Research Article

Exploiting Narrowband Efficiency for Broadband Convolutive Blind Source Separation

Robert Aichner, Herbert Buchner, and Walter Kellermann

Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Cauerstraße 7, 91058 Erlangen, Germany

Based on a recently presented generic broadband blind source separation (BSS) algorithm for convolutive mixtures, we propose in this paper a novel algorithm combining advantages of broadband algorithms with the computational efficiency of narrowband techniques. By selective application of the Szegö theorem which relates properties of Toeplitz and circulant matrices, a new normalization is derived as a special case of the generic broadband algorithm. This results in a computationally efficient and fast converging algorithm without introducing typical narrowband problems such as the internal permutation problem or circularity effects. Moreover, a novel regularization method for the generic broadband algorithm is proposed and subsequently also derived for the proposed algorithm. Experimental results in realistic acoustic environments show improved performance of the novel algorithm compared to previous approximations.

1. INTRODUCTION

Blind source separation (BSS) refers to the problem of recovering signals from several observed linear mixtures [1]. In this paper we deal with the convolutive mixing case as encountered, for example, in acoustic environments. Therefore, we are interested in finding a corresponding demixing system, where the output signals $y_q(n)$, $q = 1, \ldots, P$, are described by

$$y_q(n) = \sum_{p=1}^{P} \sum_{\kappa=0}^{L-1} w_{pq,\kappa} x_p(n - \kappa), \qquad (1)$$

and where $w_{pq,\kappa}$, $\kappa = 0, \ldots, L-1$, denote the current weights of the MIMO filter taps from the pth sensor channel $x_p(n)$ to the qth output channel. In this paper the number of *active source* signals Q is less than or equal to the number of microphones P. BSS algorithms are solely based on the fundamental assumption of mutual statistical independence of the different source signals. The separation is achieved by forcing the output signals y_q to be mutually statistically decoupled up to joint moments of a certain order.

In [2] a generic framework called TRINICON (Triple-N ICA for convolutive mixtures) has been introduced for multichannel blind signal processing, such as BSS or

dereverberation based on multichannel blind deconvolution (MCBD). In [3, 4] we have also shown that based on this framework many seemingly different BSS algorithms can be treated in a unified way. Apart from these existing BSS algorithms, also several novel broadband convolutive BSS algorithms for both the time and frequency domains have been derived. In this paper we exemplarily use a second-order BSS algorithm resulting from the broadband time-domain derivation in [3, 4]. This yields an algorithm which possesses an inherent normalization of the coefficient update leading to fast convergence also for colored signals such as speech. However, for realistic acoustic environments large correlation matrices have to be inverted for every output channel. An approximation of this matrix by a diagonal matrix led to a very efficient algorithm which allows real-time implementation using a block-online update structure [5]. In Section 2 the generic broadband algorithm combined with the block-online update is briefly summarized. In Section 3 a novel normalization strategy is presented which is obtained by the application of the Szegö theorem and constitutes a better approximation of the inverse autocorrelation matrix. In general, the Szegö theorem relates the eigenvalues of circulant and Toeplitz matrices which can here be interpreted as the relation between broadband and narrowband signal models. The novel normalization leads to an algorithm where the

main parts of the algorithm are still implemented in a broadband manner and thus avoid the internal permutation problem and circularity effects as experienced in purely narrowband BSS algorithms. Due to the *selective* application of the Szegö theorem only the normalization is implemented using the narrowband approximation which leads to a computationally efficient algorithm as the matrix inverse can be replaced by a scalar inversion in each frequency bin. Another important aspect for robust implementations is the regularization of the possibly ill-conditioned correlation matrices prior to inversion. This issue is discussed in Section 4 and a novel regularization strategy is presented for the generic broadband algorithm. An analogous regularization method is then derived for the proposed algorithm. Finally, experimental results show the improved performance of the new algorithm.

2. GENERIC BROADBAND ALGORITHM

2.1. Cost function and block-online update

A block processing broadband algorithm simultaneously exploiting nonwhiteness and nonstationarity of the source signals is derived from the following matrix formulation [3]. First, we introduce a block output signal matrix

$$
\mathbf{Y}_q(m) = \begin{bmatrix} y_q(mL) & \cdots & y_q(mL - D + 1) \\ y_q(mL + 1) & \ddots & y_q(mL - D + 2) \\ \vdots & \ddots & \vdots \\ y_q(mL + N - 1) & \cdots & y_q(mL - D + N) \end{bmatrix} \quad (2)
$$

and reformulate the convolution (1) as

$$
\mathbf{Y}_q(m) = \sum_{p=1}^{P} \mathbf{X}_p(m)\mathbf{W}_{pq}, \quad (3)
$$

with m being the block time index and N denoting the block length. The $N \times D$ matrix $\mathbf{Y}_q(m)$ incorporates D *time lags* into the correlation matrices in the cost function, as is necessary for the exploitation of the nonwhiteness property. To ensure linear convolutions for all elements of $\mathbf{Y}_q(m)$, the $N \times 2L$ matrices $\mathbf{X}_p(m)$ and $2L \times D$ matrices $\mathbf{W}_{pq}$ are given as

$$
\mathbf{X}_p(m) = \begin{bmatrix} x_p(mL) & \cdots & x_p(mL - 2L + 1) \\ x_p(mL + 1) & \ddots & x_p(mL - 2L + 2) \\ \vdots & \ddots & \vdots \\ x_p(mL + N - 1) & \cdots & x_p(mL - 2L + N) \end{bmatrix}, \quad (4)
$$

$$
\mathbf{W}_{pq}(m) = \begin{bmatrix} w_{pq,0} & 0 & \cdots & 0 \\ w_{pq,1} & w_{pq,0} & \ddots & \vdots \\ \vdots & w_{pq,1} & \ddots & 0 \\ w_{pq,L-1} & \vdots & \ddots & w_{pq,0} \\ 0 & w_{pq,L-1} & \ddots & w_{pq,1} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & w_{pq,L-1} \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (5)
$$

where the matrices $\mathbf{X}_p(m)$, $p = 1, \ldots, P$, in (3) are Toeplitz matrices due to the shift of subsequent rows by one sample each. The matrices $\mathbf{W}_{pq}$ exhibit a Sylvester structure which is a special form of a Toeplitz matrix, where each column is shifted by one sample containing the current weights $\mathbf{w}_{pq} = [w_{pq,0}, w_{pq,1}, \ldots, w_{pq,L-1}]^T$ of the MIMO sub-filter of length L from the pth sensor channel to the qth output channel. Superscript T denotes transposition of a vector or a matrix. It can be seen that for the general case $1 \leq D \leq L$ the last $L - D + 1$ rows are padded with zeros to ensure compatibility with $\mathbf{X}_p$. To allow a convenient notation of the algorithm combining all channels, we write (3) for all channels simultaneously as

$$
\mathbf{Y}(m) = \mathbf{X}(m)\mathbf{W}, \quad (6)
$$

with the matrices

$$
\mathbf{Y}(m) = [\mathbf{Y}_1(m), \ldots, \mathbf{Y}_P(m)],
$$
$$
\mathbf{X}(m) = [\mathbf{X}_1(m), \ldots, \mathbf{X}_P(m)], \quad (7)
$$
$$
\mathbf{W} = \begin{bmatrix} \mathbf{W}_{11} & \cdots & \mathbf{W}_{1P} \\ \vdots & \ddots & \vdots \\ \mathbf{W}_{P1} & \cdots & \mathbf{W}_{PP} \end{bmatrix}.
$$

The definition of $\mathbf{Y}_q$ in (2) leads to the short-time correlation matrix $\mathbf{R}_{yy}(m) = \mathbf{Y}^H(m)\mathbf{Y}(m)$ of size $PD \times PD$ which is composed of channelwise $D \times D$ submatrices $\mathbf{R}_{y_p y_q}(m) = \mathbf{Y}_p^H(m)\mathbf{Y}_q(m)$ each containing D time lags. Here $\cdot^H$ denotes conjugate transposition. In [3] a cost function based on these correlation matrices has been presented which inherently includes all D time lags of all autocorrelations and cross-correlations of the BSS output signals

$$
\mathcal{J}(m, \mathbf{W}) = \sum_{i=0}^{\infty} \beta(i, m)\{\log \det \mathrm{bdiag}\, \mathbf{R}_{yy}(i) - \log \det \mathbf{R}_{yy}(i)\}, \quad (8)
$$

where $\mathrm{bdiag}\, \mathbf{R}_{yy}$ creates a $PD \times PD$ block-diagonal matrix with the channelwise $D \times D$ submatrices $\mathbf{R}_{y_q y_q}$, $q = 1, \ldots, P$, on the main diagonal and zeros elsewhere. The variable β denotes a weighting function with finite support that is normalized according to $\sum_{i=0}^{m} \beta(i, m) = 1$ allowing offline, online, or block-online realizations of the algorithm. The concept of a

general weighting function is already well known from supervised adaptive filtering [6]. There it was shown that, for example, the weighting function $\sum_{i=0}^{\infty} \beta(i, m) = \sum_{i=0}^{m}(1 - \lambda)\lambda^{m-i}$ leads to a recursive online algorithm. The parameter λ denotes the exponential forgetting factor ($0 < \lambda < 1$) and i is the summation index of all blocks up to the current block m. The cost function becomes zero if and only if $\mathbf{R}_{\mathbf{y}_p \mathbf{y}_q}$, $p \neq q$, that is, all *output cross-correlations over all time lags* become zero. Thus, (8) explicitly exploits the nonwhiteness property of the output signals.

In [3] a coefficient update based on (8) was derived and in [5] a block-online update rule was derived for the coefficient update by specifying $\beta(i, m)$ such that it leads to a combination of an online update and an offline update. In the block-online update scheme the *offline* part is calculated iteratively for the current block m containing KN samples as

$$\widetilde{\mathbf{W}}^j(m) = \widetilde{\mathbf{W}}^{j-1}(m) - \mu \widetilde{\mathcal{Q}}(m, \widetilde{\mathbf{W}}^{j-1}(m)), \qquad (9)$$

$$\widetilde{\mathcal{Q}}(m, \widetilde{\mathbf{W}}^{j-1}(m)) = \frac{1}{K} \sum_{i=mK}^{mK+K-1} \mathcal{Q}(i, \widetilde{\mathbf{W}}^{j-1}(m)), \qquad (10)$$

where $j = 1, \ldots, j_{\max}$ denotes the current iteration, μ is the stepsize, and $\widetilde{\mathbf{W}}^j(m)$ is the demixing filter matrix after j iterations based on data of the mth block. Equation (10) performs a simultaneous optimization for K blocks of length N which allows to exploit the nonstationarity of the source signals as for each block the source statistics change and thus new conditions are generated. Thus, (10) contains K update terms $\mathcal{Q}(i, \widetilde{\mathbf{W}}^{j-1}(m))$ which are determined as the natural gradient of the cost function (8) [3]

$$\mathcal{Q}(i, \mathbf{W}) = \mathbf{W}\{\mathbf{R}_{\mathbf{yy}}(i) - \mathrm{bdiag}\,\mathbf{R}_{\mathbf{yy}}(i)\}\,\mathrm{bdiag}^{-1}\mathbf{R}_{\mathbf{yy}}(i). \quad (11)$$

A high number of offline iterations $j_{\max}$ allows a fast convergence without introducing an additional algorithmic delay but at the cost of an increased computational complexity. The demixing filter matrix $\widetilde{\mathbf{W}}^{j_{\max}}(m)$ of the current block m which is obtained from the offline part after $j_{\max}$ iterations is then used as input to the *online* part of the block-online algorithm which is written recursively as

$$\mathbf{W}(m) = \lambda \mathbf{W}(m-1) + (1-\lambda)\widetilde{\mathbf{W}}^{j_{\max}}(m), \qquad (12)$$

with the forgetting factor λ. This yields the final demixing filter matrix $\mathbf{W}(m)$ of the current block m containing the filter weights $\mathbf{w}_{pq}(m)$ used for separation. The demixing filter weights $\mathbf{w}_{pq}(m)$ of the current block are then used as initial values for the offline algorithm (9) of the next block. An overview of the block-online update procedure can also be found in the pseudocode given in Table 1.

It should be pointed out that the natural gradient (11) obtained from the cost function (8) can similarly be derived using the Kullback-Leibler divergence based on multivariate probability density functions [4]. The second-order BSS algorithm is then obtained by using the multivariate Gaussian probability density function.

TABLE 1: Pseudocode of the block-online algorithm with improved normalization according to Section 3.3 exemplarily shown for the update $\Delta \mathbf{w}_{11}(m)$ in the 2×2 case.

	Online part
(1)	Get $KL + N$ new samples $x_p(mKL), \ldots,$ $x_p((m+1)KL + N - 1)$ of the sensors x_p, $p = 1, 2$, and online block index $m = 0, 1, 2, \ldots$
	Offline part
	Compute for each iteration $j = 1, \ldots, j_{\max}$
(2)	Compute output signals $y_q(mKL), \ldots,$ $y_q((m+1)KL + N - L - 1)$, $q = 1, 2$ by convolving x_p with filter weights $\widetilde{w}_{pq}^{j-1}(m)$ from previous iteration.
(3)	Generate K blocks of N samples $[y_q(iL), \ldots,$ $y_q(iL + N - 1)]$ with offline block index, $i = mK, \ldots, mK + K - 1$, to exploit nonstationarity.
	Compute for each block $i = mK, \ldots, mK + K - 1$
(4)	Compute cross-correlation matrix $\mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1}(i)$ by $r_{y_2 y_1}(i, u)$ for $u = -L+1, \ldots, L-1$ according to (14).
(5)	Calculate the values on the diagonal of $\underline{\widetilde{\mathbf{Y}}}_1$ by computing the DFT of length R of the ith output signal block of length N of Step (3).
(6)	Calculate the signal energy of each block i $\sigma_{y_1}^2(i) = r_{y_1 y_1}(i, 0) = \sum_{n=iL}^{iL+N-1} y_1^2(n)$.
(7)	Calculate $\underline{\widetilde{\mathbf{Y}}}_1^H \underline{\widetilde{\mathbf{Y}}}_1$ in (33) by scalar multiplication in each frequency bin and perform narrowband regularization according to (33) by using the signal energy $\sigma_{y_1}^2$ $\mathbf{S}_{\mathbf{y}_1 \mathbf{y}_1}(i) = \rho \underline{\widetilde{\mathbf{Y}}}_1^H(i)\underline{\widetilde{\mathbf{Y}}}_1(i) + (1 - \rho)\sigma_{y_1}^2(i)\mathbf{I}$.
(8)	Perform scalar inversion of the frequency-domain values on the main diagonal of $\mathbf{S}_{\mathbf{y}_1 \mathbf{y}_1}(i)$ as given in (26) and apply the inverse DFT to the resulting vector to obtain the first column of the circulant matrix $\mathbf{C}_{\underline{\widetilde{\mathbf{Y}}}_1 \underline{\widetilde{\mathbf{Y}}}_1}^{-1}(i)$.
(9)	In (27) the circulant matrix $\mathbf{C}_{\underline{\widetilde{\mathbf{Y}}}_1 \underline{\widetilde{\mathbf{Y}}}_1}^{-1}(i)$ is constrained to yield the approximation of the inverse of the Toeplitz matrix $\mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1}^{-1}(i)$. Matrix $\mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1}^{-1}(i)$ can be generated by picking the first L and last $L - 1$ values of the resulting vector from Step (8).
(10)	Compute the matrix product $\mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1}(i)\mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1}^{-1}(i)$ in (11) by fast convolution techniques exploiting the Toeplitz structure of both matrices. The result $\mathbf{A}_{\mathbf{y}_2 \mathbf{y}_1}(i)$ of the matrix product may be approximated due to complexity reasons by calculating only the entries $[a(i, 0), \ldots, a(i, -L+1)]$ in the first column and the entries $[a(i, 0), \ldots, a(i, L-1)]$ in the first row and generate a Toeplitz structure from these values.
(11)	Compute the matrix product $\widetilde{\mathbf{W}}_{12}^{j-1}(m)\mathbf{A}_{\mathbf{y}_2 \mathbf{y}_1}(i)$ as a convolution using Sylvester constraint $\mathcal{SC_R}$. Each filter weight update $\Delta \widetilde{w}_{11,\kappa}^j$, $\kappa = 0, \ldots, L-1$, is thus calculated as $\widetilde{\mathcal{Q}}(m, \widetilde{\mathbf{W}}_{11}^{j-1}(m)) = \frac{1}{K} \sum_i \sum_{n=0}^{L-1} \widetilde{w}_{12,n}^{j-1}(m)a(i, n - \kappa)$.

Table 1: Continued.

(12)	Update equation for the offline part (note that also an adaptive stepsize according to [5] can be applied): $\widetilde{\mathbf{W}}_{11}^{j}(m) = \widetilde{\mathbf{W}}_{11}^{j-1}(m) - \mu\widetilde{\mathcal{Q}}(m, \widetilde{\mathbf{W}}_{11}^{j-1}(m))$.
	Online part
(13)	Compute the recursive update of the online part yielding the demixing filter $\mathbf{W}_{11}(m)$ used for separation: $\mathbf{W}_{11}(m) = \lambda\mathbf{W}_{11}(m-1) + (1-\lambda)\widetilde{\mathbf{W}}_{11}^{j_{max}}(m)$.
(14)	Compute Steps (4)–(13) analogously for the other channels and use the demixing filter $\mathbf{W}_{pq}(m)$ as initial filter for the offline part $\widetilde{\mathbf{W}}_{pq}^{0}(m+1) = \mathbf{W}_{pq}(m)$.

2.2. Estimation of the correlation matrices and Sylvester constraint $\mathcal{SC}$

In principle, there are two basic methods for the block-based estimation of the short-time output correlation matrices $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}(i)$ for nonstationary signals: the so-called *covariance method* and the *correlation method*, as they are known from linear prediction problems [7].[1] In [3] the more accurate co-variance method was introduced by the definition $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}(i) = \mathbf{Y}_p^H(i)\mathbf{Y}_q(i)$. In [5] the computationally less complex correlation method was used which is obtained by assuming stationarity within each block i. This leads to a *Toeplitz structure* of the $D \times D$ matrix $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}(i)$ which can be expressed as

$$\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}(i) = \begin{bmatrix} r_{y_py_q}(i,0) & \cdots & r_{y_py_q}(i,D-1) \\ r_{y_py_q}(i,-1) & \ddots & r_{y_py_q}(i,D-2) \\ \vdots & \ddots & \vdots \\ r_{y_py_q}(i,-D+1) & \cdots & r_{y_py_q}(i,0) \end{bmatrix}, \quad (13)$$

$$r_{y_py_q}(i,u) = \begin{cases} \sum_{n=iL}^{iL+N-u-1} y_p(n+u)y_q(n) & \text{for } u \geq 0, \\ \sum_{n=iL-u}^{iL+N-1} y_p(n+u)y_q(n) & \text{for } u < 0. \end{cases} \quad (14)$$

Using the correlation method, the Toeplitz matrix $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}$ can also be written as a matrix product

$$\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}(i) = \widetilde{\mathbf{Y}}_p^H(i)\widetilde{\mathbf{Y}}_q(i), \quad (15)$$

where $\widetilde{\mathbf{Y}}_p$ denotes $(N+D) \times D$ matrix exhibiting a Sylvester structure as shown for the coefficient matrix in (5). The first column vector of $\widetilde{\mathbf{Y}}_p(i)$ contains the output signal values $y_q(iL), \ldots, y_q(iL+N-1)$ analogously to the first column vector of (2). In contrast to the covariance method using the matrix defined in (2) now additionally D zeros are appended to the output signal values. For each subsequent column this vector is shifted by one sample as shown in (5).

---

[1] It should be emphasized that the terms *covariance method* and *correlation method* are *not* based upon the standard usage of the covariance function as the correlation function with the means removed.

In [3] the coefficient update was derived by taking the derivative with respect to the Sylvester matrix $\mathbf{W}$. There, it was shown that the Sylvester structure of the update $\mathcal{Q}$ in (11) has to be ensured by a Sylvester constraint ($\mathcal{SC}$). In [5, 8] two efficient versions have been discussed. They allow to implement the matrix multiplication of $\mathbf{W}$ with the remaining Toeplitz matrix in (11) as a fast convolution reducing the complexity from $\mathcal{O}(L^3)$ to $\mathcal{O}(\log(L))$. A detailed analysis of the computational complexity of the algorithm (9)–(12) can be found in [5]. In the present paper we apply the row Sylvester constraint $\mathcal{SC}_{\mathcal{R}}$ which calculates only the Lth row of the update $\mathcal{Q}$ and then replicates the elements to obtain the Sylvester structure of $\mathbf{W}$. A detailed discussion of the Sylvester constraints can be found in [8].

3. NORMALIZATION STRATEGIES

The update of the generic algorithm given by (11) exhibits an inherent normalization by the inverse of a block-diagonal matrix. This is an advantage compared to algorithms based on Frobenius norm cost functions as, for example, [9] where heuristic normalizations have to be introduced. Moreover, (11) allows for several normalization strategies by applying certain approximations as shown in the following.

3.1. Exact normalization based on matrix inverse

When using the correlation method, the $D \times D$ Toeplitz matrices $\mathbf{R}_{\mathbf{y}_q\mathbf{y}_q}$, $q = 1, \ldots, P$, given by (15), have to be inverted in (11). This is similar to the matrix inversion occurring in the recursive least-squares (RLS) algorithm in supervised adaptive filtering [6]. The complexity of a Toeplitz matrix inversion is $\mathcal{O}(D^2)$. For realistic acoustic environments large values for D (e.g., 1024) are required which are prohibitive for a real-time implementation of the exact normalization on most current hardware platforms.

3.2. Normalization based on diagonal matrices in the time domain

In [5] an approximation of the matrix inverse has been used to obtain an efficient algorithm suitable for real-time implementations. There, the off-diagonals of the autocorrelation submatrices have been neglected, so that for the correlation method it can be approximated by a diagonal matrix with the output signal powers, that is,

$$\mathbf{R}_{\mathbf{y}_q\mathbf{y}_q}(i) \approx \text{diag}\{\mathbf{R}_{\mathbf{y}_q\mathbf{y}_q}(i)\} = \sigma_{y_q}^2(i)\mathbf{I} \quad (16)$$

for $q = 1, \ldots, P$, where the diag operator applied to a matrix sets all off-diagonal elements to zero. Thus, the matrix inversion is replaced by an element-wise division. This is comparable to the normalization in the well-known normalized least mean squares (NLMS) algorithm in supervised adaptive filtering approximating the RLS algorithm [6].

3.3. Novel approximation of exact normalization based on the Szegö theorem

The broadband algorithm given by (9)–(12) can also be formulated equivalently in the frequency domain as has been presented in [3]. Additionally it has been shown that by certain approximations to this frequency-domain formulation a purely narrowband version of the broadband algorithm can be obtained. In this section we will derive a novel algorithm combining broadband and narrowband techniques by using two steps. First, the exact normalization is formulated equivalently in the frequency domain (Section 3.3.1). In a second step the Szegö theorem is applied to the normalization to obtain an efficient version of the exact normalization (Section 3.3.2). The Szegö theorem allows a *selective* introduction of narrowband approximations to specific parts of the algorithm. This approach allows to combine both the advantages of the broadband algorithm (e.g., avoiding internal permutation ambiguity and circularity problem) and the low complexity of a narrowband approach.

3.3.1. Exact normalization expressed in the frequency domain

In [10] it was shown that any Toeplitz matrix can be expressed equivalently in the frequency domain by first generating a circulant matrix by proper extension of the Toeplitz matrix. Then the circulant matrix is diagonalized by using the discrete Fourier transform (DFT) matrix $\mathbf{F}_R$ of size $R \times R$ where $R \geq N + D$ denotes the transformation length. These two steps are given for the Toeplitz output signal matrix $\tilde{\mathbf{Y}}_q$ as

$$\tilde{\mathbf{Y}}_q = \mathbf{W}_{N+D\times R}^{01_{N+D}} \mathbf{C}_{\tilde{\mathbf{Y}}_q} \mathbf{W}_{R\times D}^{1_D 0} \tag{17}$$

$$= \mathbf{W}_{N+D\times R}^{01_{N+D}} \mathbf{F}_R^{-1} \underline{\tilde{\mathbf{Y}}}_q \mathbf{F}_R \mathbf{W}_{R\times D}^{1_D 0}, \tag{18}$$

where $\mathbf{C}_{\tilde{\mathbf{Y}}_q}$ is a $R \times R$ circulant matrix and the window matrices are given as

$$\mathbf{W}_{N+D\times R}^{01_{N+D}} = \left[\mathbf{0}_{N+D\times R-N-D}, \mathbf{I}_{N+D\times N+D}\right],$$
$$\mathbf{W}_{R\times D}^{1_D 0} = \left[\mathbf{I}_{D\times D}, \mathbf{0}_{R-D\times D}\right]. \tag{19}$$

Here the convention is used that the lower index of a matrix denotes its dimensions and the upper index describes the positions of ones and zeros. The size of the unity submatrices is indicated in subscript (e.g., "01_D"). The matrix $\underline{\tilde{\mathbf{Y}}}_q$ exhibits a diagonal structure containing the eigenvalues of the circulant matrix $\mathbf{C}_{\tilde{\mathbf{Y}}_q}$ on the main diagonal. The eigenvalues are calculated by the DFT of the first column of $\mathbf{C}_{\tilde{\mathbf{Y}}_q}$ and thus $\underline{\tilde{\mathbf{Y}}}_q$ can be interpreted as the frequency-domain counterpart of $\tilde{\mathbf{Y}}_q$:

$$\underline{\tilde{\mathbf{Y}}}_q$$
$$= \text{Diag}\left\{\mathbf{F}_R[0,\ldots,0,y_q(iL),\ldots,y_q(iL+N-1),0,\ldots,0]^T\right\}. \tag{20}$$
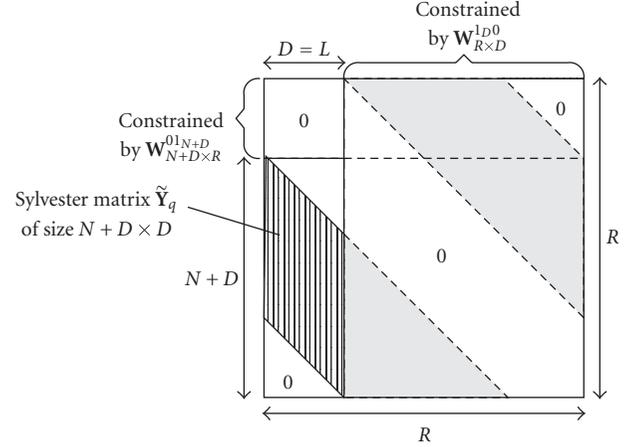
FIGURE 1: Illustration of (17) showing the relation between circulant matrix $\mathbf{C}_{\tilde{\mathbf{Y}}_q}$ and Toeplitz matrix $\tilde{\mathbf{Y}}_q$.

The operator Diag$\{\mathbf{a}\}$ denotes a square matrix with the elements of vector $\mathbf{a}$ on its main diagonal. An illustration of the circulant matrix $\mathbf{C}_{\tilde{\mathbf{Y}}_q}$ and the window matrices, which constrain the circular matrix to the original matrix $\tilde{\mathbf{Y}}_q$, is given in Figure 1. With (18) we can now write $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}$ as

$$\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q} = \mathbf{W}_{D\times R}^{1_D 0} \mathbf{F}_R^{-1} \underline{\tilde{\mathbf{Y}}}_p^H \mathbf{F}_R \mathbf{W}_{R\times N+D}^{01_{N+D}}$$
$$\cdot \mathbf{W}_{N+D\times R}^{01_{N+D}} \mathbf{F}_R^{-1} \underline{\tilde{\mathbf{Y}}}_q \mathbf{F}_R \mathbf{W}_{R\times D}^{1_D 0}. \tag{21}$$

It can be seen in the upper left corner of the illustration in Figure 1 that by extending the window matrix $\mathbf{W}_{N+D\times R}^{01_{N+D}}$ to $\mathbf{W}_{R\times R}^{01_R} = \mathbf{I}_{R\times R}$ only rows of zeros are introduced at the beginning of the matrix $\tilde{\mathbf{Y}}_q$, that is, (17) is now of the form

$$\begin{bmatrix} \mathbf{0}_{R-N-D\times D} \\ \tilde{\mathbf{Y}}_q \end{bmatrix} = \mathbf{C}_{\tilde{\mathbf{Y}}_q} \mathbf{W}_{R\times D}^{1_D 0}. \tag{22}$$

These appended rows of zeros have no effect on the calculation of the correlation matrix $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q}$ and thus we can replace the multiplication of the window matrices in (21) by

$$\mathbf{W}_{R\times R}^{01_R} \mathbf{W}_{R\times R}^{01_R} = \mathbf{I}_{R\times R}. \tag{23}$$

This leads to

$$\mathbf{R}_{\mathbf{y}_p\mathbf{y}_q} = \mathbf{W}_{D\times R}^{1_D 0} \mathbf{F}_R^{-1} \underline{\tilde{\mathbf{Y}}}_p^H \underline{\tilde{\mathbf{Y}}}_q \mathbf{F}_R \mathbf{W}_{R\times D}^{1_D 0} \tag{24}$$

$$= \mathbf{W}_{D\times R}^{1_D 0} \mathbf{C}_{\tilde{\mathbf{Y}}_p\tilde{\mathbf{Y}}_q} \mathbf{W}_{R\times D}^{1_D 0}. \tag{25}$$

The correlation matrix in (24) is an equivalent expression to (15) in the frequency domain. Thus, the normalization based on the inversion of (24) or (25) for $p = q = 1,\ldots,P$ still corresponds to the exact normalization based on the matrix inverse of a Toeplitz matrix as described in Section 3.1. In the following it is shown how the inverse of (25) can be approximated to obtain an efficient implementation.

3.3.2. Application of the Szegö theorem

In the tutorial paper [10] the Szegö theorem is formulated and proven for finite-order Toeplitz matrices. A finite-order Toeplitz matrix is defined as an $R \times R$ Toeplitz matrix where a finite D exists such that all elements of the matrix with the row or column index greater than D are equal to zero. It was shown in [10] that the $R \times R$ Toeplitz matrix of order D is asymptotically equivalent to the $R \times R$ circulant matrix generated from an appropriately complemented $D \times D$ Toeplitz matrix. If the two matrices are also of Hermitian structure, then the Szegö theorem on the asymptotic eigenvalue distribution states the following.

(1) The eigenvalues of both matrices lie between a lower bound and an upper bound.
(2) The arithmetic means of the eigenvalues of both matrices are equal if the size R of both matrices approaches infinity.

Then, the eigenvalues of both matrices are said to be asymptotically equally distributed.

It can be seen in (25) that the autocorrelation matrix necessary for the normalization can be expressed as a $D \times D$ Toeplitz matrix $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}$ or an $R \times R$ circulant matrix $\mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}$ generated from the Toeplitz matrix by extending it appropriately and multiplying it with some window matrices. According to [10] both matrices are asymptotically equivalent. As both the Toeplitz and the circulant matrices are Hermitian, it is possible to apply the Szegö theorem. The eigenvalues of $\mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}$ are given in (24) as the elements on the main diagonal of the diagonal matrix $\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q$. The Szegö theorem states that the eigenvalues of the $R \times R$ Toeplitz matrix generated by appending zeros to $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}$ can be asymptotically approximated by $\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q$ for $R \to \infty$. The benefit of this approximation becomes clear if we take a look at the inverse of a circulant matrix. The inverse of a circulant matrix can be easily calculated by inverting its eigenvalues

$$\mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}^{-1} = \mathbf{F}_R^{-1} \left(\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q \right)^{-1} \mathbf{F}_R. \tag{26}$$

By using the Szegö theorem we can now approximate the inverse of the Toeplitz matrix $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}$ by the inverse of the circulant matrix (26) for $R \to \infty$,

$$\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}^{-1} \approx \mathbf{W}_{D \times R}^{1_D 0} \mathbf{F}_R^{-1} \left(\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q \right)^{-1} \mathbf{F}_R \mathbf{W}_{R \times D}^{1_D 0}. \tag{27}$$

This can also be denoted as narrowband approximation because the eigenvalues $\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q$ can easily be determined as the DFT of the first column of the circulant matrix $\mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}$. The inverse in (27) can now be efficiently implemented as a scalar inversion because $\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q$ denotes a diagonal matrix. Moreover, it is important to note that the inverse of a circulant matrix is also circulant. Thus, after the windowing by $\mathbf{W}_{\cdots}^{1_D 0}$ the resulting matrix $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}^{-1}$ exhibits again a Toeplitz structure.

The error which is introduced by the narrowband approximation has been examined in [11] for the case of stationary random processes. The error has been measured as the difference between the exact inversion of the Toeplitz matrix given in (24) and the approximated inverse given in (27). The results obtained in [11] show that for $R \gg D$ the narrowband approximation is well justified.

In summary, (27) can be efficiently implemented as a DFT of the first column of $\mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}$ followed by a scalar inversion of the frequency-domain values and then applying the inverse DFT. After the windowing operation these values are then replicated to generate the Toeplitz structure of $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}^{-1}$. This approach reduces the complexity from $\mathcal{O}(D^2)$ to $\mathcal{O}(R \log R)$ (e.g., experiments in Section 5: $D = L$, $R = 4L$). Obtaining a Toeplitz matrix after the inversion has the advantage that in the update equation (11) again a product of Toeplitz matrices has to be calculated which can be efficiently implemented using fast convolutions. For more details see [5].

4. REGULARIZATION OF THE MATRIX INVERSE

Prior to the inversion of the autocorrelation Toeplitz matrices according to (15) a regularization is necessary as these matrices may be ill-conditioned. Here we propose to attenuate the off-diagonals of $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}$ by multiplying them with the factor ρ:

$$\begin{aligned} \check{\mathbf{R}}_{\mathbf{y}_q \mathbf{y}_q} &= \rho \mathbf{R}_{\mathbf{y}_q \mathbf{y}_q} + (1 - \rho) \operatorname{diag} \{ \mathbf{R}_{\mathbf{y}_q \mathbf{y}_q} \} \\ &= \rho \mathbf{R}_{\mathbf{y}_q \mathbf{y}_q} + (1 - \rho) \sigma_{y_q}^2 \mathbf{I}. \end{aligned} \tag{28}$$

The attenuation factor ρ has to be within the range $0 \le \rho \le 1$. Using this regularization, the algorithm performs also well even if there is just one active source. It should be noted that for $\rho = 0$ the previous approximation of the normalization in [5] and Section 3.2 can be seen as a special case of the regularized version of the novel normalization presented in Section 3.3.

The selective narrowband approximation of Section 3.3 leads to an inversion of circulant matrices $\mathbf{C}_{\widetilde{\mathbf{Y}}_p \widetilde{\mathbf{Y}}_q}$ instead of Toeplitz matrices $\mathbf{R}_{\mathbf{y}_q \mathbf{y}_q}$. Thus, analogously to (28) it is desirable for the proposed algorithm to also regularize $\mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}$ prior to inversion:

$$\check{\mathbf{C}}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q} = \rho \mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q} + (1 - \rho) \operatorname{diag} \{ \mathbf{C}_{\widetilde{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q} \}. \tag{29}$$

In Section 3.3 it was pointed out that every circulant matrix can be expressed using the DFT, inverse DFT matrix, and a diagonal matrix

$$\mathbf{C}_{\check{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q} = \mathbf{F}_R^{-1} \underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q \mathbf{F}_R. \tag{30}$$

The diagonal matrix $\underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q$ contains the DFT transformed elements of the first column of the circulant matrix on its diagonal. Thus, by applying the diag operator on $\mathbf{C}_{\check{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q}$ we can write

$$\begin{aligned} \operatorname{diag} \{ \mathbf{C}_{\check{\mathbf{Y}}_q \widetilde{\mathbf{Y}}_q} \} &= r_{y_q y_q}(0) \cdot \mathbf{I} = \sigma_{y_q}^2 \cdot \mathbf{I} \\ &= \mathbf{F}_R^{-1} \sigma_{y_q}^2 \cdot \mathbf{I} \cdot \mathbf{F}_R. \end{aligned} \tag{31}$$

Thus, (29) can be simplified to a narrowband regularization in each frequency bin as

$$\check{\mathbf{C}}_{\widetilde{\mathbf{Y}}_p \widetilde{\mathbf{Y}}_q} = \rho \mathbf{F}_R^{-1} \underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q \mathbf{F}_R + (1 - \rho)\sigma_{y_q}^2 \mathbf{I} \tag{32}$$

$$= \mathbf{F}_R^{-1}(\rho \underline{\widetilde{\mathbf{Y}}}_q^H \underline{\widetilde{\mathbf{Y}}}_q + (1 - \rho)\sigma_{y_q}^2 \mathbf{I})\mathbf{F}_R. \tag{33}$$

Note that the second term in (32) is equivalent to the second term in (28). This time-frequency equivalence can be explained by the Parseval theorem. It should be noted that the regularization in (32) can also be applied to purely narrowband algorithms (e.g., [3, Section IV-C]). There, considerable separation performance improvements compared to a regularization by adding a constant have been observed too.

A pseudocode of the efficient implementation of the proposed algorithm based on (9)–(12) together with the novel normalization presented in Section 3.3 and the new regularization in Section 4 is given in Table 1. There, the implementation is exemplarily shown for the update $\Delta \mathbf{w}_{11}(m)$ for $P = 2, D = L$ and application of the Sylvester constraint $\mathcal{SC}_R$.

5. EXPERIMENTS

The experiments were conducted using speech data convolved with measured impulse responses of speakers in two different environments: (a) in a real room (580 cm × 590 cm × 310 cm) with reverberation time $T_{60} = 250$ ms at ±45° and 2 m distance of the sources to the array, and (b) impulse responses of a driver and codriver in a car ($T_{60} = 50$ ms) with the array mounted to the rear mirror. In the car scenario also recorded background noise with 0 dB SNR was added. The sampling frequency was $f_s = 16$ kHz. A two-element microphone array with an interelement spacing of 20 cm was used for both recordings. The demixing filter length L was chosen to 1024 taps, the block length $N = 2L$, and the number of time lags considered in the correlation matrices was set to $D = L$. The frameshift was L samples, $K = 8$ blocks have been used to exploit nonstationarity, and $j_{max} = 5$ iterations have been used as number of iterations for the offline update. The adaptive stepsize proposed in [5] has been used with the minimum and maximum values $\mu_{min} = 0.0001$, $\mu_{max} = 0.01$, respectively, and the forgetting factor $\lambda = 0.2$. The factor ρ for the novel regularization has been set to $\rho = 0.5$. The demixing filters were initialized with a shifted unit impulse where $w_{qq,20} = 1$ for $q = 1, \ldots, P$ and zeros elsewhere.

To evaluate the performance, the signal-to-interference ratio (SIR) was calculated which is defined for the qth channel as the ratio of the signal power of the target source signal $y_{s,q}(n)$ to the signal power from the crosstalk signal $y_{c,q}(n)$ given by

$$\text{SIR}_q(n) = 10 \log_{10} \frac{\hat{E}\{y_{s,q}^2(n)\}}{\hat{E}\{y_{c,q}^2(n)\}}, \tag{34}$$

where the estimate $\hat{E}$ of the expectation operator is implemented as a moving average. To obtain the target and
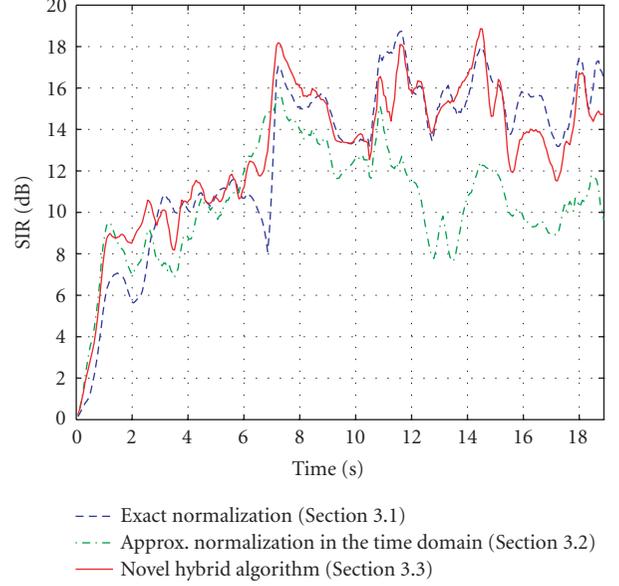
FIGURE 2: SIR results for reverberant room.

- - - Exact normalization (Section 3.1)
-·-· Approx. normalization in the time domain (Section 3.2)
— Novel hybrid algorithm (Section 3.3)
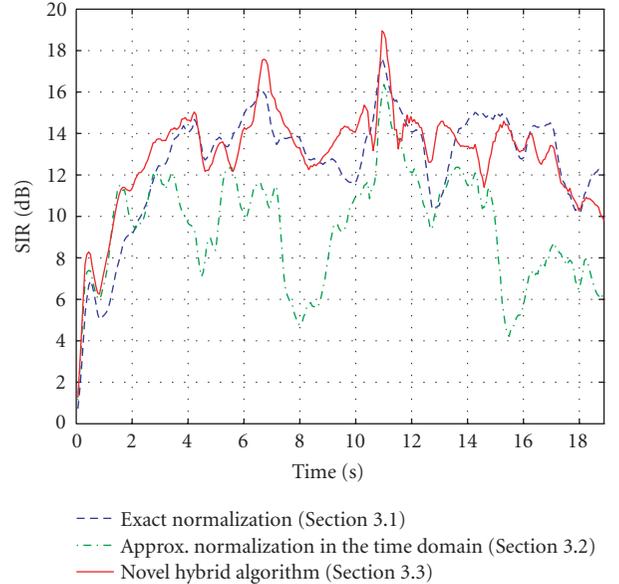
FIGURE 3: SIR results for car environment (0 dB car noise).

- - - Exact normalization (Section 3.1)
-·-· Approx. normalization in the time domain (Section 3.2)
— Novel hybrid algorithm (Section 3.3)

crosstalk signal component for the SIR calculation, each signal component at the microphone signals is processed individually by the demixing system obtained by the BSS algorithm. A possible external permutation, that is, if the source signal $s_p(n)$ is obtained at a BSS output channel $y_q(n)$ with $p \neq q$, is corrected before the SIR calculation. In the experiments the channelwise SIR_q defined in (34) has been averaged over both channels $q = 1, 2$.

In Figures 2 and 3 the results of the broadband algorithm with the three different normalization schemes presented in Section 3 are shown. The dashed line represents the exact normalization by the inverse of the Toeplitz matrix which

is estimated using the correlation method. It can be seen that the novel normalization scheme (solid) obtained by the narrowband approximation corresponding to the inversion of a circulant matrix approximates the exact normalization very well. Moreover, the novel normalization yields improved performance compared to the time-domain approximation (dash-dotted) resulting in a normalization by the output signal power. Sometimes the novel algorithm even seems to slightly outperform the exact normalization. This can be explained by the usage of an adaptive stepsize [5] which may result in slightly different convergence speeds for all three algorithms. It should also be noted that the fluctuation of the SIR is due to the nonstationarity of the speech signals.

6. CONCLUSION

In this paper a novel efficient normalization scheme was presented resulting in a novel algorithm combining advantages of broadband algorithms with the efficiency of narrowband techniques. Moreover, a regularization method was proposed leading to improved convergence behavior. Experimental results in realistic acoustic environments confirm the efficiency of the proposed approach.

ACKNOWLEDGMENT

REFERENCES

[1] A. Hyvaerinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, NY, USA, 2001.

[2] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON: a versatile framework for multichannel blind signal processing," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 3, pp. 889–892, Montreal, Quebec, Canada, May 2004.

[3] H. Buchner, R. Aichner, and W. Kellermann, "A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 120–134, 2005.

[4] H. Buchner, R. Aichner, and W. Kellermann, "Blind source separation for convolutive mixtures: a unified treatment," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y. Huang and J. Benesty, Eds., pp. 255–293, Kluwer Academic, Boston, Mass, USA, 2004.

[5] R. Aichner, H. Buchner, F. Yan, and W. Kellermann, "A real-time blind source separation scheme and its application to reverberant and noisy acoustic environments," *Signal Processing*, vol. 86, no. 6, pp. 1260–1277, 2006.

[6] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 4th edition, 2002.

[7] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, Springer, Berlin, Germany, 1976.

[8] R. Aichner, H. Buchner, and W. Kellermann, "On the causality problem in time-domain blind source separation and deconvolution algorithms," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 5, pp. 181–184, Philadelphia, Pa, USA, March 2005.

[9] L. Parra and C. Spence, "Convolutive blind separation of nonstationary sources," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 320–327, 2000.

[10] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Transactions on Information Theory*, vol. 18, no. 6, pp. 725–730, 1972.

[11] P. J. Sherman, "Circulant approximations of the inverses of Toeplitz matrices and related quantities with applications to stationary random processes," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 6, pp. 1630–1632, 1985.

Robert Aichner received the Dipl.-Ing. (FH) degree in electrical engineering from the University of Applied Sciences, Regensburg, Germany, in 2002. In 2000 he was an intern at Siemens Energy and Automation, Atlanta, Ga, USA. From 2001 to 2002, he did research at the Speech Open Lab of the R&D Division of the Nippon Telegraph and Telephone Corporation (NTT) in Kyoto, Japan. There he was working on time-domain blind source separation of audio signals. Since 2002, he is a member of the research staff at the Chair of Multimedia Communications and Signal Processing at the University of Erlangen-Nuremberg, Germany. His current research interests include multichannel adaptive algorithms for hands-free human-machine interfaces and their application to blind source separation, noise reduction, source localization, adaptive beamforming, and acoustic echo cancellation. In 2004, he was a visiting Researcher at the Sound and Image Processing Lab at the Royal Institute of Technology (KTH), Stockholm, Sweden. He received the Stanglmeier Award for his intermediate diploma from the University of Applied Sciences, Regensburg, in 1999 and the Best Student Paper Award at the IEEE International Conference on Acoustics, Speech, and Signal Processing in 2006.

Herbert Buchner is a member of the research staff at the Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Germany. He received the Dipl.-Ing. (FH) and the Dipl.-Ing. university degrees in electrical engineering from the University of Applied Sciences, Regensburg, in 1997, and the University of Erlangen-Nuremberg in 2000, respectively. In 1995, he was a visiting Researcher at the Colorado Optoelectronic Computing Systems Center (OCS), Boulder/Fort Collins, Colo, USA, where he worked in the field of microwave technology. From 1996 to 1997, he did research at the R&D Division of Nippon Telegraph and Telephone Corporation (NTT), Tokyo, Japan, working on adaptive filtering for teleconferencing. In 1997/1998 he was with the Driver Information Systems Department of Siemens Automotive in Regensburg, Germany. His current areas of interest include efficient multichannel algorithms for adaptive digital filtering, and their applications for acoustic human-machine interfaces, such as multichannel acoustic echo cancellation, beamforming, blind source separation, source localization, and dereverberation. He has authored or coauthored over 50 journal articles, book chapters, and conference papers in his field, and he received the VDI Award in 1998 for his Dipl.-Ing. (FH) thesis from the Verein Deutscher Ingenieure and a Best Student Paper Award in 2001.

Walter Kellermann is a Professor for communications at the Chair of Multimedia Communications and Signal Processing of the University of Erlangen-Nuremberg, Germany. He received the Dipl.-Ing. (Univ.) degree in electrical engineering from the University of Erlangen-Nuremberg in 1983, and the Dr.-Ing. degree from the Technical University Darmstadt, Germany, in 1988. From 1989 to 1990, he was a Postdoctoral Member of technical staff at AT&T Bell Laboratories, Murray Hill, NJ. In 1990, he joined Philips Kommunikations Industrie, Nuremberg, Germany. From 1993 to 1999, he was a Professor at the Fachhochschule Regensburg, before he had joined the University of Erlangen-Nuremberg as a Professor and Head of the Audio Research Laboratory in 1999. He authored or coauthored seven book chapters and more than 70 refereed papers in journals and conference proceedings. He served as a Guest Editor to various journals, as an Associate Editor and Guest Editor to IEEE Transactions on Speech and Audio Processing from 2000 to 2004, and presently serves as an Associate Editor to the EURASIP Journal on Signal Processing and EURASIP Journal on Advances in Signal Processing. He was the General Chair of the 5th International Workshop on Microphone Arrays in 2003 and the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics in 2005. His current research interests include speech signal processing, array signal processing, adaptive filtering, and its applications to acoustic human/machine interfaces.