Delayed AMUSE – A Tool for Blind Source Separation and Denoising

Ana R. Teixeira[1], Ana Maria Tomé[1], Elmar W. Lang[2], and Kurt Stadlthanner[2]

[1] Departamento de Electrónica e Telecomunicações/IEETA
Universidade de Aveiro, P-3810 Aveiro, Portugal
ana@ieeta.pt

[2] Institute of Biophysics, Neuro- and Bioinformatics Group
University of Regensburg, D-93040 Regensburg, Germany
elmar.lang@biologie.uni-regensburg.de

Abstract. In this work we propose a generalized eigendecomposition (GEVD) of a matrix pencil computed after embedding the data into a high-dim feature space of delayed coordinates. The matrix pencil is computed like in AMUSE but in the feature space of delayed coordinates. Its GEVD yields filtered versions of the source signals as output signals. The algorithm is implemented in two EVD steps. Numerical simulations study the influence of the number of delays and the noise level on the performance.

1 Introduction

Blind Source Separation (BSS) methods consider the separation of observed sensor signals into their underlying independent source signals knowing neither these source signals nor the mixing process. BSS methods using second order statistics only can be based on a generalized eigendecomposition (GEVD) of a matrix pencil. They are exact and efficient but sensitive to noise [1].

There are several proposals to improve efficiency and robustness of these algorithms when noise is present[1], [2] which mostly rely on an approximative *joint* diagonalization of a set of correlation or cumulant matrices. Also there exist local projective de-noising techniques which in a first step increase the dimension of the data by joining delayed versions of the signals [3], [4], [5] hence projecting them into a high-dimensional feature space. A similar strategy is used in Singular Spectrum Analysis (SSA) [6] 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 was used to extract information from short and noisy time series and then provide insight into the underlying system that generates the series [7].

In this work we combine the ideas of solving BSS problems algebraically using a GEVD with local projective denoising techniques. We propose, like in AMUSE, a GEVD of two correlation matrices i.e, the *simultaneous* diagonalization of a

C.G. Puntonet and A. Prieto (Eds.): ICA 2004, LNCS 3195, pp. 287–294, 2004.

matrix pencil formed with a correlation matrix and a matrix of time-delayed correlations. But the proposed algorithm, called dAMUSE, computes the pencil in a high-dimensional feature space of time-delayed coordinates.

In the following section we show, starting from a noise-free model of linearly mixed sensor signals, that the estimated independent signals correspond to filtered versions of the underlying source signals. We also present an algorithm to compute the eigenvector matrix of the pencil which involves a two step procedure based on the standard eigendecomposition (EVD) approach. The advantage of this procedure is concerned with a dimension reduction between the two steps as well as a reduction in the number of independent signals, thus performing a denoising of the estimated underlying source signals. Finally, simulations with artificially mixed signals are discussed to illustrate the proposed method.

2 Generalized Eigendecomposition Using Time-Delayed Signals

Considering the sensor signals x_i, the trajectory matrix [6] of the sensor signals computed for a set of L samples is given by

$$X_i = \begin{bmatrix} x_i[M-1] & x_i[M] & x_i[M+1] & \cdots & x_i[L-1] \\ x_i[M-2] & x_i[M-1] & x_i[M] & \cdots & x_i[L-2] \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ x_i[0] & x_i[1] & x_i[2] & \cdots & x_i[L-M] \end{bmatrix} \tag{1}$$

and encompasses M delayed versions of the signal x_i. Given a group of N sensor signals, $x_i,\ i = 1 \dots N$, the trajectory matrix of the set will be a concatenation of the component trajectory matrices X_i computed for each sensor, i.e

$$X = \left[X_1, X_2, \dots X_N \right]^T \tag{2}$$

Assuming that each sensor signal is a linear combination $X = HS$ of N underlying but unknown source signals (s_i), a source signal trajectory matrix S can be written in analogy to eqn(1) and eqn(2). Then the mixing matrix (H) is a block matrix with a diagonal matrix in each block

$$H = \begin{bmatrix} h_{11} I_{M\times M} & h_{12} I_{M\times M} & \cdots & h_{1N} I_{M\times M} \\ h_{21} I_{M\times M} & h_{22} I_{M\times M} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1} I_{M\times M} & \cdots & \cdots & h_{NN} I_{M\times M} \end{bmatrix} \tag{3}$$

The matrix I_{MxM} represents the identity matrix and in accord with an instantaneous mixing model the mixing coefficient h_{ij} relates the sensor signal i with the source signal j.

The time-delayed correlation matrices of the matrix pencil are computed with one matrix (X_r) obtained by eliminating the first k_i columns of X and another matrix, (X_l), obtained by eliminating the last k_i columns. Then, the time-delayed correlation matrix $R_x(k_i) = X_r X_l^T$ will be an $NM \times NM$ matrix. Each of these two matrices can be related with a corresponding matrix in the source signal domain

$$R_x(k_i) = HR_s(k_i)H^T = HS_r S_l^T H^T \tag{4}$$

Then the two pairs of matrices $(R_x(k_1), R_x(k_2))$ and $(R_s(k_1), R_s(k_2))$ represent congruent pencils [8] with the following properties:

- Their eigenvalues are the same, i.e., the eigenvalue matrices of both pencils are identical: $D_x = D_s$.
- If the eigenvalues are non-degenerate (distinct values in the diagonal of the matrix $D_x = D_s$), the corresponding eigenvectors are related by the transformation $E_s = H^T E_x$.

Assuming that all sources are uncorrelated, the matrices $R_s(k_i)$ are block diagonal, having block matrices $R_{mm}(k_i) = S_{ri} S_{li}^T$ along the diagonal

$$R_s(k_i) = \begin{bmatrix} R_{11}(k_i) & 0 & \cdots & 0 \\ 0 & R_{22}(k_i) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{NN}(k_i) \end{bmatrix}$$

The eigenvector matrix of the GEVD of the pencil $(R_s(k_1), R_s(k_2))$ can be written as

$$E_s = \begin{bmatrix} E_{11} & 0 & \cdots & 0 \\ 0 & E_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{NN} \end{bmatrix} \tag{5}$$

where E_{mm} is the $M \times M$ eigenvector matrix of the GEVD of the pencil $(R_{mm}(k_1), R_{mm}(k_2))$. The independent components can be estimated from linearly transformed sensor signals via

$$Y = E_x^T X = E_x^T HS = E_s^T S$$

hence turn out to be filtered versions of the underlying source signals. As the eigenvector matrix E_s (eqn. 5) is a block diagonal matrix, there are M signals in each column of Y which are a linear combination of one of the source signals and its delayed versions. For instance, the block m depends on the source signal m via

$$\sum_{k=1}^{M} E_{mm}(k, j) s_m(n - k + 1) \tag{6}$$

Equation (6) defines a convolution operation between column j of E_{mm} and source signal s_m. Then the columns of the matrix E_{mm} represent impulse responses of finite impulse response (FIR) filters. Considering that all the columns of E_{mm} are different, their frequency response might provide different spectral densities of the source signal spectra. Then NM output signals y encompass M filtered versions of each of the N estimated source signals.

3 Implementation of the Algorithm

There are several ways to compute the generalized eigendecomposition. We resume a procedure valid if one of the matrices of the pencil is symmetric positive definite. Thus, we consider the pencil $(R_x(0), R_x(k_2))$ and perform the following steps:

Step 1: Compute a standard eigendecomposition of $R_x(0) = \vee \wedge \vee^T$, i.e, compute the eigenvectors (ν_i) and eigenvalues (λ_i). As the matrix is symmetric positive definite, the eigenvalues can be arranged in descending order $(\lambda_1 > \lambda_2 > ... > \lambda_{NM})$. In AMUSE (and other algorithms) this procedure is used to estimate the number of sources. But it can also be considered a strategy to reduce noise. Dropping small eigenvalues amounts to a projection from a high-dim signal plus noise feature space onto a lower dimensional manifold representing the signal+noise 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, those related with the embedded deterministic signal, according to

$$\frac{\lambda_1 + \lambda_2 + ... + \lambda_l}{\lambda_1 + \lambda_2 + ...\lambda_{NM}} \geq TH \tag{7}$$

If we are interested in the eigenvectors corresponding to directions of high variance of the signals, the threshold (TH) should be chosen such that their maximum energy is preserved. The transformation matrix can thus be computed using either the l most significant eigenvalues, in which case denoising is achieved, or all eigenvalues and respective eigenvectors. Similar to the whitening phase in many BSS algorithms the data matrix X can be transformed using

$$Q = \wedge^{-\frac{1}{2}} \vee^T \tag{8}$$

to calculate a transformed matrix of time-delayed correlations $C(k_2)$ to be used in the next step. Also note, that Q represents a $l \times NM$ matrix if denoising is considered.

Step 2: Using the transformed time-delayed correlation matrix $C(k_2) = QR_x(k_2)Q^T$ and its standard eigendecomposition: the eigenvector matrix (U) and eigenvalue matrix (D_x).

The eigenvectors of the pencil $(R_x(0), R_x(k_2))$, which are not normalized, form the columns of the eigenvector matrix $E_x = Q^T U = \vee \wedge^{-\frac{1}{2}} U$. The independent components of the time-delayed sensor signals can then be estimated via the transformation given below, yielding l (or NM) signals, one signal per row of Y

$$Y = E_x^T X = U^T Q X = U^T \wedge^{-\frac{1}{2}} \vee^T X \tag{9}$$

The first step of this algorithm is thus equivalent to a PCA in a high-dimensional feature space [4], [7] where a matrix similar to Q is used to project the data onto the signal manifold.

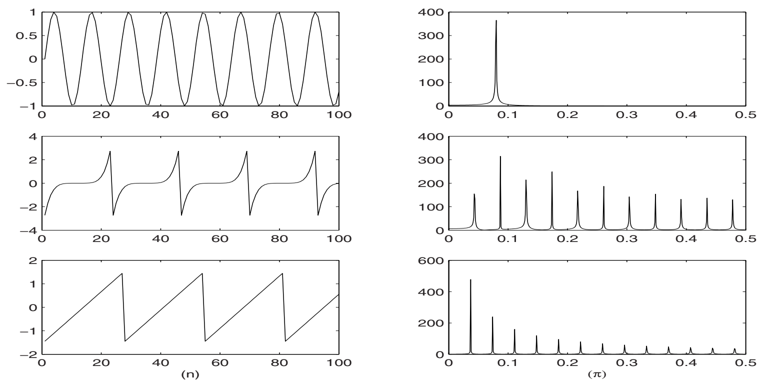

Fig. 1. Artificial signals (left column) and their frequency contents (right column).

4 Numerical Simulations

A group of three source signals with different frequency contents was 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. 1). The simulations were designed to illustrate the method and to study the influence of its parameters (especially M and TH) on the respective performance. As the separation process also involves a linear filtering operation (eqn. (6), each computed independent component has its maximum correlation with one of the source signals for a non-zero delay, besides the usual indeterminacy in order and amplitude. Two experiments will be discussed: a) one changing the number of delays M and the other b) adding noise with different levels. In what concerns noise we also try to find out if there is any advantage of using a GEVD instead of a PCA analysis. Hence the signals at the output of the first step of the algorithm (using the matrix Q to project the data) are also compared with the output signals.

After randomly mixing the source signals, the algorithm was applied for different values of M, with $TH = 0.95$, $k_1 = 0$ and $k_2 = 1$ fixed. In that case for any value of M the number of signals after the first step (or the dimension of matrix C) is $l = 6 < NM$, because the threshold TH eliminates the very low eigenvalues. Even though, the number of output signals is higher than the number of source signals, thus only 3 output signals which have the highest correlations (in the frequency domain) with any of the source signals will be considered in the following. It can be verified easily that upon increasing the number M of delays the estimated independent signals decrease their bandwidth (except for the sinusoid). Fig. 2 shows that source 2 has less components when M increases. The effect is also visible in source 3 but here the time domain characteristics of the wave are less affected as is to be expected.

The second experiment is related with the influence of the threshold parameter TH when noisy signals considered. First random noise was added to the

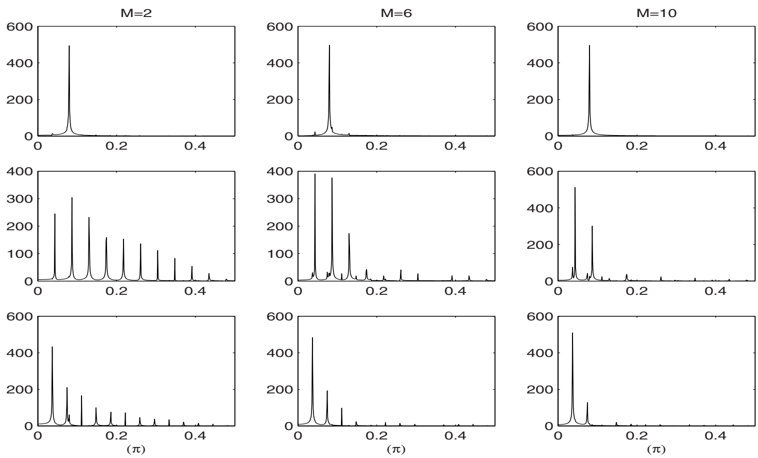

Fig. 2. Frequency contents of the output signals considering different time-delays M to form the input data matrix, hence the embedding dimension of the feature space.

Table 1. Number of output signals correlated with noise or source signals in the steps of the algoritm dAMUSE.

		1st step		*2nd step*		
SNR	**NM**	**Sources**	**Noise**	**Sources**	**Noise**	**Total**
20dB	12	6	0	6	0	6
15dB	12	5	2	6	1	7
10dB	12	6	2	7	1	8
5dB	12	6	3	7	2	9
0dB	12	7	4	8	3	11

sensor signals yielding a SNR in the range of $[0, 20]dB$. The parameters $M = 4$ and $TH = 0.95$ 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 also increases (last column of table 1). As the noise increases an increasing number of independent components will be available at the output of the two steps. Computing, in the frequency domain, the correlation coefficients between the output signals of each step of the algorithm and noise or source signals we confirm that some are 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. We can see that the number of signals correlated with noise (table 1) is always higher in the first level. Results show that for low noise levels the first step (which is mainly a principal component analysis in a space of dimension NM) achieves good solutions already. However, we can also see (for narrow-band signals and/or M low) that the time domain characteristics of the signals resemble the original source signals only after a

GEVD, i.e. at the output of the second step rather than with a PCA, i.e. at the output of first step. Fig. 3 shows examples of signals that have been obtained in the two steps of the algorithm for $SNR = 10dB$. At the output of first level the 3 signals with highest frequency correlation were chosen among the 8 output signals. Using a similar criterium to choose 3 signals at the output of the 2nd step (last column of Fig. 3), we can see that their time course is more similar to the source signals than after the first step (middle column of Fig. 3)

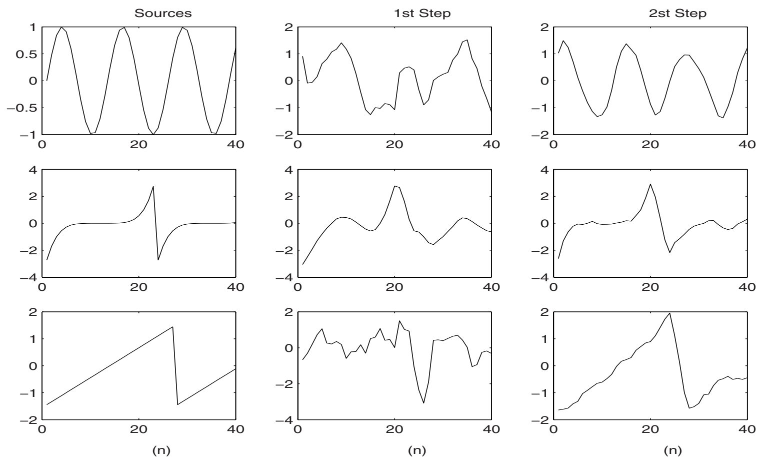

Fig. 3. Comparison of output signals resulting after the first step (second column) and the second step (last column) of dAMUSE.

5 Conclusions

In this work we propose dAMUSE, a modified version of the algorithm AMUSE. The new algorithm is based on a GEVD of a matrix pencil computed with time-delayed sensor signals. It was shown that the independent components estimated with dAMUSE represent filtered versions of the underlying source signals. The algorithm has a set of parameters whose proper choice must be further studied, particularly the number of time-delays M used to build the embedding feature space. Its choice naturally constrains the linear filtering operation that characterizes the method as it was shown in the simulations. The simulations also reveal that denoising cannot completely be achieved by PCA alone. Having at the output a high number of estimated independent components it is also important to find an automatic procedure to choose the relevant components related to the signal as some components might be related with noise only. Another aspect is to find out how the different signals can be joined together if a wide-band signal needs to be reconstructed. The choice of time-delayed matrices to compute the pencil was done according to the algorithm AMUSE as well as the values for time-delays (k_1 and k_2)[1]. Nevertheless other second-order statistics algorithms [9] [10] might be considered as well as they achieved better results when compared with AMUSE [9]. The algorithm was used to study Electrocardiograms and some of the independent signals could be related with the LF (low frequency) and HF (high frequency) fluctuations used to characterize HRV (heart rate variability) studies [11].

References

1. A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing.* Wiley, 2002.
2. R. Gharieb and A. Cichocki, "Second-order statistics based blind source separation using a bank of subband filters," *Digital Signal Processing*, vol. 13, pp. 252–274, 2003.
3. T. Schreiber and H. Kantz, "Nonlinear projective filtering II: Application to real time series," in *Proceedings of Nolta*, 1998.
4. R. Vetter, J. Vesin, P. Celka, P. Renevey, and J. Krauss, "Automatic nonlinear noise reduction using local principal component analysis and MDL parameter selection," in *IASTED International Conference on Signal Processing, Pattern Recognition&Applications*, (Crete, Greece), pp. 290–294, 2002.
5. P. Gruber, F. J. Theis, A. M. Tomé, and E. W. Lang, "Automatic denoising using local independent component analysis," in *Proceed. EIS'2004*, (Madeira, Portugal), 2004.
6. V. Moskvina and K. M. Schmidt, "Approximate projectors in singular spectrum analysis," *SIAM Journal Mat. Anal. Appl.*, vol. 24, no. 4, pp. 932–942, 2003.
7. M. Ghil, M. Allen, M. D. Dettinger, K. Ide, and e. al, "Advanced spectral methods for climatic time series," *Reviews of Geophysics*, vol. 40, no. 1, pp. 3.1–3.41, 2002.
8. A. M. Tomé and N. Ferreira, "On-line source separation of temporally correlated signals," in *European Signal Processing Conference, EUSIPCO2002*, (Toulouse, France), 2002.
9. A. M. Tomé, "Separation of a mixture of signals using linear filtering and second order statistics," in *10th European Symposium on Artificial Neural Networks*, (Brugges), pp. 307–312, 2002.
10. J. V. Stone, "Blind source separation using temporal predictibility," *Neural Computation*, vol. 13, no. 7, pp. 1557–1574, 2001.
11. A. R. Teixeira, A. P. Rocha, R. Almeida, and A. M. Tomé, "The analysis of heart rate variability using independent component signals," in *Second International Conference on Biomedical Engineering*, (Innsbruck, Austria), pp. 240–243, IASTED, 2004.