig.6c); in this case the increase of  $S_X$  from SU to UP is significant also for intermediate time scales ( $0.05 \text{ Hz} < f_\tau < 0.1 \text{ Hz}$ , Fig. 6d)). This behavior is less evident when long range correlations are modeled, as  $S_X$  increases during UP only at scale 1, and is even reverted at longer time scales, as  $S_X$  decreases during UP for  $0.01 \text{ Hz} < f_\tau < 0.1 \text{ Hz}$  (Fig.6e,f). This behavior, documenting that the complexity of heart rate variability *increases* during tilt if observed at long time scales, has been previously observed using multiscale entropy Turianikova et al. (2011). Here, it becomes visible only modeling long-range correlations through the eARFI approach and indicates that long-range correlations are likely less important during head-up tilt. Thus, the utilization of the modeling approach proposed in this study suggests that postural stress augments the capability of HP to store information at low time scales but also diminish such capability at longer time scales.

Fig. 7 reports the results of multiscale information storage analysis for the SAP time series. According to the eAR method (Fig. 7a,b), moving from SU to UP the index  $S_X$  increases significantly at scale 1 ( $f_\tau = 0.5$  Hz) and decreases

significantly at scale 2 ( $f_\tau = 0.25$  Hz), confirming in terms of information storage the results reported in Ref. Faes et al. (2017c) based on a linear complexity measure. These two opposite behaviors of the information stored in the SAP process are here explained in terms of long-range correlations, which are removed or explicitly considered respectively through detrending or through performing ARFI identification. In fact, according to the eARd method (Fig. 7c,d),  $S_X$  is still significantly higher during UP when  $f_\tau = 0.5$  Hz, but is not significantly different from SU for any other value of  $f_\tau$ . On the contrary, according to the eARFI method (Fig. 7e,f),  $S_X$  does not show significant differences between SU and UP when  $f_\tau = 0.5$  Hz, but is significantly smaller when  $f_\tau = 0.25$  Hz. These results suggest that the higher capability of SAP to store information during tilt observed at scale 1 is related exclusively to short-term dynamics, while the lower storage capability observed at intermediate scales ( $f_\tau \sim 0.25$  Hz, where respiration-related components are suppressed) is driven by long-range correlations. Thus, head-up tilt induces scale-dependent variations in the complexity of arterial pressure, with higher complexity (lower  $S_X$ ) associated with slow oscillations, and lower complexity (higher  $S_X$ ) associated to the effects of respiration.

Fig. 8 reports the results of multiscale information storage analysis for the RESP time series. In this case we find that, using all methods, at short time scales the respiration process stores more information during UP than during SU, while at longer time scales the amount of information stored in the process does not change significantly with head-up tilt. This larger regularity of the respiration dynamics and its multiscale behavior confirm previous findings Valente et al. (2018); Faes et al. (2017c), further suggesting that long-range correlations do not have significant influence on the complexity of respiratory patterns. These results may be expected since respiration is usually strongly evident in the so-called high-frequency band ( $> 0.15$  Hz) Camm et al. (1996) and is thus filtered out almost entirely for time scales  $\geq 2$ .

To further elucidate the role played by long-range correlations in determining the information stored in the considered processes, we analyze the values of the differencing parameter  $d$  computed in the various conditions using the Whittle semi-parametric estimator Beran et al. (2012). Fig. 9 reports, for any given process and condition, the individual values of  $d$  plotted for each of the analyzed subjects, together with their 95<sup>th</sup> confidence intervals relevant to the zero level (derived from the asymptotic statistic given by Eq. (6) of Ref. Leite et al. (2013)). We find that the differencing parameter computed for the HP series and for the SAP series is statistically significant (i.e., outside of the confidence intervals) in a lower number of subjects in the UP condition (Fig. 9b,d) compared with SU (Fig. 9a,c); this seems not to be the case for the RESP time series (Fig. 9e vs. Fig. 9f). These results are confirmed by the application of a Student t-test for paired data applied to the distributions of  $S_X$  computed during SU and during UP, which returns  $p$ -values lower than the critical 0.05 level for HP and SAP (respectively  $p = 0.0037$  and  $p = 0.0192$ ), while the distributions were not significantly different for RESP ( $p = 0.377$ ). These changes correspond to situations in which the eARFI method detects a

statistically significant decrease of the information storage at intermediate/long time scales (Fig. 6 and Fig. 7). This suggests that when the importance of long-range correlations decreases for a time series (lower  $d$ ), the slow dynamics of the series become more complex (lower  $S_X$ ), supporting the argument that long-range correlations play a regularizing role for the process dynamics Xiong et al. (2017).

## 5 Conclusions

In this study, we have introduced a computationally reliable approach for the practical calculation of the information stored in a dynamic process when the process is observed at specific time scales. Quantification across time scales of the information storage and of its complementary complexity measure, i.e. the conditional entropy of the process measuring the new information produced by the process at each moment in time, is performed in the present and many other studies according to a specifically defined concept of “multiscale entropy” Costa et al. (2002); Humeau-Heurtier (2015). Such concept envisages the multiscale analysis of a dynamic process as a procedure based on (i) filtering the process to remove the fast temporal scales and highlight slower dynamics, and (ii) computing the complexity of the filtered process to identify the presence of repetitive patterns (low complexity) or unpredictable patterns (high complexity). This procedure provides a multiscale representation that is somewhat superficial in its theoretical and physical meaning, as it does not lead to a well-defined entropy decomposition whereby the information of the process results from the information of the process components relevant to different temporal scales. Nevertheless, such representation turns out to be very useful for practical purposes. In fact, in physiology and neuroscience as well as in other fields, quantification of multiscale entropy and information storage has rapidly become an extremely popular way to assess the contributions of different regulatory mechanisms to the dynamics of the observed process, and to infer from them the effects of physical, physiological or pathological alterations Costa et al. (2003, 2005); Ahmed and Mandic (2011); Wang et al. (2013); Faes et al. (2017a); Escudero et al. (2006); Catarino et al. (2011); Martina et al. (2011).

Thanks to its parametric formulation, the approach proposed in this work inherits the computational reliability of linear multiscale entropy Faes et al. (2017c), exploiting it for the assessment of regularity and – most importantly – allowing the simultaneous description of short-term dynamics and long memory properties. Our simulations show that the state space formulation implemented here, though being restricted to the description of linear dynamics, outperforms model-free multiscale complexity analysis Valencia et al. (2009) and, thanks to the incorporation of long-range correlations, leads to a more reliable evaluation of the information storage at long time scales if compared with linear multiscale entropy Faes et al. (2017c). Since long-range correlations are a fundamental aspect of multiscale dynamics, the present work opens the ways to a reliable computation of the dynamical complexity of several natural and man-made pro-

cesses where different mechanisms coexist, operating across multiple temporal scales. Here, the application to cardiovascular dynamics led to unprecedented physiological results, such as the observation that, at temporal scales compatible with sympathetic neural activity, postural stress blunts the capability of heart rate and arterial pressure variability to actively store information.

The main feature of the proposed framework for multiscale analysis is that it allows, thanks to the parametric formulation, to compute exact values of information storage based on the knowledge of the model parameters. On the other hand, such a feature also relates to a main limitation of our approach, i.e. the fact that our computations hold exactly only for linear Gaussian processes. In fact, departures from linearity inducing non-Gaussian distributions may generate dynamics which can be described only partially by an ARFI model, and this may lead to miss the multiscale properties to some extent. Even though the suitability of parametric approaches and ARFI models has been widely demonstrated in cardiovascular variability analysis Camm et al. (1996); Leite et al. (2013); Almeida et al. (2017), future studies should compare linear model-based and nonparametric model-free methods for multiscale analysis in order to clarify the role played by nonlinear dynamics under different experimental conditions and even in different applicative fields. Moreover, since a linear approach to multiscale analysis evidently has a relation with the spectral representation of the observed process, future studies should also define multiscale complexity measures in the frequency domain, and compare them with the time domain approach pursued by linear multiscale entropy and information storage in order to understand similarities and compatibilities.

As regards possible generalizations of the approach proposed in this work, the framework based on ARFI models may be extended to non-stationary settings by exploring nonstationary ranges for the differencing parameter or implementing time-varying formulations of AR or ARFI linear parametric models Sharman and Friedlander (1984); Faes et al. (2009). Moreover, future developments should also focus on extending the formulations proposed in this work to the multiscale representation of vector ARFI models, in order to attain a complete decomposition across time scales of the other constitutive elements of information processing in dynamical networks, i.e. information transfer and information modification Lizier et al. (2012); Wibral et al. (2014). Employed together, these generalizations would allow a combined multivariate, time-variant and multiscale assessment of the dynamics, opening the way for a truly complete characterization of the properties of coupled dynamical systems.

## Acknowledgements

Work partially supported by UID/MAT/00144/2019 (CMUP), UID/MAT/04106/2019 (CIDMA), which are funded by FCT with national (MEC) and European structural funds through the program FEDER, under PT2020, and project STRIDE-NORTE-01-0145-FEDER-000033 funded by ERDF-NORTE 2020)

## References

- Ahmed, M.U., Mandic, D.P., 2011. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Physical Review E* 84, 061918.
- Akselrod, S., Gordon, D., Ubel, F.A., Shannon, D.C., Berger, A., Cohen, R.J., 1981. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *science* 213, 220–222.
- Almeida, R., Dias, C., Silva, M.E., Rocha, A.P., 2017. Arfima-garch modeling of hrv: clinical application in acute brain injury, in: *Complexity and Nonlinearity in Cardiovascular Signals*. Springer, pp. 451–468.
- Aoki, M., Havenner, A., 1991. State space modeling of multiple time series. *Econ. Rev.* 10, 1–59.
- Ay, N., Bertschinger, N., Der, R., Güttler, F., Olbrich, E., 2008. Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B* 63, 329–339.
- Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Taqqu, M.S., 2003. Generators of long-range dependent processes: a survey. *Theory and applications of long-range dependence* , 579–623.
- Barnett, L., Seth, A.K., 2015. Granger causality for state-space models. *Phys. Rev. E* 91, 040101.
- Beran, J., Feng, Y., Ghosh, S., Kulik, R., 2012. *Statistics for Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer.
- Bernaola-Galván, P., Ivanov, P.C., Amaral, L.A.N., Stanley, H.E., 2001. Scale invariance in the nonstationarity of human heart rate. *Physical review letters* 87, 168105.
- Boedecker, J., Obst, O., Mayer, N.M., Asada, M., 2009. Initialization and self-organized optimization of recurrent neural network connectivity. *HFSP journal* 3, 340–349.
- Camm, A., Malik, M., Bigger, J., Breithardt, G., Cerutti, S., Cohen, R., Coumel, P., Fallen, E., Kennedy, H., Kleiger, R., et al., 1996. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *task force of the european society of cardiology and the north american society of pacing and electrophysiology. Circulation* 93, 1043–1065.
- Catarino, A., Churches, O., Baron-Cohen, S., Andrade, A., Ring, H., 2011. Atypical eeg complexity in autism spectrum conditions: a multiscale entropy analysis. *Clinical neurophysiology* 122, 2375–2383.
- Chen, Z., Hu, K., Carpena, P., Bernaola-Galvan, P., Stanley, H.E., Ivanov, P.C., 2005. Effect of nonlinear filters on detrended fluctuation analysis. *Physical Review E* 71, 011104.

- Chialvo, D.R., 2010. Emergent complex neural dynamics. *Nature physics* 6, 744.
- Costa, M., Goldberger, A.L., Peng, C.K., 2002. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* 89, 068102.
- Costa, M., Goldberger, A.L., Peng, C.K., 2005. Multiscale entropy analysis of biological signals. *Physical review E* 71, 021906.
- Costa, M., Peng, C.K., Goldberger, A.L., Hausdorff, J.M., 2003. Multiscale entropy analysis of human gait dynamics. *Physica A: Statistical Mechanics and its applications* 330, 53–60.
- Cover, T.M., Thomas, J.A., Kieffer, J., 1994. Elements of information theory. *SIAM Review* 36, 509–510.
- Donges, J.F., Zou, Y., Marwan, N., Kurths, J., 2009. Complex networks in climate dynamics. *The European Physical Journal Special Topics* 174, 157–179.
- Eckberg, D.L., Eckberg, M.J., 1982. Human sinus node responses to repetitive, ramped carotid baroreceptor stimuli. *American Journal of Physiology-Heart and Circulatory Physiology* 242, H638–H644.
- Erla, S., Faes, L., Tranquillini, E., Orrico, D., Nollo, G., 2011. k-nearest neighbour local linear prediction of scalp eeg activity during intermittent photic stimulation. *Medical engineering & physics* 33, 504–512.
- Escudero, J., Abásolo, D., Hornero, R., Espino, P., López, M., 2006. Analysis of electroencephalograms in alzheimer’s disease patients with multiscale entropy. *Physiological measurement* 27, 1091.
- Faes, L., Chon, K.H., Nollo, G., et al., 2009. A method for the time-varying nonlinear prediction of complex nonstationary biomedical signals. *IEEE transactions on biomedical engineering* 56, 205–209.
- Faes, L., Erla, S., Nollo, G., 2012. Measuring connectivity in linear multivariate processes: definitions, interpretation, and practical analysis. *Computational and mathematical methods in medicine* 2012.
- Faes, L., Marinazzo, D., Stramaglia, S., 2017a. Multiscale information decomposition: exact computation for multivariate gaussian processes. *Entropy* 19, 408.
- Faes, L., Nollo, G., Stramaglia, S., Marinazzo, D., 2017b. Multiscale granger causality. *Physical Review E* 96, 042150.
- Faes, L., Porta, A., Javorka, M., Nollo, G., 2017c. Efficient computation of multiscale entropy over short biomedical time series based on linear state-space models. *Complexity* 2017.

- Faes, L., Porta, A., Nollo, G., 2015. Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics. *Entropy* 17, 277–303.
- Faes, L., Porta, A., Nollo, G., Javorka, M., 2016. Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks. *Entropy* 19, 5.
- Faes, L., Porta, A., Rossato, G., Adami, A., Tonon, D., Corica, A., Nollo, G., 2013. Investigating the mechanisms of cardiovascular and cerebrovascular regulation in orthostatic syncope through an information decomposition strategy. *Autonomic Neuroscience* 178, 76–82.
- Goldberger, A.L., Peng, C.K., Lipsitz, L.A., 2002. What is physiologic complexity and how does it change with aging and disease? *Neurobiology of aging* 23, 23–26.
- Goldstein, D.S., Benthoo, O., Park, M.Y., Sharabi, Y., 2011. Low-frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes. *Experimental physiology* 96, 1255–1261.
- Granger, C.W., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis* 1, 15–29.
- Grossman, P., Taylor, E.W., 2007. Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological psychology* 74, 263–285.
- Humeau-Heurtier, A., 2015. The multiscale entropy algorithm and its variants: A review. *Entropy* 17, 3110–3123.
- Ivanov, P.C., Amaral, L.A.N., Goldberger, A.L., Havlin, S., Rosenblum, M.G., Struzik, Z.R., Stanley, H.E., 1999. Multifractality in human heartbeat dynamics. *Nature* 399, 461.
- Javorka, M., Czippelova, B., Turianikova, Z., Lazarova, Z., Tonhajzerova, I., Faes, L., 2017. Causal analysis of short-term cardiovascular variability: state-dependent contribution of feedback and feedforward mechanisms. *Medical & biological engineering & computing* 55, 179–190.
- Kaplan, D., Furman, M., Pincus, S., Ryan, S., Lipsitz, L., Goldberger, A., 1991. Aging and the complexity of cardiovascular dynamics. *Biophysical journal* 59, 945.
- Kitzbichler, M.G., Smith, M.L., Christensen, S.R., Bullmore, E., 2009. Broadband criticality of human brain network synchronization. *PLoS computational biology* 5, e1000314.

- Kleiger, R.E., Stein, P.K., Bigger Jr, J.T., 2005. Heart rate variability: measurement and clinical utility. *Annals of Noninvasive Electrocardiology* 10, 88–101.
- Leite, A., Paula Rocha, A., Eduarda Silva, M., 2013. Beyond long memory in heart rate variability: an approach based on fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 23, 023103.
- Lizier, J.T., Prokopenko, M., Zomaya, A.Y., 2012. Local measures of information storage in complex distributed computation. *Information Sciences* 208, 39 – 54.
- Martina, E., Rodriguez, E., Escarela-Perez, R., Alvarez-Ramirez, J., 2011. Multiscale entropy analysis of crude oil price dynamics. *Energy Economics* 33, 936–947.
- Pincus, S.M., 1991. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences* 88, 2297–2301.
- Pincus, S.M., 1994. Greater signal regularity may indicate increased system isolation. *Mathematical biosciences* 122, 161–181.
- Porta, A., Baselli, G., Liberati, D., Montano, N., Cogliati, C., Gneccchi-Ruscione, T., Malliani, A., Cerutti, S., 1998. Measuring regularity by means of a corrected conditional entropy in sympathetic outflow. *Biological cybernetics* 78, 71–78.
- Porta, A., Castiglioni, P., Rienzo, M.D., Bari, V., Bassani, T., Marchi, A., Takahashi, A.C., Tobaldini, E., Montano, N., Catai, A.M., et al., 2012. Short-term complexity indexes of heart period and systolic arterial pressure variabilities provide complementary information. *Journal of Applied Physiology* 113, 1810–1820.
- Porta, A., De Maria, B., Bari, V., Marchi, A., Faes, L., 2017. Are nonlinear model-free conditional entropy approaches for the assessment of cardiac control complexity superior to the linear model-based one? *IEEE Transactions on Biomedical Engineering* 64, 1287–1296.
- Porta, A., Gneccchi-Ruscione, T., Tobaldini, E., Guzzetti, S., Furlan, R., Montano, N., 2007a. Progressive decrease of heart period variability entropy-based complexity during graded head-up tilt. *Journal of applied physiology* 103, 1143–1149.
- Porta, A., Guzzetti, S., Furlan, R., Gneccchi-Ruscione, T., Montano, N., Malliani, A., 2007b. Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction. *IEEE Transactions on Biomedical Engineering* 54, 94–106.



- Richman, J.S., Moorman, J.R., 2000. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology* 278, H2039–H2049.
- Rosser, J.B., 1999. On the complexities of complex economic dynamics. *Journal of economic Perspectives* 13, 169–192.
- Shaffer, F., Ginsberg, J., 2017. An overview of heart rate variability metrics and norms. *Frontiers in public health* 5, 258.
- Sharman, K., Friedlander, B., 1984. Time-varying autoregressive modeling of a class of nonstationary signals, in: *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, IEEE. pp. 227–230.
- Solo, V., 2016. State-space analysis of granger-geweke causality measures with application to fmri. *Neural computation* 28, 914–949.
- Stoica, P., Selen, Y., 2004. Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine* 21, 36–47.
- Taylor, J.A., Carr, D.L., Myers, C.W., Eckberg, D.L., 1998. Mechanisms underlying very-low-frequency rr-interval oscillations in humans. *Circulation* 98, 547–555.
- Turianikova, Z., Javorka, K., Baumert, M., Calkovska, A., Javorka, M., 2011. The effect of orthostatic stress on multiscale entropy of heart rate and blood pressure. *Physiological measurement* 32, 1425.
- Valencia, J., Porta, A., Vallverdú, M., Clariá, F., Baranowski, R., Orłowska-Baranowska, E., Caminal, P., 2009. Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects. *IEEE Trans. Biomed. Eng.* 56, 2202–2213.
- Valente, M., Javorka, M., Porta, A., Bari, V., Krohova, J., Czippelova, B., Turianikova, Z., Nollo, G., Faes, L., 2018. Univariate and multivariate conditional entropy measures for the characterization of short-term cardiovascular complexity under physiological stress. *Physiological measurement* 39, 014002.
- Wang, J., Shang, P., Zhao, X., Xia, J., 2013. Multiscale entropy analysis of traffic time series. *International Journal of Modern Physics C* 24, 1350006.
- Wibral, M., Lizier, J., Vögler, S., Priesemann, V., Galuske, R., 2014. Local active information storage as a tool to understand distributed neural information processing. *Frontiers in neuroinformatics* 8, 1.
- Widjaja, D., Montalto, A., Vlemincx, E., Marinazzo, D., Van Huffel, S., Faes, L., 2015. Cardiorespiratory information dynamics during mental arithmetic and sustained attention. *PLoS One* 10, e0129112.
- Xiong, W., Faes, L., Ivanov, P.C., 2017. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Physical Review E* 95, 062114.

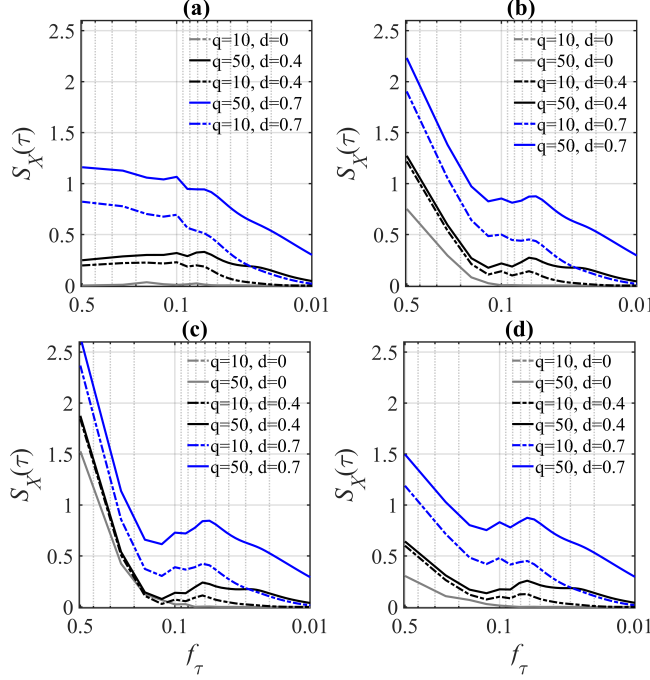


Figure 3: Dependence of the theoretical profiles of multiscale information storage on the approximation of a simulated ARFI process with a finite-order AR process. Plots depict the information storage  $S_X$  computed as a function of the cutoff frequency  $f_\tau$  of the lowpass filter used to change the time scale for an ARFI process characterized by two complex conjugate poles with phase  $\phi = 2\pi 0.1$  and modulus  $\rho = 0$  (a) or  $\rho = 0.8$  (b), or by two pairs of complex conjugate poles with modulus  $\rho_1 = \rho_2 = 0.8$  and phases  $\phi_1 = 2\pi 0.1, \phi_2 = 2\pi 0.15$  (c) or  $\phi_1 = 2\pi 0.1, \phi_2 = 2\pi 0.3$  (d). In each panel, results are plotted for values of the differencing parameter  $d = 0, 0.4, 0.7$  and two values for the truncation parameter:  $q = 10$  (dashed lines) and  $q = 50$  (solid lines).

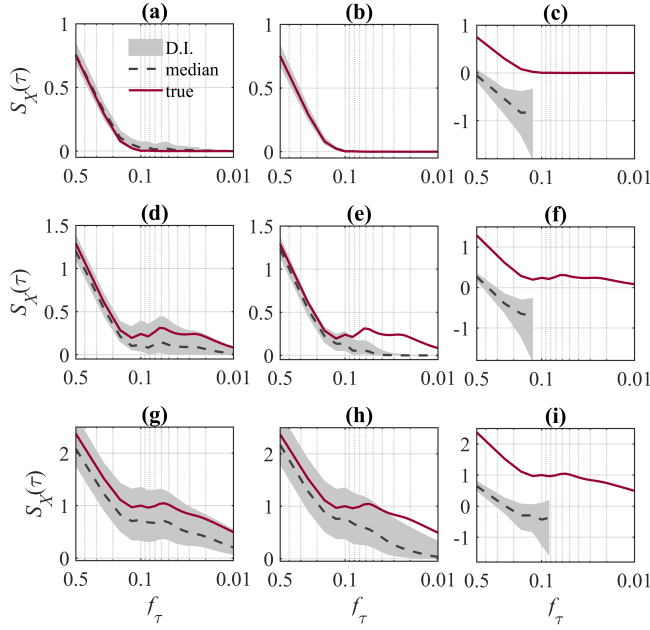


Figure 4: Estimation of multiscale information storage over finite length realizations of simulated ARFI processes. Plots depict the theoretical values (red) and the distributions (median and  $10^{th} - 90^{th}$  percentiles (dispersion interval, D.I.) over 100 realizations) of the information storage  $S_X$  computed as a function of the cutoff frequency  $f_\tau$  of the lowpass filter used to change the time scale for an ARFI process characterized by two complex conjugate poles with modulus and phase  $\rho = 0.8, \phi = 2\pi 0.1$  for values of the differencing parameters  $d = 0$  (a,b,c),  $d = 0.4$  (d,e,f) and  $d = 0.7$  (g,h,i), using the proposed ARFI method (a,d,g), the linear multiscale AR method Faes et al. (2017c) (b,e,h), and the refined multiscale complexity approach Valencia et al. (2009) (c,f,i).

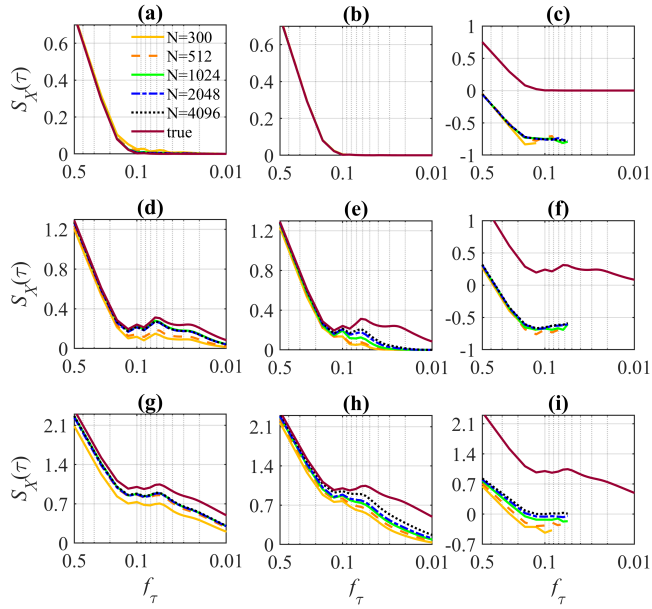


Figure 5: Estimation of multiscale information storage as a function of the length of simulated ARFI processes. Plots depict the theoretical values (red) and the average estimated values (median over 100 realizations, other colors) of the information storage  $S_X$  computed as a function of the cutoff frequency  $f_\tau$  of the lowpass filter used to change the time scale for an ARFI process characterized by two complex conjugate poles with modulus and phase  $\rho = 0.8, \phi = 2\pi 0.1$  for values of the differencing parameters  $d = 0$  (a,b,c),  $d = 0.4$  (d,e,f) and  $d = 0.7$  (g,h,i), using the proposed ARFI method (a,d,g), the linear multiscale AR method Faes et al. (2017c) (b,e,h), and the refined multiscale complexity approach Valencia et al. (2009) (c,f,i).

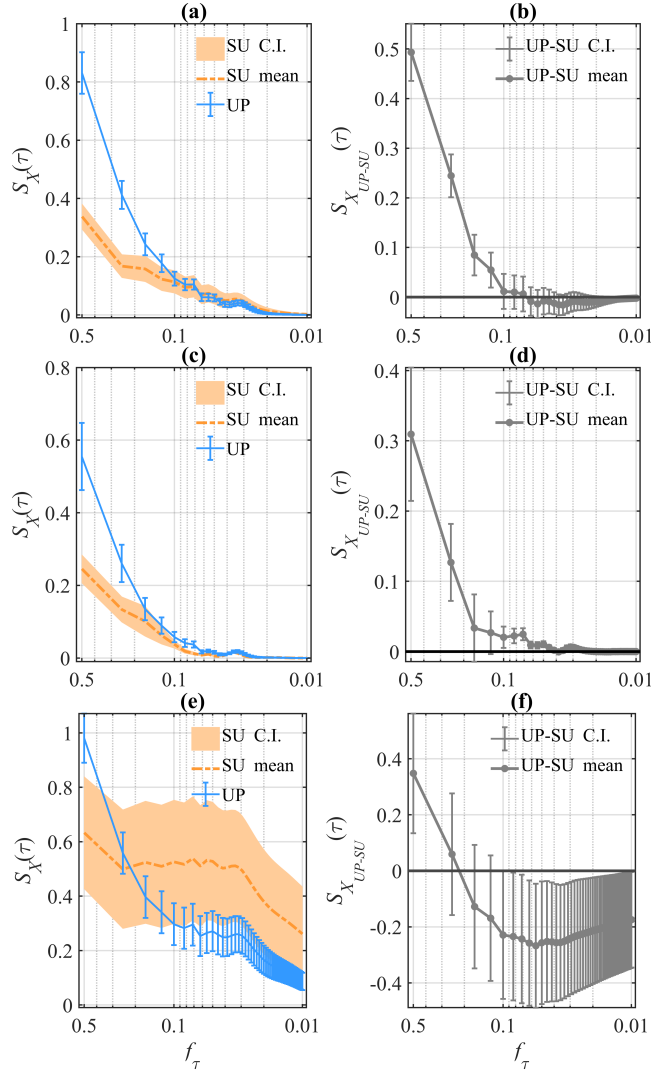


Figure 6: Multiscale Information Storage for the heart period time series. Plots report the confidence interval (C.I.) for the mean of the index of information storage computed across subjects using the eAR approach (a), the eARd method (c) and the eARFI method (e) as a function of the cutoff frequency of the rescaling filter in the supine (SU) and upright (UP) body positions. For each estimation method, the paired C.I. of UP-SU are also plotted in panels (b,d,f).

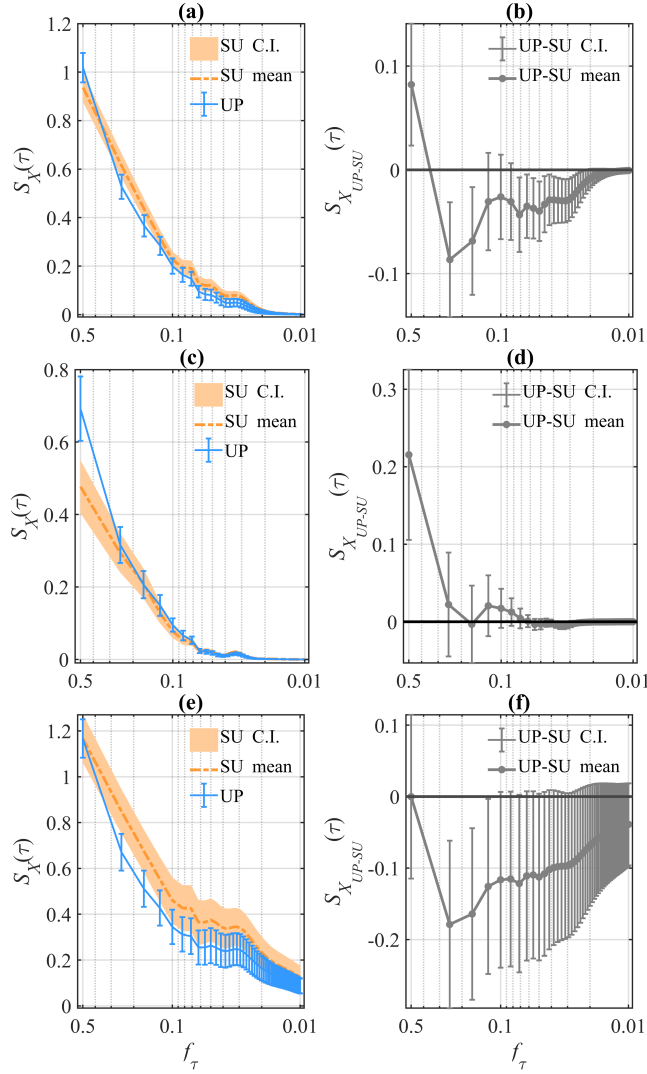


Figure 7: Multiscale Information Storage for the systolic pressure time series. Plots report the confidence interval (CI) for the mean of the index of information storage computed across subjects using the eAR approach (a), the eARd method (c) and the eARFI method (e) as a function of the cutoff frequency of the rescaling filter in the supine (SU) and upright (UP) body positions. For each estimation method, the paired CI of UP–SU are also plotted in panels (b,d,f).

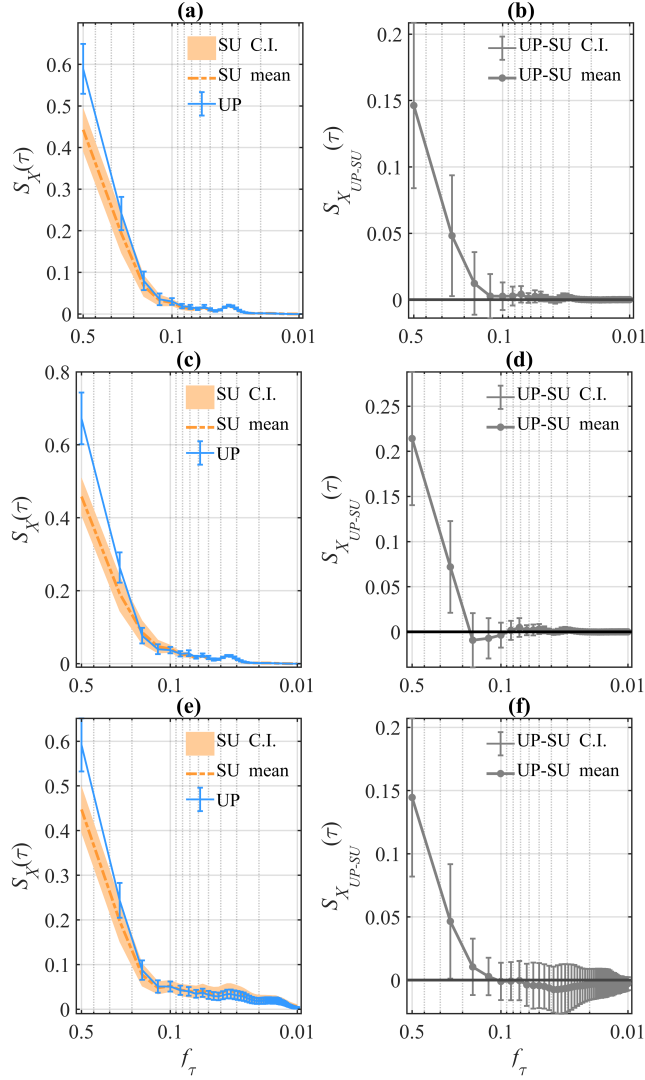


Figure 8: Multiscale Information Storage for the respiratory time series. Plots report the confidence interval (CI) for the mean of the index of information storage computed across subjects using the eAR approach (a), the eARd method (c) and the eARFI method (e) as a function of the cutoff frequency of the rescaling filter in the supine (SU) and upright (UP) body positions. For each estimation method, the paired CI of UP–SU are also plotted in panels (b,d,f).

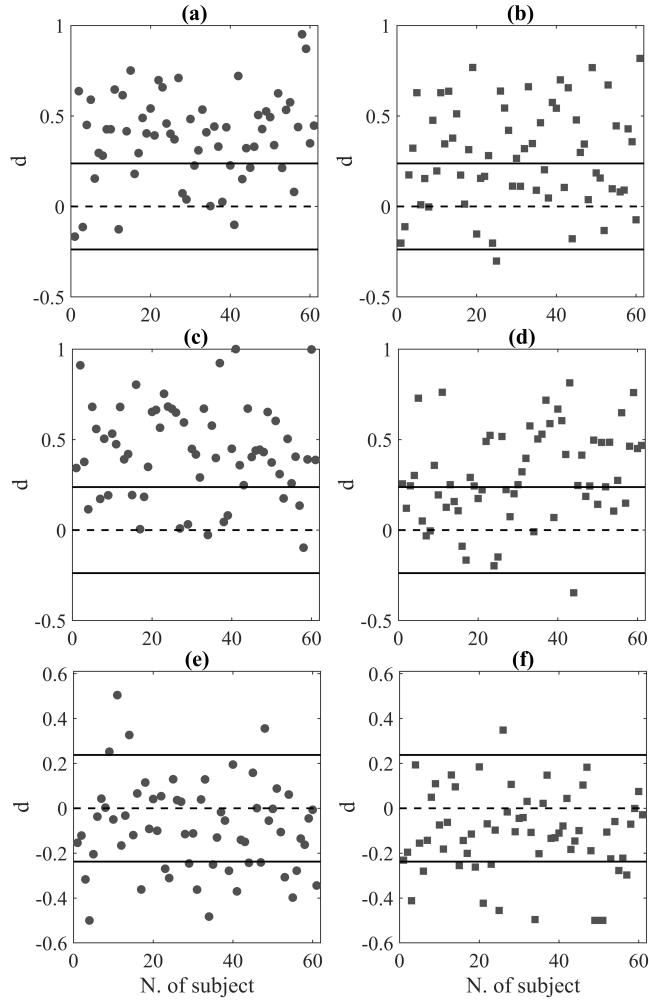


Figure 9: Values of the differencing parameter  $d$  estimated for each subject (index 1,...,61) for the heart period (a,b), systolic pressure (c,d) and respiration (e,f) time series in the supine (SU, a,c,e) and upright (UP, b,d,f) body positions. Each panel reports also the 95<sup>th</sup> confidence intervals (solid lines) of the distribution of  $d$  computed around the zero level (dashed line).