

Geometrically Constrained Independent Component Analysis

Mirko Knaak, Shoko Araki, *Member, IEEE*, and Shoji Makino, *Fellow, IEEE*

Abstract—Acoustical signals are often corrupted by other speeches, sources, and background noise. This makes it necessary to use some form of preprocessing so that signal processing systems such as a speech recognizer or machine diagnosis can be effectively employed. In this contribution, we introduce and evaluate a new algorithm that uses independent component analysis (ICA) with a geometrical constraint [constrained ICA (CICA)]. It is based on the fundamental similarity between an adaptive beamformer and blind source separation with ICA, and does not suffer the permutation problem of ICA-algorithms. Unlike conventional ICA algorithms, CICA needs prior knowledge about the rough direction of the target signal. However, it is more robust against an erroneous estimation of the target direction than adaptive beamformers: CICA converges to the right solution as long as its look direction is closer to the target signal than to the jammer signal. A high degree of robustness is very important since the geometrical prior of an adaptive beamformer is always roughly estimated in a reverberant environment, even when the look direction is precise. The effectiveness and robustness of the new algorithms is proven theoretically, and shown experimentally for three sources and three microphones with several sets of real-world data.

Index Terms—Blind source separation (BSS), independent component analysis (ICA), machine diagnosis, minimum variance beamforming, signal enhancement, statistical signal processing.

I. INTRODUCTION

FOR MANY signal processing tasks, such as speech recognition, transmission, or signal classification, the target signal must be very well reconstructed when it is disturbed by other sources. Adaptive beamformers (ABF) and blind source separation (BSS) are effective tools for multichannel signal reconstruction.

Adaptive beamformers were introduced by Frost [1] and have been expanded by several contributors, e.g., [2], [3]. It can be shown formally that they provide optimal suppression of the jammer under certain conditions. However, as shown in this paper they have limited robustness against erroneous parameters, particularly the delay vector in each frequency bin [4]. This is problematic since the delay vector is always roughly estimated in a reverberant environment, as shown in this paper. The methods traditionally used to overcome this sensitivity mostly

Manuscript received June 7, 2004; revised January 24, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Maurizio Omologo.

M. Knaak is with the Measurement Technology Laboratory, University of Technology, D-10587 Berlin, Germany and also with NTT Communication Science Laboratories, NTT Corporation, Kyoto 619-0237, Japan (e-mail: Mirko.Knaak@iav.de).

S. Araki and S. Makino are with NTT Communication Science Laboratories, NTT Corporation, Kyoto 619-0237, Japan.

Digital Object Identifier 10.1109/TASL.2006.876730

broaden and, thereby, flatten the directivity pattern, resulting in a trade-off between the signal suppression performance and the parameter sensitivity (e.g., [5], [6]).

Independent component analysis (ICA) is an emerging technique for finding statistically independent components in a multi-channel signal. The main application is BSS which has been shown to be capable of recovering multiple sources from their linear mixture if the sources are independent [7].

Due to the time-delayed superposition of the sources and reflections from the walls, the mixtures are convolutive in the field of acoustics. The separation of convolutive mixtures involves estimating significantly more parameters (see [8] for references) than when separating instantaneous mixtures. Frequency-domain approaches simplify the problem into instantaneous separation problems for the frequency components (e.g., [9], [10]). The scaling and permutation ambiguities left in the recovered frequency components become a serious problem. Different permutations of the frequency components lead to mixed outputs and degraded separation results.

There are several approaches to overcome this problem, e.g., [11], [12]. These algorithms are very time consuming or reduce the spectral resolution of the unmixing filter, therefore the scaling and permutation problem is still the main obstacle to the method. Moreover, when there are more than two sources and two microphones, it is hard to solve the permutation problem. Mainly due to these problems, the number of real applications in the acoustics field is still very limited and the separation performance is not sufficient [12].

Recent research has indicated an equivalence between ABF and BSS, e.g., [13]. BSS is an intelligent set of ABFs with an adaptive null directivity aimed in the direction of the unnecessary sounds [14]. This equivalence suggests the application of a geometrical constraint to ICA to solve the permutation and scaling problem.

In many applications the source position is roughly known in advance since it can be estimated by sound localization methods (e.g., MUSIC [15]), determined by image processing or simply known by geometry. The high sensitivity of the ABF to parameter changes however requires a very precise estimation of the source position that cannot be normally achieved by the aforementioned methods. To make it worse, even a moderate reverberation makes a delay vector estimated by the accurate position of the jammer source unreliable.

Geometrically constrained algorithms have been proposed by [16] and [17]. Their contribution does not assess in detail how a rough estimation affects the performance of the algorithm. In [18] a geometrically constrained ICA algorithm (CICA) was introduced that employs the fast convergence properties of the

FastICA [7]. In this paper, we analyze the behavior of CICA formally, and compare its performance to other ICA and beamforming algorithms with a large set of machine sounds.

In Section II, we make basic assumptions about the mixture, and discuss the standard technologies for multichannel signal processing, namely the ABF and ICA algorithms. Section III analyzes possible reasons for the estimation of the delay vector being imprecise and evaluates the consequences for the adaptive beamformers. The new algorithm is introduced and its robustness and effectiveness are assessed theoretically in Section IV. The experimental results are given in Section V.

II. ASSUMPTIONS ABOUT THE MIXTURES

In a set $\mathbf{s}^b(t) = [s_{\text{target}}^b(t), s_2^b(t), \dots, s_M^b(t)]^T$ of broadband sources, the first source is the target sound and the others are interfering sources. The sound is measured with an array of N microphones $\mathbf{x}^b(t) = [x_1^b(t), \dots, x_N^b(t)]^T$. The observed signals are filtered and mixed because the room acoustics impose a different impulse response h_{nm}^b between each source s_m^b and each microphone x_n^b

$$x_n^b = \sum_{m=1}^M h_{nm}^b * s_m^b.$$

In the frequency domain, a convolutive mixture can be written as $\mathbf{x}^f = \mathbf{H}^f \cdot \mathbf{s}^f + \mathbf{n}^f$, where \mathbf{x}^f is a narrowband signal component filtered from \mathbf{x}^b with a bandpass centered at f . For simplicity, the index f is omitted hereafter. Small bold letters mean vectors, capitalized bold letter represent matrices. $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_M]$ consists of the delay vectors \mathbf{h}_m , where \mathbf{h}_1 is the delay vector of the target sound. Only under anechoic conditions, they can be approximated by the phase shifts caused by the time delays τ_{mn} with $\mathbf{h}_m = [e^{j2\pi f \tau_{m1}}, \dots, e^{j2\pi f \tau_{mN}}]^T$. When considering echoes and reverberation, \mathbf{h}_m is the sum of all echo paths. The covariance matrix of observed signals \mathbf{x} is

$$\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^H\} \quad (1)$$

where $\langle \rangle^H$ is the Hermitian (conjugate transposed) and E is the expectation operator. Since \mathbf{R}_x is a Hermitian matrix, it has N nonzero eigenvalues $\mathbf{\Lambda} = \text{diag}[\lambda_1, \dots, \lambda_N]$ and N orthogonal eigenvectors $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N]$ belonging to them.

The goal of the algorithms discussed here is to find an optimal estimation $y_1(t)$ of the target signal s_{target} . To achieve this goal, a coefficient vector $\mathbf{w}_1 = [w_{11}, \dots, w_{1N}]^T$, or an unmixing matrix \mathbf{W}

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1^H \\ \vdots \\ \mathbf{w}_M^H \end{bmatrix} \quad (2)$$

is applied to the vector of observations as follows:

$$\mathbf{y} = \mathbf{W} \cdot \mathbf{x} \quad \text{or} \quad y_1 = \mathbf{w}_1^H \cdot \mathbf{x}. \quad (3)$$

A. Blind Source Separation With ICA

Blind source separation uses ICA to estimate the unmixing matrix \mathbf{W} . The estimation consists of two steps: sphering with \mathbf{V} and rotating with \mathbf{T}

$$\mathbf{W} = \mathbf{T}^H \cdot \mathbf{V}. \quad (4)$$

In the first step (sphering), the matrix \mathbf{V} is determined by principal component analysis (PCA)

$$\mathbf{z} = \mathbf{\Lambda}^{-\frac{1}{2}} \cdot \mathbf{U}^H \cdot \mathbf{x} = \mathbf{V} \cdot \mathbf{x} \quad (5)$$

where \mathbf{z} are the sphered signals (the intermediary result of the sphering step).

In the second step, spatio-temporal decorrelation [19], nonlinear decorrelation [20], [21], maximization of nongaussianity [22] or nonstationary decorrelation [11] can be used to determine the rotation matrix \mathbf{T} . Similar to \mathbf{W} in (2), \mathbf{T} can be written as $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_N]^H$.

B. Adaptive Beamformer

In the following, we restrict ourselves to the class of minimum variance (MV) beamformers. The optimization problem for the MV beamformer is to minimize the energy of the output signal

$$J = E\{y_1^2\} \quad (6)$$

with a constraint which ensures that the energy of a signal coming from the direction of the target is passed without changes

$$\mathbf{w}_1^H \hat{\mathbf{h}}_1 = c_1 \quad (7)$$

where c_1 is an arbitral constant and $\hat{\langle \rangle}$ is an estimation and hence $\hat{\mathbf{h}}_1$ is the estimated and not necessarily correct delay vector. Using Lagrangian multipliers, the cost function is obtained as

$$J_{\text{ABF}} = E\{y_1^2\} + \lambda \left(\mathbf{w}_1^H \hat{\mathbf{h}}_1 - c_1 \right). \quad (8)$$

Fig. 1 shows an example of a scatter plot of the observed signals x_1 and x_2 when $N = M = 2$. The random vector is distributed around the mixing vectors \mathbf{h}_1 and \mathbf{h}_2 which are added in this figure. It is possible to draw potential coefficient vectors in this scatter plot since they are in the same space. The coefficient vectors with the same output energy (6) are connected in this figure.

The energy function whose contour lines are also drawn in Fig. 1 is driven by the quadratic form $\mathbf{w}_1^H \mathbf{R}_x \mathbf{w}_1$. Hence, the cost function has a bowl shape with one global minimum at the point of origin, and the constraint in (7) is necessary to prevent the algorithm from going to the global (unwanted) minimum ($w_{11} = w_{12} = 0$). The constraint is a line (or in the general case a hyperplane), drawn in Fig. 1. It has a strong impact on the performance of the beamformer as shown in [15].

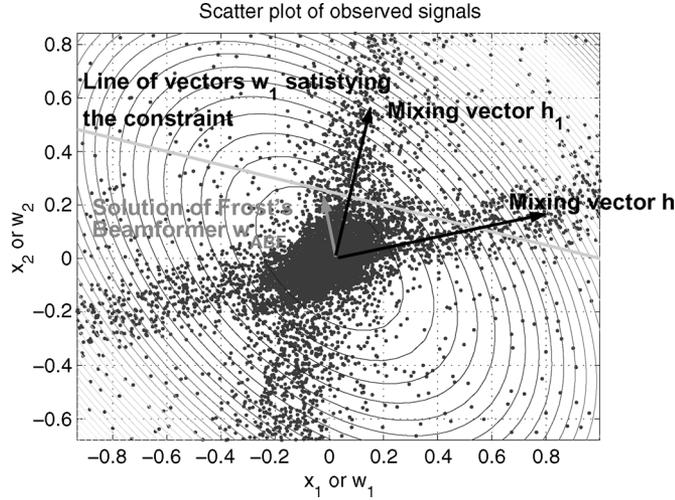


Fig. 1. Scatter plots of the observed signals with the contour lines of Frost's cost function (6).

The optimal coefficient vector \mathbf{w}_{ABF} is shown in Fig. 1

$$\mathbf{w}_{\text{ABF}} = c_1 \frac{\mathbf{R}_x^{-1} \hat{\mathbf{h}}_1}{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1}. \quad (9)$$

Due to the equivalence between ABF and ICA under ideal conditions as shown in [13], the solution in (9) is also the optimal ICA solution. An important consequence is that the vector \mathbf{t}_{opt} of the rotation matrix that reconstructs the target signal is simply the delay vector of the target signal transformed with a sphering matrix (except for an amplification factor δ)

$$\mathbf{t}_{\text{ABF}} = \mathbf{t}_{\text{opt}} = \delta \mathbf{V} \mathbf{h}_1. \quad (10)$$

III. EFFECT OF A ROUGH ESTIMATION OF THE DELAY VECTOR

A. Reasons for an Imprecise Estimation of $\hat{\mathbf{h}}_1$

A major drawback of ABFs is that they rely on the correct estimation of the delay vector. Since the impulse response of a room is normally not available, the delay vector $\hat{\mathbf{h}}_1$ is estimated solely from the time delays of the direct sound by

$$\hat{\mathbf{h}}_1(f) = \begin{bmatrix} 1 \\ \vdots \\ e^{j2\pi \cdot (N-1)d \cdot \cos(\hat{\theta}) \frac{f}{c}} \end{bmatrix} \quad (11)$$

where d is the distance between the microphones, c is the speed of sound and $\hat{\theta}$ is the estimated direction of arrival.

An estimation error of $\hat{\mathbf{h}}_1$ has two causes: a wrong direction of arrival $\hat{\theta}$, or the existence of reverberation. The latter is due to the multiple directions of arrival while only the direct sound is used in (11) for the estimation. When reverberation is present, the estimation becomes imprecise even when the source position is well known.

To determine the error included in a rough estimation of $\hat{\mathbf{h}}_1$, we introduce a new measure: the delay or steering vector error angle (SVA) μ . For its definition, we use a generalized cosine

$$\cos(\mathbf{x}, \mathbf{y}) = \frac{\|\langle \mathbf{x}, \mathbf{y} \rangle\|}{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}} \quad (12)$$

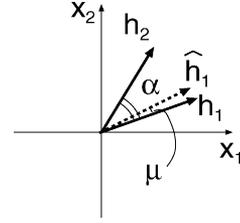


Fig. 2. Definition of angles α and μ . The axes are the same as those of Fig. 1.

to define the angle between two complex vectors \mathbf{x} and \mathbf{y} . In (12), $\langle \mathbf{x}, \mathbf{y} \rangle$ is the inner product of the two complex valued vectors \mathbf{x} and \mathbf{y} . Let $\mu(f)$ be the angle between the roughly estimated delay vector $\hat{\mathbf{h}}_1(f)$ and the actual mixing vector $\mathbf{h}_1(f)$ and let $\alpha(f)$ be the angle between the mixing vectors $\mathbf{h}_1(f)$ of the target signal and $\mathbf{h}_2(f)$ of the jammer signal. For the visualization of the definition see Fig. 2. Since μ and α are defined in the (complex) space of demixing vectors, they are not identical with the spatial angles, e.g., between the locations of the sources

$$\mu(f) = \arccos\left(\frac{\mathbf{h}_1^H \hat{\mathbf{h}}_1}{|\mathbf{h}_1| |\hat{\mathbf{h}}_1|}\right) \quad \alpha(f) = \arccos\left(\frac{\mathbf{h}_1^H \mathbf{h}_2}{|\mathbf{h}_1| |\mathbf{h}_2|}\right). \quad (13)$$

Fig. 3 shows the SVA in two rooms with and without reverberation and with two levels of preciseness of the look direction. The two levels of reverberation were achieved by using impulse responses from an anechoic room ($t_R \approx 0$ ms) and a normal room ($t_R = 128$ ms). The impulse responses for the anechoic room were derived by using the first portion of a measured impulse response. The real vector \mathbf{h}_1 was calculated by applying a Fourier transform to the impulse response. It is clearly seen that in the reverberant case an SVA may exist even though the look direction is precise [Fig. 3(c)].

B. Consequences for the Adaptive Beamformer

In order to assess the signal suppression properties of the ABF when the estimation of $\hat{\mathbf{h}}_1$ is rough, this subsection evaluates the eigenstructure of the covariance matrices used for the considered beamformers. In the derivation we restrict ourselves to the case where $N = M = 2$. The results are also valid in a general case.

The output signal consists of two components; p_1 describes the contribution of the target signal source and p_2 of the jammer source to the output signal y_1

$$\begin{aligned} y_1 &= \mathbf{w}^H (\mathbf{h}_1 s_1 + \mathbf{h}_2 s_2) \\ &= \frac{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1}}{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1} (\mathbf{h}_1 s_1 + \mathbf{h}_2 s_2) \\ &= \frac{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_1}{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1} s_1 + \frac{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2}{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1} s_2 \\ &= p_1 s_1 + p_2 s_2. \end{aligned} \quad (14)$$

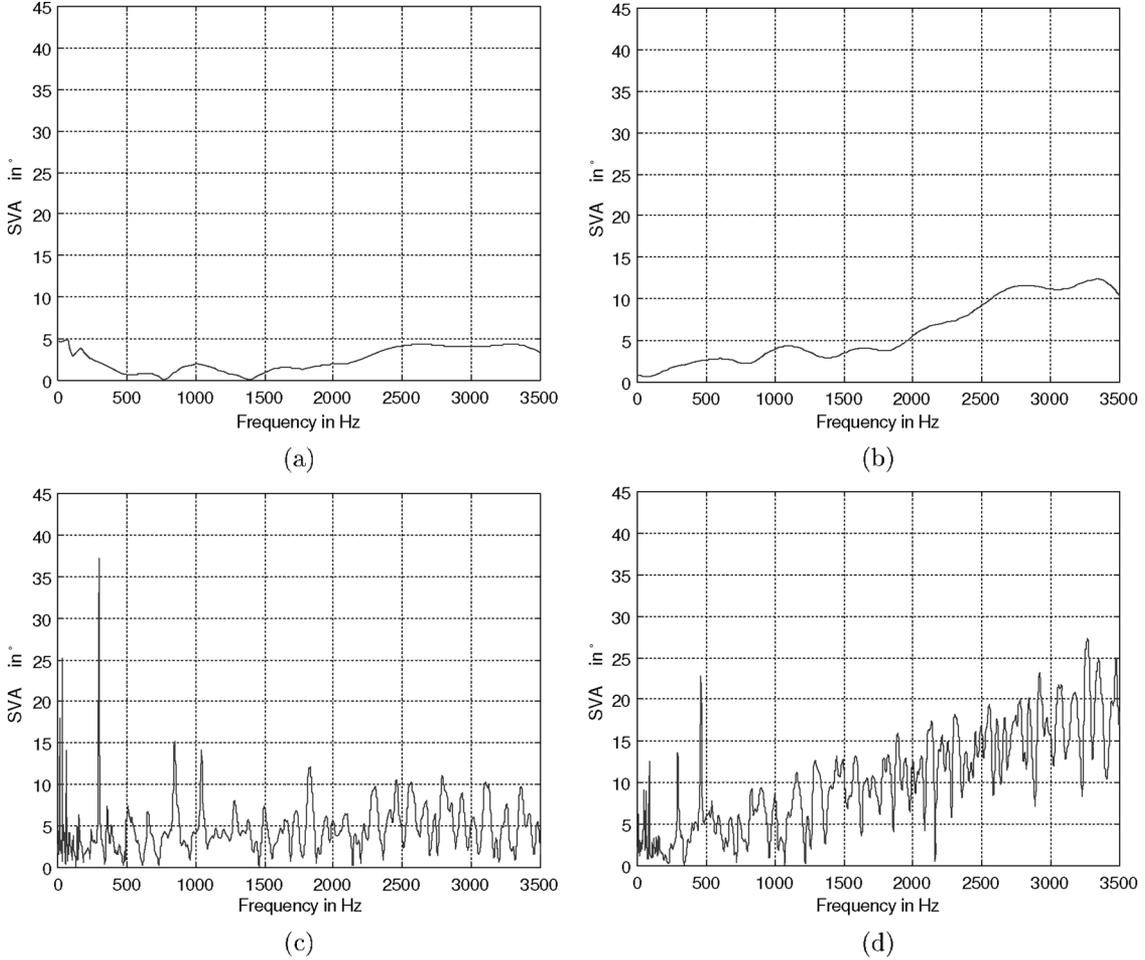


Fig. 3. SVA $\mu(f)$ induced by a rough estimation and reverberation. The effect of reverberation is greater than the effect of an imprecise estimation. $\Delta\theta$ is the difference between the estimated look direction and the actual look direction. (a) Correct estimation $\Delta\theta = 0^\circ$, $t_R \approx 0$ ms. (b) Incorrect estimation $\Delta\theta = 10^\circ$, $t_R \approx 0$ ms. (c) Correct estimation $\Delta\theta = 0^\circ$, $t_R = 128$ ms. (d) Incorrect estimation $\Delta\theta = 10^\circ$, $t_R = 128$ ms.

The eigenanalysis of the covariance matrix \mathbf{R}_x of the observed signals allows a further insight into these coefficients. The covariance matrix can be written as

$$\mathbf{R}_x = \sigma_{s_1}^2 \mathbf{h}_1 \mathbf{h}_1^H + \sigma_{s_2}^2 \mathbf{h}_2 \mathbf{h}_2^H + \sigma_n^2 \mathbf{I} \quad (15)$$

with $\sigma_{s_1}^2$ and $\sigma_{s_2}^2$ indicating the powers of the two sources. In order to simplify the eigenanalysis, we assume that both source components contribute the same power. This is satisfied if both sources have similar intensities and are nearly in the same distance from the array. Then, the variances are considered to be equal $\sigma_{s_1}^2 = \sigma_{s_2}^2 = \sigma_s^2$ and the mixing vectors to have the same norm $|\mathbf{h}_1| = |\mathbf{h}_2|$. Additionally the noise is neglected ($\sigma_s^2 \gg \sigma_n^2$).

Under these assumptions, \mathbf{R}_x has the following eigenvectors and eigenvalues:

$$\mathbf{u}_1 = \frac{\mathbf{h}_1 + \kappa \mathbf{h}_2}{|\mathbf{h}_1 + \kappa \mathbf{h}_2|} \quad \lambda_1 = \sigma_s^2 (|\mathbf{h}_1| + |\mathbf{h}_1^H \mathbf{h}_2|) \quad (16)$$

$$\mathbf{u}_2 = \frac{\mathbf{h}_1 - \kappa \mathbf{h}_2}{|\mathbf{h}_1 - \kappa \mathbf{h}_2|} \quad \lambda_2 = \sigma_s^2 (|\mathbf{h}_1| - |\mathbf{h}_1^H \mathbf{h}_2|) \quad (17)$$

with $\kappa = \mathbf{h}_2^H \mathbf{h}_1 / |\mathbf{h}_2^H \mathbf{h}_1|$. The proof is given in the Appendix.

To reduce the complexity, we treat the nominator and the denominator of p_1 and p_2 separately resulting in

$$\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_1 = \frac{\cos \mu - \cot \alpha \sin \mu}{\sigma_s^2} \quad (18)$$

$$\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2 = \frac{\sin \mu}{\sigma_s^2 \sin \alpha} \quad (19)$$

$$\begin{aligned} \hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1 &= \frac{1}{\sigma_s^2} \left(\frac{\cos \mu + \cos(\alpha - \mu)}{2 \cdot (1 + \cos \alpha)} \right)^2 \\ &+ \frac{1}{\sigma_s^2} \left(\frac{\cos \mu - \cos(\alpha - \mu)}{2 \cdot (1 - \cos \alpha)} \right)^2. \end{aligned} \quad (20)$$

The complete derivation is shown in the Appendix. Equations (18)–(20) describe the influence of the jammer and the target signal on the output signals with the angles μ and α defined in (13). They make it possible to treat the robustness of the output signal as a function of the roughness of the delay vector $\hat{\mathbf{h}}_1$. If (and only if) $\hat{\mathbf{h}}_1$ is estimated correctly ($\hat{\mathbf{h}}_1 = \mathbf{h}_1$ or $\mu = 0$), the

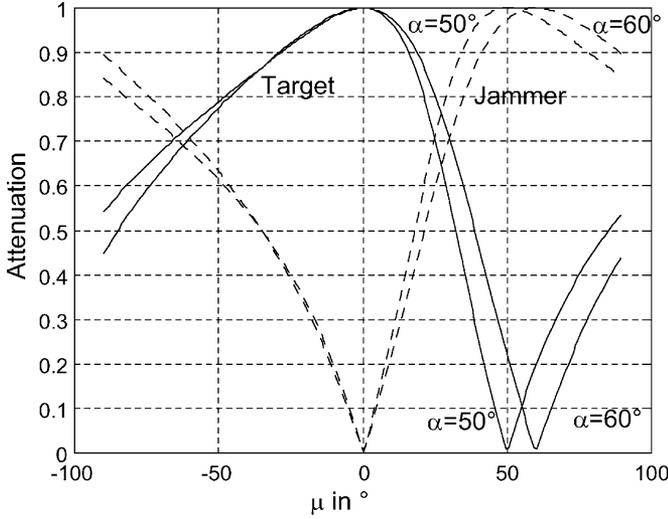


Fig. 4. Attenuation of the target signal \mathbf{p}_1 (solid) and the jammer \mathbf{p}_2 (dashed) as a function of μ . α is 50° and 60° .

coefficient vector is orthogonal to the jammer signal. The target signal is not attenuated

$$p_1 = \frac{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_1}{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_1} = \frac{\frac{1}{2\sigma_s^2} + \frac{1}{2\sigma_s^2}}{\frac{1}{\sigma_s^2}} = 1$$

$$p_2 = \frac{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2}{\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_1} = \frac{\frac{1}{2\sigma_s^2} - \frac{1}{2\sigma_s^2}}{\frac{1}{\sigma_s^2}} = 0.$$

If $\hat{\mathbf{h}}_1$ is roughly estimated, the ABF degrades drastically. Fig. 4 shows the effect of a rough estimation of the delay vector when α and μ vary. For a precise estimation ($\mu = 0$), the beamformer works properly: it passes the target signal without changing it and minimizes the jammer signal. When the estimation error increases, the performance is reduced. For higher values of μ the target signal experiences the same attenuation as the jammer signal. The actual shape of the curves is influenced by α ; when α is smaller, the decline of the attenuation starts earlier.

IV. PROPOSED CONSTRAINED ICA

A. Derivation

To avoid the undesirable high sensitivity of the estimation of $\hat{\mathbf{h}}_1$ which is very detracting as shown in Section III-B, we are employing ICA to make the algorithm more robust. As mentioned above, several ICA approaches are possible. We consider two cost functions to be applicable to the development of a constrained algorithm. The first is based on time-delayed decorrelation [10], [11] and the second is based on negentropy maximization (FastICA) [22]. As is usual in ICA approaches, sphering (5) is employed. Fig. 5 shows a scatter plot of the sphered signals. After sphering, we have the following equations based on time-delayed decorrelation

$$\arg \min_{\mathbf{T}} \sum_{\tau} |\text{off} - \text{diag} \mathbf{T} \mathbf{R}_z(\tau) \mathbf{T}^H|^2 \quad (21)$$

with a constraint to the target signal :

$$\mathbf{w}_1^T \hat{\mathbf{h}}_1 = \mathbf{t}_1^H \mathbf{V} \hat{\mathbf{h}}_1 = c_1 \quad (22)$$

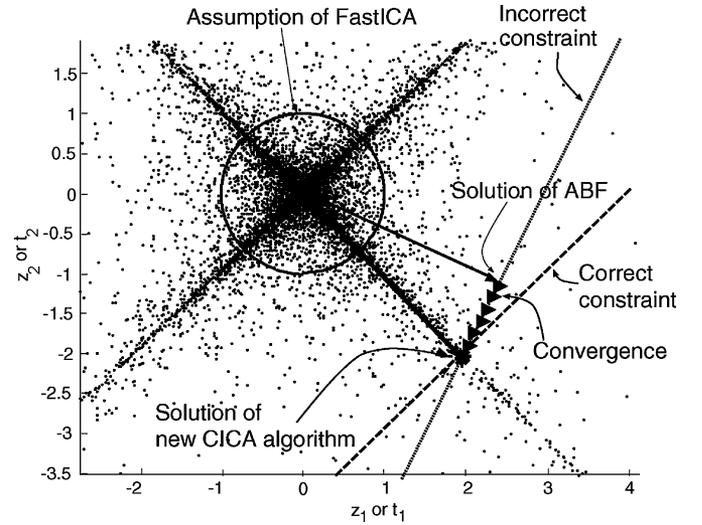


Fig. 5. Convergence of the proposed constrained ICA algorithm. The algorithm starts at the solution of the ABF and converges on the constraint line to the correct solution.

or similarly based on negentropy maximization with an arbitrary nonlinear function $G()$ with its derivation $g()$

$$\arg \min_{\mathbf{w}_1} E \left\{ G \left(|\mathbf{w}_1^H \mathbf{x}|^2 \right) \right\}$$

$$\equiv \arg \min_{\mathbf{t}_1} E \left\{ G \left(|\mathbf{t}_1^H \mathbf{z}|^2 \right) \right\} \quad (23)$$

with a constraint to the target signal :

$$\mathbf{w}_1^H \hat{\mathbf{h}}_1 = \mathbf{t}_1^H \mathbf{V} \hat{\mathbf{h}}_1 = c_1 \quad (24)$$

where \mathbf{V} is a sphering matrix, \mathbf{T} is a rotation matrix and c_1 is a constant (see Section II-A). The function that was used in the experiments for this paper is $g(z) = 1/(a+z)$.

Although most ICA algorithms claim to be unconstrained they normally employ the constraint that \mathbf{T} is unitary. It is necessary to avoid convergence to the point of origin and reduce the dimensionality of the optimization problem since it can be accomplished on the unit hyper sphere, shown for two microphones in Fig. 5. Further, the constraint significantly simplifies the derivation of the optimization function.

However, the constraint that the matrix \mathbf{T} is unitary implicitly considers source signals with a unit variance

$$\mathbf{I} = E\{\mathbf{z}^T \mathbf{z}\} = \mathbf{V} \mathbf{H} \cdot E\{\mathbf{s}^T \mathbf{s}\} \cdot \mathbf{H}^H \mathbf{V}^H$$

$$= \mathbf{V} \mathbf{H} \mathbf{H}^H \mathbf{V}^H = \mathbf{V} \mathbf{H} (\mathbf{V} \mathbf{H})^H$$

$$\mathbf{V} \mathbf{H} \text{ is unitary}$$

$$\tilde{\mathbf{T}} = \mathbf{V} \mathbf{H}^{-1} = \mathbf{V} \mathbf{H}^H \text{ is also unitary.} \quad (25)$$

The underlying assumption of convolutive BSS is obviously that $\mathbf{R}_s = \mathbf{I}$ and, therefore, $\sigma_i^2 = 1$. The algorithms tend to flatten the power spectrum. Several methods have been proposed to avoid this problem, e.g., [21].

When using ICA with the geometrical constraint of the ABF, we have to weaken the assumption and use $\tilde{\mathbf{T}}$ since the strict deployment of the constraint collides with the constraint in (22) or

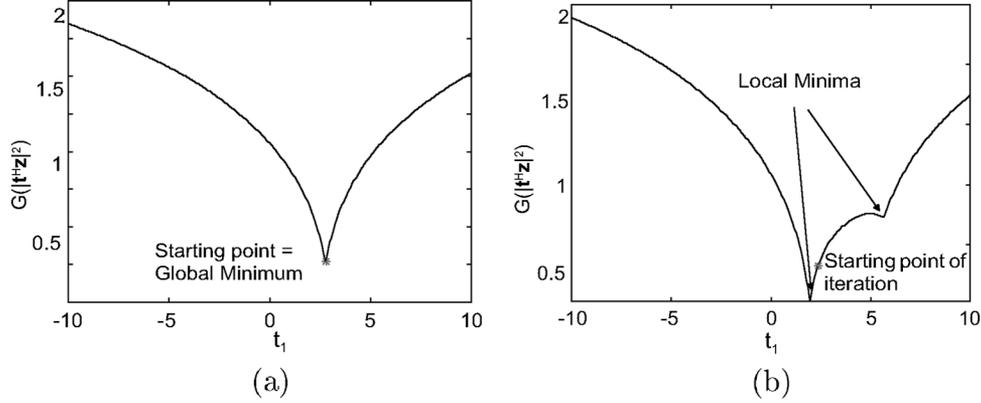


Fig. 6. Cost function on the constraint. (a) Precise estimation $\mu = 0^\circ$, $\alpha = 27^\circ$. (b) Rough estimation $\mu = -12^\circ$, $\alpha = 27^\circ$.

(24) as shown in Fig. 5. The column vectors need to be perpendicular, but they do not have to have a unit length. The degraded solution is avoided by the constraint.

Unfortunately, weakening the constraint makes the derivation of the optimization rule less simple and the employment of the time-delayed decorrelation approach impossible. Hence, only the FastICA algorithm is taken into account in the further discussion. We investigate the deflationary algorithm proposed by Bingham [22] as a basis. According to [23], the Kuhn–Tucker points of (23) are

$$E\{\mathbf{z} \cdot g(\mathbf{t}_1^H \mathbf{z})\} = \beta \mathbf{t}_1 \quad (26)$$

when the derivation is made on the unit circle $|\mathbf{t}_1|^2 = 1$.

We can also use this condition for points of \mathbf{t}_1 that are not on the unit circle. If \mathbf{t}_1 is also a Kuhn–Tucker point on the unit hypersphere, then $\mathbf{t}'_1 = \gamma \mathbf{t}_1$ (for any γ) is a Kuhn–Tucker point of the minimization problem that constrains the solution to vectors of the same norm. According the theory of FastICA, the maximal nongaussianity only says something about the direction of the unmixing vector while the norm is not decisive. We are looking for a vector that satisfies the constraint (24) and that has the highest negentropy of all vectors with the same norm.

Hence, we do not change the solution of (26) by projecting it to our constraint (24). We obtain the following algorithm, with the convergence shown in Fig. 5

$$\mathbf{t}_{1,k+1} = \mathbf{t}_{1,k} - \frac{E\{\mathbf{z}g(\mathbf{t}_{1,k}^H \mathbf{z})\} + \beta \mathbf{t}_{1,k}}{E\{\mathbf{z}g(\mathbf{t}_{1,k}^H \mathbf{z})\} + \beta} \quad (27)$$

$$\mathbf{t}_{1,k+1_{new}} = \frac{\mathbf{t}_{1,k+1}}{|\mathbf{t}_{1,k+1}^H \mathbf{V} \hat{\mathbf{h}}_1|} \quad (28)$$

$$\beta = \frac{E\{\mathbf{t}_{1,k}^H \mathbf{z} \cdot g(\mathbf{t}_{1,k}^H \mathbf{z})\}}{|\mathbf{t}_{1,k}|^2}. \quad (29)$$

Equations (27)–(29) describe the iteration of our new CICA algorithms. The index k is the iteration index. The new algorithm starts with $\mathbf{t}_{1,1} = \hat{\mathbf{t}}_1 = \mathbf{V} \hat{\mathbf{h}}_1$. If the estimation of $\hat{\mathbf{h}}_1$ is correct, the estimation itself is already the correct solution $\mathbf{t}_{1,opt} = \mathbf{V} \hat{\mathbf{h}}_1$. Then, the algorithm converges according to (27)–(29) to a stable point of the minimization problem \mathbf{t}_{opt} .

B. Theoretical Assessment

The new algorithm should be evaluated in terms of its convergence properties. The Newton method guarantees that the algorithms will converge to a stable point. Fig. 6(a) shows that there is only one global minimum in the case of a precise estimation of the delay vector $\hat{\mathbf{h}}_1$.

When the estimation is rough, two local minima appear [Fig. 6(b)]. The two minima belong to the two signals. This means that since convergence is ensured, the algorithm converges either towards the target or the jammer. A convergence towards the jammer signal amplifies the jammer signal and suppresses the target signal. This is equivalent to the permutation problem of unconstrained BSS.

The shape of the minimization of the constraint and the starting point of the iteration control the convergence. Convergence can be guaranteed when $\hat{\mathbf{t}}_1 = \mathbf{V} \hat{\mathbf{h}}_1$ is closer to the correct solution \mathbf{t}_1 than to the permuted solution \mathbf{t}_2 because it is then in the convergence range of \mathbf{t}_1 . Both parameters are influenced by the SVA in the PCA space μ_t . By using earlier results (18) and lemma 1 in the Appendix, it can be assessed as follows:

$$\begin{aligned} \cos(\mu_{t_1}) &= \frac{\hat{\mathbf{t}}_1^H \mathbf{t}_1}{|\hat{\mathbf{t}}_1| |\mathbf{t}_1|} = \frac{\hat{\mathbf{h}}_1^H \mathbf{V}^H \mathbf{V} \mathbf{h}_1}{|\hat{\mathbf{t}}_1| |\mathbf{t}_1|} \\ &= \frac{\hat{\mathbf{h}}_1^H \mathbf{U} \mathbf{\Lambda}^{-1/2} \mathbf{\Lambda}^{-1/2} \mathbf{U}^H \mathbf{h}_1}{|\hat{\mathbf{t}}_1| |\mathbf{t}_1|} \\ &= \frac{\hat{\mathbf{h}}_1^H \mathbf{U} \mathbf{\Lambda}^{-1} \mathbf{U}^H \mathbf{h}_1}{|\hat{\mathbf{t}}_1| |\mathbf{t}_1|} \\ &= \frac{\hat{\mathbf{h}}_1^H \mathbf{R}^{-1} \mathbf{h}_1}{|\hat{\mathbf{t}}_1| |\mathbf{t}_1|} \\ &= \frac{\cos \mu - \cot \alpha \sin \mu}{|\hat{\mathbf{t}}_1| |\mathbf{t}_1|}. \end{aligned} \quad (30)$$

And similarly we can define μ_{t_2} as the angle between $\hat{\mathbf{t}}_1$ and \mathbf{t}_2 in the PCA space

$$\begin{aligned} \cos(\mu_{t_2}) &= \frac{\hat{\mathbf{t}}_1^H \mathbf{t}_2}{|\hat{\mathbf{t}}_1| |\mathbf{t}_2|} \\ &= \frac{\sin \mu}{\sin \alpha}. \end{aligned} \quad (31)$$

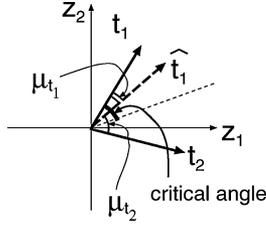


Fig. 7. Vectors in the space of the sphered signals. The axes are the same as those of Fig. 5.

Convergence can be guaranteed as long as $\cos(\mu_{t_2}) > \cos(\mu_{t_1})$ (Fig. 7)

$$\begin{aligned} \cos(\mu_{t_1}) &< \cos(\mu_{t_2}) \\ \frac{\cos \mu - \cot \alpha \sin \mu}{|\hat{\mathbf{t}}_1||\mathbf{t}_1|} &< \frac{\frac{\sin \mu}{\sin \alpha}}{|\hat{\mathbf{t}}_1||\mathbf{t}_2|} \\ \cos \mu - \cot \alpha \sin \mu &< \frac{\sin \mu}{\sin \alpha} \\ \cos \mu &< \cot \alpha \sin \mu + \frac{\sin \mu}{\sin \alpha} \\ \cot \mu &< \frac{1 + \cos \alpha}{\sin \alpha} \quad (32) \\ \cot \mu_{\text{critical}} &= \frac{1 + \cos \alpha}{\sin \alpha} \\ \Rightarrow \mu_{\text{critical}} &= \frac{\alpha}{2}. \quad (33) \end{aligned}$$

With (32), we can define the convergence range of our proposed CICA algorithm; it will converge to the target when the misestimation is smaller than the critical angle μ_{critical} in (33). Otherwise it will converge to the permuted solution (jammer signal). Since the critical angle is $\alpha/2$, where α is the angle between \mathbf{h}_1 and \mathbf{h}_2 [see Fig. 2 and (13)], the algorithm converges to the correct solution as long as its look direction is closer to the target signal than to the jammer signal.

It should be noted that this condition always holds when $\hat{\mathbf{h}}_1$ is not in the angle spanned by \mathbf{h}_1 and \mathbf{h}_2 . Since μ is defined by the absolute value of the scalar product, μ can only have a positive sign. Hence, the algorithm may also converge to the correct solution if μ is larger than the critical angle. Equation (33) is a sufficient condition for the convergence.

This condition allows us to compare the CICA algorithm with the ABF (see Fig. 8). CICA is significantly more robust against a misestimation of $\hat{\mathbf{t}}_1 = \mathbf{V}\hat{\mathbf{h}}_1$. The separation performance of the ABF decreases continuously with a growth of the estimation error (dashed curve), while the CICA algorithm has a constantly good performance until the critical angle in (33).

For a real-world acoustical application, it is important to analyze (32) with realistic numbers for μ and α in each frequency bin. As shown earlier in Section III-A, the main error is introduced by the reverberation. A similar result can be seen here (Fig. 9). In both cases the SVA is in almost all frequency bins under the critical angle.

The frequency range in the reverberant case needs to be addressed separately. The angle α , and thereby the critical angle $\alpha/2$, are very small in the low frequency range. Even a small estimation error causes a wrong permutation. In that frequency

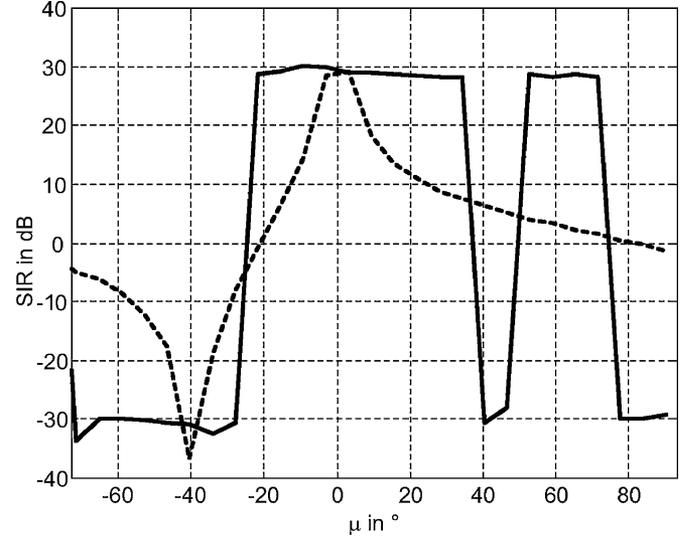


Fig. 8. Comparison of the robustness of the CICA algorithm (solid line) and the adaptive beamformer (dashed line). SIR (signal to interference ratio) is defined as the ratio of the target and jammer signal contributions to the output.

range, the authors recommend the use of the conventional FastICA algorithm and solve the permutation with maximizing the correlation to the solutions in the other frequency bins.

V. REAL-WORLD ASSESSMENT

We compared the new CICA algorithm with the ICA algorithms, Infomax (INFO), and the minimum variance adaptive beamformer (MVB) by using several tests. We measured impulse responses in a real room, and made mixtures by convolving these with the sources. The real room setup is depicted in Fig. 10. Its reverberation time t_R is 150 ms. Two ($N = M = 2$) or three ($N = M = 3$) observation channels (microphones) and sources were used.

As the signals of interest, we utilized the sound probes of electrical motors. The interfering sound here was the typical background noises of an assembly hall (like other motors, sewing and drilling noise, etc.). When the number of used microphones was two, a sound probe was interfered by one interferer, and when the case of three microphones, a sound probe was interfered by two interference sources. Investigations were carried out for ten probes: Four probes are of fault-free motors: six probes consisted an unusual sound caused by a fault (commutator fault). Such a setup is common in acoustical machine diagnosis with the objective to assess the state of machine by its sound. In an earlier publication the algorithm has also been tested with Japanese speech signals successfully [18].

First the algorithms are evaluated by the improvement in the signal to interference ratio (SIR). The SIR is defined according to [20] and [24] as $10 \log[\sum_t y_{\text{target}}^2(t) / \sum_t y_{\text{jammer}}^2(t)]$ with y_{target} the portion of the target and $y_{\text{jammer}} = y_1 - y_{\text{target}}$ the portion of the jammer in the output signal. Since this measure is based on the energy portions contributed by the different sources, it does not say anything about a potential distortion of the target signal. Therefore, a second measure is employed which measures how well the sound probes from the fault-free motors can be distinguished from the faulty ones. Two classes

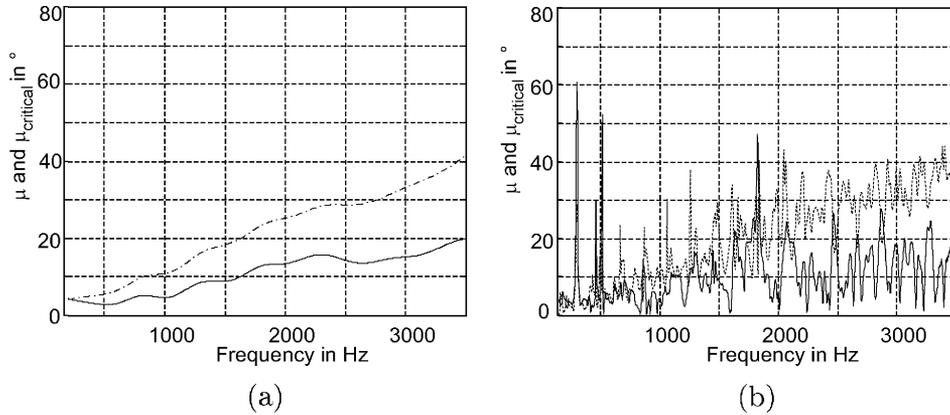


Fig. 9. Comparing the SVA (solid line) for a real room with the critical angle (dotted line) of this mixture. (a) Without reverberation, $t_R \approx 0$ ms, $\Delta\theta = 20^\circ$; (b) with reverberation $t_R = 150$ ms, $\Delta\theta = 20^\circ$.

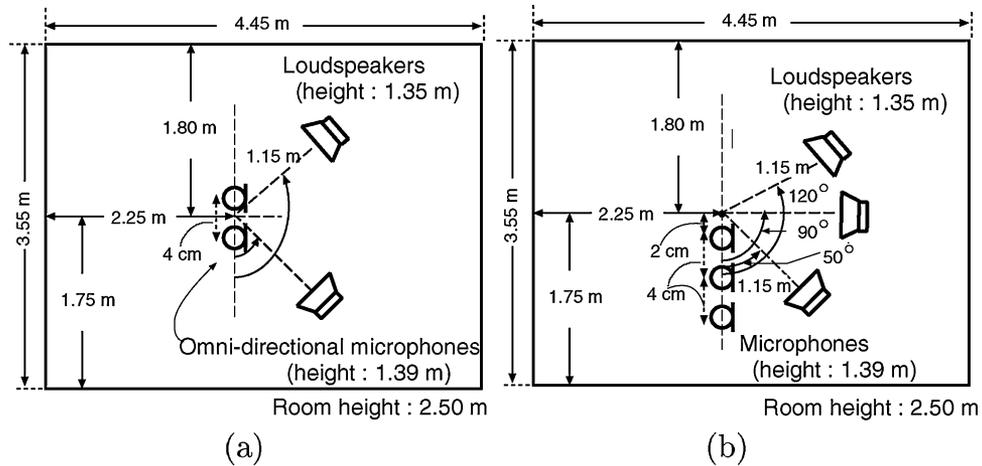


Fig. 10. Diagram of the real room and the setup of the microphones and loudspeakers with (a) two and (b) three sources and observation channels. The reverberation time is $t_R = 150$ ms.

can be separated in a certain feature space (typical features of electrical motors are loudness, roughness, energy at some frequencies, etc.) if the distance between the centers of the two classes is larger than the standard deviation of each class in the feature space. Since the Mahalanobis distance is defined as the distance between these centers normalized by the standard deviation it is a good measure of the distinctness of the classes and it must be greater than one to classify the classes automatically. In the following, the Mahalanobis distance is calculated for the original, disturbed, and reconstructed signals [25].

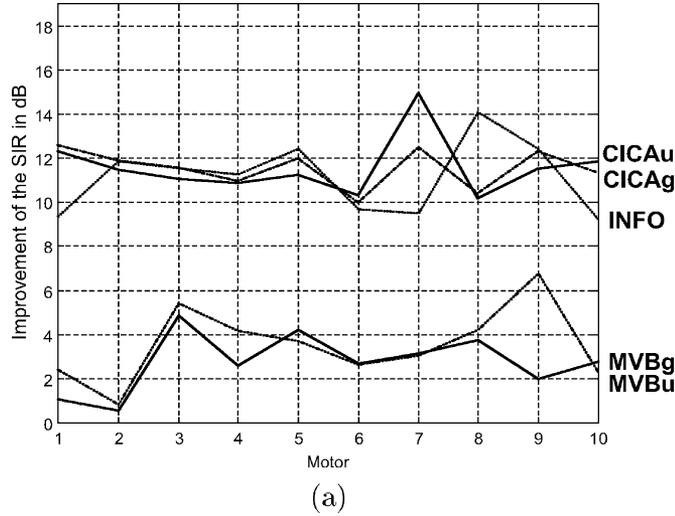
When two microphones are used, the sound probes are interfered with one interferer. Infomax does not require the direction of the motor sound, while MVB and CICA need this information. However, the ICA algorithms face the permutation problem, what MVB and CICA do not. When $N = M = 2$ the permutation problem in Infomax is solved by estimating the directivity patterns as proposed in [26]. Of course, CICA and MVB do not need to solve the permutation problem. To assess the robustness of CICA and MVB, the look direction corresponds to the direction of the motor sound ($\Delta\theta = 0^\circ$, 'g') in the first case and differs from the actual direction by 20° ($\Delta\theta = 20^\circ$, 'u') in the second case.

In the test with measured impulse responses [Fig. 11(a) with two microphones and two sources], the delay vector is always erroneous, since it is estimated by the phase shifts of the direct sound while the actual delay vector is also influenced by the reverberation. Therefore, the MVB fails completely in this environment. The blind algorithms and CICA improve the SIR sufficiently. The same result can be seen with the Mahalanobis classifier (Table I): only signals reconstructed with the ICA-based algorithms have a Mahalanobis distance larger than one that makes it possible to distinguish between the motor sets.

Finally, the tests with measured impulse responses were repeated with target signals interfered with two jammer sources ($N = M = 3$) where CICA was compared to the ABF. The ABF faces the same obstacle described above. Hence, only CICA shows good results here. Fig. 12 shows the SIR improvement for all the motors.

VI. CONCLUSION

This paper introduced and assessed a new geometrically constrained ICA algorithm. It showed the algorithm's effectiveness



Algorithms	Δ -SIR
MVBU	2.7 dB
MVBg	3.5 dB
INFO	11.1 dB
CICAu	11.6 dB
CICAg	11.7 dB

(b)

Fig. 11. Improvement in the SIR for $N = M = 2$ with measured impulse responses ($t_R = 150$ ms). (a) For all motors, (b) average SIR improvement.

TABLE I
MAHALANOBIS DISTANCES

Algorithms	Distance
Original	5.1
Interfered	0.1
MVBU	0.5
MVBg	0.4
INFO	3.0
CICAu	1.2
CICAg	3.1

and robustness theoretically and experimentally. The new algorithm solves the permutation problem of BSS of acoustic mixtures, particularly when the number of sources and microphones is greater than two. Although prior knowledge is needed, the information does not need to be precise.

APPENDIX PROOFS

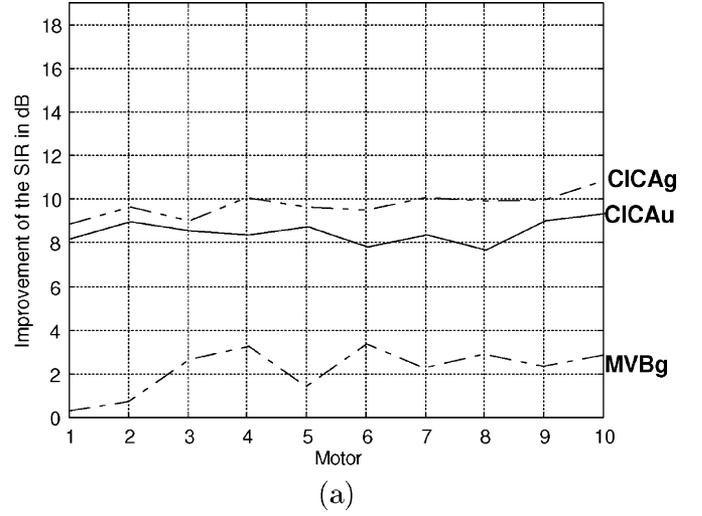
Lemma 1: The covariance matrix \mathbf{R}_x has the eigenvectors and eigenvalues

$$\mathbf{u}_1 = \frac{\mathbf{h}_1 + \kappa \mathbf{h}_2}{|\mathbf{h}_1 + \kappa \mathbf{h}_2|} \quad \lambda_1 = \sigma_s^2 (|\mathbf{h}_1| + |\mathbf{h}_1^H \mathbf{h}_2|)$$

$$\mathbf{u}_2 = \frac{\mathbf{h}_1 - \kappa \mathbf{h}_2}{|\mathbf{h}_1 - \kappa \mathbf{h}_2|} \quad \lambda_2 = \sigma_s^2 (|\mathbf{h}_1| - |\mathbf{h}_1^H \mathbf{h}_2|)$$

with $\kappa = \mathbf{h}_2^H \mathbf{h}_1 / |\mathbf{h}_2^H \mathbf{h}_1|$.

Proof: A potential eigenvector \mathbf{u} is a linear combination of \mathbf{h}_1 and \mathbf{h}_2 : $\mathbf{u} = \alpha \mathbf{h}_1 + \beta \mathbf{h}_2$. Let $b = |\mathbf{h}_1|^2 = \mathbf{h}_1^H \mathbf{h}_1 = |\mathbf{h}_2|^2$ and $d = \mathbf{h}_1^H \mathbf{h}_2$. It is worth to mention that the inner product



Algorithms	Δ -SIR
MVBg	2.2 dB
CICAu	8.5 dB
CICAg	9.7 dB

(b)

Fig. 12. Improvement in the SIR for $N = M = 3$ with measured impulse responses ($t_R = 150$ ms) for all motors (a), average SIR improvement (b).

of complex valued vectors is not commutative, in opposite the real valued vectors: $\mathbf{a}^H \mathbf{b} = (\mathbf{b}^H \mathbf{a})^*$, where $\langle \rangle^*$ is the conjugate-complex. And, therefore $d^* = \mathbf{h}_2^H \mathbf{h}_1$. An eigenvector is defined as follows:

$$\lambda \mathbf{u} = \mathbf{R}_x \mathbf{u}$$

$$\lambda(\alpha \mathbf{h}_1 + \beta \mathbf{h}_2) = (\sigma_{s_1}^2 \mathbf{h}_1 \mathbf{h}_1^H + \sigma_{s_2}^2 \mathbf{h}_2 \mathbf{h}_2^H) (\alpha \mathbf{h}_1 + \beta \mathbf{h}_2)$$

$$\lambda \alpha \mathbf{h}_1 + \lambda \beta \mathbf{h}_2 = \alpha \sigma_{s_1}^2 \mathbf{h}_1 \mathbf{h}_1^H \mathbf{h}_1 + \alpha \sigma_{s_2}^2 \mathbf{h}_2 \mathbf{h}_2^H \mathbf{h}_1$$

$$+ \beta \sigma_{s_1}^2 \mathbf{h}_1 \mathbf{h}_1^H \mathbf{h}_2 + \beta \sigma_{s_2}^2 \mathbf{h}_2 \mathbf{h}_2^H \mathbf{h}_2$$

$$= \alpha \sigma_s^2 b \mathbf{h}_1 + \alpha \sigma_s^2 d^* \mathbf{h}_2 + \beta \sigma_s^2 d \mathbf{h}_1 + \beta \sigma_s^2 b \mathbf{h}_2$$

$$= \sigma_s^2 (\alpha b + \beta d) \mathbf{h}_1 + \sigma_s^2 (\alpha d^* + \beta b) \mathbf{h}_2.$$

A comparison of coefficients yields the following equations:

$$\sigma_s^2 (\alpha b + \beta d) = \lambda \alpha$$

$$\sigma_s^2 (\alpha d^* + \beta b) = \lambda \beta.$$

Let $\alpha = 1$, then

$$\beta = \frac{\lambda - \sigma_s^2 b}{\sigma_s^2 d} \quad 0 = \sigma_s^2 d^* + \frac{(\lambda - \sigma_s^2 b)}{\sigma_s^2 d} (\sigma_s^2 b - \lambda).$$

The solution of the quadratic equation is $\lambda_i (i = 1, 2) = \sigma_s^2 (b \pm |d|)$, and hence the eigenvalues

$$\lambda_1 = \sigma_s^2 (|\mathbf{h}_1|^2 + |\mathbf{h}_1^H \mathbf{h}_2|) \quad \lambda_2 = \sigma_s^2 (|\mathbf{h}_1|^2 - |\mathbf{h}_1^H \mathbf{h}_2|).$$

β is

$$\beta = \frac{\lambda - \sigma_s^2 b}{\sigma_s^2 d} = \frac{\sigma_s^2 (b \pm |d|) - \sigma_s^2 b}{\sigma_s^2 d} = \frac{\pm |d|}{d}.$$

When $\kappa = |d|/d = d^*/|d|$ and with a normalization the

$$\mathbf{u}_1 = \frac{\mathbf{h}_1 + \kappa\mathbf{h}_2}{|\mathbf{h}_1 + \kappa\mathbf{h}_2|} \quad \mathbf{u}_1 = \frac{\mathbf{h}_1 - \kappa\mathbf{h}_2}{|\mathbf{h}_1 - \kappa\mathbf{h}_2|}.$$

q.e.d.

Corollary 1: As a special case, the eigenvectors and eigenvalues for a real valued covariance matrix are

$$\mathbf{u}_1 = \frac{\mathbf{h}_1 + \mathbf{h}_2}{|\mathbf{h}_1 + \mathbf{h}_2|} \quad \lambda_1 = \sigma_s^2 \mathbf{h}_1^H (\mathbf{h}_1 + \mathbf{h}_2) \quad (34)$$

$$\mathbf{u}_2 = \frac{\mathbf{h}_1 - \mathbf{h}_2}{|\mathbf{h}_1 - \mathbf{h}_2|} \quad \lambda_2 = \sigma_s^2 \mathbf{h}_1^H (\mathbf{h}_1 - \mathbf{h}_2). \quad (35)$$

Lemma 2: The nominator of p_2 is

$$\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2 = \frac{\sin \mu}{\sigma_s^2 \sin \alpha}.$$

The nominator of p_1 is

$$\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_1 = \frac{\cos \mu - \cot \alpha \sin \mu}{\sigma_s^2}.$$

The denominator of both components is

$$\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1 = \frac{1}{\sigma_s^2} \left(\frac{\cos \mu + \cos(\alpha - \mu)}{2 \cdot (1 + \cos \alpha)} \right)^2 + \frac{1}{\sigma_s^2} \left(\frac{\cos \mu - \cos(\alpha - \mu)}{2 \cdot (1 - \cos \alpha)} \right)^2.$$

Proof: The inverse of a hermitian matrix can be written using the eigenvalues λ_i and eigenvectors \mathbf{u}_i

$$\mathbf{K}^{-1} = \sum \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^H. \quad (36)$$

To reduce the complexity, we treat the nominator of p_1 and p_2 and the denominator separately, starting with the nominator of p_2

$$\begin{aligned} \hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2 &= \hat{\mathbf{h}}_1^H \left(\frac{\mathbf{u}_1 \mathbf{u}_1^H}{\lambda_1} + \frac{\mathbf{u}_2 \mathbf{u}_2^H}{\lambda_2} \right) \mathbf{h}_2 \\ &= \frac{\hat{\mathbf{h}}_1^H \mathbf{u}_1 \mathbf{u}_1^H \mathbf{h}_2}{\lambda_1} + \frac{\hat{\mathbf{h}}_1^H \mathbf{u}_2 \mathbf{u}_2^H \mathbf{h}_2}{\lambda_2}. \end{aligned}$$

The first addend is with the eigenvectors of lemma 1

$$\frac{\hat{\mathbf{h}}_1^H \mathbf{u}_1 \mathbf{u}_1^H \mathbf{h}_2}{\lambda_1} = \frac{\hat{\mathbf{h}}_1^H (\mathbf{h}_1 + \kappa\mathbf{h}_2) (\mathbf{h}_1 + \kappa\mathbf{h}_2)^H \mathbf{h}_2}{\sigma_s^2 (\mathbf{h}_1 + \kappa\mathbf{h}_2)^H (\mathbf{h}_1 + \kappa\mathbf{h}_2) (|\mathbf{h}|^2 + |\mathbf{h}_1^H \mathbf{h}_2|)}.$$

It is

$$(\mathbf{h}_1 + \kappa\mathbf{h}_2)^H \mathbf{h}_2 = \mathbf{h}_1^H \mathbf{h}_2 + \kappa^* |\mathbf{h}|^2 = \kappa^* |\mathbf{h}_1^H \mathbf{h}_2| + \kappa^* |\mathbf{h}|^2.$$

The last terms are scalar and can be cancelled

$$\frac{\hat{\mathbf{h}}_1^H \mathbf{u}_1 \mathbf{u}_1^H \mathbf{h}_2}{\lambda_1} = \frac{\kappa^* \hat{\mathbf{h}}_1^H (\mathbf{h}_1 + \kappa\mathbf{h}_2)}{\sigma_s^2 (\mathbf{h}_1 + \kappa\mathbf{h}_2)^H (\mathbf{h}_1 + \kappa\mathbf{h}_2)}.$$

Additionally, the following holds:

$$\begin{aligned} (\mathbf{h}_1 + \kappa\mathbf{h}_2)^H (\mathbf{h}_1 + \kappa\mathbf{h}_2) &= |\mathbf{h}|^2 + \kappa^* \mathbf{h}_2^H \mathbf{h}_1 + \kappa \mathbf{h}_1^H \mathbf{h}_2 + |\mathbf{h}|^2 \\ &= 2 \cdot |\mathbf{h}|^2 + 2 \cdot |\mathbf{h}_1^H \mathbf{h}_2|. \end{aligned}$$

Therefore,

$$\frac{\hat{\mathbf{h}}_1^H \mathbf{u}_1 \mathbf{u}_1^H \mathbf{h}_2}{\lambda_1} = \frac{\kappa^* \hat{\mathbf{h}}_1^H \mathbf{h}_1 + \hat{\mathbf{h}}_1^H \mathbf{h}_2}{2\sigma_s^2 (|\mathbf{h}|^2 + |\mathbf{h}_1^H \mathbf{h}_2|)}.$$

The same is applicable to the second addend

$$\begin{aligned} \frac{\hat{\mathbf{h}}_1^H \mathbf{u}_2 \mathbf{u}_2^H \mathbf{h}_2}{\lambda_2} &= -\frac{\kappa^* \hat{\mathbf{h}}_1^H \mathbf{h}_1 - \hat{\mathbf{h}}_1^H \mathbf{h}_2}{2\sigma_s^2 (|\mathbf{h}|^2 - |\mathbf{h}_1^H \mathbf{h}_2|)} \\ \hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2 &= \frac{\kappa^* \hat{\mathbf{h}}_1^H \mathbf{h}_1 + \hat{\mathbf{h}}_1^H \mathbf{h}_2}{2\sigma_s^2 (|\mathbf{h}|^2 + |\mathbf{h}_1^H \mathbf{h}_2|)} \\ &\quad - \frac{\kappa^* \hat{\mathbf{h}}_1^H \mathbf{h}_1 - \hat{\mathbf{h}}_1^H \mathbf{h}_2}{2\sigma_s^2 (|\mathbf{h}|^2 - |\mathbf{h}_1^H \mathbf{h}_2|)}. \end{aligned}$$

Using the geometrical meaning of the scalar product p_2 can be simplified significantly

$$\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2 = \frac{\cos \mu + \cos(\alpha - \mu)}{2\sigma_s^2 \cdot (1 + \cos \alpha)} - \frac{\cos \mu - \cos(\alpha - \mu)}{2\sigma_s^2 \cdot (1 - \cos \alpha)}.$$

Further simplification can be made using the sine addition formula

$$\begin{aligned} \hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \mathbf{h}_2 &= \frac{\cos \mu + \cos(\alpha - \mu)}{2\sigma_s^2 \cdot (1 + \cos \alpha)} + \frac{\cos \mu - \cos(\alpha - \mu)}{2\sigma_s^2 \cdot (1 - \cos \alpha)} \\ &= \frac{\cos \mu + \cos \alpha \cos \mu + \sin \alpha \sin \mu}{2\sigma_s^2 \cdot (1 + \cos \alpha)} \\ &\quad - \frac{\cos \mu - \cos \alpha \cos \mu - \sin \alpha \sin \mu}{2\sigma_s^2 \cdot (1 - \cos \alpha)} \\ &= \frac{\cos \mu(1 + \cos \alpha)}{2\sigma_s^2 \cdot (1 + \cos \alpha)} - \frac{\sin \alpha \sin \mu}{2\sigma_s^2 \cdot (1 + \cos \alpha)} \\ &\quad - \frac{\cos \mu(1 - \cos \alpha)}{2\sigma_s^2 \cdot (1 - \cos \alpha)} + \frac{\sin \alpha \sin \mu}{2\sigma_s^2 \cdot (1 - \cos \alpha)} \\ &= \frac{\sin \alpha \sin \mu}{2\sigma_s^2} \left(\frac{1}{1 + \cos \alpha} + \frac{1}{1 - \cos \alpha} \right) \\ &= \frac{\sin \alpha \sin \mu}{2\sigma_s^2} \left(\frac{1 - \cos \alpha + (1 + \cos \alpha)}{(1 + \cos \alpha)(1 - \cos \alpha)} \right) \\ &= \frac{\sin \alpha \sin \mu}{\sigma_s^2 (1 - \cos^2 \alpha)} = \frac{\sin \alpha \sin \mu}{\sigma_s^2 \sin^2 \alpha} \\ &= \frac{\sin \mu}{\sigma_s^2 \sin \alpha}. \end{aligned}$$

Similarly, for the nominator of p_1

$$\begin{aligned}
& \hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1 \\
&= \frac{\hat{\mathbf{h}}_1^H \mathbf{h}_1 + \hat{\mathbf{h}}_1^H \mathbf{h}_2}{2\sigma_s^2 \cdot (\mathbf{h}_1^H \mathbf{h}_1 + \mathbf{h}_1^H \mathbf{h}_2)} \\
&+ \frac{\hat{\mathbf{h}}_1^H \mathbf{h}_1 - \hat{\mathbf{h}}_1^H \mathbf{h}_2}{2\sigma_s^2 \cdot (\mathbf{h}_1^H \mathbf{h}_1 - \mathbf{h}_1^H \mathbf{h}_2)} \\
&= \frac{\cos \mu + \cos \alpha \cos \mu + \sin \alpha \sin \mu}{2\sigma_s^2 \cdot (1 + \cos \alpha)} \\
&+ \frac{\cos \mu - \cos \alpha \cos \mu - \sin \alpha \sin \mu}{2\sigma_s^2 \cdot (1 - \cos \alpha)} \\
&= \frac{\cos \mu}{\sigma_s^2} + \frac{\sin \alpha \sin \mu}{2\sigma_s^2} \left(\frac{1}{1 + \cos \alpha} - \frac{1}{1 - \cos \alpha} \right) \\
&= \frac{\cos \mu}{\sigma_s^2} + \frac{\sin \alpha \sin \mu}{2\sigma_s^2} \left(\frac{1 - \cos \alpha - (1 + \cos \alpha)}{(1 + \cos \alpha)(1 - \cos \alpha)} \right) \\
&= \frac{\cos \mu}{\sigma_s^2} - \frac{\sin \alpha \sin \mu \cos \alpha}{\sigma_s^2 (1 - \cos \alpha^2)} \\
&= \frac{\cos \mu - \cot \alpha \sin \mu}{\sigma_s^2}.
\end{aligned}$$

Finally, for the denominator

$$\begin{aligned}
\hat{\mathbf{h}}_1^H \mathbf{R}_x^{-1} \hat{\mathbf{h}}_1 &= \frac{1}{\sigma_s^2} \left(\frac{\cos \mu + \cos(\alpha - \mu)}{2 \cdot (1 + \cos \alpha)} \right)^2 \\
&+ \frac{1}{\sigma_s^2} \left(\frac{\cos \mu - \cos(\alpha - \mu)}{2 \cdot (1 - \cos \alpha)} \right)^2.
\end{aligned}$$

q.e.d.

ACKNOWLEDGMENT

The authors would like to thank S. Winter for revising the algorithms, Dr. H. Sawada and R. Mukai for daily cooperation and valuable discussions, and Dr. D. Filbert and Dr. S. Katagiri for their continuous encouragement.

REFERENCES

- [1] O. L. Frost, "An algorithm for linearly constrained adaptive antenna array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.
- [2] L. J. Griffith and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propag.*, vol. AP-30, no. 1, pp. 27–34, Jan. 1982.
- [3] K. M. Buckley and L. J. Griffith, "An adaptive generalized sidelobe canceller with derivative constraints," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 311–319, Mar. 1986.
- [4] J. Bitzer, K. U. Simmer, and K.-D. Kammeyer, "Theoretical noise reduction limits of the generalized sidelobe canceller (GSC) for speech signals," in *Proc. ICASSP*, 1999, vol. 5, pp. 2965–2968.
- [5] I. Thng, A. Cantoni, and Y. H. Leung, "Constraints for maximally flat optimum broadband antenna arrays," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1334–1347, Jun. 1995.
- [6] Z. Shutao and I. Thng, "Pre-steering derivative constraint for robust broadband antenna arrays," in *Proc. ICASSP*, Salt Lake City, UT, May 2001, pp. 2913–2916.
- [7] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.

- [8] K. Torkkola, "Blind separation of delayed and convolved mixtures," in *Unsupervised Adaptive Filtering*, S. Haykin, Ed. New York: Wiley, 2000, vol. 1.
- [9] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomput.*, vol. 22, pp. 21–34, 1998.
- [10] C. Mejuto and J. Principe, "A second-order method for blind source separation of convolutive mixtures," in *Proc. Int. Workshop of ICA*, Aussois, France, Jan. 1999, pp. 395–400.
- [11] L. Parra and C. Spence, "Convolutive blind separation of nonstationary sources," *IEEE Trans. Speech Audio Process.*, vol. 8, pp. 320–327, 2000.
- [12] M. Z. Ikram and D. R. Morgan, "Exploring permutation inconsistency in blind separation of speech signals in a reverberant environment," in *Proc. ICASSP*, Jun. 2000, pp. 1041–1044.
- [13] S. Araki, S. Makino, Y. Hinamoto, R. Mukai, T. Nishikawa, and H. Saruwatari, "Equivalence between frequency domain blind source separation and frequency domain adaptive beamforming for convolutive mixtures," *EURASIP J. Appl. Signal Process.*, no. 11, pp. 1157–1166, 2003.
- [14] H. Saruwatari, S. Kurita, and K. Takeda, "Blind source separation combining frequency-domain ICA and beamforming," in *Proc. ICASSP*, Salt Lake City, UT, May 2001, pp. 2733–2736.
- [15] D. Johnson and D. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [16] M. Knaak and D. Filbert, "Acoustical semi-blind source separation for machine monitoring," in *Proc. 3rd. Int. Conf. ICA*, San Diego, CA, Dec. 2001, pp. 361–366.
- [17] L. Parra and C. Alvino, "Geometric source separation: Merging convolutive source separation with geometric beamforming," in *Proc. IEEE Workshop on NNSP*, 2001, pp. 273–282.
- [18] M. Knaak, S. Araki, and S. Makino, "Geometrically constrained ICA for convolutive mixtures of sound," in *Proc. ICASSP*, Hong Kong, Apr. 2003, pp. 725–728.
- [19] F. Asano, Y. Motomura, H. Asoh, and T. Matsui, "Effect of PCA filter in blind source separation," in *Proc. 2nd. Int. Workshop of ICA*, Helsinki, Finland, Jul. 2000, pp. 57–62.
- [20] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Polar coordinate based nonlinear function for frequency domain blind source separation," in *Proc. ICASSP*, Orlando, FL, May 2002, pp. 1001–1004.
- [21] S. Ikeda and N. Murata, "A method of ICA in time-frequency domain," in *Int. Workshop ICA*, Aussois, France, Jan. 1999, pp. 365–370.
- [22] E. Bingham and A. Hyvärinen, "A fast fixed point algorithm for independent component analysis of complex valued signals," *Int. J. Neural Networks*, vol. 10, no. 1, pp. 1–8, 2000.
- [23] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.
- [24] R. Mukai, S. Araki, and S. Makino, "Separation and dereverberation performance of frequency domain blind source separation," in *Proc. 3rd. Int. Conf. ICA*, San Diego, CA, Dec. 2001, pp. 230–235.
- [25] M. Knaak, M. Kunter, and D. Filbert, "Blind source separation for acoustical machine diagnosis," in *Proc. 4th IEEE Int. Conf. Digital Signal Processing*, Santorini, Greece, Jun. 2002, pp. 401–404.
- [26] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust approach to the permutation problem of frequency-domain blind source separation," in *Proc. ICASSP*, Hong Kong, 2003, pp. 381–384.



Mirko Knaak received the M.E. (Dipl.-Ing.) and Ph.D. degrees in electrical engineering from the Technical University of Berlin, Berlin, Germany, in 1997 and 2003, respectively.

His research interests include blind source separation and beamforming. During the time of the research for this article, he was with Measurement Technology Lab, University of Technology Berlin, and NTT Communication Science Laboratories, NTT Corporation, Kyoto. He is now with IAV GmbH, Berlin.



Shoko Araki (M'01) received the B.E. and the M.E. degrees in mathematical engineering and information physics from the University of Tokyo, Tokyo, Japan, in 1998 and 2000, respectively.

She joined NTT Communication Science Laboratories, Kyoto, Japan, in 2000. She is now with the NTT Communication Science Laboratories. Her research interests include array signal processing, blind source separation applied to speech signals.

Ms. Araki received the TELECOM System Technology Award from the Telecommunications Advancement Foundation in 2004, the Best Paper Award of the IWAENC in 2003, and the 19th Awaya Prize from Acoustical Society of Japan (ASJ) in 2001. She is a member of the Institute of Electronics, Information, and Communication Engineers of Japan, and a member of the Acoustical Society of Japan.

200 articles in journals and conference proceedings and has been responsible for more than 150 patents.

Dr. Makino received the TELECOM System Technology Award of the TAF in 2004, the Best Paper Award of the IWAENC in 2003, the Paper Award of the IEICE in 2005 and 2002, the Paper Award of the ASJ in 2005 and 2002, the Achievement Award of the IEICE in 1997, and the Outstanding Technological Development Award of the ASJ in 1995. He is a member of both the Awards Board and the Conference Board of the IEEE SP Society. He is an Associate Editor of the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING and an Associate Editor of the *EURASIP Journal on Applied Signal Processing*. He is a member of the Technical Committee on Audio and Electroacoustics of the IEEE SP Society as well as the Technical Committee on Blind Signal Processing of the IEEE Circuits and Systems Society. He is also a member of the International ICA Steering Committee and the Organizing Chair of the ICA2003 in Nara. He is the General Chair of the WASPAA2007 in Mohonk and the General Chair of the IWAENC2003 in Kyoto. He is the Chair of the Technical Committee on Engineering Acoustics of the IEICE and the ASJ. He is a council member of the ASJ, and a member of the EURASIP, and a member of the IEICE.



Shoji Makino (A'89–M'90–SM'99–F'04) received the B.E., M.E., and Ph.D. degrees from Tohoku University, Japan, in 1979, 1981, and 1993, respectively.

He joined NTT Corporation, Kyoto, Japan, in 1981. He is now an Executive Manager at the NTT Communication Science Laboratories. He is also a Guest Professor at the Hokkaido University. His research interests include adaptive filtering technologies and realization of acoustic echo cancellation, blind source separation of convolutive mixtures of speech. He is the author or co-author of more than