

# Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis

Binbin Xu, Sabir Jacquir, Stéphane Binczak, Hussein M. Yahia, Rémi Dubois

▶ **To cite this version:**

Binbin Xu, Sabir Jacquir, Stéphane Binczak, Hussein M. Yahia, Rémi Dubois. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis. 42nd Annual Conference of Computing in Cardiology, Sep 2015, Nice, France. pp.4, 10.1109/CIC.2015.7410995 . hal-01163088v2

**HAL Id: hal-01163088**

**https://inria.hal.science/hal-01163088v2**

Submitted on 8 Sep 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis

B. Xu[1,2], S. Jacquir[3], S. Binczak[3,*], H. Yahia[2], R. Dubois[1]

## Abstract

***Background:*** *The animal models (in vitro or in vivo) provide an excellent tool to study heart diseases, among them the arrhythmia remains one of the most active research subjects.* ***Problems:*** *However, the arrhythmia inducing or treating effects in cardiac culture often happened long after the initial applications or in some relatively short time windows. Human-assisted monitoring is time-consuming and less efficient to capture rapidly the events.* ***Methods:*** *Electrocardiological signals are features by repetitive or similar patterns reflecting their intrinsic dynamics. Analyzing these patterns is of considerable interest to monitor/evaluate these dynamics' changes.* ***Aims:*** *Find appropriate (complexity) index which allows monitoring and classifying the arrhythmic events during the real-time signal acquisition in vitro or in clinical applications.*

## 1. Introduction

Cardiovascular diseases are one of the leading death causes. The electrophysiological signal of the heart provide an excellent tool to study the phenomena of cardiovascular diseases and then to help better understanding of their mechanisms. Animal models provide a good platform to model specific functional information of human heart, for example, the arrhythmia inducing and treatment. However, the events in the inducing / treating process would not appear immediately and often happen in some relatively short time windows or even long after the initial applications. It is then necessary to monitor the signal to capture these events in time. Human-assisted monitoring is time-consuming and less efficient. So an automatic monitoring/classification method for real-time use would be useful and necessary.

It is known that electrophysiological signals reflect essential biological information (different pathological conditions etc.). Repetitive or similar patterns or shifts present often in these signals, especially in the context of cardiac studies. Classification of these patterns would help not only to explore the physiological mechanisms but also to improve the diagnosis. Conventional signal processing methods from time-domain, frequency-domain could be used to analyze these patterns but with major issue of robustness. Because these signals could be strongly nonlinear, it is thus hard to apply blindly the conventional methods for pure engineer processing purpose.

The analysis should be based on their own properties. Since this type of signals represents often the output of a complicated network with both stochastic and deterministic components, the methods based on concepts from nonlinear dynamics and theory of deterministic chaos have been of considerable interest. However, since the reliability of methods like dimensions / exponents (correlation dimension, fractal dimensions, Lyapunov exponents, etc. ) require large data sets. The observations from these techniques on short experimental data are often questionable. For example, to get a relatively reliable Lyapunov exponents, at least $10^m - 30^m$ points ($m$ is the dimension) are needed [1]. The dimension $m$ for typical cardiac electrophysiological time series is generally no less than 4-6 [2]. So, application of these methods for short and noisy experimental data seemed questionable.

Methods from complexity analysis are developed to better process short and noisy data, like Approximate Entropy [3], Sample Entropy [4], Detrended Fluctuation Analysis [5], Hurst Exponent [6], Time Lagging [7] etc. Considering the fact that each method is based on certain mathematical conditions, it is in consequence advisable to compare these methods in order to find appropriate method allowing discriminating pathological signals from normal ones. Here comes then the motivation of this study: compare different complexity analysis methods and find the more appropriate ones which allows monitoring and classifying the arrhythmic events during the real-time signal acquisition from in vitro culture on the platform multi-electrodes array. This work-flow can also be extended to clinical applications.

## 2. Models and Methods

### Data preparation

The in vitro model used here is monolayer cardiac cell

[1]LIRYC, L'Institut de RYthmologie et modélisation Cardiaque, Bordeaux France. [2]Team GéoStat, INRIA Bordeaux Sud-Ouest, France. [3]LE2I UMR6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France. *stbinc@u-bourgogne.fr

culture which is prepared with cardiomyocytes (CM) from new-born (less than 4 days) rats' heart (culture preparing details in [8]). Different from other rats' CM models in literature, this model shares very similar physiological characteristic to those in situ. This CM culture allows reproducing in vitro a wide range of cardiac pathological conditions such as ischemia reperfusion, the radical stress or thermal shock, and any combination of these conditions.

To study the electrophysiological properties of the culture, a multi-electrodes array (MEA) system is used (platform details in [9]). The CM culture is directly prepared on the MEA dishes, the real-time recorded extracellular field potential (EFP) has a rather good spatial-temporal solution (60 electrodes in form of $8 \times 8$ array, of diameter $30\,\mu$m, inter-electrodes distance $100\,\mu$m; cell's size : of diameter $15\,\mu$m, length $100\,\mu$m). The EFP signals correlate highly with the action potential at depolarization and repolarizing phase which makes interpreting the result at *action potential level* possible [10].

**Complexity analysis methods**

• Approximate Entropy (ApEn) is designed to examine the regularity or fluctuations of a time series, especially for short, noisy times series [3]. To calculate ApEn, the time series $x(t)$ is first converted into $m$-dimensional vector $X(i,m)$. $C_i^m(r) = [\text{number of } x(j) \text{ such that } d[x(i),x(j)] \leq r]/(N-m+1)$ measures the regularity within the scale level $r$ in given dimension $m$ (in an optimal way, $m$ can be estimated by the method *false nearest neighbor*). Defining $\phi^m(r) = (N-m+1)^{-1}\sum_{i=1}^{N-m+1}\log(C_i^m(r))$, the ApEn can then be obtained by $\mathrm{ApEn} = \lim_{\mathrm{N}\to\infty}\left[\phi^{\mathrm{m}}(\mathrm{r}) - \phi^{\mathrm{m+1}}(\mathrm{r})\right]$ Generally, a time series containing repetitive patterns has relatively small ApEn values; otherwise, the ApEn values would be higher.

• Sample Entropy (SampEn) is a slight modification of ApEn. ApEn uses a template-wise approach to calculating this average logarithmic probability. When calculating $C_i^m(r)$, ApEn requires that each template contributes a defined, nonzero probability which is not always possible. To avoid this constraint, self-matching is thus allowed at each sequence which creates biased estimation. SampEn allows reducing this bias by excluding all self-matches in the calculation of the probability [4]. Compared to ApEn, SampEn is defined as $\mathrm{SampEn} = \lim_{\mathrm{N}\to\infty}\left[-\ln[\phi^{\mathrm{m}}(\mathrm{r})/\phi^{\mathrm{m+1}}(\mathrm{r})]\right]$.

• Detrended Fluctuation Analysis (DFA) was initially developed to quantify the long-range power-law correlation of DNA sequences [5]. It is today one of the most used methods to determine signal's statistical self-similarity. The original time series $x(t)$ is first converted into unbounded series $y(k) = \sum_{i=1}^{t}(x_i - \langle x_i\rangle)$ Then the obtained series (length $N$) are divided into $\frac{N}{L}$ boxes with different length $L$. Then the "local trend" $y_m(k)$ is calculated by a local least squares straight-line fit in each box [11]. Detrending is carried out by subtracting the local trend $y_m(k)$ in each segment. The root-mean-square fluctuation $F(m)$ of the resulting series is then obtained. Using log-log plot ($\log F(m)$ and $n$), the fluctuations $\alpha$ can be characterized by the slope of the line relating $\log F(m) \propto n^{\alpha}$.

• Hurst Exponent (HEx) is another general method to measure long term memory and correlation properties of time series. It quantifies the signal's relative tendency either to regress strongly to the mean or to cluster in a direction, in other words : in terms of the asymptotic behavior of the rescaled range ($\frac{R(n)}{S(n)}$, $R(n)$ : the range of the first $n$ values; $S(n)$ : their standard deviation) as a function of the time span $T$ of a time series [6]: $H = \log(R(n)/S(n))\log(T)$.

• Modified Hurst Exponent ($\mathrm{HEx_M}$). Like the action potential or electrocardiogram, the extracellular potentials are characterized by strong periodicity and fluctuations. In case of arrhythmia, this periodicity is disturbed or even broken. Cumulative sum (*cumsum*) allows detecting the change in a time series, especially for extracellular potentials. The obtained time series by *cumsum* are smoother but with the same statistical properties. So, instead of using the original time series $x(t)$ as input for the method HEx, we propose here to use the cumulative sum of the time series.

• The method Time Lagging (TLag) plays an essential role in phase space reconstruction allowing transforming one-dimensional analysis into topological problem [7]. Since TLag determines the coordinates of the reconstructed space, it can serve as a complexity index in this sense. The general method to obtain TLag is by auto-correlation function. According to the correlation theorem, under certain conditions, two observations of the same process $(x(t_1), x(t_2))$ can be regarded as identical, if $x(t_1) = x(t_2)$. To examine the self-similarity of a signal and the $\tau$ shifted signal with auto-correlation function : $C(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)\mathrm{d}t$ . The bigger is $C(\tau)$, more $x(t+\tau)$ looks like $x(t)$. $\tau$ is chosen as $C(\tau) = \frac{C(\tau)_0}{\mathrm{e}}$, $C(\tau)_0$ is the initial value of $C(\tau)$.

## 3. Results

Two types of EPF signals are considered : healthy and arrhythmic ones. To simulate arrhythmia with this model in vitro, several methods can be used : injection of specific drugs such as aconitine and acetylcholine; by electrical stimulation or by hypothermia. In this study, we use single electrical stimulation. As shown in Figure 1, after stimulation, the EFP signals become arrhythmic : both the periods and amplitudes are no more regular which are sim-

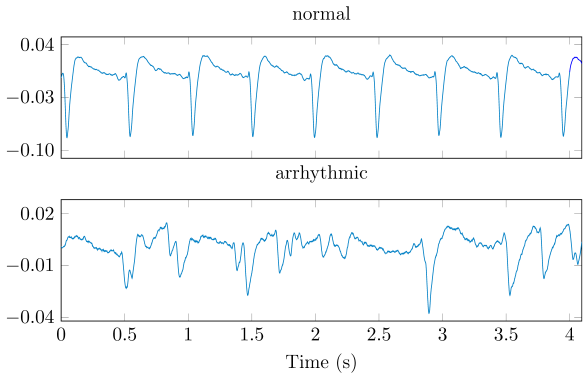



Figure 1: Extracellular potential, in healthy and arrhythmic cases. (y-axis unit : mV; sampling frequency 1 kHz)

ilar to clinical data. One of the main advantages of the platform MEA is that the data is acquired from 60 channels, giving a possibility to examine statistically the global performance of each method. So all results shown here are based on analysis of more than 40 EFP signals (signals not in accordance with quality standard are eliminated). To better show the discrimination performance, ANOVA tests were performed for the six methods.

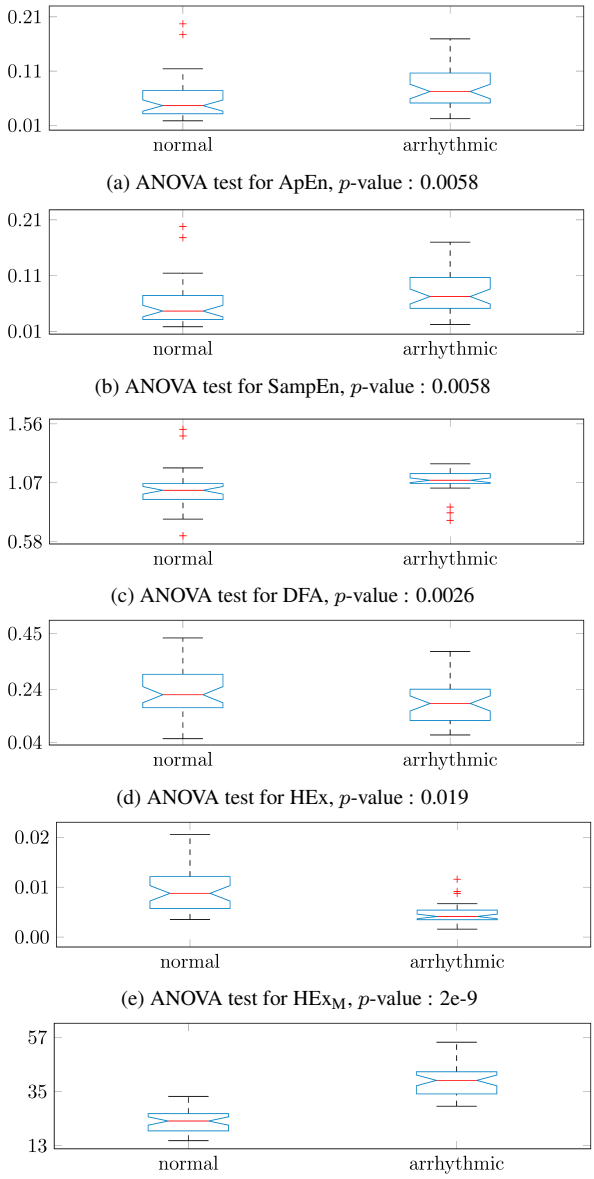
ApEn and SampEn showed almost the same results. The small P-values of both methods (5.8e-3) showed that this result is statistically significant. Compared to normal signals, the medians of ApEn and SampEn for arrhythmic ones are $56\%$ higher, which means that the latter signals are less repetitive.

DFA allows a slightly better result in terms of $p$-value (2.6e-3). Its values corresponded well to that DFA method looked for : generally $\alpha > 1$ signifies that the tested data would be most probably non-stationary / unbounded, which is the case for electro-cardiological signals. What's more, since the exponent $\alpha$ for arrhythmic signals have smaller variance, these signals are more likely similar.

Hurst Exponent measures the relative tendency of a time series. When this exponent is in the range $(0 \sim 0.5)$, it indicates that the time series has a tendency to switch between high and low values lasting a long time into the future, which is true for cardiac signals (either for normal signals or for arrhythmic ones).

As shown in Figure 2d, all these exponents are in the range $(0 \sim 0.5)$, which proves the high-low-values-switching tendency of EFP signals. This is more likely to be true for healthy signals which keep in general a relatively constant rhythm (periodic depolarization / repolarization to ensure the propagation of action potential in the heart). In case of arrhythmia, this rhythm is strongly disturbed, that's why their exponents have smaller values.

The cumulatively sum affects more the arrhythmic signals. For signals having periodic positive / negative values, it will not break too much these signals' rhythm. This




(a) ANOVA test for ApEn, $p$-value : 0.0058

(b) ANOVA test for SampEn, $p$-value : 0.0058

(c) ANOVA test for DFA, $p$-value : 0.0026

(d) ANOVA test for HEx, $p$-value : 0.019

(e) ANOVA test for $\text{HEx}_{\text{M}}$, $p$-value : 2e-9

(f) ANOVA test for TLag, $p$-value : 9e-23. y-axis unit : millisecond

Figure 2: ANOVA tests for different methods. (except TLag having unit of millisecond, other parameters have no IS units.)

explains the result in Figure 2e : arrhythmic signals have smaller hurst exponents and a very compact confident interval.What's more, the $p$-value is much better in $\text{HEx}_{\text{M}}$ than in the original HEx : improved from 1.9e-2 to 2.0e-9.

Time lag is the determining parameter in phase space reconstruction by revealing if the time series may be produced by a large number of independent random variables or be generated by a deterministic process having less independent variables. Too large or too small time lag could make the final phase space set uncorrelated or ill-

determined. As shown in Figure 2f, the arrhythmic signals have almost twice time lag than normal signals. The two datasets are completely separated ($p$-value : 9e-23).

| | $p$-value | median values | |
|---|---|---|---|
| | | normal | arrhythmic |
| ApEn | 5.8e-03 | 0.0463 | 0.0723 |
| SampEn | 5.8e-03 | 0.0464 | 0.0725 |
| DFA | 2.6e-03 | 1.0072 | 1.0905 |
| HEx | 1.9e-02 | 0.2196 | 0.1866 |
| $HEx_M$ | 2.0e-09 | 0.0087 | 0.0041 |
| TLag | 9.0e-23 | 23 | 39.5 |

Table 1: Median index values and $p$-values in ANOVA.

Globally, these methods allowed discriminating the normal and arrhythmic EFP signals. If only the median values of these index are considered (Table 1), the classification by ApEn and TLag showed good performance. The $HEx_M$

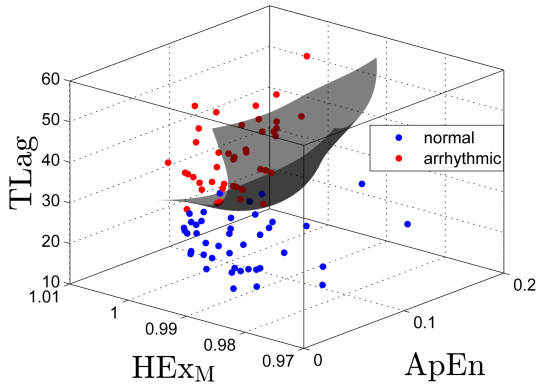



Figure 3: Classification of normal and arrhythmic EFP signals by ApEn, $HEx_M$ and TLag.

improved the robustness of classification, compared to the original exponent. By combining them, as shown in Figure 3, a separatrix plane can be found in space ApEn-$HEx_M$-TLag, which provides a more robust classifier.

## 4. Conclusion

In this study, we compared six complexity analysis methods for electro-cardiological signals in vitro in order to obtain the most appropriate index to discriminate the normal cardiac signals from arrhythmic ones. All results from each method confirmed their theoretical meaning. It is the Time Lagging method allowing obtaining the best discrimination index. The ApEn and SampEn gave a similar result as TLag. However their $p$-values are larger than in TLag. The results with DFA and HEx are less significant. They would not be probably adaptable to this type of signals. Though $HEx_M$ did not give a better arrhythmic/normal ratio than the original HEx, the $p$-value is improved from 1.9e-2 to 2.0e-9 which does show the potential of this method. Further investigation is required. Classification based on these indexes provides a more robust method to discriminate the normal cardiac signal and arrhythmic ones. This part is essential in the development of a work-flow of real-time arrhythmic events monitoring/classification for in vitro studies and clinical applications as well.

## Acknowledgements

We acknowledge the Institute of Cardiovascular Research (Dijon, France) and the NVH Medicinal (Dijon, France) for our collaboration in the data acquisition. We acknowledge the Burgundy & Aquitanian regional council for financial support of this project.

## References

[1] Wolf A, Swift JB, Swinney HL, Vastano JA. Determining lyapunov exponents from a time series. Physica D Nonlinear Phenomena 1985;16(3):285–317. ISSN 0167-2789.

[2] Xu B, Jacquir S, Laurent G, Bilbault JM, Binczak S. Analysis of an experimental model of in vitro cardiac tissue using phase space reconstruction. Biomedical Signal Processing and Control September 2014;13(0):313–326.

[3] Pincus SM. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci U S A Mar 1991; 88(6):2297–2301.

[4] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;278(6):H2039–H2049.

[5] Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of dna nucleotides. Phys Rev E February 1994;49(2):1685–1689.

[6] Rajendra Acharya U, Paul Joseph K, Kannathal N, Lim C, Suri J. Heart rate variability: a review. Med Biol Eng Comput 2006;44(12):1031–1051–. ISSN 0140-0118.

[7] Takens F. Detecting strange attractors in turbulence. In Rand D, Young LS (eds.), Dynamical Systems and Turbulence, Lecture Notes in Mathematics, volume 898. Springer Berlin / Heidelberg, 1981; 366–381.

[8] Athias P, Vandroux D, Tissier C, Rochette L. Development of cardiac physiopathological models from cultured cardiomyocytes. Annales de Cardiologie et dAngeiologie apr 2006;55(2):90–99. ISSN 0003-3928.

[9] Jacquir S, Binczak S, Xu B, Laurent G, Vandroux D, Athias P, Bilbault J. Investigation of micro spiral waves at cellular level using a microelectrode array technology. Int J Bifurcation Chaos 2011;21:1–15.

[10] Reppel M, Pillekamp F, Lu ZJ, Halbach M, Brockmeier K, Fleischmann BK, Hescheler J. Microelectrode arrays: A new tool to measure embryonic heart activity. J Electrocardiol October 2004;37, Supplement(0):104–109. ISSN 0022-0736.

[11] Akay M (ed.). Nonlinear Biomedical Signal Processing, Volume 2, Dynamic Analysis and Modeling. Wiley-IEEE Press, 2000.