Entropy-Based Detection of Complexity and Nonlinearity in Short-Term Heart Period Variability under different Physiopathological States*

Luca Faes, *Senior Member, IEEE*, Riccardo Pernice, and Giandomenico Nollo, *Member, IEEE*

***Abstract*— We compare different estimators of a popular entropy-based nonlinear dynamic measure, i.e. the conditional entropy (CE), as regards their ability to assess the complexity and nonlinearity of short-term heart rate variability (HRV). The CE is computed using binning, kernel and nearest neighbor entropy estimators in HRV time series measured from young, old and post-myocardial infarction patients studied at rest and during orthostatic stress. We find that the three estimators yield similar patterns of CE, but different patterns of nonlinear dynamics, across groups and conditions. These results suggest that the strategy for CE estimation is not crucial for the quantification of complexity, but has a remarkable impact on the detection of nonlinear HRV dynamics.**

I. Introduction

Short-term heart rate variability (HRV), commonly assessed measuring the beat-to-beat durations of the heart period (RR interval of the electrocardiogram, ECG) over few minutes, reflects the activity of several simultaneously active physiological mechanisms [1]. As a result, the heart period time series is often analyzed with methods able to capture the complex nonlinear dynamics underlying HRV. Among these methods, information-theoretic approaches, typically exploiting the conditional entropy (CE) measure [2], have become very popular to assess the complexity of the HRV time series.

Besides the quantification of complexity, nonlinear analysis methods can be exploited also to detect the presence of nonlinear dynamics underlying HRV [3]. Previous works have suggested that distinct types of nonlinearity may be present in different physiopathological states, which can be revealed by different nonlinear analysis methods [4]. In this context, the present work is aimed at testing the effect that different estimation approaches have on the evaluation of complexity and nonlinearity in short-term HRV. To this end, we compare binning, kernel and nearest neighbor estimators of the CE applied to HRV time series measured during different experimental conditions (rest vs. orthostatic stress) and physiopathological states (healthy aging and myocardial infarction).

II. Methods

A. Estimation of Conditional Entropy

In a stochastic process $X=\{X_n,\ n=1,2,...\}$ describing the states visited by a dynamical system over time, the conditional entropy (CE) quantifies the new information generated by the system in the transition from its past m states $X_n^m=[X_{n-1},...,X_{n-m}]$ to its present state X_n, and is defined as

$$H(X_n|X_n^m) = H(X_n, X_n^m) - H(X_n^m) = -E[log\ p(x_n|x_n^m)] \quad (1)$$

where $p(x_n|x_n^m) = p(x_n, x_n^m)/p(x_n^m)$ is the conditional probability that X takes the value x_n at time n given that the values taken in the previous m states were $x_n^m=[x_{n-1},...,x_{n-m}]$, $E[\cdot]$ is the expectation, and $H(\cdot)$ denotes entropy [2]. The CE reflects the dynamical complexity of the process inasmuch as it quantifies the uncertainty about its present state which cannot be resolved by learning its past states.

In practical analysis, the CE is computed from a time series $x=\{x_n,\ n=1,...,N\}$ using appropriate entropy estimators [2]. The most intuitive of them is the binning estimator (*bin*), which is based on coarse graining the time series through uniform quantization and on computing entropy after approximating the probabilities with the frequency of visitation of the quantized states. If the amplitude range of the series is covered using Q quantization levels, each sample x_n is mapped into a symbol corresponding to the level to which it belongs. For a vector of dimension m, the m-dimensional space is partitioned into Q^m disjoint bins q_i, i=1,…, Q^m, and the probability of each bin is estimated as $p_i=N_i/N$, where N_i is the number of outcomes of the vector variable falling into the i-th bin q_i. Finally, each of the two entropy terms in (1) is computed averaging over all bins.

An alternative estimator computes the probability of a vector random variable (e.g., X_n or X_n^m) by centering kernel functions at each outcome of the variable, and then plugs the estimated probabilities into the desired entropy equation. Accordingly, the CE is computed as

$$H(X_n \mid X_n^m) = -\ln\left\langle p(x_n, x_n^m)\right\rangle / \left\langle p(x_n^m)\right\rangle, \quad (2)$$

where <·> denotes average over all outcomes. In this work, we use the Heaviside kernel with parameter r and the maximum norm to compute distances, so that probabilities are estimated counting the number of data points which fall within a cell of fixed size r from the reference point. With this choice, the kernel estimator (*ker*) of CE corresponds to the well-known Sample Entropy measure [5].

The third adopted method uses the k-nearest neighbor (*knn*) entropy estimator, whereby probabilities are computed from the statistics of the distances between neighboring points in the state space. This method adopts cells of variable size around each data point, where the size is determined by the distance between the reference point and its k-th nearest neighbor. Exploiting the theoretical derivations and the distance-projection methods proposed in [6] to limit the bias due

* Research supported by PRIN 2017, PRJ-0167, "Stochastic forecasting in complex systems", and PON R&I 2014-2020 AIM project no. AIM1851228-2, University of Palermo, Italy.

L. F. and R. P. are with the Department of Engineering, University of Palermo, Italy (corresponding author e-mail: luca.faes@unipa.it).

G. N. is with the Department of Industrial Engineering, University of Trento, 38123 Trento, Italy (e-mail: giandomenico.nollo@unitn.it).

to the different dimension of the variables involved in the estimation of different entropy terms, the CE is obtained as:

$$H(X_n \mid X_n^m) = -\psi(k) + \langle \psi(N_m) \rangle + \ln \langle \varepsilon_n \rangle, \qquad (3)$$

where ε_n is twice the distance from (x_n, x_n^m) to its k-th neighbor, N_m is the number of points whose distance from x_n^m is smaller than $\varepsilon_n/2$ and is ψ the digamma function.

In this study, the presence of nonlinear dynamics was assessed using the method of surrogate data. Specifically, for each entropy estimator, the CE computed for a given RR time series was compared with its distribution assessed from 100 surrogate time series under the null hypothesis of linear Gaussian dynamics with the same power spectrum and probability density of the original series. Each surrogate time series was obtained through the iterative amplitude adjusted Fourier Transform (IAAFT) method [3]. Then, the original time series was regarded as described by nonlinear dynamics if its estimated CE was lower than the 5-th percentile of the surrogate CE distribution.

B. Patients and Experimental Protocol

The analyzed dataset, described in detail in [7], consists of heart period time series measured in a group of myocardial infarction patients (*M*, 35 patients, 58.5±10.2 yrs), and in two control groups of healthy subjects: 19 young subjects (*Y*, 25.1±2.6 yrs) and 12 old subjects (*O*, 63.1±8.3 yrs). All subjects underwent a head-up tilt protocol such that they were monitored first in the resting supine position (R) and then in the 70º upright position reached by passive tilt (T). In the two conditions, an artifact-free stationary window of 300 consecutive RR intervals, measured from the lead II ECG, was considered for the analysis. The series were normalized to zero mean and unit variance prior to entropy computation. The CE measures were computed using standard settings for the free estimation parameters (dimension m=2; binning: Q=6; kernel: r=0.2; nearest neighbor: k=10) [2,4,5].

III. Results and Discussion

The distribution across subjects and conditions of the three CE measures is depicted in Fig. 1. Distributions were compared statistically using the two-way ANOVA followed by Student t-tests for pairwise comparison. All estimators detected a statistically significant decrease of the CE measured for the *Y* group moving from R to T, and a significantly higher CE during T for the *O* group compared to *Y*. A significant decrease of the CE during T was detected also for the *A* group using the *ker* and *knn* estimators. These results are in agreement with the known cardiovascular physiology documenting a decreased complexity of HRV during postural stress, related to sympathetic activation [3].

The relevance of nonlinearities, measured counting the subjects for which the CE was outside its distribution on surrogates, varied markedly depending on the estimator adopted. Indeed, nonlinear dynamics were detected in more than 50% of the subjects at R, and decreased below 50% during T, using the *bin* method; the percentage of nonlinear dynamics was lower and not evidently different across groups and between conditions using the *ker* method; a higher percentage of nonlinear dynamics was detected in all groups using the *knn* estimator, with a higher amount of nonlinearities detected in *Y* compared to *O* and *A*. These findings indicate that the type of entropy estimator has a big impact on the detection of

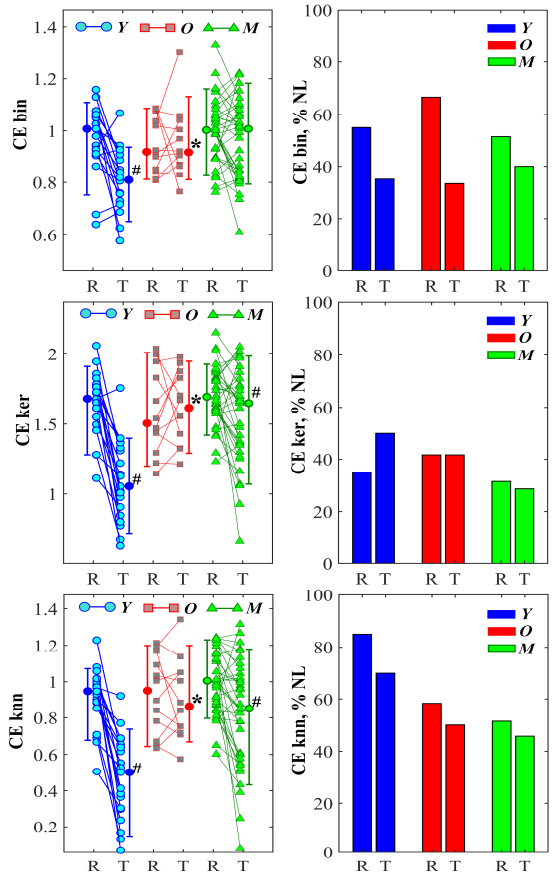

Figure 1. Distribution (median-percentiles and individual values) of the CE (a) and percentage of recordings with significant nonlinear dynamics (b) obtained using the three entropy estimators (bin, ker, knn) in the Young (*Y*), Old (*O*) and myocardial infarction (*M*) patients monitored at rest (R) and during tilt (T). #, p<0.05 R vs. T; *, p<0.05 *Y* vs. *O*

nonlinear HRV dynamics, with a role likely played by how the state space is partitioned (in disjoint cells of fixed size using binning, in overlapping cells of fixed size by kernels, and in overlapping cells of variable size by nearest neighbors) [2,3].

References

[1] M. A. Cohen and J. A. Taylor, "Short-term cardiovascular oscillations in man: measuring and modelling the physiologies," *J. Physiol.*, vol. 542, no. 3, pp. 669–683, 2002.

[2] W. Xiong, L. Faes, and P. C. Ivanov, "Entropy measures, entropy estimators and their performance in quantifying complex dynamics: effects of artifacts, nonstationarity and long-range correlations," *Phys. Rev. E*, vol. 95, p. 62114, 2017.

[3] A. Porta, S. Guzzetti, R. Furlan, T. Gnecchi-Ruscone, N. Montano, and A. Malliani, "Complexity and nonlinearity in short-term heart period variability: Comparison of methods based on local nonlinear prediction," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 1, pp. 94–106, 2007.

[4] A. Porta *et al.*, "Complexity analyses show two distinct types of nonlinear dynamics in short heart period variability recordings," *Front. Physiol.*, vol. 6, no. MAR, 2015.

[5] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 278, no. 6, pp. H2039-49, 2000.

[6] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, no. 6, p. 16, 2004.

[7] G. Nollo *et al.*, "Evidence of unbalanced regulatory mechanism of heart rate and systolic pressure after acute myocardial infarction," *Am. J. Physiol. - Hear. Circ. Physiol.*, vol. 283, no. 3 52-3, 2002.