# Source Classification in Atrial Fibrillation Using a Machine Learning Approach

Pedro Marinho R. de Oliveira[1], Vicente Zarzoso[1], Carlos Alexandre R. Fernandes[2]

[1] Université Côte d'Azur, CNRS, I3S Laboratory, Sophia Antipolis, France
[2] Universidade Federal do Ceará, Sobral, Brazil

## Abstract

*A precise analysis of the atrial activity (AA) signal in electrocardiogram (ECG) recordings is necessary for a better understanding of the mechanisms behind atrial fibrillation (AF). Blind source separation (BSS) techniques have proven useful in extracting the AA source from ECG recordings. However, the automated selection of the AA source among the other sources after BSS is still an issue. In this scenario, the present work proposes two contributions: i) the use of the normalized mean square error of the TQ segment (NMSE-TQ) as a new feature to quantify the AA content of a source, and ii) an automated classification of AA and non-AA sources using three well-known machine learning algorithms. The tested classifiers outperform the techniques present in literature. A pattern in the mean and standard deviation of the used features, for AA and non-AA sources, is also observed.*

## 1. Introduction

Decreasing life quality and increasing healthcare costs, atrial fibrillation (AF) is a supraventricular tachyarrhythmia characterized by an uncoordinated and irregular atrial activation [1]. Persistent AF represents a particularly complex case of this arrhythmia, where extensive atrial remodeling has taken place due to sustained AF, significantly affecting atrial activity (AA) and AF perpetuation itself. This challenging cardiac condition represents a major health, social and economical concern. Also, the mechanisms behind AF are complex and not completely understood, increasing the intensity of clinical research into this cardiac rhythm disturbance in the past few years.

Signal processing techniques are important and necessary tools to noninvasively separate the AA from the standard 12-lead electrocardiogram (ECG), for a precise analysis and characterization of the f waves, in order to better understand the complex mechanisms behind AF. The extraction of AA from multi-lead ECGs accepts a blind source separation (BSS) formulation [2] and methods to solve BSS problems, such as principal component analysis (PCA) [3], independent component analysis (ICA) [4]

and block term decomposition (BTD) [5] are reported in the literature as useful tools for noninvasive AA extraction [2], [6–10].

However, after separating the sources that compose the ECG recording, one needs to select the atrial source estimate among the other sources. Even with the help of some parameters for AA content measurement, there is still no optimal automated method for atrial source selection, which often needs to be done by visual inspection to achieve optimality.

To provide more information about the AA content of an estimated source and improve the automated atrial source selection accuracy, the present work proposes the normalized mean square error of the TQ segment (NMSE-TQ) between the original recording and the estimated source as a new criterion to quantify AA content. Furthermore, an automated approach for classification of AA and non-AA sources is put forward, using three different machine learning algorithms: linear and quadratic discriminant analysis (LDA and QDA) [11], and support vector machine (SVM) [12], showing that this approach provides better accuracy than the techniques reported in the literature.

## 2. Methods

### 2.1. Blind Source Separation Techniques

Signal processing techniques that solve BSS problems separate the observed ECG signal matrix $\mathbf{Y}$ in a linear combination of a mixing matrix $\mathbf{M}$ and a source matrix $\mathbf{S}$:

$$\mathbf{Y} = \mathbf{MS} \in \mathbb{R}^{\mathbb{K} \times \mathbb{N}}. \tag{1}$$

In the present case of study, $\mathbf{Y} \in \mathbb{R}^{\mathbb{K} \times \mathbb{N}}$ is the AF ECG data matrix, composed of $K$ signals (leads) and $N$ samples, $\mathbf{M} \in \mathbb{R}^{\mathbb{K} \times \mathbb{R}}$ is the mixing matrix, modeling the propagation of the $R$ cardiac electrical sources from the heart to the $K$ leads in the body surface, and $\mathbf{S} \in \mathbb{R}^{\mathbb{R} \times \mathbb{N}}$ is the source matrix that contains $R$ sources, mainly atrial, ventricular and noise sources.

In this work, three different BSS methods are applied in the observed recordings in order to compute an estimate

of the sources that compose the original AF ECG, as described below.

1) PCA: This matrix-based method performs an orthogonal linear transformation in the original data, resulting in principal components that are mutually uncorrelated [3]. The orthogonal linear transformation is given by the diagonalization of $\hat{\mathbf{R}}_y = \frac{1}{N}\mathbf{YY}^T$.

When applied in multi-lead AF ECG recordings, PCA exploits inter-lead correlation, decomposing the recording so that the first principal components represent the ventricular activity (VA) source, whereas the next ones correspond to the AA source, noises, muscular activity, etc.

2) RobustICA-f: ICA decomposes the observed data into statistically independent sources. RobustICA-f is a variant of ICA that performs exact line search optimization of the kurtosis contrast function in the frequency domain [13]. This method does not require sources to be spatially orthogonal and, since AA and VA during AF episodes are assumed to be statistically independent, such technique is more suitable for AA extraction compared to PCA.

3) Hankel-based BTD: This tensor factorization technique is based on a third-order tensor $\mathcal{Y}$ built from Hankel matrices that are constructed from each row of the observed data matrix. The tensor is then decomposed as:

$$\mathcal{Y} = \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_{.r} \tag{2}$$

where $\circ$ represents the outer product, $\mathbf{c}_{.r}$ is a nonzero vector, $\mathbf{E}_r$ is a Hankel matrix built from each source and $R$ is the number of sources. This technique suits the characteristics of AA during AF, since atrial signals can be approximated by all-pole models and mapped onto Hankel matrices with rank equal to the number of poles [7].

## 2.2. Classification Algorithms

Classifiers are machine learning algorithms that operate on labeled data, predicting in which class (or category) the data belong to. Such algorithms are also defined as supervised learning models, as they need labeled data (training set) to learn and perform the classification of new data points (testing set). Three well-known classifiers are used in this work to classify the source estimates provided by the BSS techniques, introduced in the previous section, into AA and non-AA sources.

1) LDA: Commonly used for data classification and dimensionality reduction, LDA guarantees maximal separability with a linear decision region between the classes by computing discriminant scores for each observation obtained by a linear combination of the training data. LDA assumes that the data are normally distributed and all classes identically distributed.

2) QDA: Unlike LDA, this method separates the given classes with a quadratic decision region. In this method, the discriminant scores for each observation are obtained by a non-linear combination of the training data. Since QDA has more degrees of freedom, it tends to perform better than LDA. However, if the database is too small, LDA may provide a better performance.

3) SVM: This popular classifier aims to separate the classes by creating a linear (or non-linear) decision hyperplane. An optimal separating hyperplane can be found by minimizing the distance between misclassified data points and the decision margin. These misclassified observations are called support vectors, as they determine how SVM discriminates between classes, supporting the classification. In the present work, a Gaussian kernel is used.

Three parameters are extracted and used as features for classification: spectral concentration (SC) in % [6], kurtosis of the signal in frequency domain [8] and the proposed NMSE-TQ, given by:

$$\text{NMSE-TQ} = \left[ \frac{||m_r^{(V1)}\mathbf{s}_{r.} - \mathbf{y}^{(V1)}||_F^2}{||\mathbf{y}^{(V1)}||_F^2} \right]_{TQ} \tag{3}$$

where $|| \cdot ||_F$ stands for the Frobenius norm, $m_r^{(V1)}$ is the contribution of the $r$th source estimate to lead V1, $\mathbf{s}_{r.}$ is the $r$th source estimate and $\mathbf{y}^{(V1)}$ is the original recording on lead V1. Note that lead V1 is chosen as reference for this parameter, since this lead is the one that typically best reflects AA in AF ECGs.

## 3. Experimental Results

## 3.1. Database and Experimental Setup

All ECG recordings belong to a database provided by the Cardiology Department of Princess Grace Hospital Center, Monaco. The recordings are acquired at a 977 Hz sampling rate and are preprocessed by a zero-phase forward-backward type-II Chebyshev bandpass filter with cutoff frequencies of 0.5 and 40 Hz, in order to suppress high-frequency noise and baseline wandering. Experiments are performed on different segments of ECG recordings from 30 different patients suffering from persistent AF, processing all the 12 leads.

The Hankel-based BTD is implemented using the non-linear least squares (NLS) method available in Tensorlab MATLAB toolbox [14] choosing $R = 6$ and $L_r = L$, for $r = 1, 2, ..., R$, with $L$ taking values in the set $\{17, 48, 95\}$. The choice of $R$ is based on the singular value decomposition (SVD) of the observed data matrix, taking into account the most significant singular values. The choice of $L_r$ is based on the work [7], which showed that such values provided satisfactory results for the heartbeat with the

largest TQ segment of one of the patients in the considered database. Also, in [9] it was showed that $L_r = 48$ provides satisfactory results for consecutive segments of the whole ECG recording of the same patient.

The randomly chosen recordings are from 0.82 to 1.75 seconds in length and are downsampled by a factor of two, since the 3rd-order tensors built from the original sample rate pose some difficulties to Tensorlab. For the matrix-based techniques PCA and RobustICA-f, no downsampling is needed. Monte Carlo runs with Gaussian random initialization for the spacial and temporal factors of BTD at each run are used. Monte Carlo runs are needed since the performance of BTD depends strongly on the initialization of its factors and a suitable initialization is still an open challenge. These three BSS methods are applied to short ECG recordings from our AF patient database, generating 1283 sources that are visually labeled as 551 AA sources and 732 non-AA sources.

## 3.2. Classification

After performing BSS using the techniques mentioned in Section 2.1, the three features mentioned in the previous section are extracted from the estimated sources. Using all of these three features, LDA, QDA and SVM are applied to classify the database in AA sources and non-AA sources. Figure 1 illustrates these two classes for one of the patients, as it can be seen the original recording in gray and some sources estimates by the Hankel-based BTD. The AA source is shown in blue, followed by non-AA sources. In this case the VA source estimate and noises of different amplitudes can be seen.

From the 1283 generated source estimates, 1046 are used for training the classifiers, while 237 are used for testing. A cross-validation with 5 partitions, i.e., 5-fold cross-validation, is performed in all classifiers. Figure 2 illustrates the accuracy performance of all the methods compared in this work. The parameter SC by itself provides an accuracy of 49.90% (the signal with the highest SC is chosen as AA source), being outperformed by the two techniques PM1 and PM2 proposed in [8], which provide 65.66% and 74.53%, respectively. The proposed parameter NMSE-TQ on its own provides 73.07% of accuracy (the signal with the lowest NMSE-TQ is chosen as AA source), almost the same performance as PM2, while the LDA, QDA and SVM classifiers provide 84.39%, 87.34% and 91.98% of accuracy, respectively. As expected, QDA performs better than LDA, as it uses a non-linear region to separate the classes, which seems to suit better the AA source classification problem in AF.

It is valid to state that if the AA is represented by more than one source, in a particular ECG recording, the classifiers are able to discriminate all of them from the non-AA sources. This is not possible with the existing techniques
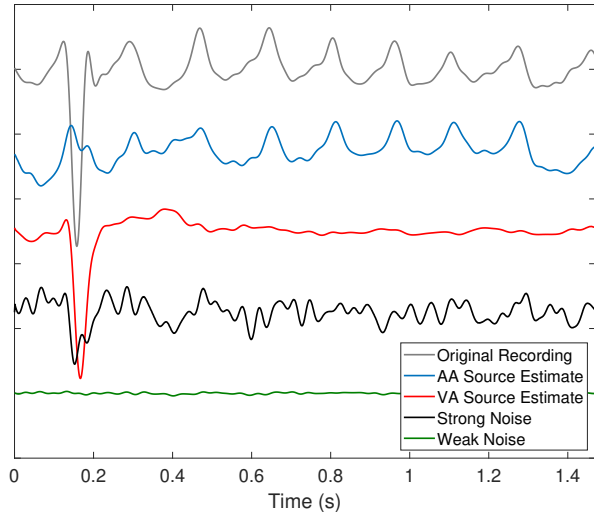


Figure 1. Original ECG recording, AA source estimate and non-AA source estimates (VA and noises) on lead V1. The signals are vertically shifted for clarity.

Table 1. Confusion matrix of LDA, QDA and SVM. AcA: actual AA sources, AcN: actual non-AA sources, PrA: predicted AA sources and PrN: predicted non-AA sources.

| N = 237 | LDA | | QDA | | SVM | |
|---|---|---|---|---|---|---|
| | PrA | PrN | PrA | PrN | PrA | PrN |
| AcA | 82 | 26 | 108 | 0 | 99 | 9 |
| AcN | 11 | 118 | 30 | 99 | 10 | 119 |

here compared, since they are based on the assumption that the AA is represented by a single source.

Table 1 shows the confusion matrix of the three classifiers. It can be seen that LDA misclassifies the largest number of sources, with 37 misclassified sources (26 AA and 11 non-AA sources). QDA correctly classifies all the AA sources, but misclassifies more non-AA sources than LDA (30 non-AA sources). SVM, the classifier that provides the highest accuracy, misclassifies the shortest number of sources, with 19 misclassified sources (9 AA and 10 non-AA sources).

It was also observed that the mean and standard deviation ($\mu \pm \sigma$) of SC, kurtosis and NMSE-TQ for the AA sources was $61.8 \pm 15.1$, $141.3 \pm 65.4$ and $1.6 \pm 3.0$, while for the non-AA sources was $37.8 \pm 14.5$, $41.7 \pm 29.1$ and $34.9 \pm 111.0$, respectively. It can be seen that AA sources have very small NMSE-TQ values concentrated in a short range, while non-AA sources have big NMSE-TQ values varying in a long range. This shows that the proposed parameter discriminates very well these signals, which explains why it provides a satisfactory performance when used alone for classification in this challenging applica-
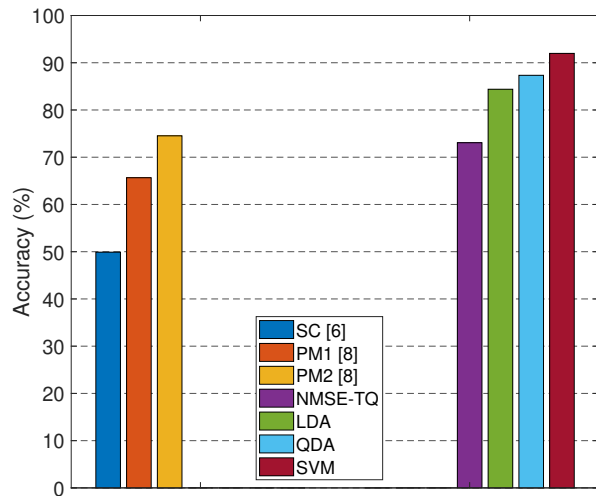
Figure 2. AA source selection accuracy by different techniques: SC proposed in [6]; PM1 and PM2 proposed in [8]; NMSE-TQ proposed in the present work; and the three classifiers: LDA, QDA and SVM.

tion.

## 4.    Conclusions

In order to better quantify the AA content of a source estimate and improve the automated atrial source selection accuracy, the present work has proposed a new criterion called NMSE-TQ and the use of machine learning classifiers to discriminate the AA signals from the other estimates. The NMSE-TQ was shown to be a satisfactory measure of AA content and also a relevant feature for classification. The three classifiers used in the present work provided better accuracy than the existing techniques here compared.

Future works will aim to assess the performance of other classifiers as well as carry out an unsupervised classification, i.e., classify the data without labels. Increasing the database of AF patients in order to provide more relevant clinical results is another a perspective.

## Acknowledgments

## References

[1]    C. T. January, L. S. Wann, J. S. Alpert, H. Calkins, J. C. Cleveland, J. E. Cigarroa, J. B. Conti, *et al.*, "2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society", *Circulation*, vol. 64, no. 21, pp. 2246–2280, Dec. 2014.

[2]    J. J. Rieta, F. Castells, C. Sánchez, V. Zarzoso, and J. Millet, "Atrial activity extraction for atrial fibrillation analysis using blind source separation", *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1176–1186, Jul. 2004.

[3]    I.T. Jolliffe, *Principal Component Analysis*, Wiley Online Library, 2005.

[4]    P. Comon, "Independent component analysis, A new concept?", *Signal Processing*, no. 36, pp. 287–314, 1994.

[5]    L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-$(l_r, l_r, 1)$ terms", *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1451–1474, 2011.

[6]    F. Castells, J. J. Rieta, J. Millet, and V. Zarzoso, "Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias", *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 2, pp. 258–267, Feb. 2005.

[7]    V. Zarzoso, "Parameter estimation in block term decomposition for noninvasive atrial fibrillation analysis", in *Proc. CAMSAP-2017, IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Curaçao, Dutch Antilles, pp. 1–5, Dec. 10-13, 2017.

[8]    P. M. R. de Oliveira and V. Zarzoso, "Source analysis and selection using block term decomposition in atrial fibrillation", in *Proc. LVA/ICA-2018, 14th International Conference on Latent Variable Analysis and Signal Separation*, Guildford, U.K., pp. 46–56, Jul. 2-6, 2018.

[9]    P. M. R. de Oliveira and V. Zarzoso, "Temporal stability of block term decomposition in noninvasive atrial fibrillation analysis", in *Proc. Asilomar-2018, 52nd Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, U.S.A, pp. 816–820, Oct. 28-31, 2018.

[10]    P. M. R. de Oliveira and V. Zarzoso, "Block Term Decomposition of ECG Recordings for Atrial Fibrillation Analysis: Temporal and Inter-Patient Variability", *Journal of Communication and Information Systems*, vol. 34, no. 1, pp. 111-119, Apr. 2019

[11]    R. A. Fisher, "The use of multiple measurements in taxonomic problems", *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[12]    B. E. Boser, I. M. Guyon, V. N. Vapnik, "A training algorithm for optimal margin classifiers", in *Proc. COLT-1992, 5th Annual Workshop on Computational Learning Theory*, Pittsburgh, U.S.A, pp. 144–152, Ju. 27–29, 1992.

[13]    V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size", *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 248–261, 2010.

[14]    N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, *Tensorlab 3.0*, Available online, Mar. 2016. URL: https://www.tensorlab.net/

Address for correspondence:

Pedro Marinho R. de Oliveira
Université Côte d'Azur, CNRS, I3S Laboratory
06903 Sophia Antipolis Cedex
marinho@i3s.unice.fr