HIGH FIDELITY BLIND SOURCE SEPARATION OF SPEECH SIGNALS

*A. K. Kattepur *, J. P. Lie †, F. Sattar *, C. M. S. See‡*

* Sch. of EEE, Nanyang Tech. Univ.
50 Nanyang Ave., Singapore 639798.
phone: +(65)67904525
fax: +(65)67920415
email: {aj0001ur,efsattar}@ntu.edu.sg

†TL@NTU, Nanyang Tech. Univ.
50 Nanyang Drv., Singapore 637553.
phone: +(65)67905447
fax:+(65)67900215
email: jonilie@ntu.edu.sg

‡DSO National Labs.
20 Science Park Drv.
Singapore, 118230.
phone: +(65)68712413
email: schongme@dso.org.sg

ABSTRACT

This paper addresses blind source separation (BSS) problem of multiple speech signals in low signal-to-interference-noise ratio (SINR) environment. We consider an over-determined case so that we can form multiple sub-arrays (of which there are as many sensors as speech signals), and propose a novel hybrid scheme to obtain high fidelity speech signals after separation. Firstly, the proposed method applies the commonly-used BSS technique at each sub-array to separate the speech signals. Next, the outputs of the same speech signal from different sub-arrays are grouped to form a new sub-array. We can then exploit the spatial diversity of the new sub-array to achieve high fidelity source separation. This configuration is the key innovation of this paper. Another contribution of the paper is on the justification of using the hybrid configuration to further increase the output SINR. From numerical analysis, it is demonstrated that 12 dB SINR improvement can be achieved using 5-element sensor array in the presence of two other interfering speech signals over a range of input SNR values. A significant improvement can also be seen from the output signal-to-artifact ratio (SAR) of the recovered signals.

1. INTRODUCTION

Blind source separation (BSS) algorithms are used to separate individual sources from a mixture where there is minimal prior information about either the source signals or the mixing process. These techniques have been used in various fields with considerable success including speech and music processing, sonar, EEG and financial data [1].

Statistically, the concept underlying blind signal processing is to use independent component analysis (ICA) techniques, which are based on the assumptions that the original signals are independent and non-Gaussian in nature [2]. The instantaneous mixing model for source separation in the time domain can be described by:

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \tag{1}$$

Here, $\mathbf{X}$ is the observed mixed signal, $\mathbf{A}$ is the mixing matrix, $\mathbf{S}$ is the source signal and $\mathbf{N}$ is the additive white Gaussian noise. The objective of any blind source separation algorithm is to develop an un-mixing matrix $\mathbf{W}$ such that the resulting signal $\hat{\mathbf{S}}$ is a close estimate of the original source signal $\mathbf{S}$.

$$\hat{\mathbf{S}} = \mathbf{WX} \tag{2}$$

A number of algorithms that have been used for blind source separation include FastICA [3], Infomax [4], JADE [5] and RobustICA [6]. These make use of second or higher order statistics to estimate the unmixing matrix $\mathbf{W}$. Measures of non-Gaussianity, include using higher order statistics (kurtosis and negentropy), are used to minimize the mutual information between the mixed components. To obtain satisfactory results, the noise is usually neglected or high input signal-to-interference plus noise ratio (SINR) is considered.

For separation of speech signals, the noise is always present at the sensors. Reduction of noise especially in low signal-to-noise ratio (SNR) conditions is crucial for accurate reconstruction. The source separation of speech signals in noisy environments have been studied in [7]. However, these techniques do not make use of spatial diversity of the sensors. Alternatively, higher-order cumulants can be used to obtain accurate de-noising and source separation. But this is achieved at the cost of high computational complexity.

In this paper, we propose a hybrid configuration for two-stage source separation and noise reduction scheme under over-determined setting by exploiting the spatial diversity. By combining the commonly used source separation techniques like FastICA and Infomax, with the minimum distortion noise reduction (MDNR) algorithm [8], we have shown the improvement in terms of the output SINR and signal-to-artifact ratio (SAR). Unlike other beamforming algorithms, the MDNR algorithm does not require the direction-of-arrival (DOA) information which would have restricted the position of the sources and sensors. The simulation results show that the MDNR algorithm provides better output when compared to the conventional Delay-and-Sum (DAS) beamforming.

2. FASTICA ALGORITHM

The FastICA algorithm [3] makes use of an efficient learning rule to maximize the non-Gaussianity of the projection. It is among the most commonly used algorithms for optimal search of the unmixing matrix $\mathbf{W}$ that is updated based on a nonlinear contrast function. The optimization techniques like gradient search or Newton optimization are used for updating the contrast function $G(\mathbf{WX})$, where $\mathbf{X}$ is the observed matrix of the mixed source signals.

The fixed point FastICA method makes use of batch processing of the observed data such that at each step, one row vector $\mathbf{w}$ of the unmixing matrix $\mathbf{W}$ can be estimated. The optimization of the objective function $G(\mathbf{w}^T\mathbf{x})$ ($\mathbf{x}$ refers to each row of the observed mixture $\mathbf{X}$) is subject to the constraint $E[(\mathbf{w}^T\mathbf{x})(\mathbf{w}^T\mathbf{x})^T] = 1$. Defining $g(.)$ and $g'(.)$ as the first and second derivatives of the contrast function, at the optimal value $\mathbf{w}_0$ with $\|\mathbf{w}_0\| = 1$ yields:

$$\phi = E[\mathbf{w}_0^T \mathbf{x} g(\mathbf{w}_0^T \mathbf{x})] \tag{3}$$

The unmixing process can then be optimized based on the Newton optimization method with the update for the n^{th} iteration given as

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \eta \left[(E[\mathbf{x} g(\mathbf{w}_n^T \mathbf{x})] - \varphi \mathbf{w}_n)/E[g'(\mathbf{w}_n^T \mathbf{x})] - \phi \right] \tag{4}$$

$$\mathbf{w}_{n+1} = \mathbf{w}_{n+1}/\|\mathbf{w}_{n+1}\| \tag{5}$$

where $\mathbf{w}_{n+1}$ is the new estimated value for every n^{th} iteration and φ is the step size.

As shown by [1], the FastICA algorithm can also be compared to the stochastic gradient method for maximizing likelihood like the infomax method [4]. However, the convergence of FastICA is cubic or quadratic, which is much faster than the linearly converging gradient descent methods. It can also be used to estimate both sub-Gaussian and super-Gaussian independent components. Due to these advantages, it has been chosen as the source separation algorithm in our scheme.

The performance bounds of noisy linear ICA has been studied in [10, 11]. The optimal solution in the case of noisy ICA is close to the minimum mean square error (MMSE) solution given by:

$$\mathbf{W}^{MMSE} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \sigma^2 I)^{-1} \tag{6}$$

where σ^2 is the noise variance. This leads to the minimum attainable signal to interference plus noise ratio for the k^{th} estimated signal characterized by:

$$\Psi = (I + \sigma^2(\mathbf{A}^T\mathbf{A})^{-1})^{-1} \tag{7}$$

$$\min \mathrm{SINR}_k = \frac{\Psi_{kk}^2}{\sum_{i \neq k}^{d} \Psi_{ki}^2 + \sigma^2 \sum_{i=1}^{d} (\Psi \mathbf{A}^{-1})_{ki}^2} \tag{8}$$

where i and k represent the rows and columns of the observed mixed signals, respectively and d refers to the total number of signals observed in the mixture. This shows that the bound is dependent only on the mixing matrix $\mathbf{A}$ and the noise variance σ^2.

3. MINIMUM DISTORTION NOISE REDUCTION

The MDNR algorithm proposed in [8] addresses the problem of estimating *one* source signal given the received signals at the microphone array. Let $\{\mathbf{y}_1(k), \cdots, \mathbf{y}_N(k)\}$ be the discretized received signals of L samples. By exploiting the spatio-temporal diversity, the source signal of m-th sensor at k-th sample $x_m(k)$ can be obtained by passing the received signals at N sensors (of which there are L samples) through N temporal filters of length L

$$\hat{x}_m(k) = \mathbf{h}_m^T \mathbf{y}(k) = \mathbf{h}_m^T \mathbf{x}(k) + \mathbf{h}_m^T \mathbf{v}(k) \tag{9}$$

where $\mathbf{h}_m = [\mathbf{h}_{1m}^T, \cdots, \mathbf{h}_{Nm}^T]^T$, $\mathbf{h}_{nm}$ is the column vector of L coefficients of the temporal filter for the n^{th} received signal. $\mathbf{y}(k) = [\mathbf{y}_1^T(k), \cdots, \mathbf{y}_N^T(k)]^T$, $\mathbf{x}(k) = [\mathbf{x}_1^T(k), \cdots, \mathbf{x}_N^T(k)]^T$, and $\mathbf{v}(k) = [\mathbf{v}_1^T(k), \cdots, \mathbf{v}_N^T(k)]^T$ are the received signal, clean signal and noise signal column vectors, respectively. Notice that we have grouped the signal term and noise term separately.

Using this form shown in (9), the task of the estimator is to find $\mathbf{h}_m$ by minimizing the mean-square-error due to the noise term under the constraints that the error due to the signal term $(\mathbf{h}_m^T \mathbf{x}(k) - x_m(k))$ is zero. That is, by solving the following optimization

$$\mathbf{h}_{m,o} = \arg\min_{\mathbf{h}_m} \mathbf{h}_m^T \mathbf{R}_{vv} \mathbf{h}_m \quad \text{s.t.} \quad \mathbf{Q}_m \mathbf{h}_m = \mathbf{u}_1 \tag{10}$$

where $\mathbf{Q}_m = [\mathbf{Q}_{1m}^T, \cdots, \mathbf{Q}_{Nm}^T]$ is the spatial-temporal prediction matrix, which relates the signal at one microphone to others: $\mathbf{x}_n(k) = \mathbf{Q}_{nm}\mathbf{x}_m(k)$. $\mathbf{R}_{vv} = E[\mathbf{v}(k)\mathbf{v}^T(k)]$. $\mathbf{u}_1 = [1, 0, \cdots, 0]^T$.

Solving (10) using Lagrangian multiplier method, the optimum $\mathbf{h}_m$ can be computed given the spatial-temporal prediction matrix. Instead of using the true $\mathbf{Q}_m$, which is usually unknown, an estimate can be obtained easily as

$$\mathbf{Q}_{nm,o} = (\mathbf{R}_{y_n,y_m} - \mathbf{R}_{v_n,v_m})(\mathbf{R}_{y_m,y_m} - \mathbf{R}_{v_m,v_m})^{-1} \tag{11}$$

where $\mathbf{R}_{v_n,v_m} = E[\mathbf{v}_n(k)\mathbf{v}_m(k)]$. The same definition applies similarly to $\mathbf{R}_{y_n,y_m}$. Therefore, the final expression of $\mathbf{h}_{m,o}$ is obtained by solving (10) and substituting (11) into the solution

$$\mathbf{h}_{m,o} = \mathbf{R}_{vv}^{-1}\mathbf{Q}_{m,o}^T[\mathbf{Q}_{m,o}\mathbf{R}_{vv}^{-1}\mathbf{Q}_{m,o}^T]\mathbf{u}_1 \tag{12}$$

where $\mathbf{Q}_{m,o}$ is arranged the same way as $\mathbf{Q}_m$.

It is stated in [8] that the worst-case performance of the MDNR algorithm will be that of the delay-and-sum beamforming [9] which is the case when only spatial diversity can be exploited for noise reduction. In this case, the noise power will be reduced by a factor of $1/N$ while the signal power remains unchanged. Given that the signal and noise power are σ_s^2 and σ_n^2, respectively. The worst-case output SINR for the MDNR algorithm can be expressed as

$$\mathrm{SINR}_{mdnr} = \frac{N\sigma_s^2}{\sigma_n^2} \tag{13}$$

4. SYSTEM MODEL

The problem of separating speech sources, or the typical cocktail party problem, has been investigated in previous literature [2]. Both convolutive and instantaneous mixtures for separating speech sources have also been studied. However, the efficient separation performance is limited to the case when either the noise is ignored or the input SNR is high. Gaussian noise causes deterioration of second order cumulants which the source separation algorithms can depend on. In case of FastICA, the assumptions regarding the covariance of the observed signals can be distorted especially under low SNR conditions. The mixing matrix can also become ill-conditioned leading to poor separation and de-noising capabilities.

In order to improve both the noise reduction and separation performance, we propose a hybrid approach. With limited pre-processing, the observed noisy data is passed through the blind source separation algorithm. By making use of the overdetermined condition when the number of sensors is more than the number of sources, the diversity in each of the separated outputs is used for noise reduction. The minimum distortion noise reduction algorithm makes use of the outputs from these multiple channels to achieve noise reduction.

Unlike conventional preprocessing/post processing noise reduction techniques used by common source separation algorithms, this scheme exploits the spatial diversity of the sensor locations for multiple channel noise reduction. Under low input SNR, other noise reduction schemes may distort the speech signal output. By exploiting the spatial diversity of the BSS algorithms and applying the MDNR technique, high fidelity in the speech output is ensured which is advantageous in its application to low SNR conditions.

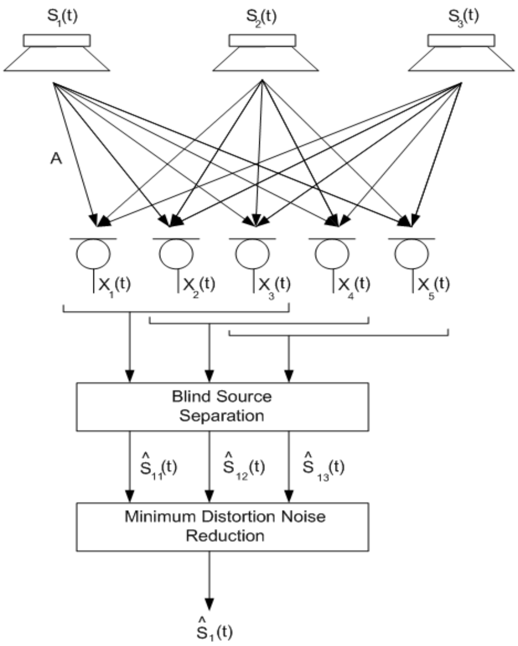

Figure 1: Scenario used for testing the proposed algorithm.

As illustrated in Fig. 1, three speech source signals $S_1(t)$, $S_2(t)$ and $S_3(t)$ are received by an array of 5 sensors after passing through a channel mixing matrix. Because of the over-determined condition, we form 3 sub-arrays and perform source separation for each sub-array. This will provide the spatial diversity required for multiple channel noise reduction. The next stage is to utilize the output of the BSS algorithm $\hat{S}_{11}(t)$, $\hat{S}_{12}(t)$ and $\hat{S}_{13}(t)$ (from the first, second and third sub-arrays respectively) as the input to the MDNR algorithm. This multiple channel noise reduction process, in turn, provides a high fidelity output $\hat{S}_1(t)$ of the target source signal $S_1(t)$. This procedure can be further repeated for the sources $S_2(t)$ and $S_3(t)$ to similarly obtain high fidelity outputs $\hat{S}_2(t)$ ans $\hat{S}_3(t)$, respectively.

A problem with most BSS algorithms like FastICA is the ordering of sources (permutation problem). In our technique, as we make use of outputs from each sub-array, this ordering is critical to provide accurate input to the MDNR stage. The correlation between the separated signals is used to solve this. As seen from Fig. 2, the highest correlation values r_1 and r_2 are used as the basis for matching the separated outputs.

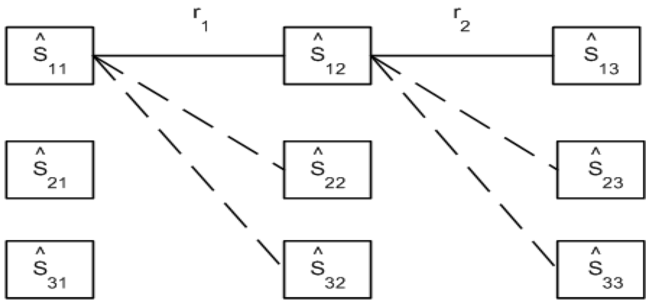

Figure 2: Example of using correlation to solve the permutation problem. The solid lines indicate highest correlation matching the separated output of each sub-array to a particular source.

5. PERFORMANCE ANALYSIS

The performance of the proposed hybrid approach can be evaluated in terms of the overall SINR output. Let S_k denote the desired speech signal to be separated. Given that there are d speech signals, the rest of the speech signals (S_i where $i \neq k$ and $i = 1, 2, \cdots, d$) are considered as interferences.

The minimum attainable output SINR of noisy linear ICA has been given in (8). The expression contains three different terms for the desired signal, interferences and noise power:

$$\begin{aligned} \sigma_s^2 &= \Psi_{kk}^2 \\ \sigma_i^2 &= \sum_{i \neq k}^{d} \Psi_{ki}^2 \\ \sigma_n^2 &= \sigma^2 \sum_{i=1}^{d} (\Psi \mathbf{A}^{-1})_{ki}^2 \end{aligned} \tag{14}$$

From Fig. 1, it can be seen that the output of the noisy linear ICA is also the input of the MDNR algorithm. Therefore, the output of the hybrid approach can be expressed as

$$\mathrm{SINR}_{hybrid,min} \leq \mathrm{SINR}_{hybrid} \leq \mathrm{SINR}_{hybrid,max} \tag{15}$$

where $\{\mathrm{SINR}_{hybrid,min}, \mathrm{SINR}_{hybrid,max}\}$ are the minimum and maximum attainable SINR output which can be written as

$$\begin{aligned} \mathrm{SINR}_{hybrid,min} &= \frac{\Psi_{kk}^2}{\sum_{i \neq k}^{d} \Psi_{ki}^2 + \frac{\sigma^2}{N_{sub}} \sum_{i=1}^{d} (\Psi \mathbf{A}^{-1})_{ki}^2} \\ \mathrm{SINR}_{hybrid,max} &= \frac{\Psi_{kk}^2}{\frac{\sigma^2}{N_{sub}} \sum_{i=1}^{d} (\Psi \mathbf{A}^{-1})_{ki}^2} \end{aligned} \tag{16}$$

where $N_{sub} = N - d + 1$ is the number of subarrays formed after the BSS and N is the total sensors used. Notice that the above inequality is used to express the output SINR of the proposed hybrid approach, because the MDNR algorithm is not formulated for suppressing the interferences. Thus, the expression for $\mathrm{SINR}_{hybrid,min}$ relates to only reduction of noise by the MDNR algorithm with no interference suppression by BSS. The $\mathrm{SINR}_{hybrid,max}$ is achieved when the BSS technique has a perfect signal separation with improved noise suppression by the MDNR algorithm.

As compared to using a standard direct approach by just applying the BSS technique, the hybrid approach offers additional noise reduction capability. This is reflected on the noise power expression at the output of the hybrid approach. It is clear that the noise power has been reduced to $(1/N)$ fraction of the noise power at the BSS intermediate output. Applying the standard direct BSS approach does not effectively exploit the extra sensor outputs as the performance is similar to the critically determined case. By effectively using the overdetermined criterion, the hybrid approach offers additional SINR improvement by reducing the output noise power. This produces a significantly better output than the direct BSS approach.

6. RESULTS AND DISCUSSIONS

Based on the scenario shown in Fig. 1 the speech sources are mixed based with no reverberation considered. The array processing toolbox developed by [12] is used for the mixing process based on the position of the speech sources and the distribution of sensors (uniform linear array). The sources and sensors are placed at distances of 1 and 0.1 meters apart, respectively[1]. The source separation of the observed mixed signals at the sensor array is performed based on the FastICA algorithm. The permutation problem is seen in all the BSS algorithms, specially those that operate in the frequency domain. In this application, back correlation with respect to the input signal is used as the solution to the permutation problem. That is, the highest correlated BSS output with respect to the input signal is used as the corresponding estimate of that particular input signal. Further denoising of the separated output based on the sub-array structure is achieved by either the minimum distortion noise reduction (MDNR) or the delay-and-sum (DAS) beamforming algorithms.

Fig. 3 shows the performance of the algorithm when applied to three noisy mixed speech signals. The outputs of the BSS algorithm with input SNR of -10 dB is input to both the MDNR and the DAS algorithms. The MDNR algorithm is able to successfully recover the denoised version of the original signal. For listening test, a demonstration of the proposed separation process is available in [14].

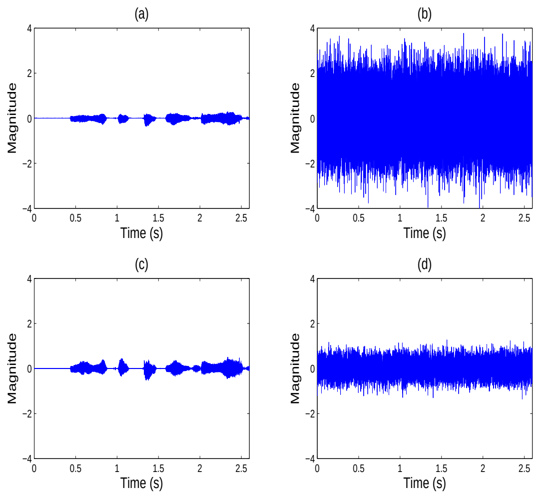

Figure 3: Example of using the algorithm to separate and denoise the signal (a) Original signal (b) Separated signal before denoising (c) Estimated signal after MDNR (d) Estimated signal after DAS.

[1]Note that the geometry of the array can be made arbitrary

The output of the proposed algorithm has been tested based on the output SINR and SAR improvements using the toolbox developed by [13]. The separation performance is computed for each estimated source $\hat{\mathbf{s}}_j$ and compared with the true source $\mathbf{s}_j$. The first step is to decompose the estimated unmixed signal as shown.

$$\hat{\mathbf{s}}_j = \mathbf{s}_{target} + e_{interf} + e_{noise} + e_{artif} \tag{17}$$

where $\mathbf{s}_{target}$ is a version of $\mathbf{s}_j$ modified by an allowable distortion, e_{interf} is an allowed deformation of the sources which accounts for the interferences of the unwanted sources, e_{noise} is an allowed deformation of the perturbating noise and e_{artif} is an artifact term that corresponds to artifacts of the separation algorithm such as musical noise or to deformations induced by the separation algorithm that are not allowed.

The next step is to compute energy ratios to evaluate the relative amount of each of the four terms in (17) either on the whole signal duration or on local frames. The computation of the SINR and SAR follows from the equations given below.

$$SINR = 10\log_{10}\frac{\left\|\mathbf{s}_{target}\right\|^2}{\left\|e_{interf}+e_{noise}\right\|^2} \tag{18}$$

$$SAR = 10\log_{10}\frac{\left\|\mathbf{s}_{target}+e_{interf}+e_{noise}\right\|^2}{\left\|e_{artif}\right\|^2} \tag{19}$$

While SINR is a measure of the separation performance, SAR measures the distortions caused by the source separation algorithm on the signals of interest.

As shown in Fig. 4, the MDNR algorithm provides better SINR output when compared to the DAS beamforming, specially at low input SNR. The cases for two and three mixed sources have also been considered. At higher SNR, the noise suppression performance of both DS and MDNR techniques converge. Due to this, we notice an overlap of the graphs at high SNR, specially for the case of three sources. This is because, at higher SNR, the performance is dependent mainly on the separation performance. A reference to the minimum attainable SINR as shown in eq. (16) is provided for a mixture of three sources.

Similarly, the improvements in SAR for two and three sources are presented. As seen from Figs. 5 and 6, the SAR improvements for both two and three sources are considerable especially under low input SNR. This demonstrates that both source separation and noise reduction have been successfully incorporated assuring the high fidelity of the output speech signals. The mean values of SINR and SAR have been used in all the above cases with approximations based on the toolbox in [13].

7. CONCLUSIONS

We have proposed a two-stage approach to extract high fidelity speech signals after BSS. The proposed method considers an over-determined setting, where the number of sensors used is more than the speech signals to be separated. In this setting, the spatial diversity is exploited to provide higher SINR improvement using MDNR algorithm. The theoretical performance analysis as well as simulation results confirmed that the proposed method is able to achieve higher SINR improvement, particularly in low input SNR condition. Compared to the DAS algorithm with the same settings, the proposed method achieves almost 3 dB additional SINR improvement. It is also able to provide high SAR outputs in case of two or three interfering sources. In future, this scheme can be applied to reverberant and convolutive source mixtures in order to evaluate its efficacy in restoring the original sources.

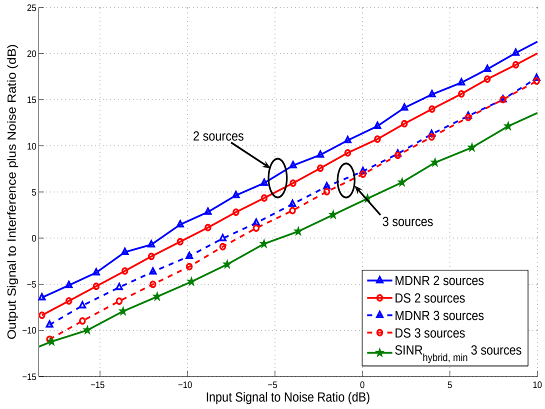

Figure 4: Output SINR for various settings of input SINR based on a mixture of two and three sources.

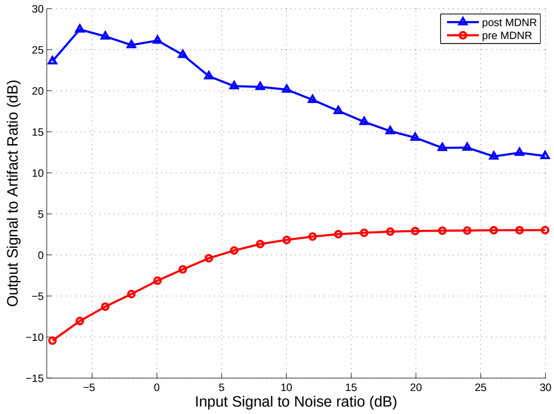

Figure 5: Output SAR for various input settings based on a mixture of two sources.

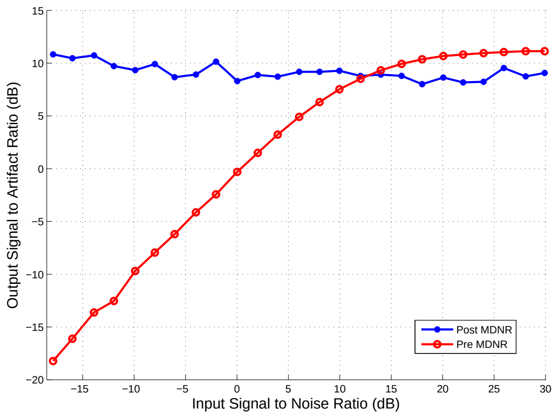

Figure 6: Output SAR for various input settings based on a mixture of three sources.

REFERENCES

[1] A. Hyvarinen and E. Oja, "Independent Component Analysis: Algorithms and Applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

[2] M. S. Pedersen, J. Larsen, U. Kjems, and L. C. Parra, "A Survey of Convolutive Blind Source Separation Methods," *Springer Handbook on Speech Processing and Speech Communication*, 2007.

[3] A. Hyvarinen and E. Oja, "A Fast Fixed-point Algorithm for Independent Component Analysis," *Neural Computation*, vol. 9, pp. 483–492, 1997.

[4] A. J. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, pp. 129–159, 1995.

[5] J. F. Cardoso and A. Souloumiac, "Jacobi angles for Simultaneous Diagonalization," *SIAM J. of Matrix Analysis and Applications*, vol. 17, no. 1, pp. 161–165, 1996.

[6] J. Cao and N. Murata, "A Stable and Robust ICA Algorithm Based on t-distribution and Generalized Gaussian Distribution Models," *IEEE Conf. on Neural Networks for Signal Processing*, Madison, 1999.

[7] Z. Zhang and M. Etoh, "ICA-based Noise Reduction for Mobile Phone Speech Communication," *16th Intl. Conf. on Computer Communications and Networks*, Honolulu, pp. 470–473, 2007.

[8] J. Chen, J. Benesty and Y. Huang, "A Minimum Distortion Noise Reduction Algorithm with Multiple Microphones," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 16, no. 3, pp. 481–493, 2008.

[9] B.D. Van Veen and K.M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE Acoustics, Speech and Signal Processing Magazine*, vol.5, no.2, pp.4–24, April 1988.

[10] Z. Koldovsky and P. Tichavsky, "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise," *IEEE Intl. Conf. on Acoustics, Speech and Signal Processing ICASSP*, France, vol. 5, 2006.

[11] P. Tichavsky, Z. Koldovsky, and E. Oja, "Performance Analysis of the FastICA Algorithm and CramerRao Bounds for Linear Independent Component Analysis," *IEEE Trans. on Signal Processing*, vol. 54, no. 4, pp. 1189–1203, 2006.

[12] K. D. Donohue, "Microphone Array Processing Toolbox," Audio Systems Laboratory, University of Kentucky, 2008. http://www.engr.uky.edu/ donohue/audio/Arrays/MAToolbox.htm.

[13] C. Fevotte, R. Gribonval and E. Vincent, "BSS EVAL Toolbox User Guide," IRISA Technical Report 1706, Rennes, France, 2005. http://www.irisa.fr/metiss/bss eval/.

[14] J. P. Lie and A. K. Kattepur, "Demonstration of High Fidelity BSS of Speech Signals," Nanyang Technological University, http://www3.ntu.edu.sg/home/jonilie/research/BSS_MDNR.htm