XVII IMEKO World Congress
Metrology in the 3rd Millennium
June 22–27, 2003, Dubrovnik, Croatia

COMPARISON OF BLIND SOURCE SEPARATION ALGORITHMS FOR CONVOLUTIVE MIXTURES

Samo Beguš, Dušan Fefer

Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Abstract – The goal of blind source separation is to recover independent sources without knowing the originals and the way they are mixed. Three algorithms for convolutive mixtures were tested and compared on anechoic room recordings.

All algorithms separate all possible combinations of signals, except when both signals are noises. The algorithm with natural gradient outperforms other algorithms. The quality of separation increases if an acoustic absorber is placed between the microphones. The mixture of noise and attenuated signal could be successfully separated if two different algorithms are applied one after another.

Keywords: blind source separation, convolutive mixtures, independent component analysis.

1. INTRODUCTION

A model of the blind source separation problem is presented in Fig. 1.

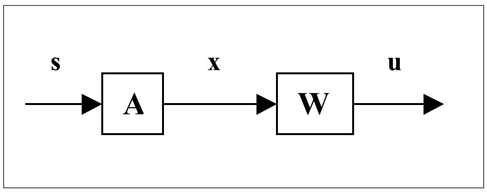

Fig. 1. The model of blind source separation problem, instantaneous mixtures

The goal of blind source separation is to find such unmixing matrix **W** that the output signals **u** are as independent as possible, **s** are the original signals, **A** is the mixing matrix and **x** are the mixtures:

$$\mathbf{u}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}\mathbf{s}(t). \tag{1}$$

Sources are separated completely when the unmixing matrix **W** is equal to the inverse of the mixing matrix **A**:

$$\mathbf{W} = \mathbf{P}\mathbf{A}^{-1}. \tag{2}$$

The unmixing matrix could be only estimated because the original signals and the mixing process are unknown. **P** is scaling and permutation matrix.

In the case of the "Cocktail party problem" where recordings are carried out in a reverberant room, the mixing process should be modelled with convolutive adding of sources:

$$x_i = \sum_{j=1}^{N}\sum_{k=1}^{M} a_{ijk} s_j(t-k), \tag{3}$$

where x_i are sensor signals, a_{ij} are elements of the mixing matrix of the length M and s_j are independent signals.

Reference [1] shows the results of convolutive mixtures separation quality with natural gradient algorithm. The algorithm performs well, but needs parameter adjustments, which affect the convergence speed and the separation quality.

2. ALGORITHMS

The first algorithm - ICA - is based on the natural gradient and minimizes the mutual information between output signals [2, 7]. The unmixing matrix is iteratively adapted by:

$$\Delta\mathbf{W} \propto [\mathbf{I} - f(\mathbf{u})\mathbf{u}^T]\mathbf{W}. \tag{4}$$

Because of the complexity of the separation of convolutive mixtures the Finite Impulse Response algebra is introduced and convolution is transformed into multiplication in frequency domain. The matrix of unmixing filters $\overrightarrow{\mathbf{W}}$ is adapted in frequency domain by:

$$\Delta\overrightarrow{\mathbf{W}} \propto [\vec{\mathbf{I}} - f(\vec{\mathbf{u}})\vec{\mathbf{u}}^T]\overrightarrow{\mathbf{W}}, \tag{5}$$

and by:

$$\overrightarrow{\mathbf{W}}(t+1) = \overrightarrow{\mathbf{W}}(t) + \alpha\Delta\overrightarrow{\mathbf{W}}(t) + \eta\Delta\overrightarrow{\mathbf{W}}(t-1), \tag{6}$$

where $\vec{\mathbf{u}}$ is the block of input data and $f(\cdot)$ is nonlinear sigmoid function. Signals are separated in time domain. The algorithm needs tuning of some parameters.

The second algorithm – JADE – Joint Approximate Diagonalization of Eigen-matrices [3, 9] does not need any parameter tuning and considers statistical properties of the signals, fourth order cumulants. The algorithm could be summarized in the following steps:

1.) decorrelation of input data
2.) calculation of the sample 4-th order cumulants
3.) joint diagonalization of the cumulants
4.) separation matrix calculation

The third algorithm Fast Fixed-point Algorithm for Independent Component Analysis – FastICA - [5, 10] is computationally very efficient and does not need any parameter tuning. Algorithm separates each independent component separately. The learning algorithm for one component is:

$$\begin{aligned}\mathbf{w}^{+} &= E\{\mathbf{x}\,conj(\mathbf{w}^{H}\mathbf{x})g(|\mathbf{w}^{H}\mathbf{x}|^{2})\} - \\ &E\{g(|\mathbf{w}^{H}\mathbf{x}|^{2}) + |\mathbf{w}^{H}\mathbf{x}|^{2} g'(|\mathbf{w}^{H}\mathbf{x}|^{2})\},\end{aligned} \tag{7}$$

where $\mathbf{w}^{+}, \mathbf{w}$ are new and old column vectors of unmixing matrix, $g(\cdot), g'(\cdot)$ are nonlinear function and its derivative respectively, $\mathbf{x}$ is the vector of input data.

For the two algorithms signals are separated in frequency domain. Signals are transformed into frequency domain with short time Fourier transformation, see (8):

$$\begin{aligned}&\mathbf{v}(\omega, t_s) = \sum_{t} e^{-j\omega t} x(t) w(t - t_s)\\ &\omega = 0, \frac{1}{N}2\pi, \ldots.., \frac{N-1}{N}2\pi\\ &t_s = 0, \Delta T, 2\Delta T, \ldots..\end{aligned} \tag{8}$$

where w is frequency, N is the number of elements of discrete Fourier transformation, t_s is window placement, $w(t)$ is the window function and ΔT is the window step.

Separated signals are reconstructed by means of separation of every frequency component and inverse Fourier transformation, but there remain the scaling and permutation problems of each separated signal frequency component. The solution is presented in [6]. The scaling problem is solved by filtering each separated signal with inverse separation matrix. The permutation problem is solved by comparison of signal envelopes in frequency domain.

Evaluation of separation quality is calculated as proposed in [7], except that the decimal logarithm is used.

$$Kj = 10\log_{10}\left[\frac{E\left\{\left(y_{j,sj}\right)^{2}\right\}}{E\left\{\left(\sum_{i\neq j} y_{j,si}\right)^{2}\right\}}\right] \tag{9}$$

Separation quality of the j-th signal Kj is evaluated according to (9), where $y_{j,si}$ is the j-th separated output when only the original signal s_i is activated.

3. SIGNAL MIXING

For the comparison of algorithms ten signals were prepared: seven speech signals, one music signal and white and pink noise, see Table 1. Note that signal pairs 1,2 and 7,8 are from the same speaker saying different phrases.

TABLE 1: Description and kurtosis of the signals used

Signal	Description	κ
1	female speech, slovenian	1,40
2	female speech, slovenian	2,13
3	male speech, slovenian	0,94
4	piano music	1,02
5	female speech, english	1,93
6	male speech, english	1,77
7	female speech, english	2,28
8	female speech, english	2,57
9	white noise	0,02
10	pink noise	0,03

They were reproduced through loudspeakers in anechoic chamber and recorded. The separation difficulty was changed by inserting a wooden plate with dimensions 0,8 m x 1,9 m x 0,02 m between the microphones and loudspeakers. Fig. 2 shows the anechoic chamber and the loudspeaker and microphone placement.

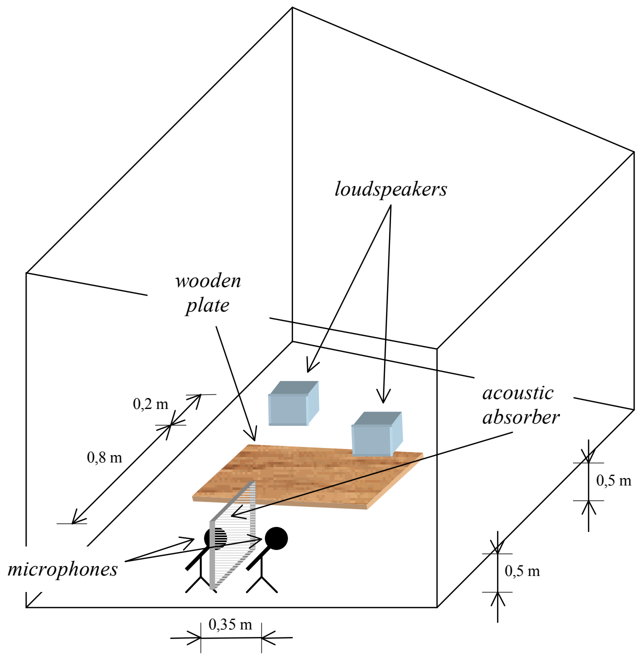

Fig. 2. The set-up for source mixing with wooden plate and acoustic absorber in the anechoic chamber

Recorded signals were sampled at 16 kHz, and filtered with a pass band filter with 200 Hz and 7500 Hz corner frequencies.

5. ACHIEVED RESULTS

Mixtures were recorded in four different set-ups. Loudspeakers and microphones were always at the same place, as shown in Fig. 2, but the wooden plate was placed on the floor between the microphones and the speakers and an acoustic absorber was placed between the microphones.

separation quality was evaluated according to (9). The separation quality in dependence of iteration number and parameter α is shown in Fig. 3. Fig. 4 shows the best separation quality in dependence of parameter α for three pairs of mixtures. Final settings for parameters were: α = 0,001, η = 0 and N_{ICA} = 25 iterations.

The separation quality for all possible pairs of signals and different set-ups for all three algorithms was evaluated. Fig. 5 shows the results. Axis labelling was omitted for reasons of clarity. Different set-ups are represented by numbers on horizontal axis, see Table 2, vertical axis

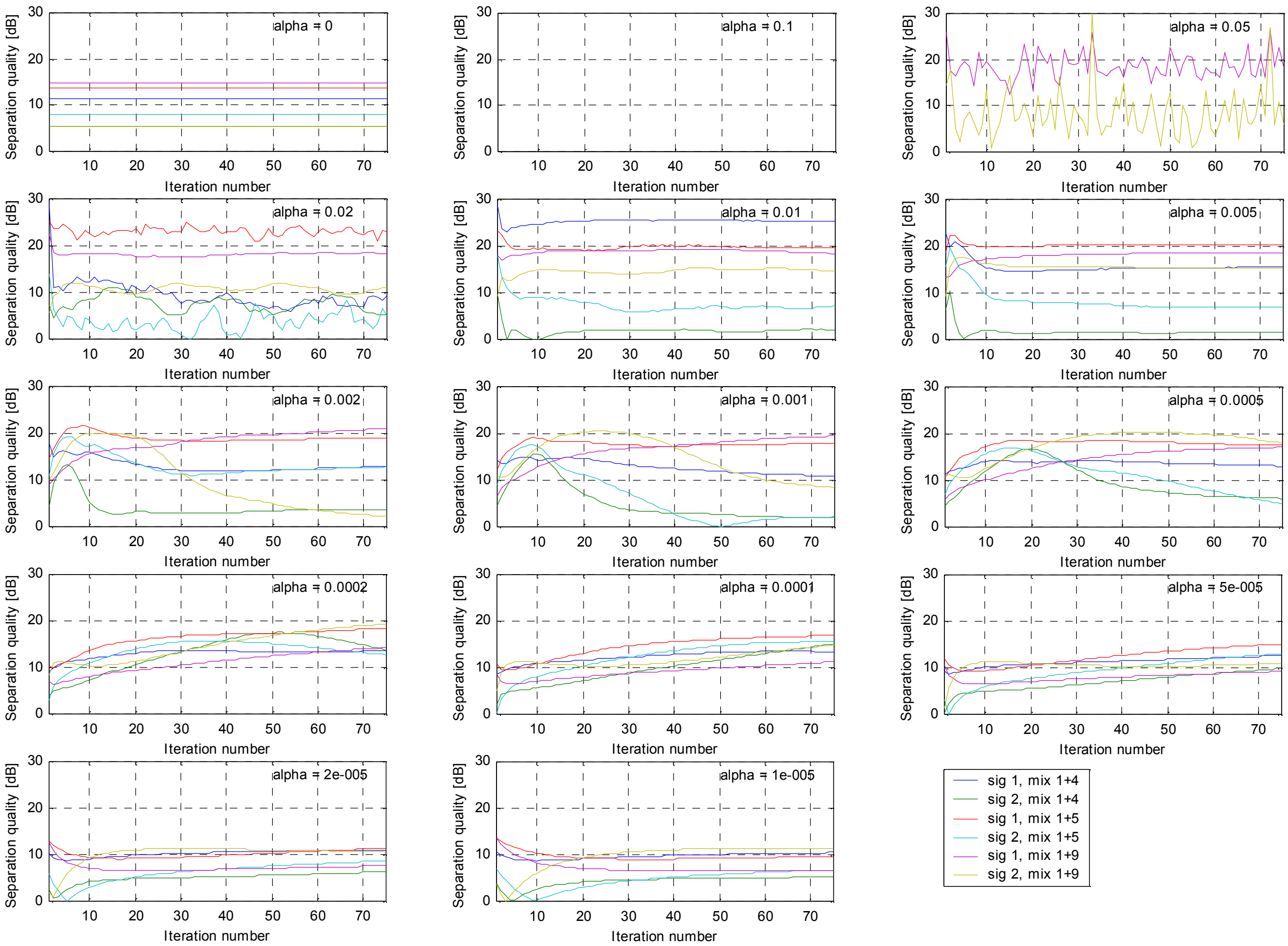

Fig. 3: Separation quality in dependence of iteration number and parameter α

Mixtures for all possible pairs of signals and different attenuations of one signal in the pair were recorded. The attenuation levels of one signal were: 0 dB, 10 dB, 15 dB, 20 dB and 25 dB for every set-up.

Kurtosis κ of original speech and music signals ranges from 0,95 to 2,57, see Table 1.

In case of the first and the second algorithm there are no parameters to be tuned. For the third algorithm (ICA) the parameters were chosen according to signal separation evaluations with different parameters.

There are three major parameters to be tuned for ICA algorithm: the learning rate α, the forgetting factor η and the number of iterations N_{ICA}. For the evaluation of these parameters they were changed and after each iteration the

represents signal separation quality in dB. The placing of graph in Fig. 5 determines which signal mixture is being separated. Signals have the same numbering as in Table 1. The legend is placed on the right bottom corner for reasons of clarity.

TABLE 2: Set-up number representation

Set-up number	Wooden plate	Acoustic absorber
1	✘	✓
2	✘	✘
3	✓	✓
4	✓	✘

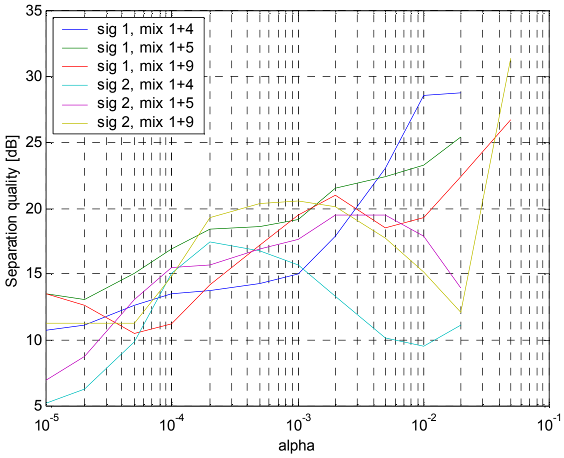

Fig. 4: The best separation quality in dependence of parameter α

Set-up numbers are chosen according to separation difficulties, so a decreasing tendency should be observed in graphs in Fig.5, but that is the case for ICA algorithm only. For JADE and FastICA algorithm the permutation problem is not always solved completely and in that case the separation quality worsens. The ICA algorithm has a high level of linear distortion: enhancement of mid- and high-audio frequencies, which on the other hand contributes to speech intelligibility.

Fig. 6 shows separated signals from the same speaker

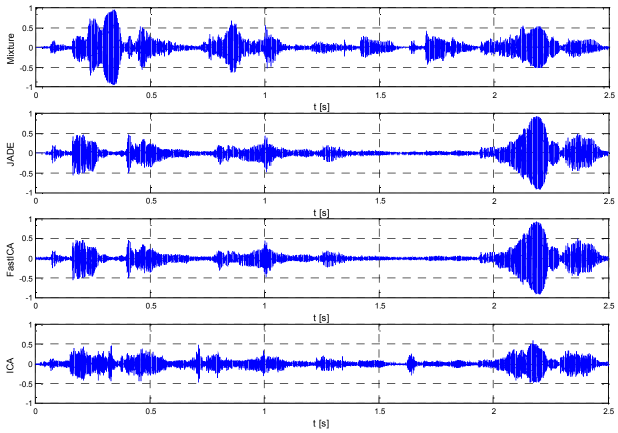

Fig. 6: Separated signals from mixtures 7 and 8

(signals 7 and 8). The upper graph represents the mixture of signals, the second one represents the separated signal with JADE algorithm, the next one shows the separated signal with FastICA algorithm and at the bottom the separated signal with ICA algorithm is shown. Separated signals with ICA algorithm are linearly distorted, therefore some equalization of output has been implemented.

Fig. 7 shows the separation results, with permutation problem not solved correctly (middle graphs).

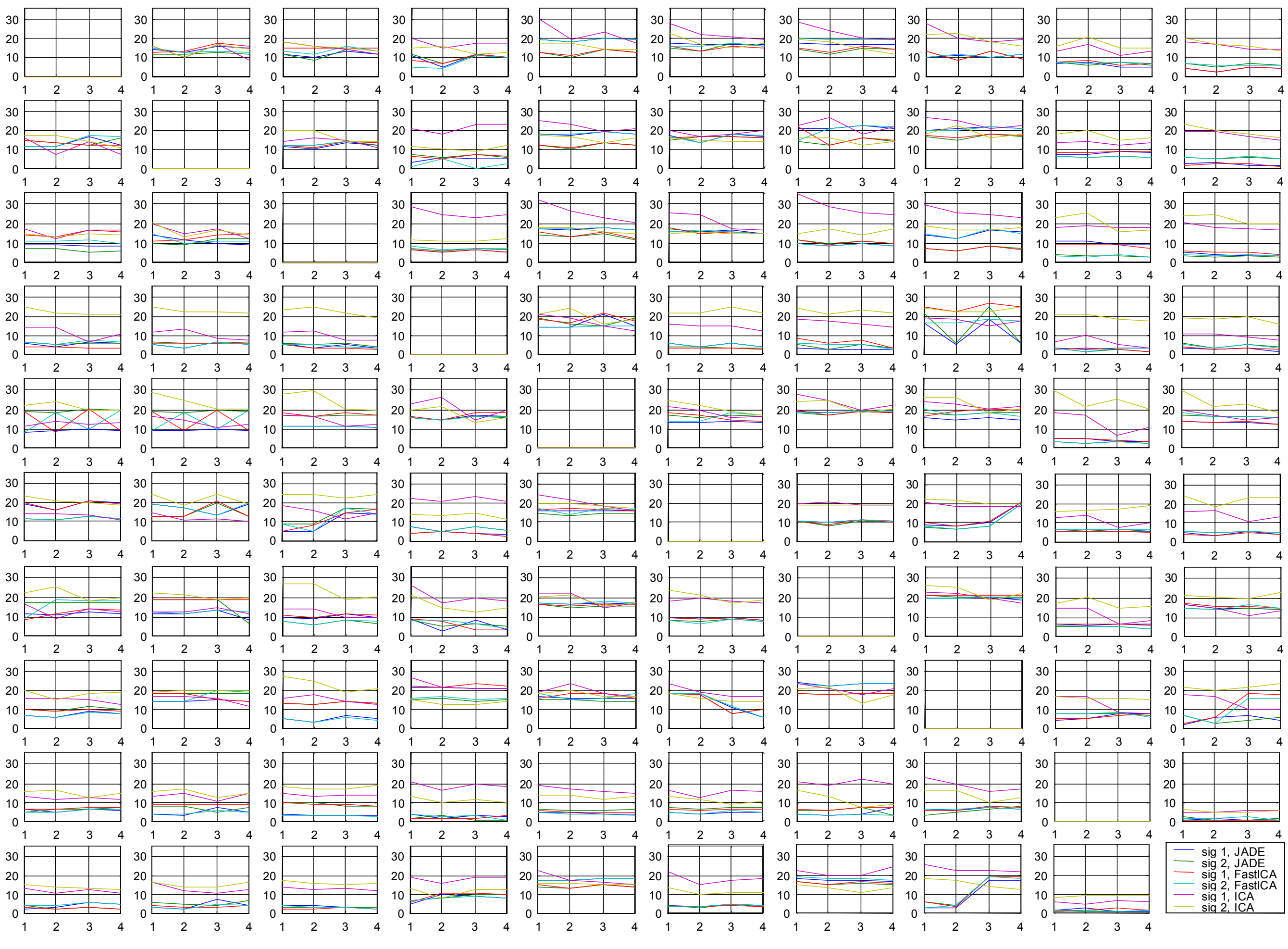

Fig. 5: Signal separation quality comparison

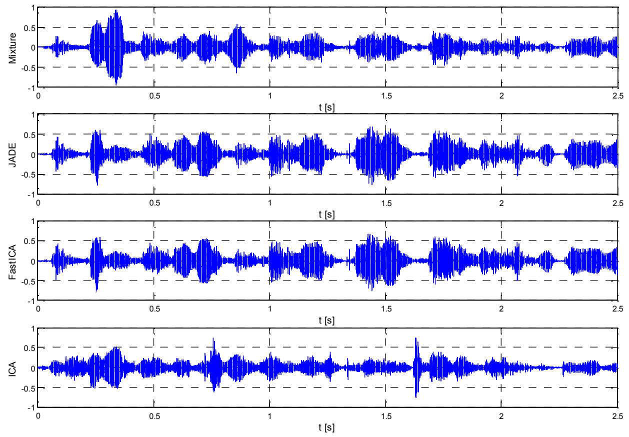

Fig. 7: Separated signals with permutation not solved correctly

The performance of algorithms could be increased if two algorithms are applied one after another. The mixture of noise and attenuated signal of 25 dB and 20 dB for white and pink noise respectively were chosen in such a way that a human being could not recognize the "hidden" signal in the noise. Signals were successfully separated so that the "hidden" signal could be recognized and understood. The algorithms applied were JADE and ICA one after another. The results are shown in Fig. 8. The top graph represents the mixture of speech and noise, the middle one represents the separated signal with JADE algorithm and at the bottom the separated signal with ICA algorithm is shown.

6. CONCLUSIONS

All three algorithms separate mixtures of signals produced by the same speaker. The quality of separation increases if an acoustic absorber is placed between the microphones. On the other hand the quality of separation decreases if a wooden plate is inserted.

The mixture of noise and attenuated signal could be successfully separated if two different algorithms are applied one after another.

The ICA algorithm outperforms other algorithms even if the parameters were the same for all mixtures. It could be used in measurements where several independent signals are present and signal scaling is permitted, like in frequency measurements.

REFERENCES

[1] D. Fefer, S. Beguš, "Blind acoustical source separation", Tenth Electrotechnical and Computer Science Conference ERK 2001, pp. 549-552, 2001.
D. Fefer, S. Beguš, "Slepo ločevanje akustičnih izvorov", *Zbornik desete Elektrotehniške in računalniške konference ERK 2001*, str. 549-552, 2001.

[2] T. Kosel, "Blind separation of independent acoustic emission sources", *doctoral dissertation*, Faculty of Mechanical Engineering, pp. 23-34, 2000.
T. Kosel, "Slepo ločevanje neodvisnih izvorov akustične emisije", *doktorsko delo*, Fakulteta za strojništvo, str. 23-34, 2000.

[3] J. F. Cardoso, A. Soloumiac, "Blind beamforming for non-gaussian signals", *IEE-Proceedings F*, vol. 140 no. 6}, pp. 362-370, 1993.

[4] R. H. Lambert, "Multichannel blind deconvolution: FIR matrix algebra and separation of multipath mixtures", *doctoral dissertation*, University of southern california, 1996.

[5] E. Bingham, A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals", Helsinki university of technology, 2000.

[6] N. Murata, S. Ikeda, A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals", Lab. for information synthesis RIKEN brain science, 1998.

[7] D. Schobben, K. Torkkola, P. Smaragdis. "Evaluation of Blind Signal Separation Methods", Proceedings int. workshop independent component analysis and blind signal separation, Aussois, France, January 11-15, 1999, pp. 261-266.

[8] T.W. Lee, "Independent component analysis theory and applications", Kluwer, 1998.

[9] http://sound.media.mit.edu/ica-bench

[10] http://www.tsi.enst.fr/icacentral/Algos/cardoso/

[11] http://www.cis.hut.fi/ella/publications/cfastica_public.m

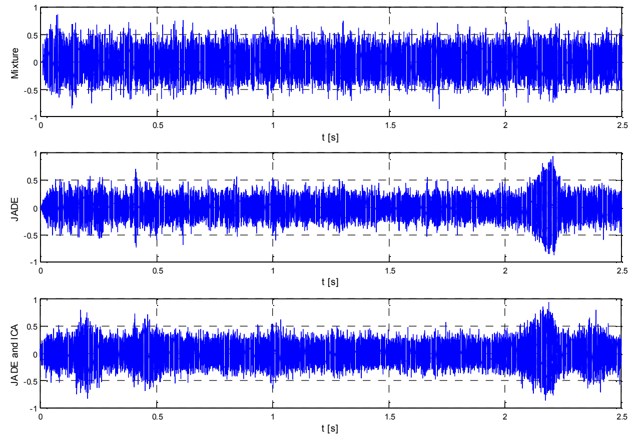

Fig. 8: Separated "hidden signal" in noise

Authors: Samo Beguš B.Sc., prof. Dušan Fefer Ph.D, Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia, tel. +386 14768421, fax. +386 14768214, samo.begus@fe.uni-lj.si, dusan.fefer@fe.uni-lj.si.