

Mixing Matrix Pseudostationarity and ECG Preprocessing Impact on ICA-Based Atrial Fibrillation Analysis

José Joaquín Rieta, César Sánchez,
Juan Manuel Sanchis, Francisco Castells, and José Millet

Bioengineering, Electronics, Telemedicine and Medical Computer
Science Research Group. Valencia University of Technology
Carretera Nazaret–Oliva s/n, 46730, Gandía (Valencia) Spain
{jjrieta,jmsanch,fcastells,jmillet}@eln.upv.es

Abstract. In this work two relevant considerations in the ICA-based estimation of atrial activity (AA) in atrial fibrillation (AF) episodes from real electrocardiogram (ECG) recordings are presented. Firstly, the impact of low-pass filtering preprocessing on the extraction quality of AA is analyzed, showing an average improvement over 17% in spectral concentration (SC) when low-pass filtering is applied after ICA with respect to the application of the same filtering before ICA. Secondly, it is demonstrated that the ICA mixing matrix obtained from one AF segment can also be used to estimate the AA present in different segments of the same recording, thus proving the pseudostationarity of the mixing matrix. Results over 32 AF segments show a mean cross-correlation of $\bar{R}_{dp} = 81.5\%$ between the directly estimated AA and the estimated using pseudostationarity. Changes in spectral concentration from one case to the other ($\Delta SC_{dp} = 1.4\%$) are negligible.

1 Introduction

One of the most important research areas where independent component analysis (ICA) techniques have proved their success is in biomedical engineering [1], with a relevant increase of novel applications during the past years. Regarding the electrocardiogram (ECG), it is well known the extraction of the fetal ECG from maternal recordings [2], the separation of breathing artifacts and other disturbances [3], analysis of ST segments for ischemia detection [4], identification of humans using the ECG [5], ventricular arrhythmia detection and classification [6] and the study of atrial fibrillation (AF).

AF is the most common sustained arrhythmia encountered by clinicians and occurs in approximately 0.4% to 1.0% of the general population. Its prevalence increases with age, and up to 10% of the population older than 80 have been diagnosed with AF [7]. ICA and methods related to blind signal separation have also been applied to AF. In this sense, principal component analysis (PCA) has been used both to extract the atrial activity (AA) from the 12-lead surface ECG in patients with AF [8] and to measure the degree of local organization of this

arrhythmia [9]. Regarding ICA, it has also been applied for the extraction of AA in AF episodes from the surface ECG [10, 11], the suppression of artifacts from internal epicardial recordings [12] and the discrimination among supraventricular arrhythmias [13]. This contribution presents the impact of traditional filtering steps, used for ECG preprocessing, that may decrease ICA performance in the estimation of AA in AF episodes. Next the paper shows the empirical demonstration of the instantaneous linear mixing model of an AF recording, through the corroboration of the ICA mixing matrix pseudostationarity (MMPS).

2 Methods

The suitability of ICA to extract the AA from the ECG in patients with AF has been already demonstrated [11] and the results compared to other AA estimation methods [14]. Nevertheless, the ICA-based optimal estimation of AA still remains as an open issue in continuous development [15]. The following sections deal with these considerations, starting with a short description of AF and its ICA model.

2.1 Atrial Fibrillation and ICA Suitability

The manifestation of AF is characterized by uncoordinated atrial activation with consequent deterioration of atrial mechanical function [7]. AF occurs when the electrical impulses in the atria degenerate into a chaotic pattern, resulting in an irregular and rapid heartbeat due to the unpredictable conduction of these impulses across the atrioventricular node [7]. On the ECG, AF is described by the replacement of P waves by fibrillatory waves that vary in size, shape, and timing, being the topic of intensive research because it is the most common sustained cardiac arrhythmia [7]. ICA can be applied to AF due to the fulfillment of these conditions [16]: independence of the sources, nongaussianity and nonorthogonal observations generated by instantaneous linear mixing of the sources [11]. These considerations can be proved through the study of the electrophysiological mechanisms regarding the generation of AF and the matrix-form solution of the forward problem of electrocardiography [17].

2.2 Impact of ECG Preprocessing

Preprocessing is used over ECG recordings to improve the later analysis or processing stages. The most widely used involves notch filtering, to cancel out mains interference, high pass filtering, to eliminate baseline wandering and low-pass filtering to reduce thermal and muscular noise [18]. Additionally, ICA works much better with low noise data, thus allowing for a better separation of the independent components [19]. On the other hand, linear filtering does not affect the fulfillment of the ICA model [19], hence, having a column vector of observations $\mathbf{x}(t)$ obtained by linearly mixing a column vector of sources $\mathbf{s}(t)$ with a mixing matrix \mathbf{A} , the ICA model is $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$ [19].

Filtering the set of observed signals to obtain a new observation vector $\mathbf{x}_f(t)$, will generate the ICA model $\mathbf{x}_f(t) = \mathbf{A}\mathbf{s}_f(t)$, where the mixing matrix \mathbf{A} is the same as before and the source column vector $\mathbf{s}_f(t)$ corresponds to the set of sources $\mathbf{s}(t)$ filtered with the same filtering applied to $\mathbf{x}(t)$.

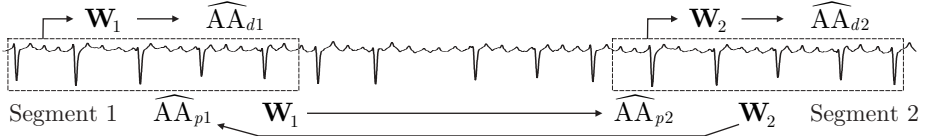


Fig. 1. Proposed methodology for the mixing matrix pseudostationarity analysis. In each AF segment the AA can be directly estimated ($\widehat{\mathbf{AA}}_{d1}$ and $\widehat{\mathbf{AA}}_{d2}$) or indirectly via the pseudostationarity property ($\widehat{\mathbf{AA}}_{p1}$ and $\widehat{\mathbf{AA}}_{p2}$).

A problem with low-pass filtering is that it reduces the information in the data, since high-frequency features of the data are lost. Hence, this information reduction may involve a reduction of independence [19]. In addition, low-pass filtering performs some kind of averaging over the data, and sums tend to increase Gaussianity [19], thus, decreasing ICA performance.

2.3 Mixing Matrix Pseudostationarity

The signal vector from the AF recording can be identified with the observations in ICA, the set of sources being composed of the independent atrial and ventricular activities and other nuisance signals. The mixing matrix entries will can be associated to the transfer coefficients relating the potentials from the heart towards the body surface [17]. Apart from the direct observation or the comparison to other techniques [14], one way to corroborate that an AF recording satisfies the ICA model, may consist of verifying the similarities of the mixing (or separation) matrices between two different segments. If the matrix is preserved then it will be reasonable to consider that AF fulfills the ICA model. Nonetheless, the study and proper electrophysiological interpretation of the variations across the 144 entries of a 12×12 matrix is not easy to perform.

As an alternative each AF recording can be divided into two non-overlapped segments (see Fig. 1). In the first segment we can define the directly estimated atrial activity, $\widehat{\mathbf{AA}}_{d1}$, as the activity obtained after performing ICA over it. This will also give us the mixing and separation matrices for the first segment, \mathbf{A}_1 and \mathbf{W}_1 , respectively, where $\mathbf{W}_1 = \mathbf{A}_1^{-1}$. With the same procedure, it is possible to obtain $\widehat{\mathbf{AA}}_{d2}$ and matrices \mathbf{A}_2 and \mathbf{W}_2 for the second segment. Hence, the verification of the MMPS will consist of applying the separation matrices \mathbf{W}_1 and \mathbf{W}_2 over the second and first ECG segments, respectively, to obtain the activities $\widehat{\mathbf{AA}}_{p2}$ and $\widehat{\mathbf{AA}}_{p1}$ (see Fig. 1). Evaluating the similarity between $\widehat{\mathbf{AA}}_{d1}$ and $\widehat{\mathbf{AA}}_{p1}$, for segment 1, and $\widehat{\mathbf{AA}}_{d2}$ and $\widehat{\mathbf{AA}}_{p2}$, for segment 2, it will be possible to assess the preservation of the matrix. Considering the real world limitations, it is reasonable to define this behavior as MMPS.

2.4 Measurement of Atrial Activity Estimation Quality

The estimation of AA from real AF recordings represents an inverse problem where true sources are impossible to observe. Considering the spectral mor-

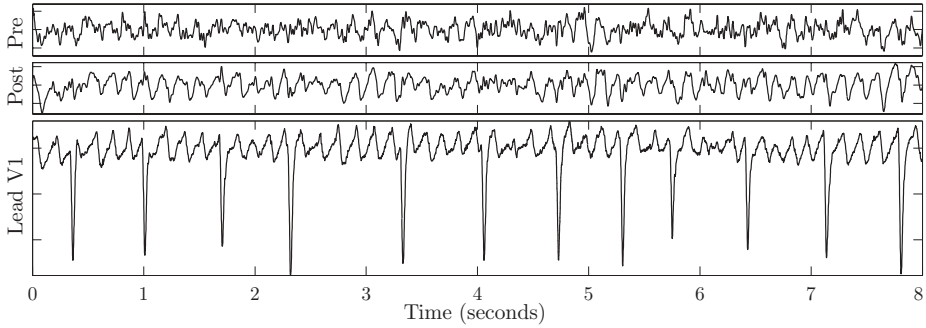


Fig. 2. Estimated AA waveforms from patient #10 for the pre-ICA and post-ICA low-pass filtering. Lead V1 from the same patient is included for comparison purposes.

phology of AA, with a very pronounced peak, no harmonics and insignificant amplitudes above 15Hz [8, 20], it is possible to define a performance extraction index to evaluate the AA extraction quality based on the spectral concentration, that can be defined as [15]

$$SC = \left(\int_{0.82f_p}^{1.17f_p} P_{AA}(f) df \right) / \left(\int_0^{\frac{f_s}{2}} P_{AA}(f) df \right) \quad (1)$$

where f_p is the frequency of the AF main peak, P_{AA} is the power spectral density (PSD) of the AA and f_s is the sampling frequency. SC can evaluate the spectral variation due to the presence of other nuisance signals outside the main peak band. Hence, for a concrete AA, a higher SC will indicate a more efficient elimination of non-AA components. The PSD was obtained via the Welch–WOSA method, discarding the content above 20Hz due to its low contribution.

3 Results and Discussion

The low-pass filtering impact was by its application before and after ICA. The study was carried out over the authors' own database comprising 12-lead AF recordings from 16 patients sampled at 1kHz with segments of 8 seconds. This recording time is long enough because it includes several cardiac cycles and other human events, like breathing, that may affect the transfer coefficients. The low-pass filtering applied in this study was a linear phase Chebyshev type II digital filter, with no ripple in the pass-band and 40dB ripple in the stop-band, the cut-off frequency being 70Hz. Next, the FastICA algorithm was used due to its robust performance and fast convergence [19]. Fig. 2 plots the AA filtering before (pre) and after (post) the ICA stage. The same Fig. also shows lead V1 from the ECG, because it is considered as the lead with higher AA contribution [7]. As can be seen, the post-ICA AA is much more approximated to the AA contained in the ECG. Fig. 3 plots the PSD of the AA signals from Fig. 2. The main atrial frequency and spectral morphology is quite similar for both signals, but

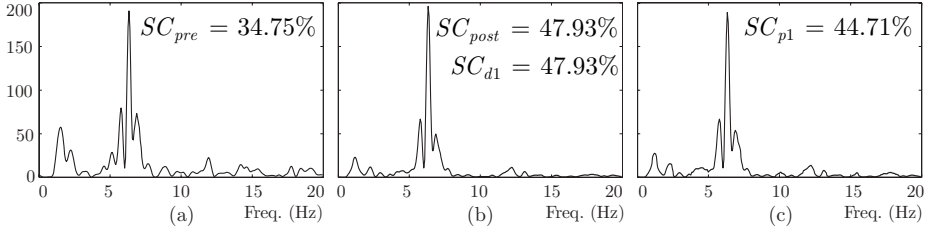


Fig. 3. PSD and SC for the AA in the first segment of patient #10 (Fig 2). (a) Pre-ICA low-pass filtering. (b) Post-ICA low-pass filtering and directly estimated activity \widehat{AA}_{d1} using \mathbf{W}_1 . (c) Indirectly estimated activity \widehat{AA}_{p1} via \mathbf{W}_2 from the second segment.

Table 1. Percentages of SC for each patient and mean values for the whole database. SC_{pre} is for low-pass filtering previous to ICA and SC_{post} for the reverse methodology.

	Pat.01	Pat.02	Pat.03	Pat.04	Pat.05	Pat.06	Pat.07	Pat.08
SC_{pre}	30.4	37.9	41.9	45.8	49.9	34.2	31.1	21.6
SC_{post}	63.0	44.9	57.7	47.1	84.5	47.0	45.4	55.8
$\Delta_{post-pre}$	32.6	6.9	15.8	1.3	34.5	12.7	11.2	34.2

	Pat.09	Pat.10	Pat.11	Pat.12	Pat.13	Pat.14	Pat.15	Pat.16
SC_{pre}	36.6	34.7	38.2	40.5	43.0	28.1	33.3	36.7
SC_{post}	43.3	47.9	49.9	65.4	61.3	42.2	44.3	62.0
$\Delta_{post-pre}$	6.7	13.1	11.6	24.9	18.3	14.1	11.0	25.3

$$\overline{SC_{pre}} = 36.7 \quad \overline{SC_{post}} = 53.9 \quad \overline{\Delta_{post-pre}} = 17.1$$

the pre-ICA estimation (Fig. 3.a) shows larger spectral components below 5Hz and above 10Hz that are notably reduced with post-ICA (Fig. 3.b). Hence, the SC for this latter case is higher and the AA extraction performance is better.

Table 1 shows preprocessing (SC_{pre}) and postprocessing (SC_{post}) values for all the patients, $\Delta_{post-pre}$ being the percentage of increase in SC . Similar analysis have been performed regarding notch filtering and high-pass filtering. The results, though are not presented in this contribution, show an irrelevant impact over the ICA separation performance.

To study the MMPS a recording of 20 seconds in length was selected, then, FastICA was applied over the first 8 seconds (segment 1), giving us \widehat{AA}_{d1} , \mathbf{A}_1 and \mathbf{W}_1 . Segment 2 was comprised of the last 8 seconds, obtaining \widehat{AA}_{d2} , \mathbf{A}_2 and \mathbf{W}_2 . The post-ICA low-pass filtering strategy was selected and applied over both segments. Next, by using the separation matrices, the AA estimation was obtained indirectly to verify the MMPS: \widehat{AA}_{p1} was extracted using \mathbf{W}_2 and \widehat{AA}_{p2} with \mathbf{W}_1 , respectively. Fig. 4 shows the result over the AF segment in Fig. 2. The similarity obtained is significant, thus corroborating the MMPS. Also for this AA pair, the cross-correlation percentage is 81.7%.

Regarding SC for the AA pairs, Fig. 3.b plots the PSD and SC for \widehat{AA}_{d1} from Fig. 4 and the result for \widehat{AA}_{p1} can be seen in Fig. 3.c. Comparing both Figs.

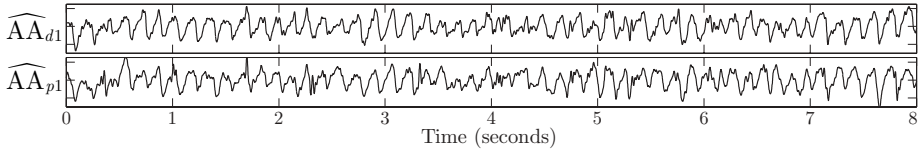


Fig. 4. Waveform of the directly estimated atrial activity \widehat{AA}_{d1} from the first segment of patient #10 (see Fig. 2) and the same activity obtained via the MMSP \widehat{AA}_{p1} .

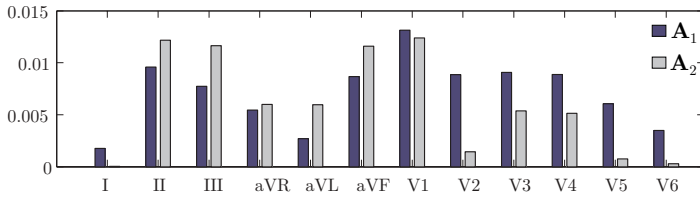


Fig. 5. Column comparison for the entries, in absolute value, of the mixing matrices A_1 and A_2 related to the AA source from the ECG recording of Fig. 2.

the differences in spectral morphology and main peak frequency are negligible. There is a variation in the main peak amplitude and the spectral concentration (SC_{d1} vs SC_{p1}) quite acceptable when dealing with real AF recordings.

Fig. 5 plots the columns of the mixing matrices A_1 and A_2 associated to the AA source. Observe the similarities in the leads with largest AA contribution. Finally Fig. 6 concentrates the MMPS results for the 16 patients in the database. Above each patient number there are six bars, the first three ones for the first AF segment and the others for the second segment. Within each bar group, the first bar indicates SC_{d1} in percentage, i.e., the spectral concentration of \widehat{AA}_{d1} . The second bar is SC_{p1} associated to \widehat{AA}_{p1} and the third bar is cross-correlation percentage, R_{dp1} . The same is applicable for the other bar group on the right, but for the second segment. The differences between the SC_d and SC_p pairs are generally small, being the cross-correlation below 75% in only 6 of the 32 analyzed cases. The mean difference between SC_d and SC_p is $\overline{\Delta SC}_{dp} = 1.47\%$ in spectral concentration. Regarding the correlation value, its mean for the same 32 situations is $\overline{R}_{dp} = 81.56 \pm 10.74\%$, thus reinforcing the large similarity between \widehat{AA}_d and \widehat{AA}_p for all the cases analyzed.

Note that this application assumes the fulfillment of the ICA model, hence, it will only be possible to derive the spatial filters (mixing matrix) and the sources from the ECG, when the physical sources associated to heart's activity are spatially stationary [1]. Contraction of the atria during fibrillation, the ventricles in the cardiac cycle or any other relative movement from sources to observations, could violate the assumption of spatial stationarity. The authors consider that these variations do not affect significantly the ICA model in AF: firstly, results of the AA have demonstrated its validity [10, 11], secondly, the main atrial frequency of AA obtained using ICA is in agreement with the results obtained via other accepted AA extraction techniques [20]. Finally, the MMPS corroboration gives the definitive support to say that the 12-lead ECG of an AF recording fulfills the instantaneous linear mixing ICA model.

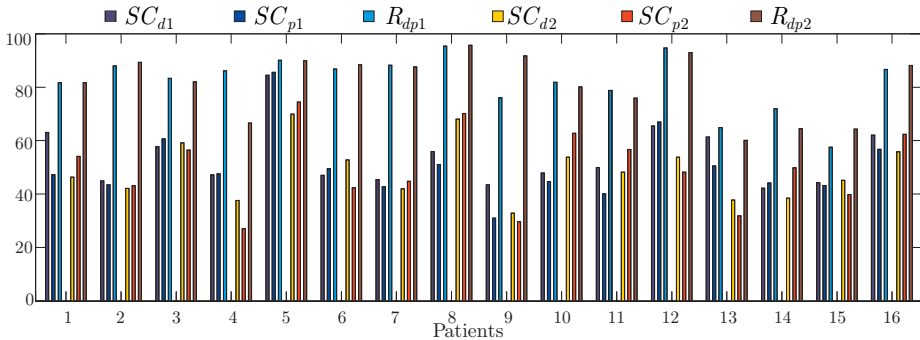


Fig. 6. MMPS analysis results for the whole patient database.

4 Conclusions

This contribution has demonstrated that preprocessing has a significant impact on the separation performance of ICA-based AA estimation. The study has proved that low-pass filtering, though itself is a linear operation, involves a data reduction that decreases the quality of the ICA-based AA estimation. To solve this problem, the ICA approach has to be applied before the low-pass filtering and then, any other post-ICA processing could be performed over the data. Besides, the impact of notch and baseline wander filtering is not relevant.

Moreover, it has been demonstrated that the AA from one segment can be recovered with the separation matrix from other segment. This observation, defined as the MMPS, gives the definitive support to the fulfillment of the ICA model for AF. In addition, this fact may imply the birth of other studies on the evolution and periodicity of the mixing matrix entries across the cardiac cycle or the patient's own breathing, not only in atrial fibrillation, but in other supraventricular arrhythmias and cardiac pathologies where atrial and ventricular activities can be regarded as decoupled or independent.

Acknowledgements. This work was partly funded by the research incentive program of the Valencia University of Technology and TIC2002-00957. The authors would like to thank the cardiologists from the Universitary Clinical Hospital of Valencia (Spain), for their clinical advice and kind help in obtaining the signals.

References

1. Jung, T.P., Makeig, S., Lee, T.W., McKeown, M.J., Brown, G., Bell, A.J., Sejnowski, T.J.: Independent component analysis of biomedical signals. *International Conference on Independent Component Analysis and Blind Signal Separation (ICA)* **2** (2000) 633–644
2. Zarzoso, V., Nandi, A.K.: Noninvasive fetal ECG extraction: blind separation versus adaptive noise cancellation. *IEEE Trans. Biomed. Eng* **48** (2001) 12–18
3. Barros, A.K., Mansour, A., Ohnishi, N.: Adaptive blind elimination of artifacts in ECG signals. *International Workshop on Independence & Artificial Neural Networks (I&ANN'98)* (1998) 1380–1386

4. Stamkopoulos, T., Diamantaras, K., Maglaveras, N., Strintzis, M.: ECG analysis using nonlinear PCA neural networks for ischemia detection. *IEEE Transactions on Signal Processing* **46** (1998) 3058–3067
5. Biel, L., Pettersson, O., Philipson, L., Wide, P.: ECG analysis: A new approach in human identification. *IEEE Transactions on Instrumentation and Measurement* **50** (2001) 808–812
6. Owis, M.I., Youssef, A.B.M., Kadam, Y.M.: Characterisation of ECG signals based on blind source separation. *Med. Biol. Eng. Comput.* **40** (2002) 557–564
7. Fuster, V., Ryden, L.E., Asinger, R.W., et al.: ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation. *European Heart Journal* **22** (2001) 1852–1923
8. Langley, P., Bourke, J.P., Murray, A.: Frequency analysis of atrial fibrillation. *IEEE Computers in Cardiology* **27** (2000) 65–68, Boston, MA, USA
9. Faes, L., Nollo, G., Kirchner, M., Olivetti, E., et al.: Principal component analysis and cluster analysis for measuring the local organisation of human atrial fibrillation. *Med. Biol. Eng. Comput.* **39** (2001) 656–663
10. Rieta, J.J., Zarzoso, V., Millet, J., Garcia, R., Ruiz, R.: Atrial activity extraction based on blind source separation as an alternative to QRST cancellation for atrial fibrillation analysis. *IEEE Computers in Cardiology* **27** (2000) 69–72, Boston.
11. Rieta, J.J., Castells, F., Sanchez, C., Igual, J.: ICA applied to atrial fibrillation analysis. *International Conference on Independent Component Analysis and Blind Signal Separation (ICA)* **4** (2003) 59–64, Nara, Japan
12. Liu, J.H., Kao, T.: Removing artifacts from atrial epicardial signals during atrial fibrillation. *International Conference on Independent Component Analysis and Blind Signal Separation (ICA)* **4** (2003) 179–183, Nara, Japan
13. Rieta, J.J., Millet, J., Zarzoso, V., Castells, F., Sanchez, C., Garcia, R., Morell, S.: Atrial fibrillation, atrial flutter and normal sinus rhythm discrimination by means of blind source separation and spectral parameters extraction. *IEEE Computers in Cardiology* **29** (2002) 25–28, Memphis, TN
14. Langley, P., Rieta, J.J., Stridh, M., Millet, J., Sörnmo, L., Murray, A.: Reconstruction of atrial signals derived from the 12-lead ECG using atrial signal extraction techniques. *IEEE Computers in Cardiology* **30** (2003) 129–132, Greece.
15. Castells, F., Ruiz, R., Rieta, J.J., Millet, J.: An integral atrial wave identification approach based on spatiotemporal source separation: Clinical validation. *IEEE Computers in Cardiology* **30** (2003) 717–720, Thessaloniki, Greece.
16. Cardoso, J.F.: Blind signal separation: Statistical principles. *Proceedings of the IEEE* **86** (1998) 2009–2025
17. Rieta, J.J., Castells, F., Sanchez, C., Moratal, D., Millet, J.: Bioelectric model of atrial fibrillation: Applicability of blind source separation techniques for atrial activity estimation in atrial fibrillation episodes. *IEEE Computers in Cardiology* **30** (2003) 525–528, Thessaloniki, Greece
18. Tompkins, W.J.: Biomedical digital signal processing: C language examples and laboratory experiments for the IBM PC. Prentice Hall, New Jersey (1993)
19. Hyvarinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. John Wiley & Sons, Inc. (2001)
20. Langley, P., Stridh, M., Rieta, J.J., Sörnmo, L., Millet, J., Murray, A.: Comparison of atrial rhythm extraction techniques for the detection of the main atrial frequency from the 12-lead ECG in atrial fibrillation. *IEEE Computers in Cardiology* **29** (2002) 29–32, Memphis, (TN)