Overdetermined Blind Source Extraction exploiting a Generalized Sidelobe Canceller structure

B.B.A.J. Bloemendal  
Eindhoven University of Technology  
Department of Electrical Engineering  
Eindhoven, The Netherlands  
b.bloemendal@tue.nl

J. van de Laar  
Philips Research Laboratories  
Digital Signal Processing Group  
Eindhoven, The Netherlands  
jakob.van.de.laar@philips.com

P.C.W. Sommen  
Eindhoven University of Technology  
Department of Electrical Engineering  
Eindhoven, The Netherlands  
p.c.w.sommen@tue.nl

Abstract—In many acoustic applications, the extraction of only one desired source signal from a mixture of signals is required. We propose a novel method to perform this extraction in the overdetermined case where more sensors than sources are available. We apply blind signal processing techniques in a structure that is similar to a Generalized Sidelobe Canceller (GSC). In a GSC the upper branch reproduces the desired source signal with a beamformer while the lower branch reduces the contribution of both noise and interferers. In this work the upper branch is used to extract the desired source from a noisy mixture of all sources and the lower branch is solely used to perform noise reduction. Therefore, the underlying optimization problem is essentially different from the conventional GSC. We validate our method by means of a computer simulation.

I. INTRODUCTION

In many practical applications, the extraction of only one desired source signal from a mixture of signals is required, e.g., in a teleconferencing setup. In such an extraction problem a number of sensors measure mixtures of several source signals, typically corrupted by noise. One specific source signal has to be extracted from the observed measurements. This extraction should be performed such that a high Signal to Noise Ratio (SNR) and, especially for audio applications, a high intelligibility of the extracted signal is obtained.

A widely used solution to the extraction problem in acoustic environments is a beamformer [1]. The development of a beamformer starts with the design of a sensor array. The sensor locations of this array have to be known exactly such that the differences between mutual sensor signals can be modeled. Based on this model and knowledge about the Direction Of Arrival (DOA) of the desired source, a beamformer is designed. The beamformer filters the sensor signals such that the desired signal is summed coherently. An extension to beamforming is a Generalized Sidelobe Canceller (GSC) structure [1], which was introduced simultaneously in [2] and [3]. A GSC uses a beamformer in the upper branch to extract the desired source signal and uses an (adaptive) filter in the lower branch to reduce the contribution of both noise and interferers to the beamformer output. A disadvantage of beamforming is that the locations of all sensors have to be known exactly and that errors in sensor positioning lead to performance degradation; furthermore, the sensor arrays have to be designed, built, and placed in order to use the beamforming techniques, which makes these systems expensive.

In this paper we use Blind Signal Processing (BSP) techniques to perform the signal extraction, which is much less dependent on sensor positioning. A lot of research has been performed on Blind Source Separation (BSS) of all sources, Blind Identification (BI) of mixing systems, and Blind Source Extraction (BSE) of potentially interesting signals based on properties such as non-whiteness, smoothness, and linear predictability [4]–[7]. However, little research has been performed on directly extracting the desired signal by exploiting some additional knowledge about the desired source that would otherwise be used by a classifier following the BSE algorithm [8].

Our approach is schematically depicted in Figure 1. We observe \( D \) noisy sensor signals \( x \), which are assumed to be an overdetermined mixture \( A \) of \( S \) source signals \( s \) contaminated by \( D \) additive spatially uncorrelated noise signals \( \nu \). The objective is to present a novel concept to extract a desired source in acoustic environments; therefore, we work with an instantaneous mixing model \( A \) to obtain insight and to extend it later towards the more practical convolutive mixing model.

First, the observed sensor signals are separated by two orthogonal transformation matrices \( P_s \) and \( P_\nu \) into an upper branch of \( S \) signals and a lower branch of \( D - S \) signals, respectively. The upper branch signals contain both source and noise signals, while the lower branch signals contain only noise. Second, the signals in the upper branch are filtered by an extraction filter \( w \) to extract the desired source signal, while the undesired source signals are suppressed. Simultaneously, the signals in the lower branch are filtered by \( a \), possibly adaptive, filter \( w_\nu \), and subtracted from the extracted signal to reduce the noise contribution.

The transformations \( P_s \) and \( P_\nu \) are identified by applying subspace techniques to correlation matrices \( C_f \) that are estimated from the sensor signals. These correlation matrices are estimated from an a priori available Noise-Free Region Of Support (NF-ROS) [8]–[10]. Subsequently, linear combinations of the correlation matrices are taken based on a mold that contains a priori information about the DOA of the desired source. This DOA information is a priori knowledge that is also used in the design of a beamformer; however,
we show that our DOA information is not required to be exact. We compose a matrix $M$ with a specific eigenstructure from the linear combinations of correlation matrices. The eigenvector that corresponds to the smallest eigenvalue is the extraction filter $w$ that extracts the source that has a DOA closest to the DOA used in the mold. Finally, the filter $w$ can be optimized to reduce the noise contribution to the extracted signal by minimizing the output power.

II. METHODS

First we introduce our assumptions on the Second Order Statistics (SOS) of the signals and subsequently we depict the structure of sensor correlation matrices. We apply subspace techniques to these correlation matrices in order to identify the matrices $P_s$ and $P_v$. Finally, we identify the extraction and noise reduction filter.

A. Assumptions on Second Order Statistics

Each sensor signal $x_i(t)$ is sampled at discrete time instances $t = nT_s$ where $n \in \mathbb{Z}$ and $T_s$ is the sampling time such that no aliasing occurs. This sampling gives us the following $D$ sensor signals $x_1[n], \ldots, x_D[n]$, which are stacked in the sensor signal column vector $x[n]$. Similarly, the $S$ source and $N$ noise signals are stacked in the column vectors $s[n]$ and $n[n]$, respectively. As discussed in Section I, the mixing system is modeled by a $D \times S$ matrix $\mathbf{A} = [\mathbf{a}^1, \ldots, \mathbf{a}^S]$, where $\mathbf{a}^1, \ldots, \mathbf{a}^S$ are mixing column vectors. The relation between the source, noise, and sensor signal vectors is then given by $x[n] = \mathbf{A}s[n] + \nu[n]$.

**Definition II.1.** The correlation function between a signal pair $(p_{i_1}[n], q_{i_2}[n-k])$ for all available $i_1, i_2$ at time $n \in \mathbb{Z}$ and with lag $k \in \mathbb{Z}$ is defined as follows:

$$r_{p_{i_1}q_{i_2}}[n, k] = \mathbb{E}\{p_{i_1}[n]q_{i_2}[n-k]\}$$

where $\mathbb{E}\{}$ is the mathematical expectation operator, which we approximate by averaging over a certain period of time. If both signals use the same variable, then we use a shorthand notation $r_{p_{i_1}q_{i_2}}[n, k] \equiv r_{p_{i_1}p_{i_2}}[n, k]$.

In Definition II.1 the time index $n$ is incorporated. If the statistics of the signals change over time, then a new measurement of the statistics is obtained that can be used explicitly by the presented algorithm. This property creates flexibility by allowing for different types of signals such as stationary non-white signals and non-stationary signals.

By replacing the signal pair $(p_{i_1}, q_{i_2})$ in Definition II.1 by the sensor, source, and noise signal pairs we obtain the following correlation functions:

$$r_{p_{i_1}q_{i_2}}[n, k] \quad \forall 1 \leq i_1, i_2 \leq D$$

$$r_{s_{i_1}q_{i_2}}[n, k] \quad \forall 1 \leq i_1, i_2 \leq S$$

$$r_{n_{i_1}q_{i_2}}[n, k] \quad \forall 1 \leq i_1, i_2 \leq N$$

Finally, we consider the crosscorrelation functions $r_{p_{i_1}s_{i_2}}[n, k]$ and $r_{p_{i_1}n_{i_2}}[n, k]$ that correspond to the source and noise signal pairs $(p_{i_1}, s_{i_2})$ for $1 \leq i_1 \leq S$ and $1 \leq i_2 \leq D$.

Our assumptions on the SOS of the source and noise signals are introduced by defining a Noise-Free Region Of Support (NF-ROS).

**Definition II.2.** The Noise-Free Region of Support (NF-ROS), also denoted by $\Omega$, consists of a set of time-lag pairs $(n, k)$ for which the following assumptions hold:

$$\forall (n,k) \in \Omega : \left\{ \begin{array}{l} r_{p_{i_1}p_{i_2}}[n,k] = 0 \forall 1 \leq i_1 \neq i_2 \leq S \\ r_{p_{i_1}s_{i_2}}[n,k] = 0 \forall 1 \leq i_1 \leq i_2 \leq S \\ r_{p_{i_1}n_{i_2}}[n,k] = 0 \forall 1 \leq i_1 \leq i_2 \leq N \\ r_{s_{i_1}s_{i_2}}[n,k] = 0 \forall 1 \leq i_1 \leq i_2 \leq D \\ r_{n_{i_1}n_{i_2}}[n,k] = 0 \forall 1 \leq i_1 \leq i_2 \leq S \\ r_{s_{i_1}n_{i_2}}[n,k] = 0 \forall 1 \leq i_1 \leq i_2 \leq D \end{array} \right.\right.$$
Applying both subspace matrices \( P_s \) and \( P_v \) to the sensor signals \( x \), as depicted in Figure 1, leads to a separation of \( S \) noisy mixtures of source signals in the upper branch \( P_s x = P_s A_s x + P_s \nu \) and \( D - S \) noise signals in the lower branch \( P_v x = P_v \nu \).

**C. Upper branch: performing BSE**

The identification of extraction filters based on the generalized eigenvalue decomposition of sensor correlation matrices \( C^s_i \) was discussed in [10]. There the number of sensors was assumed equal to the number of sources and the extraction filters could be identified randomly. In the current paper we characterize the generalized eigenvalues based on some additional a priori information to identify directly the desired extraction filter. The procedure is as follows. First we transform the sensor correlation matrices \( C^s_i \) to \( S \times S \) full rank correlation matrices. Second, we introduce the additional a priori information and its conditions. Third, we create \( S \) linear combinations of the \( D \) transformed correlation matrices. Finally, we combine the \( S \) linear combinations in a specific way such that the eigenvector that corresponds to the smallest eigenvalue of the constructed matrix is the desired extraction filter.

The matrix \( P_s \) is used to reduce from \( D \) sensor signals to \( S \) sensor signals that still contain all source signals. Similarly, the sensor correlation matrices can be transformed, without losing information, from size \( D \times N \) to size \( S \times N \) by pre-multiplying the correlation matrices by \( P_s \). Furthermore, the matrix \( V_s \) is orthogonal to the sensor correlation matrix, i.e., \( C^s_i (V_s)^T = 0 \), and the matrix \( V_s \) is orthogonal to \( V_v \). Therefore, right multiplying the sensor correlation matrix with the matrix \( Q^s = (V_s)^T \) leads to a lossless transformation from \( N \) to \( S \) columns in the sensor correlation matrices. Applying both these transformations \( P_s \) and \( Q^s \) to the sensor correlation matrices transforms them to the following full rank sensor correlation matrices of size \( S \times S \):

\[
\hat{C}^s_i \triangleq P_s C^s_i Q^s \equiv P_s A_i \text{diag}(a_i) C^s_i Q^s \quad \text{for } 1 \leq i \leq D
\]

where \( P_s A_i \) and \( C^s_i Q^s \) are full rank \( S \times S \) matrices. We use these \( D \) correlation matrices combined with additional prior information to identify the extraction filter that extracts the desired source signal from the signals \( P_s x \).

We assume to have an a priori estimation of the mixing column available that belongs to the desired source, which we call a mold. We represent this mold by \( a^i \) and we assume that the following condition holds:

\[
\left| \frac{a^{i}}{a^{d}} \right| > \left| \frac{a^{i}}{a^{i}} \right| \quad \forall 1 \leq i \neq d \leq S
\]

This assumption implies that the mold has the smallest angle with respect to the mixing column vector that corresponds to the desired source. Additionally, we assume that the mold is a vector from the signal space, i.e., \( P_s a^0 = 0 \); otherwise, we could use a projection method to project the mold onto the signal space.

Linear combinations of the correlation matrices \( \hat{C}^s_i \) are as follows:

\[
\Gamma_i \triangleq \sum_{i=1}^{D} \xi_i \hat{C}^s_i \equiv P_s A_i \text{diag}(\alpha_i, \cdots, \alpha_i^S) C^s_i Q^s
\]

where \( \xi_i \equiv \{[\xi_1, \cdots, \xi_i]\}^T \) and \( \alpha_i \equiv \langle \xi_i, a^i \rangle \). Here \( \langle \cdot, \cdot \rangle \) is the Euclidean inner product. We choose a set of \( S \) vectors \( \xi_1, \cdots, \xi_S \) based on the mold and the signal subspace matrix \( U^s \). Observe that when a vector \( \xi_i \) is chosen from the noise space \( U^s \), i.e., \( \xi_i = U^s z \) with \( z \in \mathbb{R}^{D-S} \), then the inner product \( \alpha_i = \langle \xi_i, a^i \rangle = \langle U^s z, a^i \rangle \) equals zero for all available mixing columns \( a^i \) and \( \Gamma_i \) becomes zero. Therefore, the vectors \( \xi_1, \cdots, \xi_S \) should be chosen as linear combinations of the vectors in the matrix \( U^s \). We choose the vectors as follows. The vector \( \xi_1 = a^i / ||a^i|| \) and \( \xi_2, \cdots, \xi_S \) are orthonormal to \( \xi_1 \), to the vectors in the matrix \( U^s \), and to each other. A way to compute these \( S-1 \) vectors is by taking the first \( S-1 \) left singular vectors of the following product: \((I - \xi_1^T (\xi_1)^T) U^s \), which is a projection of the matrix \( U^s \) onto the space orthogonal to \( \xi_1 \).

Based on the vectors \( \xi_1, \cdots, \xi_S \) we obtain the linear combinations of sensor correlation matrices \( \Gamma_1, \cdots, \Gamma_S \). We combine these matrices into a single matrix \( M \) from which the eigenvector that corresponds to the smallest eigenvalue is the desired extraction filter.

**Theorem II.1.** The desired extraction filter is the left eigenvector that corresponds to the smallest eigenvalue of the following matrix:

\[
M \triangleq \sum_{i=2}^{S} (\Gamma_i (\Gamma_i)^{-1})^2
\]

if \( \Gamma_1 \) is invertible. The extraction filter \( w \) is found as the solution of \( w (M - \lambda I) = 0 \) for the smallest eigenvalue \( \lambda \).

For this smallest eigenvalue problem are efficient algorithms available such as the power method [11].

**Proof:** We use a short hand notation to denote the structure of a linear combination of correlation matrices, which is given by \( \Gamma_i = B \text{diag}(\alpha_i, \cdots, \alpha_i^S) C \), where \( B = P_s A_i, C = C^s_i Q^s \), and \( \alpha_i = \langle \xi_i, a_i \rangle \). The structure in the product \( \Gamma_i (\Gamma_i)^{-1} \) is given as

\[
\Gamma_i (\Gamma_i)^{-1} \equiv B \text{diag} \left( \frac{\alpha_1}{\alpha_1^2}, \cdots, \frac{\alpha_S}{\alpha_S^2} \right) B^{-1}
\]

which is equal to an eigenvalue decomposition with an eigenvector matrix \( B \) and eigenvalues \( \alpha_i / \alpha_1^2 \). Squaring this product leads to squared eigenvalues \( \alpha_i^2 / \alpha_1^2 \) and the same eigenvector matrix \( B \). Summation of \( S-1 \) of these squared matrix products as in (8) gives

\[
M \equiv B \left( \text{diag} \left( \sum_{i=2}^{S} (\alpha_i^2 / \alpha_1^2), \cdots, \sum_{i=2}^{S} (\alpha_i^2 / \alpha_S^2) \right) \right) (B^{-1})
\]

Notice that \( \alpha_i^2 = \sum_{i=1}^{S} \alpha_i^2 \) is a decomposition of a mixing column vector into \( S \) orthonormal components. Furthermore, \( ||a^i|| = \sum_{i=1}^{S} ||\xi_i, a^i|| = \sum_{i=1}^{S} (\alpha_i^2) \). This means that \( \sum_{i=2}^{S} (\alpha_i^2) = ||a^i||^2 - (\alpha_i^2) \), thus

\[
M = B \left( \text{diag} \left( ||a^i||^2 - (\alpha_i^2), \cdots, ||a^i||^2 - (\alpha_i^2) \right) \right) (B^{-1})
\]

Because the mixing column \( a^i \) appears both in the nominator and denominator of the eigenvalues, the eigenvalues are independent of the length or norm of the mixing vectors. From (6) it follows that \( ||a^i|| = |\langle \xi_i, a^i \rangle| \) has the largest value for \( i = d \) if we assume normalized mixing column vectors; furthermore, \( ||a^i|| \leq ||a^0|| \). Therefore, \( ||a^i||^2 - (\alpha_i^2) \) is smallest and \( (\alpha_i^2) \) is largest for \( i = d \). This implies that the smallest eigenvalue of \( M \) must correspond to the desired source if (6) holds. The corresponding left eigenvector of \( M \) is the \( d \)’th row from the inverse of \( B \), which extracts the desired source from the mixture \( P_v x \). ■

**D. Lower branch: noise reduction**

The noise reduction step in a GSC structure is performed by a noise reduction filter \( w_n \), which is often an adaptive filter. One of the main concerns of such noise reduction filter is that no leakage of the desired source signal is allowed in the lower branch. This will be
the case when the subspace matrices are not identified correctly. The objective of the noise reduction filter $w_a$ is to minimize the output power over its components, which is a least squares optimization problem that is well known in the literature [1], [5]. The optimal noise reduction filter from this optimization problem is given by

$$w_a = wP_aR^T (P_a)^T \left( P_a R^T (P_a)^T \right)^{-1}$$

where $R^T \triangleq \mathbb{E} \{ x[n] (x[n])^T \}$. All components of this optimal filter are identified or estimated in our algorithm except for the sensor autocorrelation matrix $R^s$. Methods to find this optimal filter $w_a$ are based on adaptive filtering or batch processing are available [1], [5].

### III. SIMULATION RESULTS AND DISCUSSION

The objective of our work is to present the concept of using BSP techniques in order to extract one desired source. Therefore, we validate our method by simulating a mixing scenario. This simulation makes it possible to compare the results of the our method with the optimal solution of the underlying optimization problem.

In our simulations, we used a BSE scenario where five sensors measured 50000 samples of noisy mixtures of three stationary sources. The sources had unit variance and an autoregressive temporal structure of order one with a pole at $z = 0.5$, $z = 0.9$, and $z = -0.2$, respectively. The sensor noise was spatially and temporally white. The noise variances were chosen randomly and equal to $[0.081, 0.034, 0.085, 0.036, 0.264]$ for sensor one until five. All components were correlated Gaussian distribution. The mixing system and the mold were as follows:

$$A = \begin{bmatrix} -0.0773 & 0.2251 & -0.2362 \\ 0.0666 & 0.5218 & 0.7880 \\ 0.7518 & -0.5419 & -0.3498 \\ 0.3072 & -0.2026 & 0.2617 \\ 0.5745 & 0.5851 & 0.3638 \end{bmatrix} \quad a^0 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

The objective was to extract the first source, i.e., the source with a pole at $z = 0.9$, from the mixtures. Based on the properties of stationary signals the NF-ROS was chosen as the lags $k \in \{1, \ldots, 7\}$ for any time instance $n \in \mathbb{Z}$. We used these assumptions in the following algorithm:

1. Estimate $D$ sensor correlation matrices $C_i^s$ in the NF-ROS.
2. Perform a SVD of one of the correlation matrices to identify $P_a$, $P_v$, and $Q_v$.
3. Reduce the sensor correlation matrices sizes $C_i^a = P_a C_i^s Q_v^T$.
4. Create $S$ linear combinations of the correlation matrices $C_i^s$, based on the mold $a^0$ and signal subspace $P_a$.
5. Construct the matrix $M$ as in (8).
6. Calculate the left eigenvector of the smallest eigenvalue of $M$, which is the desired extraction filter $w$.
7. Use the LMS algorithm with the update step size $\mu = 5 \cdot 10^{-4}$ to minimize the output power over the components of $w_a$.
8. Perform the extraction as follows: $y = (wP_a - w_a P_v)x$.

First, the matrices $P_a$, $P_v$, and filter $w$ were calculated using the batch algorithm. Second, the LMS algorithm was used to minimize the output power by adapting the components of $w_a$. The simulation results are depicted in Figure 2. We observe that our method converges towards the optimal solution that can be obtained for the underlying optimization problem. Simultaneously, the constraints are respected, which is observed from the power contribution of the desired source and the interferences. This means that there is no leakage of the desired source into the lower branch; otherwise, the desired source would have been suppressed.

### IV. CONCLUSIONS AND FUTURE RESEARCH

We introduced a novel approach to use BSP techniques in order to extract one desired source from an overdetermined mixture of sources contaminated by additive noise. The algorithm suppresses the interferers and reduces the noise contribution with additional degrees of freedom in a structure that is similar to a GSC. From the field of BSP we know that some additional a priori knowledge has to be available in order to select the desired source; therefore, we used a mold that contains an a priori estimation of the mixing of this desired source. By means of computer simulations we validated our method and we showed that the underlying optimization problem is solved.

Future research topics are as follows. The robustness and efficiency of the proposed method should be analyzed and possibly improved. Furthermore, the use of different types of a priori information should be investigated. Finally, the approach should be applied on a convolutive mixing model, which is much more applicable in acoustic environments.

### REFERENCES