581

Measure of Quality of Source Separation for Sub- and Super-Gaussian Audio Mixtures

Ganesh R. NAIK

RMIT University
GPO BOX 2476V, Victoria-3001, Australia
e-mail: ganesh.naik@rmit.edu.au

Abstract. Conventional Blind Source Separation (BSS) algorithms separate the sources assuming the number of sources equals to that of observations. BSS algorithms have been developed based on an assumption that all sources have non-Gaussian distributions. Most of the instances, these algorithms separate speech signals with super-Gaussian distributions. However, in real world examples there exist speech signals which are sub-Gaussian. In this paper, a novel method is proposed to measure the separation qualities of both super-Gaussian and sub-Gaussian distributions. This study measures the impact of the Probability Distribution Function (PDF) of the signals on the outcomes of both sub and super-Gaussian distributions. This paper also reports the study of impact of mixing environment on the source separation. Simulation improves the results of the separated sources by 7 dB to 8 dB, and also confirms that the separated sources always have super-Gaussian characteristics irrespective of PDF of the signa ls or mixtures.

Keywords: blind source separation, probability distribution function, independent component analysis, kurtosis, signal to interference ratio, sub-Gaussian, super-Gaussian.

1. Introduction

Audio source separation is the problem of automated separation of audio sources present in a room, using a set of differently placed microphones, capturing the auditory scene (Benaroya *et al.*, 2006; Wilson, 2007; Dubnov *et al.*, 2006; Foote, 1999). The whole problem resembles the task a human can solve in a cocktail party situation, where using two sensors (ears), the brain can focus on a specific source of interest, suppressing all other sources present (cocktail party problem; Morita and Nanri, 2006; He *et al.*, 2006; Hyvarinen *et al.*, 2001).

Recently, Blind Source Separation (BSS) using Independent Component Analysis (ICA) has received a great deal of attention for its potential in acoustics, telecommunication, medical and image signal processing (He *et al.*, 2006; Stone, 2002; Cristescu *et al.*, 2000; De Martino *et al.*, 2007). BSS is an emerging technique, which enables the extraction of target speech from observed mixed speeches without the need for source positioning, spectral construction, or a mixing system. To achieve this, attention was focused on a method based on source separation technique. ICA is an unsupervised statistical method used for decomposing a complex mixture of signals into independent sources.

ICA's main focus is finding underlying components that are statistically independent from multivariate statistical data (Hyvarinen *et al.*, 2001; Bell and Sejnowski, 1997).

ICA is being used routinely to separate signals from different independent and nearly non-Gaussian sources. The applications of this include defence, surveillance, security, communication and entertainment. In the recent past, many researchers have studied the impact and quality of the sub- and super-Gaussian sources (Blaschke and Wiskott, 2003; Eriksson and Koivunen, 2004; Zarzoso *et al.*, 2006; Zarzoso and Nandi, 2002). While the assumption of independence is important for the success of ICA, the impact of Probability Distribution Function (PDF) of the sources has not yet been considered in detail. The other issue is the impact of the mixing environment on the quality of separation of the sources. In order to generate the hidden observations well, the probability density of the source has to be estimated with accuracy. Additionally, the density estimate of the source plays a significant role in the performance of the learning rule of the ICA algorithm. To make the source separation more effective for security and surveillance, there is need to determine the reliability of the use of ICA for obtaining the separated signals.

This research attempts to determine the answers for the following questions:

- What is the difference in the quality of separation for signals with different distribution functions?
- What is the impact of the mixing matrix on the separation of sources?
- Does the probability distribution of the signals remain the same after separation using ICA?
- Does the quality of the separated sources improves using ICA?

In this paper, a novel method for estimating the quality of audio mixtures with both super- and sub-Gaussian distributions is proposed. The proposed method also compares the source separation performance of the various BSS algorithms commonly used in literature.

2. Non-Gaussianity and Independence

There are several measures of non-Gaussianity that can be used. The classical one is Kurtosis value or fourth order cummulant. This value is zero, negative and positive for Gaussian, sub-Gaussian and super-Gaussian data respectively. The absolute value of Kurtosis is used frequently since it will be either zero or positive and will reach its maximum value when the signal is independent. Furthermore, Gaussianity also implies the degree of randomness of a signal and is related to information content of a signal. The less random signal (more structured signal) carries less information than the random one and vice versa. Gaussian signal is the most random signal among other signals and therefore it has the potential for the largest possible information content. According to central limit theorem the distribution of a sum of independent signals with arbitrary distributions tends toward a Gaussian distribution under certain conditions. The sum of two independent signals usually has a distribution that is closer to Gaussian than distribution of the two original signals. Thus a Gaussian signal can be considered as a linear combination

of many independent signals. This also explains that separation of independent signals from their mixtures can be achieved by finding a transformation that yields non-Gaussian distributions (Hyvärinen, 1999; Bell and Sejnowski, 1995; Cichocki and Amari, 2002).

Non-Gaussianity is an important and essential principle in ICA estimation. To use non-Gaussianity in ICA estimation, there needs to be quantitative measure of non-Gaussianity of a signal. Before using any measures of non-Gaussianity, the signals should be normalised (Cichocki and Amari, 2002; Lee, 1998). Some of the commonly used measures are kurtosis and entropy measures. Kurtosis is used as one of the measures in this paper, which is explained next.

2.1. *Kurtosis*

Kurtosis is the classical method of measuring non-Gaussianity. When data is preprocessed to have unit variance, kurtosis is equal to the fourth moment of the data. The kurtosis of signal (s), denoted by $kurt(s)$, is defined by

$$kurt(s) = E\{s^4\} - 3(E\{s^4\})^2. \tag{1}$$

This is a basic definition of kurtosis using higher order (fourth order) cumulant, this simplification is based on the assumption that the signal has zero mean. To simplify things, we can further assume that (s) has been normalised so that its variance is equal to one: $E\{s^2\} = 1$. Hence (1) can be further simplified to

$$kurt(s) = E\{s^4\} - 3. \tag{2}$$

(2) illustrates that kurtosis is a nomralised form of the fourth moment $E\{s^4\} = 1$. For Gaussian signal, $E\{s^4\} = 3(E\{s^4\})^2$ and hence its kurtosis is zero. For most non-Gaussian signals, the kurtosis is nonzero. Kurtosis can be both positive or negative. Random variables that have positive kurtosis are called as *super-Gaussian*, and those with negative kurtosis are called as *sub-Gaussian*. Non-Gaussianity is measured using the absolute value of kurtosis or the square of kurtosis.

Kurtosis has been widely used as a measure of non-Gaussianity in ICA and related fields because of its computational simplicity. Theoretically, it has a linearity property such that

$$kurt(s_1 \pm s_2) = kurt(s_1) \pm kurt(s_2), \tag{3}$$

and

$$kurt(\alpha s_1) = \alpha^4 kurt(s_1), \tag{4}$$

where α is a constant. Computationally kurtosis can be calculated using the fourth moment of the sample data, by keeping the variance of the signal constant (Lee *et al.*, 1999; Lipeika, 2010; Peric *et al.*, 2010). Kurtosis is extremely simple to calculate, however, it is very sensitive to outliers in the data set (Pham *et al.*, 1992; Svecko *et al.*, 2010).

2.2. *Sub- and Super-Gaussian Sources*

There are actually two types of non-Gaussian signals. The two non-Gaussian signals are known by various terms, such as super-Gaussian and sub-Gaussian or equivalently known as "platy kurtotic" and "lepto kurtotic" respectively. sub- and super-Gaussian density are represented using different forms. For a sub-Gaussian density, a symmetrical form in [4] is adopted as follows

$$p(u_i) = \frac{1}{2}\big(N(\mu, \sigma^2) + N(-\mu, \sigma^2)\big), \tag{5}$$

where $(N(\mu, \sigma^2)$ is the normal density with mean μ and variance σ^2. In this paper, a sub-Gaussian density with $\mu = 1$ and $\sigma^2 = 1$ is used.

- Super-Gaussian Sources

For the super-Gaussian density of speech signal, the Laplacian density can be represented as follows

$$p(u_i) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}|\mu_i|}{\sigma}}, \tag{6}$$

if the PDF of sub- and super-Gaussian functions are determined as of (5) and (6), respectively. The non-linear function $\varphi(\mu_i)$ can be represented as

$$\varphi(\mu_i; k_i) = \begin{cases} \sqrt{2}\mathrm{sign}\,(u_i) & \text{for } k_i = 1; \quad \text{super-Gaussian,} \\ \mu_i - \tanh\,(\mu_i) & \text{for } k_i = -1; \quad \text{sub-Gaussian,} \end{cases} \tag{7}$$

where $\mathrm{sign}\,(\mu_i)$ gives 1 when μ_i is positive and -1 when μ_i is negative. Hence, the nonlinearity function is represented as $\varphi(\mu_i; k_i)$ where k_i is 1 for super-Gaussian function and -1 for sub-Gaussian function. The switching condition for k_i between the sub- and super-Gaussian distributions is determined according to the sign of the kurtosis of estimated source $(\mu_i; k_i) = 1$ for positive kurtosis and $k_i = -1$ for negative kurtosis.

A signal with super-Gaussian probability distribution function (PDF) has most of its values clustered around zero. A speech signal is a typical example for a super-Gaussian source. Figure 1 shows a typical super-Gaussian source (speech signals). From the figure it is also evident that the super-Gaussian signals have PDFs that are more peaky than that of a Gaussian signals.

The signals with sub-Gaussian pdf has a wide distributed function, which is illustrated in Fig. 2. A saw-tooth signal, polyphonic music signal and white noise signals are typical sub-Gaussian sources (Cristescu *et al.*, 2000). The sub-Gaussian signals have PDFs that are less peaky than that of a Gaussian signals.
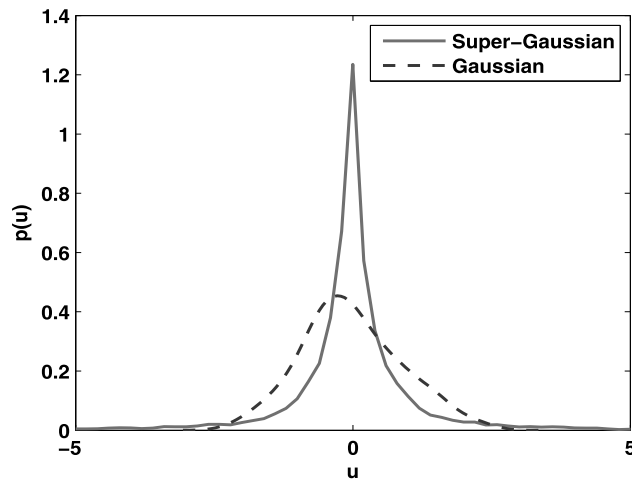
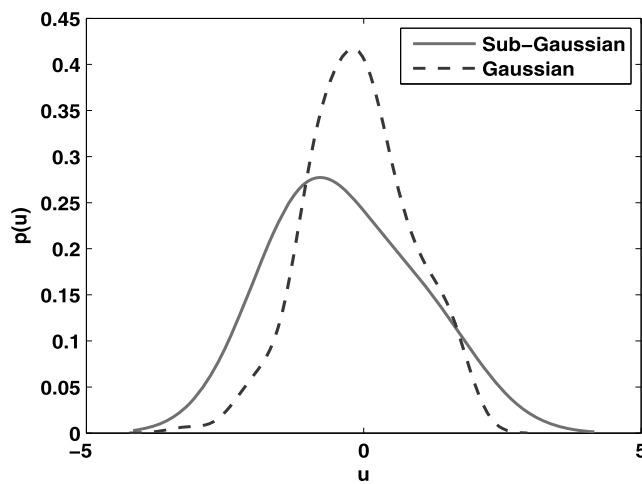Fig. 1. Pdf of super-Gaussian (speech) signal.

Fig. 2. Pdf of sub-Gaussian signal.

3. Independent Component Analysis

Independent Component Analysis is a new statistical technique that aims at transforming an input vector into a signal space in which the signals are statistically independent (Hyvarinen *et al.*, 2001; Jang and Lee, 2003; Lee *et al.*, 2000).

ICA assumes the mixing process as linear, so it can be expressed as:

$$x(t) = As(t), \tag{8}$$

where $x = [x_1(t), x_2(t), \ldots, x_n(t)]$ are the recordings, $s = [s_1(t), s_2(t), \ldots, s_n(t)]^T$
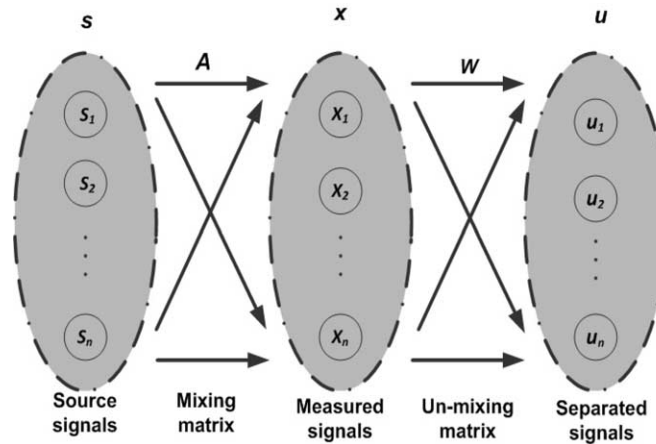
Fig. 3. Independent component analysis (ICA) block diagram. $s(t)$ are the sources. $x(t)$ are the recordings, $\hat{s}(t)$ are the estimated sources A is mixing matrix and W is un-mixing matrix.

are the original signals, and A is the $n \times n$ mixing matrix. This mixing matrix and each of the original signals are unknown. To separate the recordings to the original signals, an ICA algorithm performs a search of the un-mixing matrix W by which observations can be linearly translated to form Independent output components so that:

$$s(t) = Wx(t) = WAs(t), \tag{9}$$

for this purpose, ICA relies strongly on the statistical independence of the sources s. The block diagram approach of ICA for source separation is shown in Fig. 3. The ICA technique iteratively estimates the un-mixing matrix using the maximisation of independence of the un-mixed signals as the cost function. Signals are statistically independent if the joint probability density of those components can be expressed as a multiplication of their marginal probability density. It is important to observe the distinction between independence and uncorrelatedness, since decorrelation can always be performed by transforming the signals with a whitening matrix to get the identity covariance matrix. Independent signals are always uncorrelated but uncorrelated signals are not always independent. But in case of Gaussian signals, uncorrelatedness implies independence. Transforming of a Gaussian signal with any orthogonal un-mixing matrix or transform results in another Gaussian signal, and thus the original signals cannot be separated. Hence Gaussian signals are forbidden for ICA. Thus the key of independent component estimation is measuring the non-Gaussianity of the signals (Cardoso, 1998).

There exist many ICA methods which uses Higher order and Second order statistics for separating the sources. In this paper, the detailed ICA algorithms employed are FastICA, Joint Approximate Diagonalization of Eigen-matrices (JADE), SOBI (Second Order Blind Identification) and Second Order Non-stationary source Separation (SONS). One of the reasons for this choice is the large number of successful applications in various fields of data mining, particularly audio signal processing (Benaroya *et al.*, 2006; Wilson,

2007; Morita and Nanri, 2006). The other factors are the broad availability of the algorithms and the diversity in the way the source are estimated. The brief explanation of these methods are explained next.

- FastICA

FastICA is a fixed point ICA algorithm that employs higher order statistics for the recovery of independent sources. FastICA has two algorithmic methods available for the estimation of the unmixing matrix: one is the symmetric approach and the other is the deflationary approach where independent components are estimated one by one like in projection pursuit. The FastICA algorithm extracts independent components by separately maximizing the Negentropy $J(y)$ of each mixture FastICA uses simple estimates of Negentropy based on the maximum entropy principle, which requires the use of appropriate nonlinearities for the learning rule of the neural network (Hyvärinen and Oja, 1997; Hyvarinen *et al.*, 2001; Hyvärinen, 1999).

- JADE

The JADE is an algorithm based on the joint diagonalization of cumulant matrices under the assumption that the sources have non-Gaussian distributions. After whitening and possible dimension reduction, a set of matrices obtained from Eigen matrices of the fourth-order cumulant tensor is approximately diagonalized with a single orthogonal transformation (Cardoso and A. Souloumiac, 1996; Joho and Rahbar, 2003; Moreau, 2002; Ziehe *et al.*, 2004).

- SOBI

The SOBI (Belouchrani *et al.*, 2002, 1993; Molgedey and Schuster, 1994) takes advantage of the temporal structure in the observed data. SOBI exploits the time structure of the data for source extraction. Moreover, SOBI relies only on stationary second order statistics that are based on a simultaneous diagonalization of a set of covariance matrices for the computation of an estimate of the unmixing matrix. The basis of the SOBI algorithm is a set of time-lagged covariance matrices. To estimate the sources a joint diagonalisation of the time-lagged covariance matrices is performed similar to the JADE algorithm.

- SONS

The SONS algorithm is used to perform both ICA (for arbitrary distributed non-stationary sources with temporal structures) and BSS (for colored sources with different spectra), depending on the time delays, the number of time windows, and the size of each window (Choi, 2002; Saylani *et al.*, 2009).

The successful separation of the original signals is dependent on the fulfilment of these conditions (Hyvarinen *et al.*, 2001).

- The sources must be statistically independent.
- The sources must have non Gaussian distributions.
- The number of available recordings must be at least the same as the number of the independent sources.

- The recorded signals must be (approximately) linear combination of the independent sources and
- There should be no (little) noise common to the sources and there should be no (minimal) delay between the signals of the different sources in the recordings.

While ICA has demonstrated success in the ability to separate signals, the output of ICA suffers from the following ambiguities:

- The order of the independent components cannot be fixed and this may change for each estimate.
- The amplitude and sign of the independent components cannot be determined. While the relative amplitude within each signal is correctly estimated, relative amplitude between different signals cannot be estimated using ICA.

In most application such as the cocktail party problem, these are not serious problems. The supervisor is able to identify the different sources and determine the quality of the separation by listening to the sounds. To summarise from the above, the signals that can be separated need to be non-Gaussian and independent. For the purpose of applying ICA to audio recordings, there is a need to determine the conditions under which these signals can be considered as independent and non-Gaussian.

4. Experiment

The experiments are conducted to evaluate the performance of the sub and super-Gaussian audio mixtures. For this purpose, four recordings of speech, s_1, s_2, s_3 and s_4 are recorded. The strength of the recording corresponded to 42 dB. The recordings are made in an office environment using lapel microphones. Each recording is of average length 10 to 15 seconds, and sampled at 44,100 samples/second. The distance between the speaker and the microphone is kept very close to circumvent multi paths for each emitted source signal. The speaker read a short passage from a book. The recordings are played back to colleagues and it is found that each of the observers could easily understand what is being said by the speaker.

The experiments are evaluated in 3 steps.

- The first step is mixing the various audio sources with different mixing matrices and to separate them using both second order and higher order ICA methods.
- The second step is using these separated signals to measure quality of separation using SIR and
- The last step is to estimate the quality of the sources using the Kurtosis values and also measure the PDF of the sources.

As a first step, computer simulations are conducted to perform the source separation in sub-Gaussian and super-Gaussian mixing conditions. For this four different sub-Gaussian signals and super-Gaussian signals are considered and are shown in Fig. 4. Four different sub-Gaussian (Polyphonic music) and super-Gaussian signals (Speech signals) are mixed using randomly generated sub-Gaussian and super-Gaussian mixing matrices. One of the examples of the mixing matrices used in this research is shown in Table 1. The PDF of
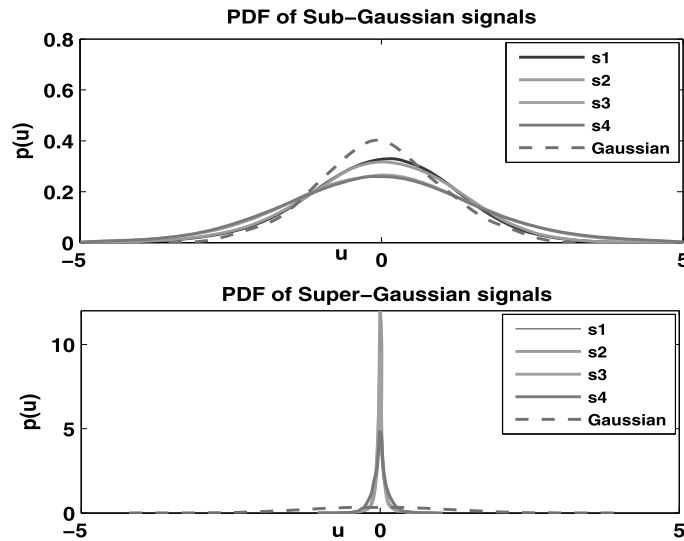
PDF of Sub–Gaussian signals

PDF of Super–Gaussian signals

Fig. 4. PDF of sub- and super-Gaussian sources.

Table 1

Sub-Gaussian and super-Gaussian mixing matrices used for the experiment

Sub-Gaussian mixing matrix				Super-Gaussian mixing matrix			
-1.0000	-1.0000	1.0000	-1.0000	0.0790	0.0309	0.0049	0.0012
-0.5250	-3.4057	0.1582	-2.6094	0.0111	0.0444	0.1000	0.1778
-1.0000	-1.0000	1.0000	1.0000	0.0309	0.2086	0.0000	0.0790
2.2949	-1.1210	0.8096	-3.9043	0.3568	0.2420	0.1494	0.0790

the signals and mixing matrices are plotted every time to confirm the validity of the sub- and super-Gaussian sources. The PDF distribution of the sub- and super-Gaussian mixing matrices are shown in Fig. 5.

Four independent audio recordings s_1, s_2, s_3 and s_4 are mixed in the following way:

- sub-Gaussian source + sub-Gaussian mixing matrix,
- sub-Gaussian source + super-Gaussian mixing matrix,
- super-Gaussian source + sub-Gaussian mixing matrix,
- super-Gaussian source + super-Gaussian mixing matrix.

The entire mixing process can be expressed in the vectors and matrix form as:

$$x_1 = a_{11}s_1 + a_{12}s_2 + a_{13}s_3 + a_{14}s_4,$$
$$x_2 = a_{21}s_1 + a_{22}s_2 + a_{23}s_3 + a_{24}s_4,$$
$$x_3 = a_{31}s_1 + a_{32}s_2 + a_{33}s_3 + a_{34}s_4,$$
$$x_4 = a_{41}s_1 + a_{42}s_2 + a_{43}s_3 + a_{44}s_4. \tag{10}$$

PDF of sub–Gaussian mixing matrix

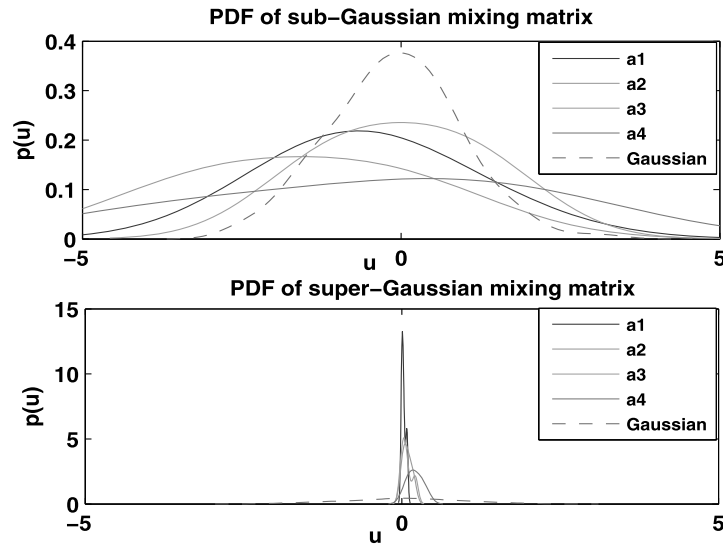**PDF of super–Gaussian mixing matrix**

Fig. 5. PDF of sub- and super-Gaussian mixing matrices.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix}.$$

An example of the mixed signal in sub- and super-Gaussian environment are shown in Fig. 6. The similar experiments are repeated for all the four above mentioned mixing process.

4.1. *Data Analysis*

Quality of the source separation can be measured using different performance indices such as:

- signal to Noise Ratio (SNR);
- signal to Distortion Ratio (SDR);
- signal to Artefacts Ratio (SAR) and
- signal to Interference Ratio (SIR).

In bio signal and audio applications, SIR is found to be a popular tool to measure the quality separation (Cichocki and Amari, 2002).

Signal to Interference Ratio is the ratio of the power of the wanted signal to the total residue power of the unwanted signals. This performance index could be used for full-rank or non-full rank analysis (Cichocki and Amari, 2002). To determine the efficacy of using SIR for an application, the authors have studied the quality of separation on
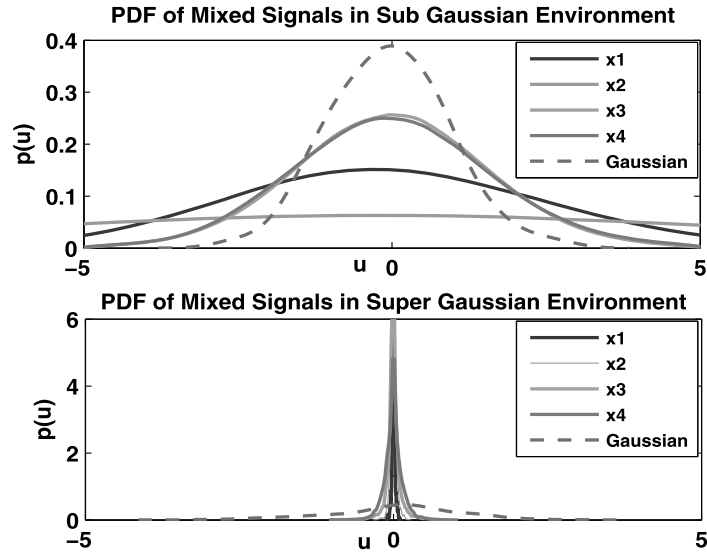
PDF of Mixed Signals in Sub Gaussian Environment

PDF of Mixed Signals in Super Gaussian Environment

Fig. 6. Examples of PDF of sub- and super-Gaussian mixed signals.

synthetically mixed data where the original source signals s_j $(j = 1, \ldots, N)$, the mixing matrix A, the observations x_i $(i = 1, \ldots, M)$, the estimated separation matrix W, and the estimated source signals y_j $(j = 1, \ldots, N)$ are all available. For real world problems, we have no access to the original signals s_j and the mixing matrix A. In this scenario the computation of the SIR_{est} would be the SIR for real world data. Taking into account that the global transfer function of the mixing-separating system can be defined as $G = W * A$, one can formulate the SIR of the estimated signals as

$$\text{SIR} = \text{SIR}_{est} = 10 \log \frac{E\{(g_{jj} * s_j)^2\}}{E\{(\sum_{k=1, k \neq j}^{N} (g_{jk} * s_k)^2\}}. \tag{11}$$

The mixed signals in sub- and super-Gaussian conditions are separated using four different ICA methods. For higher ICA methods FastICA and JADE are used, and for the Second order ICA methods SOBI and SONS ICA algorithms are used. There are a number of parameters available that can be tuned in order to optimize the performances of both higher and second order ICA algorithms. For this study, FastICA is using a tangent hyperbolic as its non-linearity instead of the default third order polynomial, which is not found to be a robust parameter, because it is only recommended when there are no outliers. In the case of JADE, all default parameters are used. On the other hand, SOBI and SONS have only the number of temporal delays as adjustable parameters (Tang, 2005). Hence for SONS and SOBI source separation, the default value of 50 delays ranging from 4 ms to 200 ms variations of the parameter between 2 and 250 delays (Naeem, 2009).

The kurtosis and SIR values are computed for each separated sources using the equation (2) and (11) respectively.

Table 2

Overall ICA separated results for sub-Gaussian and super-Gaussian sources using SIR

Sub-Gaussian sources	Super-Gaussian sources	Sub-Gaussian mixing matrix	Super-Gaussian mixing matrix	SIR (dB)	Kurtosis
X		**X**		13	2.0067
X			**X**	15	5.0845
	X	**X**		20	8.5845
	X		**X**	26	12.4867

Table 3

Overall Kurtosis results for sub-Gaussian and super-Gaussian sources, mixing matrices and mixtures

Kurtosis	Sub-Gaussian				Super-Gaussian			
	FastICA	JADE	SOBI	SONS	FastICA	JADE	SOBI	SONS
Mixed matrices	-1.6667	-1.7583	-2.0000	-1.2719	1.9333	1.7211	2.0077	1.6705
Source signals	-8.3952	-10.4959	-11.5782	-6.0666	12.3952	10.4959	10.5782	4.0666
Separated signals (s_1, s_2, s_3, s_4)	4.6438	5.6044	3.5063	3.492	11.7766	13.6044	11.8955	13.516
	5.6879	4.6475	3.9081	3.467	12.068	5.6475	4.9431	5.502
	4.6271	4.4692	3.0739	3.5573	13.5031	11.4692	11.135	11.5573
	6.658	3.9323	3.5224	4.2988	5.6049	12.9323	11.8952	11.4029
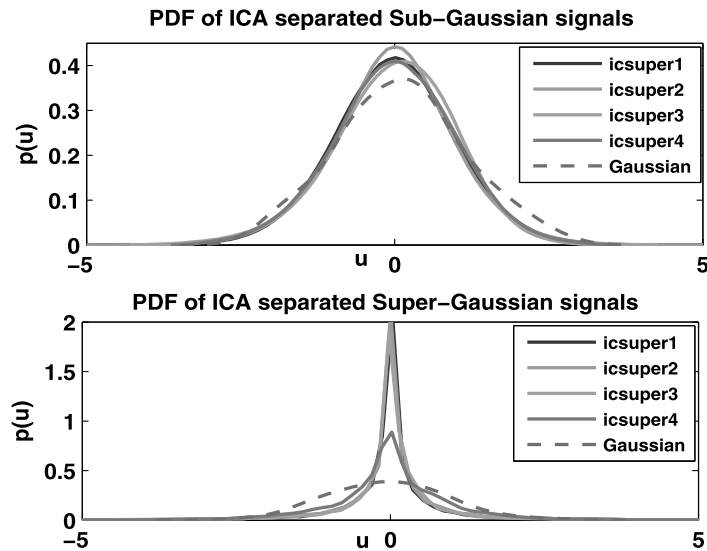
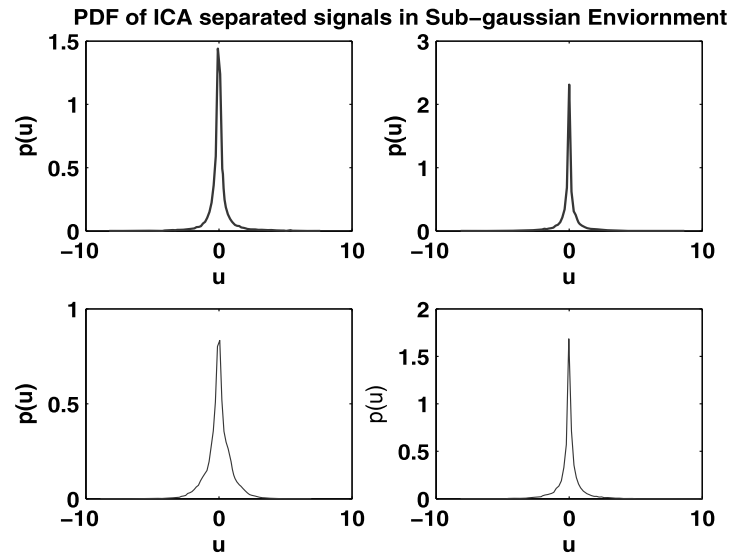Fig. 7. Examples of PDF of ICA separated sub- and super-Gaussian signals.

PDF of ICA separated signals in Sub−gaussian Enviornment

Fig. 8. PDF of ICA separated sub-Gaussian signals.

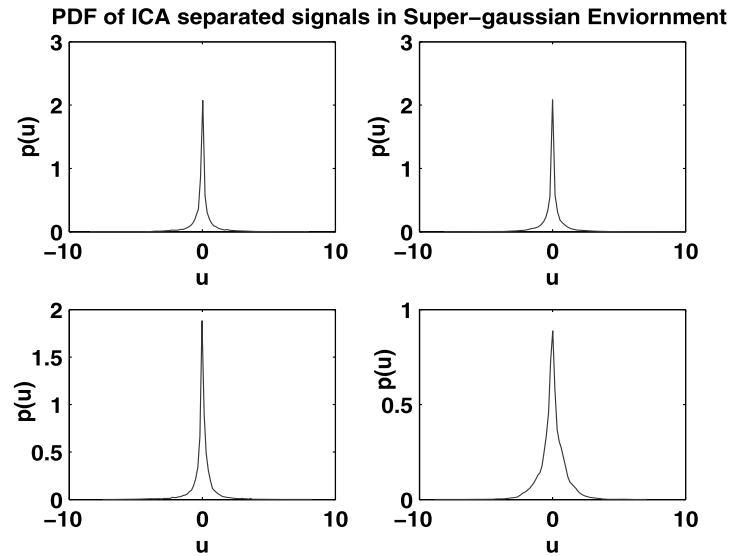**PDF of ICA separated signals in Super−gaussian Enviornment**

Fig. 9. PDF of ICA separated super-Gaussian signals.

5. Results and Observations

The overall SIR results for four different categories have been tabulated in Table 2. Figs. 7, 8, 9 and 10 represents the PDF of the separated source in sub- and super-Gaussian mixing environment. From the results, it is observed that

G.R. Naik

PDF of Signals in Sub−Gaussian and Super−Gaussian Enviornment
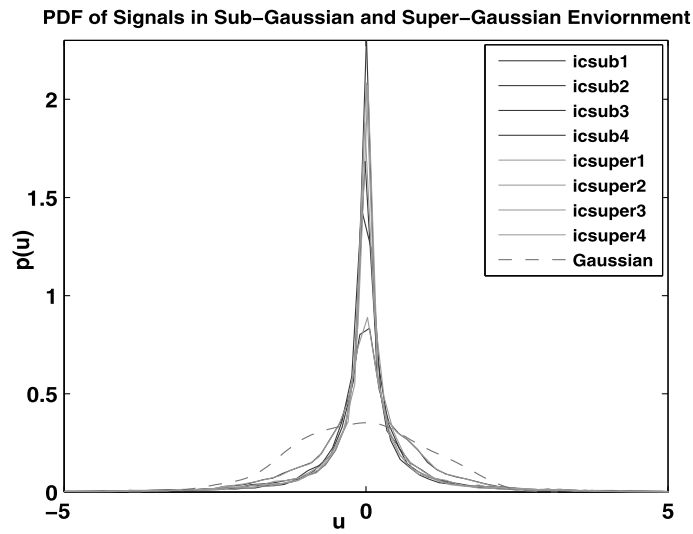

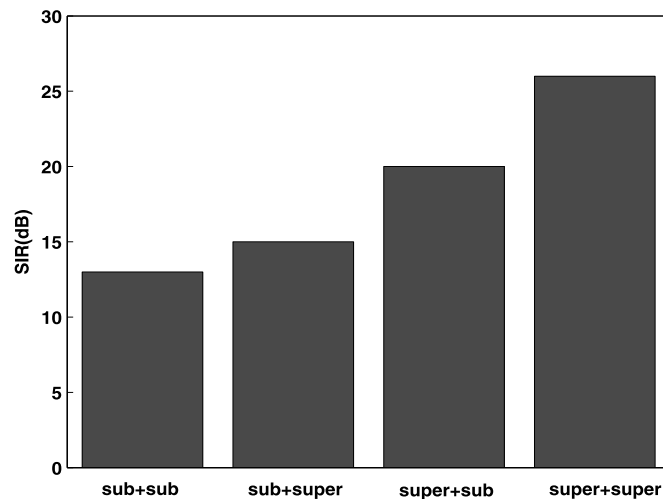Fig. 10. PDF of ICA separated sub- and super-Gaussian signals together.

Fig. 11. Overall SIR results in dB. The four categories are mentioned first one as source and the second one as mixing matrix Eg: sub + super.

- Kurtosis value is always positive, which indicates that after the source separation using ICA algorithms, all separated sources remains as super-Gaussian (refer Tables 2, 3, Figs. 7, 8, 9 and 10).
- The SIR value increases from sub-Gaussian mixtures (sub-source+sub-matrix and sub-source + super-matrix), to super-Gaussian mixtures (super-source+sub-matrix and super-source + super-matrix) by 6 dB to 7 dB (average) (refer Table 1 and Fig. 11).

Table 4

Overall SIR results for sub-Gaussian and super-Gaussian sources

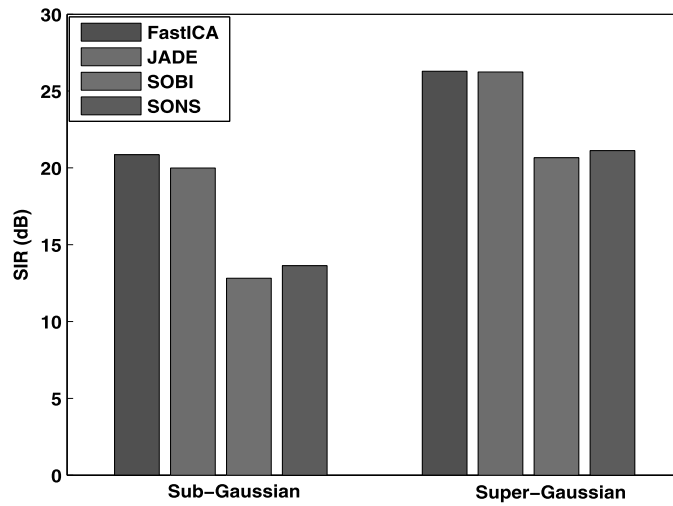| SIR	Sub-Gaussian				Super-Gaussian			
	Higher order ICA		Second order ICA		Higher order ICA		Second order ICA	
	FastICA	FJADE	SOBI	SONS	FastICA	FJADE	SOBI	SONS
SIR1	21.5603	19.2111	12.4887	13.4902	26.2026	26.2111	20.5035	20.9429
SIR2	20.5489	20.4629	12.2583	13.0029	25.7222	26.4629	21.1294	21.0907
SIR3	20.1011	19.594	13.2532	13.8534	26.5337	25.594	20.212	20.9274
SIR4	21.2085	20.6837	13.259	14.1995	26.6956	26.6837	20.8032	21.5251
SIR	**20.8547**	**19.9879**	**12.8148**	**13.6365**	**26.2885**	**26.2379**	**20.6621**	**21.1215**

Fig. 12. Overall SIR results in dB. The first class of results showing the sub-Gaussian separated sources and the second class of figure showing the super-Gaussian sources.

- The SIR value increases from 16.83 dB (average) to 23.58 dB (average) from sub-Gaussian to super-Gaussian mixtures (refer Table 4 and Fig. 12).
- The Higher order ICA methods showed significant improvement in the source separation (13.32 dB to 20.42 dB for sub-Gaussian and 20.89 dB to 26.26 dB for super-Gaussian), as compared to second order ICA methods (refer Table 4, Fig. 12).

Theoretically, higher order ICA should be useful in quite all applications as standard PCA, second order ICA and even beyond. The results of second order ICA are not impressive. This is due to the fact that in second order ICA there are no assumptions w.r.t. the non-Gaussianity of the sources. The method uses the fact that a signal containing harmonic series plus noise lies in a linear (signal) subspace, spanned by the dominant Eigen values of its correlation matrix.

6. Discussions and Conclusion

In this paper, we study the performance of four major classes of algorithms for spatial ICA, namely, maximization of non-Gaussianity, joint diagonalization of cross- cumulant matrices and second-order correlation-based methods. We compare how different ICA algorithms estimate audio mixtures in sub-Gaussian and super-Gaussian conditions. The results demonstrate that the ICA algorithms using higher-order statistical information prove to be quite consistent for audio data analysis. FastICA and JADE both yield reliable results, with each having its strengths in specific areas. On the other hand, SOBI and SONS, algorithms using second-order statistics, does not perform reliably for sub- and super-Gaussian audio mixtures.

FastICA and JADE, which all implicitly or explicitly use higher-order statistics, yield results that are consistent with each other. Among the ICA algorithms that are tested, FastICA yielded results with the highest SIR values, and SOBI yielded the lowest. FastICA maximizes the non-Gaussianity of estimated components and, thus, provides a global measure of separation, whereas SOBI uses the second-order correlation-based methods to estimate components. Using FastICA as a reference with which to compare the other algorithms proved to be reasonable since, visually, it is seen that FastICA yields results that appear to be the best. We found that FastICA performed slightly better than JADE due to higher contrast-to-noise ratio. JADE, on average, estimates smaller areas of activation in comparison to FastICA. Lowering the threshold for JADE increases the size of activation areas; however, this also causes the inclusion of a large amount of irrelevant areas (likely false positives), thus maki ng it evident that JADE yields estimates with a lower contrast to-noise ratio. JADE uses fourth-order statistics and minimizes cross-cumulants to achieve independence among estimated components in contrast to approaches such as FastICA that implicitly incorporate all higher order statistics through a nonlinear function.

SOBI and SONS are based on second-order statistics, and results suggest that it shows poor selectivity for audio analysis. It cannot differentiate between sources with similar spectra, as in the case of sub- and super-Gaussian sources in the study. It places both of these sources in the same component and also poorly estimates the time courses of these components. The iterative algorithms FastICA and JADE yield consistent results over runs, and the initialization of parameters does not cause large changes in results. For FastICA, nonlinearities tanh and pow3 yield more consistent results than the gauss nonlinearity.

The experiments shows that the most commonly used ICA approaches yield consistent results for audio data analysis. FastICA and JADE yield reliable results for sub- and super-Gaussian data and, thus, are useful algorithms for studying components on which we have no prior information. In addition, iterative ICA algorithms yield results that are quite consistent over runs for the components of interest. These results greatly increase the confidence in the use of higher order ICA for audio data analysis.

It is concluded that the signals after separation using ICA are always super-Gaussian in nature. It can be easily argued that higher order ICA methods which generally uses

Skewness and Kurtosis for source separation are responsible for this. Usage of both Higher and second order ICA methods justifies that by using any ICA methods, the separated sources always remains super-Gaussian.

While sub-Gaussian sources can also be separated, the quality of separation is poor, and the outcome is always a super-Gaussian approximation of the original signal. Hence, a signal such as polyphonic music would be separated into non-polyphonic music and there would be greater emphasis on aspects of the audio such as voice or other super-Gaussian sources. There is a strong impact of the mixing matrix distribution on the quality of source separation. If the mixing matrix is super-Gaussian in nature, the quality of separation is better than when the mixing matrix is sub-Gaussian. The experimental results also showed the significant improvement of the separated sources using Higher order ICA methods as compared to that of Second order ICA methods. In near future, it would be interesting to test the outcome of the research into noisy signals.

References

Bell, A.J., Sejnowski, T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computing*, 7(6), 1129–1159.

Bell, A.J., Sejnowski, T.J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, 37(23), 3327–3338.

Belouchrani, A., Abed-meraim, K., Cardoso, J.F., Moulines, E. (1993). Second order blind separation of temporally correlated sources.

Belouchrani, A., Abed-Meraim, K., Cardoso, J.F., Moulines, E. (2002). A blind source separation technique using second-order statistics. In: *IEEE Transactions on Signal Processing*, 45(2), pp. 434–444.

Benaroya, L., Bimbot, F.,Gribonval, R. (2006). Audio source separation with a single sensor. In: *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 14(1), pp. 191–199.

Blaschke, T., Wiskott, L. (2003). Cubica: independent component analysis by simultaneous third- and fourth-order cumulant diagonalization. *IEEE Transactions on Signal Processing*, 52.

Cardoso, J. (1998). Blind signal separation: statistical principles. *Proceedings of IEEE*, 9(10), 2009–2025.

Cardoso, J.F., Souloumiac, A. (1996). Jacobi angles for simultaneous diagonalization. *SIAM Journal, Matrix Analysis and Application*, 17(1), 161–164.

Choi, S., Cichocki, A., Beloucharni, A. (2002). Second order nonstationary source separation. *Journal of VLSI Signal Processing System*, 32(1/2), 93–104.

Cichocki, A., Amari, S.-I. (2002). *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. Wiley, New York.

Cristescu, R., Ristaniemi, T., Joutsensalo, J., Karhunen, J. (2000). Cdma delay estimation using fast ica algorithm. In: *Proceedings of IEEE Internetional Symposium on Personal, Indoor, and Mobile Communications*, Vol. 2, pp. 1117–1120.

De Martino, F., Gentile, F., Esposito, F., Balsi, M., Di Salle, F., Goebel, R., Formisano, E. (2007). Classification of fmri independent components using ic-fingerprints and support vector machine classifiers. *NeuroImage*, 34, 177–194.

Dubnov, S., Tabrikian, J., Targan, M.A. (2006). Speech source separation in convolutive environments using space-time-frequency analysis. *EURASIP Journal of Advances in Signal Processing*, 2006(1).

Eriksson, J., Koivunen, V. (2004). Identifiability, separability, and uniqueness of linear ica models. *Signal Processing Letters*, 11(7), 601–604.

Foote, J. (199). An overview of audio information retrieval. *Multimedia Systems*, 7(1), 2–10.

He, T., Clifford, G., Tarassenko, L. (2006). Application of independent component analysis in removing artefacts from the electrocardiogram. *Neural Computing & Applications*, 15(2), 105–116.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3), 626–634.

Hyvärinen, A., Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computing*, 9(7), 1483–1492.

Hyvarinen, A., Karhunen, J., Oja, E. (2001). *Independent Component Analysis*. Wiley-Interscience, New York.

Jang, G.-J., Lee, T.-W. (2003). A maximum likelihood approach to single-channel source separation. *Journal of Machine Learning Research*, 4, 1365–1392.

Joho, M., Rahbar, K. (2003). Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation. In: *Sensor Array and Multichannel Signal Processing Workshop Proceedings*, pp. 403–407.

Lee, T.W. (1998). *Independent Component Analysis: Theory and Applications*. Kluwer Academic Dordrecht.

Lee, T.W., Lewicki, M.S., Sejnowski, T.J. (1999). Unsupervised classification with non-Gaussian mixture models using ica. In: *Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems*. MIT Press, Cambridge, pp. 508–514.

Lee, T.-W., Girolami, M., Bell, A.J., Sejnowski, T.J. (2000). A unifying information-theoretic framework for independent component analysis. *Computers & Mathematics with Applications*, 39(11), 1–21.

Lipeika, A. (2010). Optimization of formant feature based speech recognition. *Informatica*, 21(3), 361–374.

Molgedey, L., Schuster, H.G. (1994). Separation of a mixture of independent signals using time delayed correlations. *Physical Review Letters*, 72(23), 3634–3637.

Moreau, E. (2002). A generalization of joint-diagonalization criteria for source separation. In: *IEEE Transactions on Signal Processing*, 49(3), pp. 530–541.

Morita, S., Nanri, Y. (2006). Sound source separation of trio using stereo musig sound signal based on independent component analysis. In: *IEEE International Conference on Multimedia and Expo*, pp. 185–188.

Naeem, M., Brunner, C., Pfurtscheller, G. (2009). Dimensionality reduction and channel selection of motor imagery electroencephalographic data. *Computational Intelligence and Neuroscience*, 5, 1–8.

Peric, Z.H., Savic, M.S., Dincic, M.R., Denic, D.B., Prascvic, M.R. (2010). Forward adaptation of novel semilogarithmic quantizer and lossless coder for speech signals compression. *Informatica*, 21(3), 375–392.

Pham, D., Garrat, P., Jutten, C. (1992). Separation of a mixture of independent sources through a maximum likelihood approach. In: *European Signal Processing Conference*, pp. 771–774.

Saylani, H., Hosseini, S., Deville, Y. (2009). Blind separation of noisy mixtures of non-stationary sources using spectral decorrelation. pp. 322–329.

Stone, J.V. (2002). Independent component analysis: an introduction. *Trends of Cognitive Science*, 6(2), 59–64.

Svecko, R., Kotnik, B., Chowdhary, A., Mezgec, Z. (2010). Gsm speech coder indirect identification algorithm. *Informatica*, 21(4), 575–596.

Tang, A.C., Sutherl, M.T., C.J.M.A. (2005). Validation of sobi components from high-density eeg, *Neuroimage*, 25(2), 539–53.

Wilson, K. (2007). Speech source separation by combining localization cues with mixture models of speech spectra. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 1, pp. 33–36.

Zarzoso, V., Nandi, A.K. (2002). Blind separation of independent sources for virtually any source probability density function. In: *IEEE Transactions on Signal Processing*, 47(9), pp. 2419–2432.

Zarzoso, V., Murillo-Fuentes, J.J., Boloix-Tortosa, R., Nandi, A.K. (2006). Optimal pairwise fourth-order independent component analysis. In: *IEEE Transactions on Signal Processing*, 54(8), pp. 3049–3063.

Ziehe, A., Laskov, P., Nolte, G., Müller, K.R. (2004). A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. *Journal of Machine Learning Research*, 5, 777–800.

G.R. Naik received BE degree in electronics and communication engineering from the University of Mysore, India, in 1997. ME degree in communication and information engineering from Griffith University, Brisbane, Australia, in 2002, and the PhD degree in the area of electronics engineering, specialised in biomedical engineering and signal processing from RMIT University, Melbourne, Australia, in 2009.

He is currently an academician and researcher at RMIT University. As an early career researcher, he has edited 6 books, authored more than 75 papers in peer reviewed journals, conferences, and book chapters over the last five years. His research interests include pattern recognition, blind source separation (BSS) techniques, audio signal processing, biosignal processing, and human-computer interface. Dr. Naik was the chair for the IEEE Computer Society CIT08 Conference, Sydney and a member of the organising committee for IEEE BRC2011, IEEE BRC 2012 and IEEE BRC 2013 conferences that are held in Brazil. He is also a reviewer and member of editorial board in several reputed journals. He was a recipient of the Baden-Württemberg Scholarship from the University of Berufsakademie, Stuttgart, Germany (2006–2007). In 2010, Dr. Naik is awarded with ISSI overseas fellowship from skilled Institute Victoria, Australia.

Šaltinių atskyrimo kokybės matas Gauso tipui artimiems mišiniams

Ganesh R. NAIK

Aklojo šaltinių atskyrimo algoritmai paprastai naudojami su prielaida, kad nagrinėjamieji šaltiniai aprašomi kitokiais nei Gauso pasiskirstymo dėsniais. Daugumoje atvejų nagrinėjami šnekos signalai, aprašomi super-Gauso pasiskirstymo dėsniu (Gauso tipui artimu dėsniu su teigiamu eksceso koeficientu). Tačiau praktiniuose uždaviniuose sutinkami ir sub-Gauso (su neigiamu eksceso koeficientu) dėsniu aprašomi signalai. Šiame straipsnyje pasiūlytas ir eksperimentiškai ištirtas šaltinių atskyrimo kokybės matas, tinkantis tiek šaltiniams su super-Gauso, tiek su sub-Gauso pasiskirstymo dėsniais. Eksperimentų rezultatai parodė, kad šaltiniams atskirti taikant nepriklausomų komponenčių analizės metodus, santykio signalas–pašalinis signalas pagerėjimas siekia 7–8 dB, o atskirtieji šaltiniai visą laiką gaunami super-Gauso tipo, nepriklausomai nuo signalų ar jų mišinio pasiskirstymo dėsnių.