# dAMUSE—A new tool for denoising and blind source separation

A.M. Tomé [a,*], A.R. Teixeira [a], E.W. Lang [b,1], K. Stadlthanner [b,2],
A.P. Rocha [c], R. Almeida [c,3]

[a] *DETUA/IEETA, Universidade de Aveiro, P-3810 Aveiro, Portugal*
[b] *Institute of Biophysics, University of Regensburg, D-93040 Regensburg, Germany*
[c] *DMA/FCUP, Universidade do Porto, P-4000 Porto, Portugal*

## Abstract

In this work a generalized version of AMUSE, called dAMUSE is proposed. The main modification consists in embedding the observed mixed signals in a high-dimensional feature space of delayed coordinates. With the embedded signals a matrix pencil is formed and its generalized eigendecomposition is computed similar to the algorithm AMUSE. We show that in this case the uncorrelated output signals are filtered versions of the unknown source signals. Further, denoising the data can be achieved conveniently in parallel with the signal separation. Numerical simulations using artificially mixed signals are presented to show the performance of the method. Further results of a heart rate variability (HRV) study are discussed showing that the output signals are related with LF (low frequency) and HF (high frequency) fluctuations. Finally, an application to separate artifacts from 2D NOESY NMR spectra and to denoise the reconstructed artefact-free spectra is presented also.
© 2005 Elsevier Inc. All rights reserved.

*Keywords:* Embedding signals; Generalized eigenvalue decomposition; Blind source separation; Denoising; Biomedical applications

* Corresponding author.
  *E-mail addresses:* ana@ieeta.pt (A.M. Tomé), elmar.lang@biologie.uni-regensburg.de (E.W. Lang), aprocha@fc.up.pt (A.P. Rocha).

## 1. Introduction

Blind source separation (BSS) methods deal with the separation of observed sensor signals into their underlying source signals knowing neither these source signals nor the mixing process. BSS methods based on a generalized eigenvalue decomposition (GEVD) of a matrix pencil rely on second order statistics only taking advantage of certain correlation structures present in the data. They are very efficient and easy to implement but they are sensitive to noise [1,2]. There are several proposals to improve efficiency and robustness of these algorithms when noise is present [1,3].

Considering noisy signals, there exist, beneath others, several projective denoising techniques, the first step of which consists in increasing the dimension of the data set by joining delayed versions of the signals [4–6]. Hence the sensor signals are embedded in a high-dimensional feature space of delayed coordinates. Then denoising is achieved by back-projection of the data into a lower-dimensional subspace corresponding to the dimension of the underlying source signals. The idea behind is that a random noise spreads all space equally while a deterministic signal lives on a sub-manifold only. Projecting onto that sub-manifold thus removes some of the noise components reducing its amplitude thereby.

A similar strategy is used in singular spectrum analysis (SSA) [7] where a matrix composed of the data and their time-delayed versions is considered. Then, a singular value decomposition (SVD) of the data matrix or a principal component analysis (PCA) of the related correlation matrix is computed. The data are then projected onto the principal directions of the eigenvectors of the SVD or PCA analysis. The SSA can be used to extract information from short and noisy time series, hence can provide insight into the underlying system that generates the series [8].

In this work we will join BSS techniques and denoising strategies to present a modified version of AMUSE [9], called dAMUSE [10], which solves the BSS problem, albeit a filtering indeterminacy remains, and in parallel provides an efficient denoising tool, hence alleviates the sensitivity of GEVD methods against noise.

AMUSE is a BSS algorithm using second-order statistics and the time structure in the data. It is based on a GEVD, i.e., the *simultaneous* diagonalization of a matrix pencil formed with a correlation matrix of zero mean sensor data and a time-delayed correlation matrix. The proposed algorithm dAMUSE also comprises this simultaneous diagonalization but the matrix pencil is formed with correlation matrices computed in a high-dimensional feature space rather than the input space, i.e., after increasing the data set dimension by joining delayed versions of each sensor signal.

In the following section we will show, considering a model of linearly mixed sensor signals, that the estimated uncorrelated signals correspond to filtered versions of the underlying source signals. We will also present an implementation of this algorithm to compute the eigenvector matrix of the matrix pencil which involves a two step procedure based on the standard eigenvalue decomposition (EVD) approach. The advantage of this procedure lies in an optional dimension reduction between the two steps with a concomitant reduction in the number of estimated underlying source signals. The assertion at this step is that part of the source components correspond to noise components only and can be removed advantageously. Hence dAMUSE elegantly combines blind source separation techniques with denoising techniques.

Finally, to illustrate the proposed method, some applications are discussed comprising simulations with artificially mixed signals of varying signal to noise ratios (SNR) [10], an artifact removal and concomitant denoising of 2D NOESY proton NMR spectra of aqueous solutions of proteins [11], and an analysis of ECG signals [12].

The outline of the paper will be the following: Section 2 presents a concise algebraic formulation of GEVD methods and the AMUSE algorithm generalized to time-delayed signals or their filtered counterparts in the canonically conjugated Fourier space. Section 3 will discuss implementation aspects of the algorithm, Section 4 will discuss applications to artificial and real world noisy sensor signals and Section 5 will present a short discussion.

## 2. The algorithm dAMUSE

### 2.1. Embedding

Consider a group of $N$ sensor signals $\vec{x}_i$, $i = 1, \ldots, N$, with each sensor signal $\vec{x}_i$ embedded in a high-dimensional feature space $\Omega$ of delayed coordinates. Then the $i$th *trajectory matrix* [7]

$$\mathbf{X}_i = \left[ \vec{x}_i^T (t_l + (K-1)\Delta t), \ldots, \vec{x}_i^T (t_l + \Delta t), \vec{x}_i^T (t_l) \right]^T \tag{1}$$

is formed with the $i$th sensor signal and its $(K-1)$ delayed versions $(k = 0, \ldots, K-1)$,

$$\vec{x}_i^T (t_l + k\Delta t) = \left( x_i(t_0 + k\Delta t), x_i(t_0 + \tau + k\Delta t), \ldots, x_i(t_0 + (L-1)\tau + k\Delta t) \right)$$

thus represents one signal or one of its $(K-1)$ delayed versions. The trajectory matrix then reads (set $t_0 = 0$ for simplicity)

$$\mathbf{X}_i = \begin{bmatrix} x_i((K-1)\Delta t) & x_i(\tau + (K-1)\Delta t) & \ldots & x_i((L-1)\tau) \\ x_i((K-2)\Delta t) & x_i(\tau + (K-2)\Delta t) & \ldots & x_i((L-1)\tau - \Delta t) \\ \vdots & \vdots & \ldots & \vdots \\ x_i(0) & x_i(\tau) & \ldots & x_i((L-1)\tau - (K-1)\Delta t) \end{bmatrix},$$

where $t_l = t_0 + l\tau$, $l = 0, \ldots, L-1$, denotes discrete time, $\tau^{-1}$ represents the sampling rate and $\Delta t = n\tau$, $n \in \mathcal{N}$, $n \ll L$, denotes a fixed delay which is an integer multiple (usually $n = 1$) of the sampling interval $\tau$. Thus each row of the trajectory matrix of each signal contains $L - n(K-1)$ samples of the signal.

The total trajectory matrix $\mathbf{X}$ of the whole set of $N$ signals will be a concatenation of the component trajectory matrices $\mathbf{X}_i$ computed for each sensor:

$$X = [X_1, X_2, \ldots, X_N]^T. \tag{2}$$

Assuming that each sensor signal is a linear combination of $M \leqslant N$ underlying but unknown and uncorrelated $L$-dim source signals $(\vec{s}_i)$, a source trajectory matrix $\mathbf{S}$ can be written in analogy to Eq. (1). To simplify notation, in the following we will only consider the case: $N = M$. Then the sensor signal trajectory matrix can be expressed as $\mathbf{X} = \mathbf{AS}$,

where the total $NK \times NK$-dimensional mixing matrix $\mathbf{A}$ is given by a Kronecker product $\mathbf{A} = \mathbf{a} \otimes \mathbf{I}_{K \times K}$. Then, the mixing matrix is a block matrix with a diagonal matrix $\mathbf{a}_{ij} = a_{ij}\mathbf{I}_{K \times K}$ in each block, i.e.,

$$
\mathbf{A} = \begin{bmatrix}
a_{11}\mathbf{I}_{K \times K} & a_{12}\mathbf{I}_{K \times K} & \dots & a_{1N}\mathbf{I}_{K \times K} \\
a_{21}\mathbf{I}_{K \times K} & a_{22}\mathbf{I}_{K \times K} & \dots & \dots \\
\vdots & \vdots & \ddots & \vdots \\
a_{N1}\mathbf{I}_{K \times K} & \dots & \dots & a_{NN}\mathbf{I}_{K \times K}
\end{bmatrix}.
\tag{3}
$$

The matrix $\mathbf{I}_{K \times K}$ represents the identity matrix and the mixing coefficient $a_{ij}$ relates the sensor signal $i$ with the source signal $j$ of the corresponding component trajectory matrices. As we are dealing with an instantaneous mixing model, all delayed versions of a signal are related by the same coefficient.

## 2.2. A GEVD using matrix pencils

Next a correlation matrix $\mathbf{R}_x = \langle \mathbf{X}\mathbf{X}^H \rangle$ (H denotes the Hermitian conjugate) of the sensor data and its time-delayed counterpart $\mathbf{R}_x(d)$ are computed to form a matrix pencil. Note that the delay $\delta_d = d \cdot \tau$, $d \in \mathcal{N}$ forms an integer multiple of the sampling interval $\tau$. The time-delayed correlation matrix of the matrix pencil is computed with one matrix $\mathbf{X}_r$ obtained by eliminating the first $d$ columns of $\mathbf{X}$ and another matrix, $\mathbf{X}_l$, obtained by eliminating the last $d$ columns. Then, the time-delayed correlation matrix $\mathbf{R}_x(d) = \langle \mathbf{X}_r\mathbf{X}_l^H \rangle$ will be an $NK \times NK$ matrix. Each of these two correlation matrices can be related with a corresponding correlation matrix in the source signal domain

$$
\mathbf{R}_x(d) = \mathbf{A}\mathbf{R}_s(d)\mathbf{A}^H = \mathbf{A}\langle \mathbf{S}_r\mathbf{S}_l^H \rangle \mathbf{A}^H.
\tag{4}
$$

Then the two pairs of matrices $(\mathbf{R}_x(0), \mathbf{R}_x(d))$ and $(\mathbf{R}_s(0), \mathbf{R}_s(d))$ represent congruent pencils [13] with the following properties:

- Their eigenvalues are the same, i.e., the eigenvalue matrices of both pencils are identical: $\mathbf{D}_x = \mathbf{D}_s$.
- If the eigenvalues are non-degenerate (distinct values in the diagonal of the matrix $\mathbf{D}_x = \mathbf{D}_s$), the corresponding eigenvectors are related by the transformation $\mathbf{E}_s = \mathbf{A}^H \mathbf{E}_x$.

Assuming that all sources are uncorrelated, the matrices $\mathbf{R}_s(d)$ are block diagonal, having block matrices $\mathbf{R}_{ii}(d) = \langle (\mathbf{S}_r)_i (\mathbf{S}_l)_i^H \rangle$ along the diagonal

$$
\mathbf{R}_s(d) = \begin{bmatrix}
\mathbf{R}_{11}(d) & 0 & \dots & 0 \\
0 & \mathbf{R}_{22}(d) & \dots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & \mathbf{R}_{NN}(d)
\end{bmatrix}.
\tag{5}
$$

The eigenvector matrix of the GEVD of the pencil $(\mathbf{R}_s(0), \mathbf{R}_s(d))$ can be written as

$$
\mathbf{E}_s = 
\begin{bmatrix}
\mathbf{E}_{11} & 0 & \ldots & 0 \\
0 & \mathbf{E}_{22} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \mathbf{E}_{NN}
\end{bmatrix},
\tag{6}
$$

where $\mathbf{E}_{ii}$ is the $K \times K$ eigenvector matrix of the GEVD of the pencil $(\mathbf{R}_{ii}(0), \mathbf{R}_{ii}(d))$. The estimated uncorrelated components $\mathbf{Y}$ can be estimated from linearly transformed sensor signals via

$$
\mathbf{Y} = \mathbf{E}_x^H \mathbf{X} = \mathbf{E}_x^H \mathbf{A} \mathbf{S} = \mathbf{E}_s^H \mathbf{S}
\tag{7}
$$

hence turn out to be filtered versions of the underlying source signals. As the eigenvector matrix $\mathbf{E}_s$ (Eq. (6)) is a block diagonal matrix, there are $K$ signals in each column of $\mathbf{Y}$ which are a linear combination of one of the source signals and its delayed versions. For instance, the block $i$ depends on the source signal $i$ via

$$
\sum_{k=1}^{K} (\mathbf{E}_{ii})_{k,j} s_i \big( t_l + (K - k)\Delta t \big).
\tag{8}
$$

Equation (8) defines a convolution operation between column $j$ of $\mathbf{E}_{ii}$ and source signal $\vec{s}_i$. Hence, the columns of the matrix $\mathbf{E}_{ii}$ represent impulse responses of finite impulse response (FIR) filters. Considering that all the columns of $\mathbf{E}_{ii}$ are different, their frequency response will provide different spectral densities of the source signal spectra. Then each of the $N$ output signals $\vec{y}_i$, $i = 1, \ldots, NK$, encompasses $K$ filtered versions of each of the $N$ estimated source signals $\vec{s}_i$.

Note that the frequency shaping introduced by this linear filtering operation depends on the length (number of delays $K$) of the filter and on the coefficients in each column of $\mathbf{E}_{ii}$. In summary, the algorithm dAMUSE yields uncorrelated component signals which are filtered versions of the underlying source signals. Hence similar to blind deconvolution methods in addition to scaling and permutation indeterminacies there appears a filtering indeterminacy here.

## 2.3. Implementation of the algorithm dAMUSE

There are several ways to compute the eigenvalues and eigenvectors of a matrix pencil if one of the matrices is symmetric positive definite [14]. One of those methods is based on standard eigenvalue decompositions (EVD) applied in two consecutive steps. One of the correlation matrices of the pair is computed for delay zero ($\delta = 0$) to ensure the symmetric positive definite condition. So, considering the pencil $(\mathbf{R}_x(0), \mathbf{R}_x(d))$ the following steps are proposed:

*Step 1.* Compute a standard eigenvalue decomposition of $\mathbf{R}_x(0) = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$, i.e., the eigenvectors ($\vec{v}_i$) and eigenvalues ($\lambda_i$). As the matrix is symmetric positive definite, the eigenvalues can be organized in descending order ($\lambda_1 > \lambda_2 > \cdots > \lambda_{NK}$). In AMUSE (and other algorithms) this procedure can be used to estimate the number of sources related

to the signal, hence can be considered a strategy to reduce noise. Dropping small eigenvalues amounts to a projection from a high-dimensional signal plus noise feature space onto a lower-dimensional manifold representing the signal subspace. Thereby it is tacitly assumed that small eigenvalues are related with noise components only. In this work, we consider a variance criterium to choose the most significant eigenvalues according to

$$\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_l}{\lambda_1 + \lambda_2 + \cdots + \lambda_{NK}} \geqslant \Theta. \tag{9}$$

If we are interested in the eigenvectors corresponding to directions of high variance of the signals, the threshold ($\Theta$) should be chosen so that the maximum energy of the signals is preserved.

A transformation matrix $\mathbf{Q}$ can then be computed using either the $l$ most significant eigenvalues $\lambda_i$ or all of the eigenvalues and respective eigenvectors $\vec{v}_i$. Considering that the matrix $\mathbf{R}_x(0)$ can be approximated with $l$ eigenvalues and eigenvectors, i.e., $\mathbf{R}_x(0) \simeq \mathbf{V}_l \mathbf{\Lambda}_l^{1/2} \mathbf{\Lambda}_l^{1/2} \mathbf{V}_l^H$, the transformation matrix [14] is then computed by

$$\mathbf{Q} = \mathbf{\Lambda}_l^{-1/2} \mathbf{V}_l^H, \tag{10}$$

where $\mathbf{Q}$ is an $l \times NK$ matrix if a reduction in dimension ($l \leqslant NK$) is considered.

*Step 2.* Compute the matrix $\mathbf{C} = \mathbf{Q} \mathbf{R}_x(d) \mathbf{Q}^H$ and its standard eigenvalue decomposition: the eigenvector matrix $\mathbf{U}$ and eigenvalue matrix $\mathbf{D}_x$. The eigenvectors (which are not normalized) of the pencil ($\mathbf{R}_x(0), \mathbf{R}_x(d)$) form the columns of the eigenvector matrix

$$\mathbf{E}_x = \mathbf{Q}^H \mathbf{U} = \mathbf{V}_l \mathbf{\Lambda}_l^{-1/2} \mathbf{U}. \tag{11}$$

Note that if the threshold $\Theta = 1$, the corresponding result can be obtained with the command eig($\mathbf{R}_x(0), \mathbf{R}_x(d)$) using the most recent version of MATLAB.

What concerns the uncorrelated components of the sensor signals and their delayed versions, they can be estimated via the transformation

$$\mathbf{Y} = \mathbf{E}_x^H \mathbf{X} = \mathbf{U}^H \mathbf{Q} \mathbf{X} = \mathbf{U}^H \mathbf{\Lambda}_l^{-1/2} \mathbf{V}_l^H \mathbf{X} \tag{12}$$

and either $l$ or $NK$ estimated uncorrelated signals are obtained representing filtered versions of the underlying source signal estimates.

## 3. Results

The proposed algorithm was applied to artificially mixed signals, RR and QT sequences of electrocardiograms (ECG) and 2D NOESY NMR signals. The artificial mixtures allow the illustration of the method once every parameter of the model is known.

The heart rate variability study (HRV), based on RR and QT sequences, illustrates a preliminary application to real data. The heart rate variability (HRV) spectrum is currently separated into three frequency bands: very low frequency band (VLF range below 0.04 Hz), low frequency band (LF range 0.04–0.15 Hz) and high frequency band (HF range 0.15–

0.4 Hz). Hence it is interesting to know if the extracted signals can be related with LF and HF fluctuations of cardiovascular signals. A similar strategy applied to RR and QT sequences was proposed in [15,16] where a principal component analysis (PCA) followed by a generalized eigendecomposition was used to extract inherent uncorrelated component signals.

The application of dAMUSE to 2D NOESY NMR signals was motivated mainly by the necessity to denoise the reconstructed spectra after the water artifact has been separated from the protein spectra using BSS techniques. The proposed algorithm dAMUSE offers an elegant and efficient way to perform both the blind signal separation and the denoising simultaneously.

### 3.1. Artificially mixed signals

Using artificial signals, the simulations were designed to illustrate the method and to study the influence of its parameters, especially the dimension of the embedding $K$ and the threshold $\Theta$, on the respective performance. As the separation process also involves a linear filtering operation, each uncorrelated component estimated has its maximum correlation with one of the source signals for a non-zero delay, besides the usual indeterminacy in order and amplitude.

### 3.1.1. Delays
In the first experiment, two artificial source signals with very similar frequency contents were considered (see Fig. 1). The signals were randomly mixed and the algorithm was applied for different numbers $K$ of delays using $\Theta = 1$, $d = 1$. Figure 2 shows the results obtained when $K = 3$. In that case six uncorrelated components corresponding to filtered versions of the two source signals are computed: three of them (Fig. 2a) are related with the first source signal, while the other three (Fig. 2b) are related with the second source signal. We can see that these estimated components represent different frequency contents of each of the corresponding source signals which each comprise a high and a low frequency component. For instance, Fig. 2a shows that:

(1) The first estimated component isolates the high frequency component HF of source signal 1.
(2) The third estimated component isolates the low frequency component LF of source signal 1.
(3) The second estimated component contains both frequencies of the source signal.

Figures 2c and 2d represent the frequency response amplitude of the filters represented by the columns of the transformed mixing matrix $\mathbf{E}_x^H \mathbf{A}$ which can be considered as containing the coefficients of the corresponding filter. These filters are then paired with the respective uncorrelated source signals. It can easily be verified that the frequency response confirms the frequency contents of each of the uncorrelated signals as the filters obtained by the algorithm represent a low-pass, a band-pass, and a high-pass filter (see Fig. 2d).

Increasing the number $K$ of delays, the uncorrelated components show either the low frequency or high frequency mode of the sensor signals. In these cases, the frequency
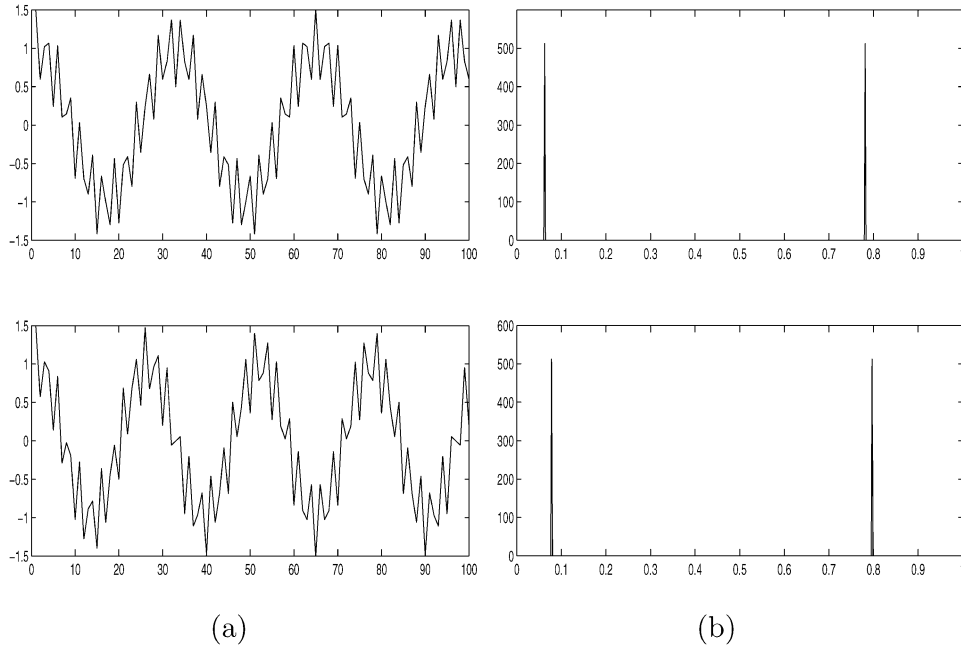
Fig. 1. Artificial signals comprising a superposition of a high-frequency and a low-frequency sinusoidal signal: (a) source signals, (b) spectral densities of the source signals.

responses of the columns of $\mathbf{E}_x^H \mathbf{A}$ do not show a clear band-pass characteristic anymore which includes the frequency contents of both source signals. Rather the filters look more similar to a multi-pass filter structure. Another problem is that the eigenvalues of the first standard eigendecomposition start to have very low values. In that case it is indicated to consider a threshold $\Theta < 1$ and also reduce the number of signals at the output of the first step from $NK$ to $l < NK$.

In a second experiment three artificially generated, narrow-band source signals with different frequency contents were chosen: one member of the group represents a narrow-band signal, a sinusoid; the second signal encompasses a wide frequency range; and the last one represents a sawtooth wave whose spectral density is concentrated in the low frequency band (see Fig. 3).

For any number $K$ of delays, the number of signals after the first step (or the dimension of matrix $\mathbf{C}$) is $l = 6 < NK$, because the threshold $\Theta = 0.95$ eliminates all very low eigenvalues. Even though, the number of output signals is still higher than the number of source signals. Thus only 3 output signals which each has a high correlation (in the frequency domain) with one of the source signals will be considered in the following. It can be verified easily that upon increasing the number $K$ of delays, the estimated independent signals decrease their bandwidth (except for the sinusoid).

For example, Fig. 4 shows that the spectrum of output signal 2 has less components when $K$ increases. The effect is also visible in the spectrum of output signal 3, but here the time domain characteristics of the wave are less affected as is to be expected. This
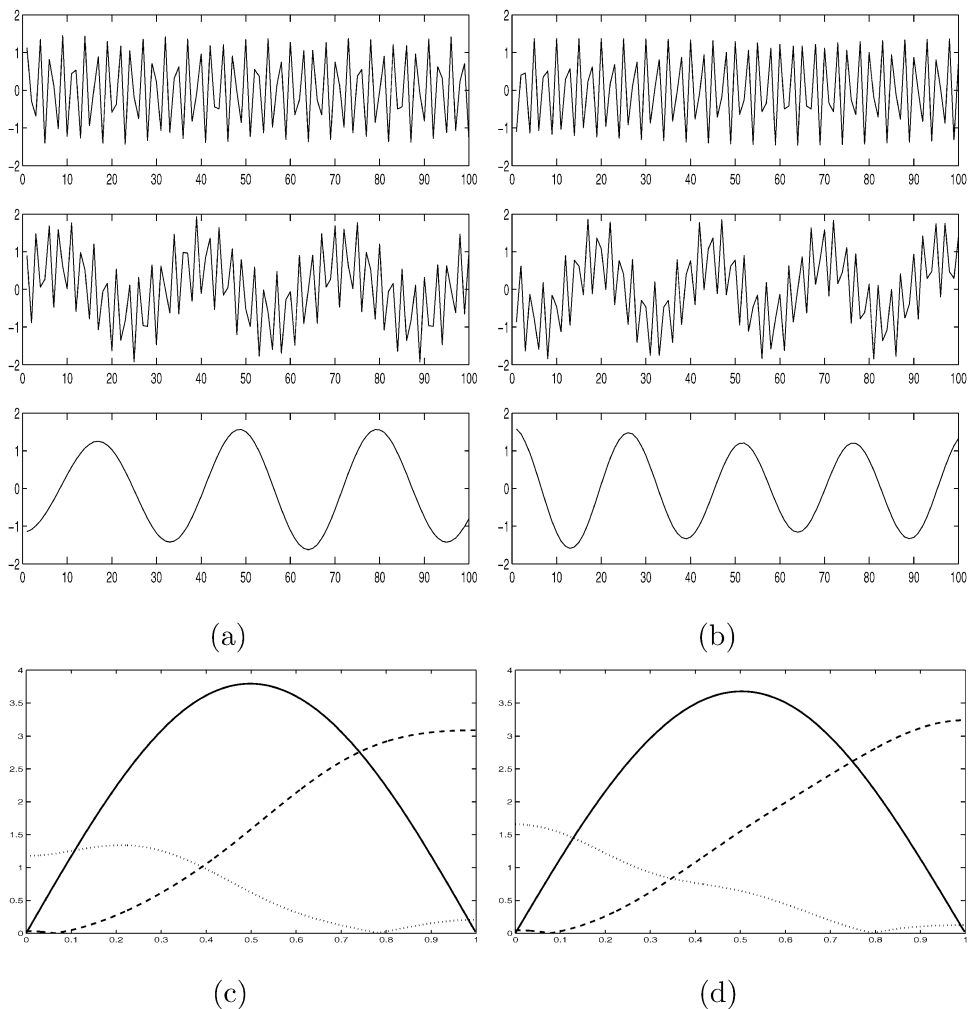
Fig. 2. (a) Independent components correlated with source signal 1, (b) independent components correlated with source signal 2, (c) frequency response amplitude of filters related with components represented in (a), (d) frequency response amplitude of filters related with components represented in (b).

artificially mixed setup shows that the generalized eigenvalue decomposition in high-dimensional space results in a filtering indeterminacy that depends on the number of delays $K$. However it has been verified that the frequency range of every source is covered by the reconstructed signals. And if the sources have a narrow band frequency contents, it is possible to separate that source from all others.

### 3.1.2. Signal-to-noise ratios and thresholds

The second group of simulations consider the influence of noise on the performance of the algorithm. The set of three artificial source signals of the last section was linearly
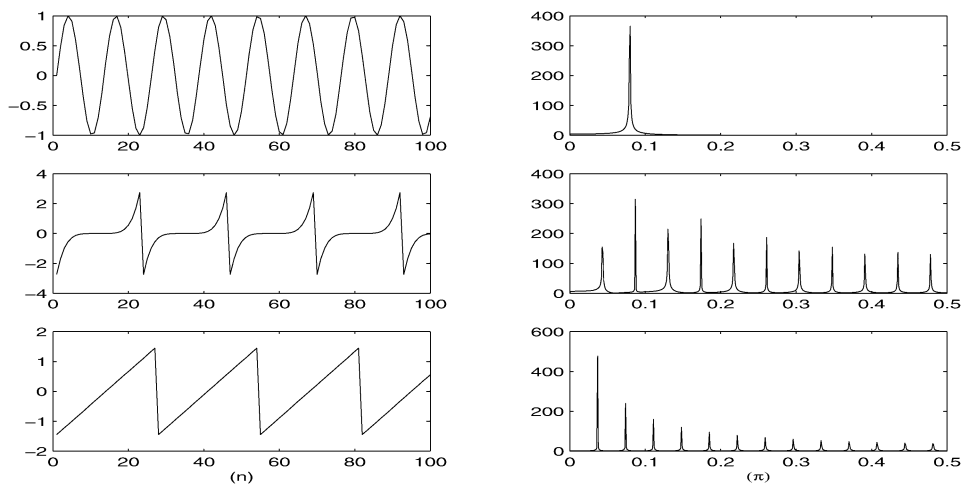
Fig. 3. Artificial signals (left column) and their frequency contents (right column).
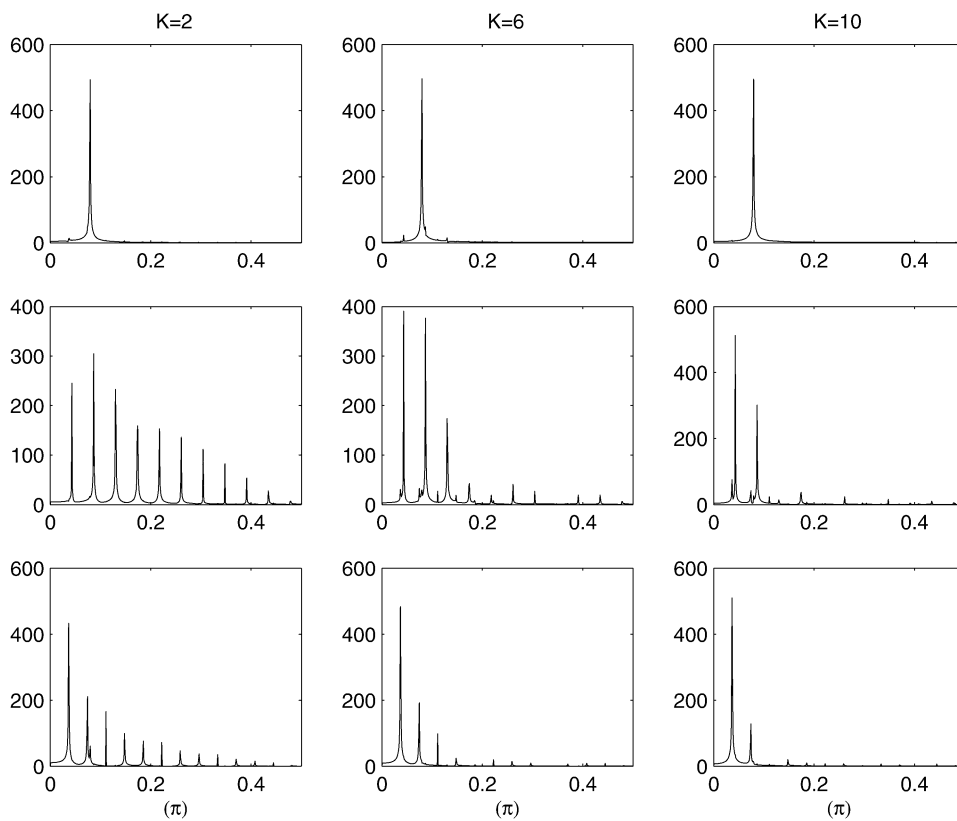


Fig. 4. Frequency contents of the output signals considering different time-delays $K$ to form the input data matrix, hence the embedding dimension of the feature space.
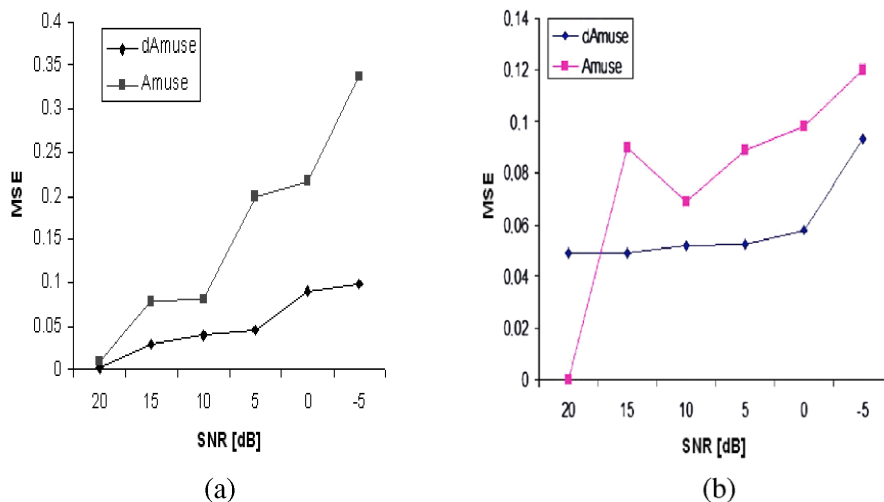
Fig. 5. Comparing AMUSE and dAMUSE using the MSE error: (a) sinusoid signal, (b) triangular signal.

mixed and random noise was added, resulting in a signal-to-noise ratio SNR in the range of $-5$ dB $\leqslant$ SNR $\leqslant 20$ dB. The mixtures then have been analyzed with the algorithm dAMUSE. The parameters of the algorithm were chosen as follows: $K = 4$, $\Theta = 0.95$, $d = 1$. In what concerns noise we also tried to figure out if there is any advantage of using a GEVD instead of a PCA analysis. Hence the signals at the out put of the first step of the algorithm (using the matrix $\mathbf{Q}$ to project the data) are also compared with the signals obtained after the second step.

The first experiment compares the performance of algorithm dAMUSE with AMUSE (the pencil has $3 \times 3$ matrices and was computed also with $d = 0$ and 1), using the mean square error (MSE) to measure the deviations of the source signals from the recovered signals for the different noise levels. Only those filtered signals, recovered at the output of dAMUSE, which retain most of their frequency content compared to the original were used in the comparison (i.e., the sinusoid and sawtooth signals). Figure 5 shows that dAMUSE performs the better the higher the noise level, i.e., for SNR $\leqslant 20$ dB for both signals. Figure 5b, however, shows a better performance for AMUSE when SNR $= 20$ dB, but this is a natural consequence of the filtering process causing a lack of high-frequency components for the filtered version of the source signal obtained at the output of dAMUSE.

The following results illustrate the role of the threshold parameter $\Theta$ selecting the dimension of the signal subspace when noisy signals are considered. The parameters $K = 4$ and $\Theta = 0.95$ were kept fixed. As the noise level increases, the number of significant eigenvalues also increases. Hence, at the output of the first step more signals need to be considered. Thus as the noise energy increases, the number of signals ($l$) or the dimension of matrix $C$ after the application of the first step increases (see last column of Table 1). As a consequence, with increasing noise level an increasing number of independent components will be available at the output of the second step also. Computing, in the frequency domain, the correlation coefficients between either the output signals of each step of the algorithm and noise or the output signals and the source signals, we confirm, that some are

Table 1
Number of output signals correlated with either noise or source signals after each step of the algorithm dAMUSE

| SNR (dB) | NK | First step | | Second step | | |
|---|---|---|---|---|---|---|
| | | Sources | Noise | Sources | Noise | Total |
| 20 | 12 | 6 | 0 | 6 | 0 | 6 |
| 15 | 12 | 5 | 2 | 6 | 1 | 7 |
| 10 | 12 | 6 | 2 | 7 | 1 | 8 |
| 5 | 12 | 6 | 3 | 7 | 2 | 9 |
| 0 | 12 | 7 | 4 | 8 | 3 | 11 |

related with the sources and others with noise. Table 1 (columns 3–6) shows that the maximal correlation coefficients are distributed between noise and source signals to a varying degree, but the number of signals correlated with noise (Table 1) is always higher in the first level.

These results show that for a low noise level, i.e., a high SNR, the first step (which is mainly a principal component analysis in a space of dimension $NK$) already achieves good solutions. But when the noise level increases, hence the SNR decreases, the additional second step, representing the GEVD of the matrix pencil improves results considerably. This was corroborated also by adding high energy noise and taking only the five most significant eigenvalues, and corresponding eigenvectors, at the output of the first step. In these cases the source signals might not be recovered.

### 3.2. Heart rate variability study

A second study to show the performance of the algorithm dAMUSE deals with real world signals. We use ECG signals of two young normal subjects from the POLY/MEDLAV database [17]. One lead of the recorded ECG signals was processed by a wavelet transform based automatic delineation system [18], and the RR and QT intervals (see Fig. 6 for an explanation) were measured between the marks obtained and form the non-uniform sequences which are processed by dAMUSE.

The ECG signals of these two subjects were chosen because the corresponding respiration signals exhibit spectral contents in distinct bands:

(1) ECG1 corresponds to a respiratory signal in a low frequency band LF = (0.05–0.1 Hz).
(2) ECG2 corresponds to a signal in a high frequency band HF = (0.2–0.4 Hz).

Each heart beat (Fig. 6) is characterized by two values: the time interval between consecutive R peaks (RR) on the one hand and the interval between the onset of the QRS complex and the end of the T wave (QT) on the other hand. Then two sequences, synchronized with the onset of the QRS complex, are constituted by the values of the RR intervals and the QT intervals. The algorithm dAMUSE was applied to $[RR(i)$ and $[RR(i), QT(i)]^T$ sequences of time intervals (to apply the formalism given in Section 2, simply set $\tau = 1$ and $\Delta t = 1$ and $t_0 = 0$ in Eq. (1)). Different numbers of delays were tested but the best results were achieved by $K = 8$, the other parameters of the algorithm are $d = 1$, and $\Theta = 0.95$. The estimated independent component signals are not correlated in the time domain and suffer
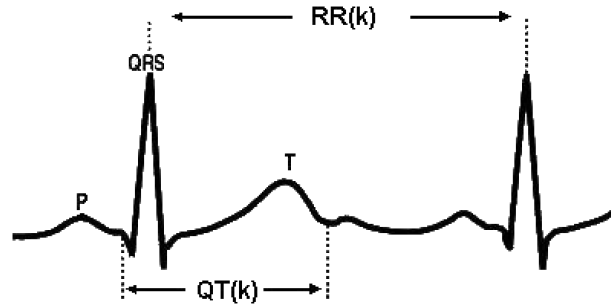
Fig. 6. Schematic representation relevant waves and intervals in a cycle of the ECG signal.

from an indeterminacy on its amplitude as it is expected from these methods. Then, in order to achieve a direct comparison, the ICs were normalized to unit variance and amplitudes in the range $[-1, 1]$. However, in HRV studies, the frequency contents in distinct frequency bands are the relevant information [19]. Hence an analysis in frequency space is performed for each independent signal using the Welch method [20]. The results in the frequency domain are compared graphically and the correlation coefficients are also computed.

### 3.2.1. RR time series

In a first experiment we considered the time intervals between the R-waves of the ECGs. For these experiments a trajectory matrix is formed with the input vectors $\vec{x}_i(t_l)$, using $K = 8$ delayed versions of RR time intervals. The number of different independent components estimated for both ECG signals studied is not the same as can been see in Figs. 7a and 7c).

The sequence of RR time intervals contains a prominent high-frequency component if the corresponding respiratory signal contains a high frequency component as well, i.e., in ECG2 signal. Nevertheless, we can see that in both cases (ECG1 and ECG2) the components computed correspond to distinct frequency ranges. Table 2 also shows that those components, which have correlation coefficients equal to zero in the time domain, also have small values when the coefficients are computed in the frequency domain.

### 3.3. RR-QT time series

For these experiments the input vectors $\vec{x}_i(t_l)$, $i = 1, 2$ were used to form a concatenated trajectory matrix using $K = 8$ delayed RR- and $K = 8$ delayed QT time intervals. The parameters of the algorithm ($\Theta = 0.95$ and $d = 1$) are the same as before, but here the number of output signals is higher (6 and 4, respectively for ECG1 and ECG2). Among all uncorrelated components estimated those three components $\vec{y}_j$, $j = 1, 2, 3$, with spectral densities in distinct frequency ranges (see Fig. 8) are chosen which show the highest contributions to the reconstruction of the original RR and QT sequences.

In Table 3 the correlations between the frequency contents of the uncorrelated components is shown. The results confirm that the correlation between the frequency contents of the uncorrelated signals estimated is very low. Also, two of those signals have a frequency content in the range of the LF and HF fluctuations considered in the HRV studies.
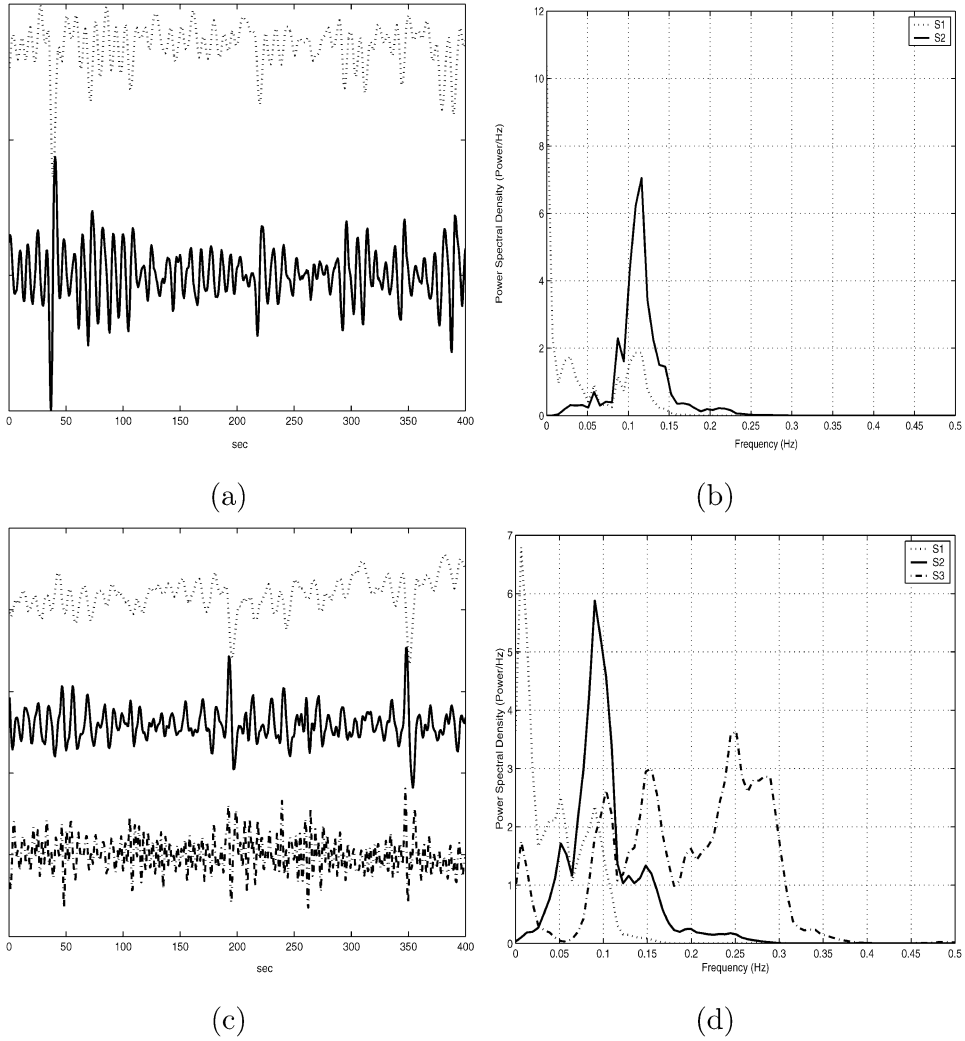
(a)



(b)



(c)



(d)

Fig. 7. ECG1: (a) independent components, (b) spectral contents of independent components; ECG2: (c) independent components, (d) spectral contents of independent components.

Table 2
Correlation coefficients between spectra of independent components considering the RR sequence

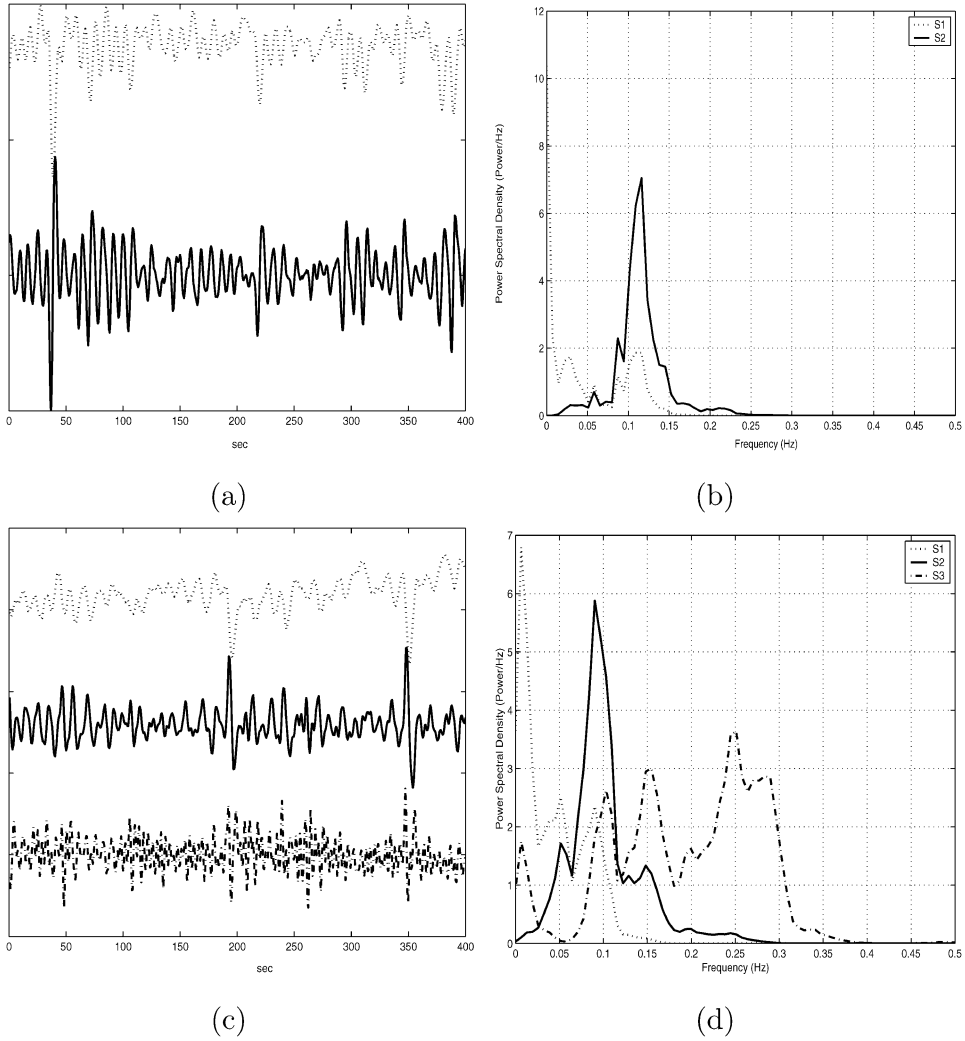|       | ECG1  |       |       | ECG2  |       |       |
|-------|-------|-------|-------|-------|-------|-------|
|       | $s_1$ | $s_2$ | $s_3$ | $s_1$ | $s_2$ | $s_3$ |
| $s_1$ | 1     | 0.26  | –     | 1     | 0.38  | 0.13  |
| $s_2$ | 0.26  | 1     | –     | 0.38  | 1     | 0.33  |
| $s_3$ | –     | –     | –     | 0.13  | 0.33  | 1     |

(a)

(b)

(c)

(d)

Fig. 8. ECG1: (a) independent components, (b) spectral densities of independent components; ECG2: (c) independent components, (d) spectral densities of independent components.

Table 3
Correlation coefficients between the uncorrelated component spectra

|  | ECG1 | | | ECG2 | | |
|---|---|---|---|---|---|---|
|  | $s_1$ | $s_2$ | $s_3$ | $s_1$ | $s_2$ | $s_3$ |
| $s_1$ | 1 | 0.10 | 0.09 | 1 | 0.38 | 0.14 |
| $s_2$ | 0.10 | 1 | 0.03 | 0.38 | 1 | 0.13 |
| $s_3$ | 0.09 | 0.03 | 1 | 0.14 | 0.13 | 1 |

For both signals the LF and VLF components are very similar to the ones computed with sequences of RR intervals as can be seen comparing Figs. 7 and 8. In fact, the correlation coefficients computed between VLF or LF components using RR or (RR, QT) sequences are $\geqslant 0.95$. The signal ECG1 now presents a third component, representing the high frequency band which was not found when the input corresponded to the RR sequence alone. This suggests, that the intervals RR and QT reflect different influences of the autonomous nervous system as also discussed in Ref. [16]. In spite of a third component appearing for ECG2, the dissimilarity found in shape and dispersion of the spectra with RR or (RR, QT) sequences as inputs reinforces the supposition of different effects. This hypothesis was already stated in Ref. [21] when studying the dynamic relation between HRV and QTV using a very different approach; furthermore our results are consistent with the ones presented in Ref. [21].

### 3.4. Water artifact removal and denoising of NMR spectra

Because of magnetic field inhomogeneities and peak suppression techniques the water signal cannot be modelled analytically. This prevents maximum likelihood estimations of model parameters within a nonlinear least squares approach to be applied successfully. Hence, in recent years post-processing algorithms have been developed to remove the water resonance (simply called water artifact) from the NMR spectra [22–29]. Early algorithmic approaches use rather crude approximations, causing changes in peak area and phase which render the structure determination process rather unreliable, hence are unacceptable in protein NMR. More sophisticated water suppression algorithms are based on second order statistics of the signals within an unsupervised approach. They usually use a state-space approach in the time domain by modelling the signal subspace using second order techniques like singular value decomposition (SVD) [30] or some variant of singular spectral analysis (SSA) [28]. With multidimensional NMR data, these approaches are computationally very demanding in general. Furthermore they have been applied so far to rather "simple" NMR spectra only, where the BSS techniques discussed above can yield rather perfect solutions. Blind source separation techniques using independent component analysis methods recently became famous for their ability to blindly decompose complicated signals into their underlying component signals without using any prior knowledge. Hence it is interesting whether BSS techniques can contribute to the removal of the water artifact in proton NMR spectra of aqueous biomolecular solutions.

We show that the algorithm dAMUSE can be used elegantly to separate the water artifact from a protein proton NMR spectrum and *simultaneously* perform denoising of the reconstructed spectrum. We demonstrate this combination of BSS and denoising with 2D NOESY proton NMR spectra of the polypeptide P11 [11].

Figure 9 shows a 1D slice (corresponding to the shortest evolution period of the NMR pulse sequence) of the original proton NMR spectrum of P11 and the corresponding artefact-free spectrum obtained with a GEVD using a matrix pencil in the frequency domain [31]. The huge water artefact could be removed successfully but only at the expense of an increased noise level of the reconstructed protein spectrum. Hence denoising as a post-processing deemed necessary. But dAMUSE can achieve both goals in one step as we will demonstrate next.

Original NMR spectrum of the P11 protein

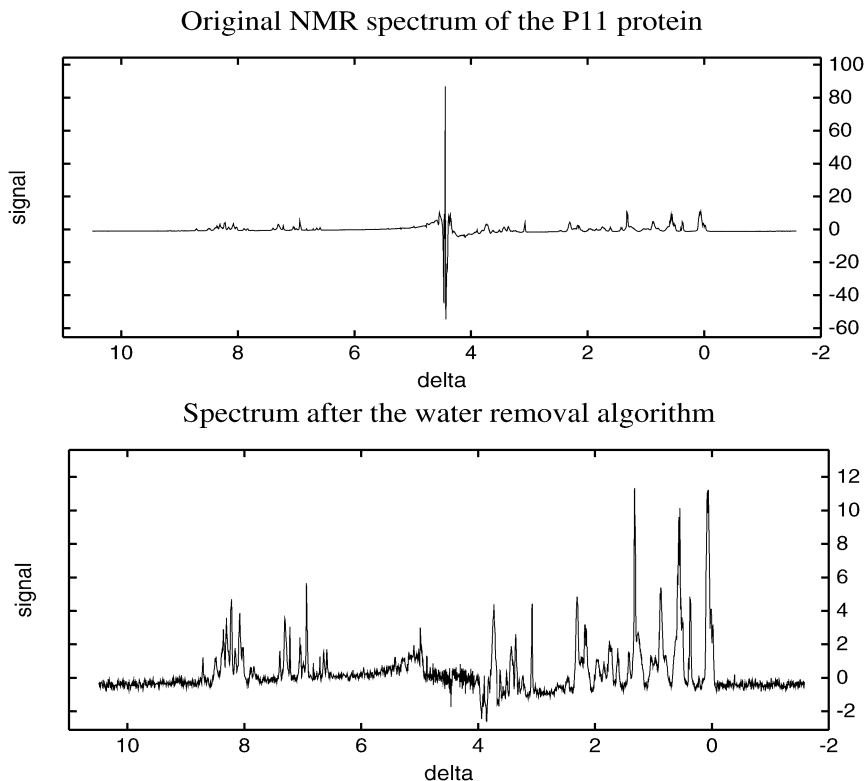Spectrum after the water removal algorithm

Fig. 9. Proton NMR spectrum of the polypeptide P11 top: original; bottom: GEVD-MP [31].

The starting point for the application of dAMUSE to P11 were 128 complex time domain signals taken from a 2D NOESY proton NMR experiment where each of the 128 signals consisted of 2048 samples taken with a sampling rate of $\tau^{-1} = (90\,\mu s)^{-1} \simeq 11$ kHz. For every signal the component trajectory matrix $\mathbf{X}_i$, $i = 1, \ldots, 128$ (cf. Eq. (2)), was formed by using only one time delay ($K = 2$) of size $\Delta t = 2\tau = 180\,\mu s$. Thus the resulting overall trajectory matrix $\mathbf{X}$ had a size $256 \times 2046$.

The data of the trajectory matrix is manipulated in the frequency domain to compute a filtered version $\mathbf{X}_f$ of the data in $\mathbf{X}$. The matrix $\hat{\mathbf{X}}$ was determined by computing the discrete Fourier transform of each row of $\mathbf{X}$. A filtered version of the matrix $\hat{\mathbf{X}}$ was generated by applying a gaussian shaped function to every row of $\hat{\mathbf{X}}$ leading to a matrix $\hat{\mathbf{X}}_f$ of bandpass-filtered sensor signals. The Gaussian filter was centered in the region of the water peak and had a width $\sigma = 1$. The filtering operation is thus achieved by computing the Hadamard product between the rows of $\hat{\mathbf{X}}$ and the frequency response function of a Gaussian bandpass filter [32]. Finally, an inverse discrete Fourier transform of each row of the matrix $\hat{\mathbf{X}}_f$ was calculated to obtain a filtered version of the time domain data matrix $\mathbf{X}_f$. The two matrices are used to compute the matrix pencil formed with $\mathbf{R} = \langle \mathbf{X}\mathbf{X}^H \rangle$ and $\mathbf{R}_f = \langle \mathbf{X}_f \mathbf{X}_f^H \rangle$.
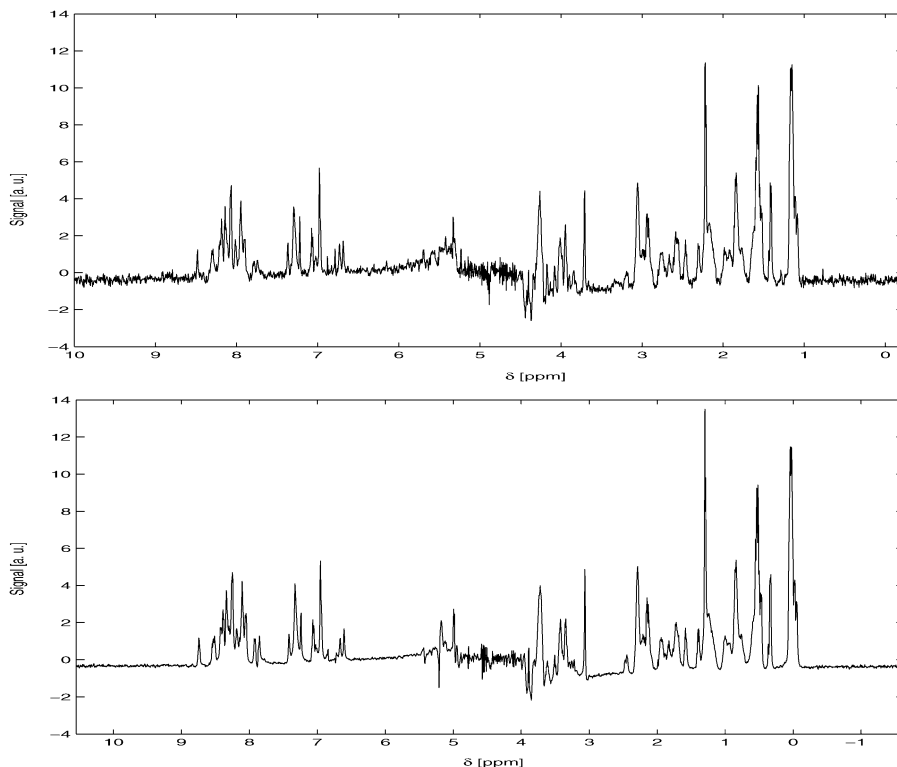
Fig. 10. Reconstructed spectrum of the polypeptide P11 using top: GEVD-MP [31]; bottom: GEVD-dAMUSE.

Next the two step procedure explained in Section 2.3 was applied. Only the 95 largest eigenvalues of the first EVD of the $256 \times 256$ correlation matrix $\mathbf{R}$ were considered after the first step in order to reduce noise. Then, the $95 \times 256$ matrix $\mathbf{Q}$ was formed and the EVD problem of the matrix $\mathbf{C} = \mathbf{Q}\mathbf{R}_f\mathbf{Q}^H$ was solved. Its eigenvector matrix $\mathbf{U}$ and the transformation matrix $\mathbf{Q}$ lead to the eigenvector matrix $\mathbf{E}_x = \mathbf{Q}^H\mathbf{U}$ of the matrix pencil according to Section 2.3. Finally, those estimated components of $\mathbf{Y} = \mathbf{E}_x^H\mathbf{X}$ which showed a significant spectral density only at the resonance frequency of the water protons were set to zero and with the remaining ICs the artifact-free protein spectra were reconstructed. Note that if less than the 95 largest eigenvalues were used the separation of the water and the protein signals failed whereas considering more than 100 of the largest eigenvalues lead to a drastic increase in noise. Figure 10 compares the results obtained by the standard GEVD-MP [31] and the dAMUSE algorithms corresponding to SNRs of 17.3 and 22.43 dB, respectively.

## 4. Conclusions

In this work we propose dAMUSE, a modified version of the algorithm AMUSE. The new algorithm is also based on a generalized eigendecomposition of a congruent matrix

pencil. But the latter is formed with correlation matrices which are computed with sensor signals embedded in a high-dimensional feature space of time-delayed coordinates much in the spirit of methods used in singular spectrum analysis [8]. It was shown that the independent components obtained at the output represent filtered versions of the unknown source signals. An implementation of the algorithm dAMUSE was also proposed which differs from the implementation of the algorithm AMUSE (and other blind source separation algorithms). dAMUSE defines a variance criterion for choosing the number of signals at the output of the first step (which is very similar to a principal component decomposition), hence paves the way for a denoising of the signals in parallel to the BSS.

The algorithm has a set of parameters, whose proper choice must be further studied, particularly the number of delays $K$ used to build the data set and the optimal choice of the threshold parameter $\Theta$. The choice of $K$ naturally constrains the linear filtering operation that characterizes the method, as was shown in the simulations with artificial signals. It was shown that additional a priori knowledge of characteristics of the spectral density of the signals considered can help in choosing $K$ appropriately.

The simulations also reveal that noise reduction cannot be achieved completely by PCA alone. Then having at the output a high number of signals, it is also important to find an automatic procedure to choose the relevant signals, as in case of noisy signals some might be related with noise only. The choice of time-delayed matrices to compute the pencil was done similar to the algorithm AMUSE, as well as the values for time-delays ($d$) [1]. With the NMR data an alternative filtering in the frequency domain was used to estimate one of the matrices of the pencil as discussed in other works [33–35] and good results were also achieved.

In what concerns the HRV study, blind source separation techniques were used to study similar time-series signals (RR and QT intervals) and to discuss the relation of the uncorrelated components extracted with the LF and HF fluctuations [16] of ECG signals. This preliminary study achieves results corroborating an earlier study [21] in spite of the use of different processing steps concerning the pre-processing of the temporal sequences and the method to estimate the uncorrelated signals. The new algorithm seems to provide a promising tool to HRV studies. Nevertheless a validation against a conventional method as well as a large data set should be used to verify the reliability and the performance of the method.

Concerning the study of 2D NOESY NMR spectra of dissolved proteins, dAMUSE combines in an elegant and efficient way blind source separation techniques with local projective denoising techniques. It offers a very fast and efficient way of removing the water artifact from the spectra and allows a denoising of the reconstructed artifact-free protein spectra to achieve noise levels at least comparable to those of the experimental spectra. Although the unknown source signals cannot be obtained straightforwardly, the method seems well suited to deal with narrow band signals where filtered versions of the signals do not differ in essential characteristics from their underlying originals. Hence the BSS problem can be solved using only the filtered signals. As in case of NMR spectra one has to deal with a large number of extracted independent components, an automatic assignment procedure is needed to assign extracted independent components to the different signals (artifact, protein, noise). Corresponding investigations are under way in our laboratories.

## Acknowledgments

## References

[1] A. Cichocki, S.-I. Amari, Adaptive Blind Signal and Image Processing, Wiley, New York, 2002.

[2] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE Trans. Signal Process. 45 (2) (1997) 434–444.

[3] R. Gharieb, A. Cichocki, Second-order statistics based blind source separation using a bank of subband filters, Digital Signal Process. 13 (2003) 252–274.

[4] T. Schreiber, H. Kantz, Nonlinear projective filtering $ii$: Application to real time series, in: Proceedings of Nolta, 1998.

[5] R. Vetter, J. Vesin, P. Celka, J.K.P. Renevey, Automatic nonlinear noise reduction using local principal component analysis and mdl parameter selection, in: Proceedings of the IASTED International Conference on Signal Processing Pattern Recognition and Applications (SPPRA '02), Crete, 2002, pp. 290–294.

[6] P. Gruber, F.J. Theis, K. Stadlthanner, E.W. Lang, A.M. Tomé, A.R. Teixeira, Denoising using local ICA and kernel-PCA, in: Proceedings of International Joint Conference on Neural Networks, IJCNN '04, Budapest, 2004.

[7] V. Moskvina, K.M. Schmidt, Approximate projectors in singular spectrum analysis, SIAM J. Math. Anal. Appl. 24 (4) (2003) 932–942.

[8] M. Ghil, M. Allen, M.D. Dettinger, K. Ide, Advanced spectral methods for climatic time series, Rev. Geophys. 40 (1) (2002) 3.1–3.41.

[9] L. Tong, R.-W. Liu, V.C. Soon, Y.-F. Huang, Indeterminacy and identifiability of blind identification, IEEE Trans. Circuits Syst. 38 (5) (1991) 499–509.

[10] A.M. Tomé, A.R. Teixeira, E.W. Lang, K. Stadlthanner, A. Rocha, Blind source separation using time-delayed signals, in: International Joint Conference on Neural Networks, IJCNN '04, vol. CD, Budapest, 2004.

[11] K. Stadlthanner, E.W. Lang, A.M. Tomé, A. Teixeira, C.G. Puntonet, Kernel-PCA denoising of artifact-free protein NMR spectra, in: Proc. IJCNN '04, Budapest, 2004.

[12] A.R. Teixeira, A.P. Rocha, R. Almeida, A.M. Tomé, The analysis of heart rate variability using independent component signals, in: Second International Conference on Biomedical Engineering, BIOMED '04, IASTED, Innsbruck, 2004, pp. 240–243.

[13] A.M. Tomé, N. Ferreira, On-line source separation of temporally correlated signals, in: European Signal Processing Conference, EUSIPCO '02, Toulouse, 2002.

[14] B.N. Parlett, The symmetric eigenvalue problem, in: SIAM's Classics in Applied Mathematics, SIAM, 1998.

[15] R. Vetter, P. Celka, J.M. Vesin, U. Scherrer, Sub-signal extraction of RR time series using independent component analysis, in: International Conference of the IEEE Engineering Medicine and Biology Society, vol. 20, 1998, pp. 286–289.

[16] R. Vetter, N. Virag, J.M. Vesin, P. Celka, U. Scherrer, Observer of autonomic outflow based on blind source separation of the ECG parameters, IEEE Trans. Biomed. Engineer. 47 (5) (2000) 578–582.

[17] F. Pinciroli, G. Pozzi, R. Rossi, M. Piovosi, A. Capo, R. Olivieri, M.D. Torre, A respiration-related ekg database, in: I.C. Society (Ed.), Computers in Cardiology, 1988, pp. 477–480.

[18] J.P. Martínez, R. Almeida, S. Olmos, A.P. Rocha, P. Laguna, A wavelet-based ECG delineator: Evaluation on standard databases, IEEE Trans. Biomed. Engineer. 51 (4) (2004) 570–581.

[19] 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, Technical report, 1996.

[20] A.V. Oppenheim, R.W. Schafer, Digital Signal Processing, Prentice Hall, New York, 1975.

[21] R. Almeida, E. Pueyo, J.P. Martinez, A.P. Rocha, P. Laguna, S. Olmos, Quantification of short term QT variability versus heart rate variability, in: Publication on Proceedings of 7th Portuguese Conference of Biomedical Engineering, BIOENG '03, Lisbon, 2003.

[22] D. Marion, M. Ikura, A. Bax, Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data, J. Magn. Reson. 84 (1989) 425–430.

[23] Y. Kuroda, A. Wada, T. Yamazaki, K. Nagayama, Postaquisition data processing method for suppression of the solvent signal, J. Magn. Reson. 84 (1989) 604–610.

[24] M. Derich, X. Hu, Elimination of water signal by postprocessing, J. Magn. Reson. 101 (1993) 229–232.

[25] K.J. Cross, Improved digital filtering technique for solvent suppression, J. Magn. Reson. 101 (1993) 220–224.

[26] J.H.J. Leclerc, Distortion-free suppression of the residual water peak in proton spectra by postprocessing, J. Magn. Reson. B 103 (1994) 64–67.

[27] C.J. Craven, J.P. Waltho, The action of time-domain convolution filters for solvent suppression, J. Magn. Reson. B 106 (1995) 40–46.

[28] G. Zhu, D. Smith, Y. Hua, Post-aquisition solvent suppression by singular-value decomposition, J. Magn. Reson. 124 (1997) 286–289.

[29] A. van den Boogaart, D. van Ormondt, W.W.F. Pijnapple, R. de Beer, M. Ala-Korpela, in: J.G. McWhriter (Ed.), Removal of the Water Resonance from [1]H Magnetic Resonance Spectra, in: Mathematics in Signal Processing, vol. III, Clarendon, Oxford, 1994.

[30] W.W.F. Pijnapple, A. van den Boogaart, R. de Beer, D. van Ormondt, SVD-based quantification of magnetic resonance signals, J. Magn. Reson. 97 (1992) 122–134.

[31] K. Stadlthanner, A.M. Tomé, F.J. Theis, W. Gronwald, H.R. Kalbitzer, E.W. Lang, On the use of independent component analysis to remove water artifacts of 2D NMR protein spectra, in: 7th Portuguese Conference on Biomedical Engineering, BIOENG '03, Lisbon, 2003.

[32] K. Stadlthanner, A. Tomé, F.J. Theis, W. Gronwald, H.R. Kalbitzer, E.W. Lang, Removing water artifacts from 2D protein spectra using GEVD with congruent pencils, in: 7th International Symposium on Signal Processing and its Applications, vol. 2, Paris, 2003, pp. 85–88.

[33] A.M. Tomé, Separation of a mixture of signals using linear filtering and second order statistics, in: 10th European Symposium on Artificial Neural Networks, ESANN '02, Brugges, 2002, pp. 307–312.

[34] J.V. Stone, Blind source separation using temporal predictibility, Neural Comput. 13 (7) (2001) 1557–1574.

[35] K. Stadlthanner, F. Theis, E.W. Lang, A.M. Tomé, W. Gronwald, H.R. Kalbitzer, A matrix pencil approach to the blind source separation of artifacts in 2D NMR spectra, Neural Inform. Process. Lett. Rev. 1 (2003) 103–110.

**Ana M. Tomé** received her PhD in electrical engineering from University of Aveiro in 1990. Currently she is associate professor of electrical engineering in the Department of Electronics and Telecomunications/IEETA of the University of Aveiro, where she teaches courses on digital signal processing for electronics and computer engineering diplomas. Her research interests include digital and statistical signal processing, independent component analysis, and blind source separation as well as classification and pattern recognition applications of neural networks.

**Ana R. Teixeira** received her diploma degree in mathematics applied to technology from University of Porto in 2003. She is currently a MsC student of electronics and telecommunications of University of Aveiro. Her research interests include biomedical digital signal processing and principal and independent component analysis.

**Elmar W. Lang** received his physics diploma in 1977 and his PhD in physics in 1980 and habilitated in biophysics in 1988 at the University of Regensburg. He is apl. professor of biophysics at the University of Regensburg, where he is heading the Neuro- and Bioinformatics Group. Currently he serves as associate editor of Neurocomputing and Neural Information Processing—Letters and

Reviews. His current research interests include biomedical signal and image processing, independent component analysis and blind source separation, neural networks for classification and pattern recognition, and stochastic process limits in queuing applications.

**Kurt Stadlthanner** received his physics diploma from the University of Regensburg in 2003. He is currently a doctoral student at the Institute of Biophysics, Neuro- and Bioinformatics Group, University of Regensburg. His scientific interests are in the fields of biological data processing and analysis by means of blind source separation and support vector machines.

**Rute Almeida** was born in Porto, Portugal, in 1979. She received a 4 years degree on Mathematics Applied to Technology in 2000, from Faculty of Sciences, University of Porto (FCUP). She was with the Autonomic Function Study Center from Hospital S. João between March and October 2000, working in methods for automatic delineation of the ECG. Since October 2001 she is a PhD student of the Applied Mathematics Department of FCUP, supported by a grant from Foundation for Science and Technology (Portugal) and European Social Fund. Her main research interests are in time-scale methods and the automatic analysis of the ECG, namely the study of ventricular repolarization.

**Ana Paula Rocha** was born in Coimbra, Portugal, in 1957. She received the Applied Mathematics degree and the Doctor degree (PhD in applied mathematics, systems theory, and signal processing) from the Faculty of Sciences, University of Porto (FCUP), in 1980 and 1993, respectively. She is currently an auxiliar professor in the Department of Applied Mathematics at FCUP, Portugal. Her research interests are in biomedical signals and system analysis (EMG, cardiovascular systems analysis and autonomic nervous system characterization), time–frequency/time-scale signal analysis, point processes spectral analysis, data treatment and interpretation.