The TRINICON framework for adaptive MIMO signal processing with focus on the generic Sylvester constraint

Herbert Buchner[1], *Robert Aichner*[2], *Walter Kellermann*[3]

[1] Deutsche Telekom Laboratories, Berlin University of Technology, Ernst-Reuter-Platz 7, 10587 Berlin, Germany
[2] Microsoft Corporation, One Microsoft Way, Redmond, WA 98052-6399, USA
[3] Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Cauerstr. 7, 91058 Erlangen, Germany
E-Mail: hb@buchner-net.com, robert.aichner@microsoft.com, wk@LNT.de

Abstract

This paper gives an overview on TRINICON, a generic framework for broadband adaptive MIMO filtering, and some of its applications in array processing for speech capture. The motivations for this framework are to bring together the various blind and supervised MIMO adaptation techniques which have been treated largely independently in the literature so far, in order to spotlight their commonalities and relationships, to derive them in a rigorous way from first principles, to facilitate the design of improved systems, and to exploit synergies. In this paper, a special focus is on new results on the so-called Sylvester constraint, an important element of the framework.

1 Introduction

In broadband signal acquisition by sensor arrays, such as in hands-free speech communication systems, the original source signals $s_q(n)$, $q = 1,\dots,Q$ are filtered by a linear multiple input and multiple output (MIMO) system (e.g., the reverberant room) before they are captured as sensor signals $x_p(n),\ p = 1,\dots,P$. The general tasks of the corresponding array signal processing are on the one hand to acquire clean source signals from the mixtures $x_p(n)$, and on the other hand to analyze the acoustic scene, the most important parameters usually being the source positions. In this paper, we describe this MIMO mixing system by FIR filter models, where $h_{qp,\kappa}$, $\kappa = 0,\dots,M-1$ denote the coefficients of the model from the q-th source signal $s_q(n)$ to the p-th sensor signal $x_p(n)$ according to Fig. 1. Moreover, we assume throughout this paper that $Q \le P$ which are known as the *overdetermined* and *determined* cases, respectively, as explained below. Note that

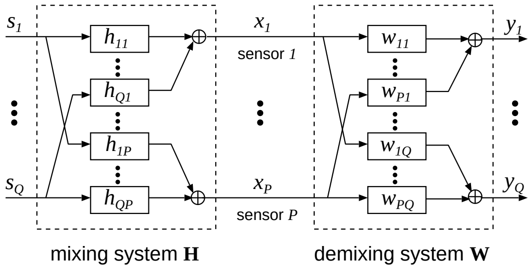


Figure 1: Setup for blind MIMO signal processing.

in general, the sources $s_q(n)$ may or may not be all simultaneously active at a particular instant of time. According to a certain optimization criterion, we are interested in finding a corresponding length-L FIR "demixing" system with coefficients $w_{pq,\kappa}$ by adaptive signal processing. This yields the output signals $y_q(n)$.

Based on this MIMO structure, we may identify the following general signal processing problems from the above-mentioned tasks:

(a) Signal separation (noise/interferers): Cancel out all cross-channels of the cascaded mixing-demixing system.

(b) Deconvolution/Dereverberation: In addition to the separation, acquire "dry" sources up to a delay and a scaling factor.

(c) Identification of the mixing system/source localization

The *ideal MIMO separation filters* to solve (a) were derived and discussed in detail in [1] for an arbitrary number of sources and sensors. It can be shown that an ideal signal-independent broadband separation solution only exists for $Q \le P$. An important property of this ideal separation solution is that it is very closely related to the ideal identification solution (c). We therefore denote (a) and (c) in this paper as the "direct" problems in contrast to the inverse problem (b) whose ideal solution was discussed in [2]. It was shown that the ideal inversion solution requires $Q < P$.

To obtain *estimates* of the solutions to (a)-(c) in practice, it is commonly distinguished between *blind, supervised, and semi-blind* adaptation algorithms. If both the propagation paths and the original source signals in Fig. 1 are unknown, the demixing system has to be estimated by a blind algorithm (e.g., for blind source separation (BSS)) for which the method of independent component analysis (ICA) is typically applied [3]. In other practical scenarios, in which some or even all interfering source signals are directly accessible and/or some side information on the propagation path is known, we can tackle the problem by *supervised* adaptation algorithms, such as the popular least-mean-square (LMS)- or the recursive least-squares (RLS)-type algorithms [4]. Well known supervised and semi-blind techniques for "direct" problems are, e.g., acoustic echo cancellation (AEC) and adaptive beamforming, respectively.

The TRINICON framework brings together these various adaptation techniques. It can be shown that both many of the well-known algorithms in the literature but also numerous novel and efficient algorithms can be derived as special cases of this generic framework. Some of these novel algorithms have already led to efficient real-time systems.

2 Adaptive MIMO Signal Processing based on TRINICON

In this section we first give a brief overview of the essential elements of TRINICON ('TRIple-N ICA for CONvolutive mixtures'), a generic concept for broadband adaptive MIMO filtering [5, 6, 7, 8]. Thereby, we restrict the presentation here to simple gradient-based coefficient updates in the time domain.

Various approaches exist to estimate the demixing matrix $\mathbf{W}$ by utilizing the following fundamental source signal properties [3] which were all combined in TRINICON as a versatile framework:

(i) Nongaussianity is exploited by using higher-order statistics for ICA. The minimization of the mutual information (MMI) among the output channels can be regarded as the most general approach to separation problems [3]. To obtain an estimator that is also suitable for the inverse problem, we use the Kullback-Leibler divergence (KLD) [9] between a certain *desired* joint pdf (essentially representing a hypothesized stochastic source model as shown below) and the joint pdf of the actually estimated output signals.

(ii) Nonwhiteness is exploited by simultaneous minimization of output cross-relations over multiple time-lags. We therefore consider multivariate pdfs, i.e., 'densities including D time-lags'.

(iii) Nonstationarity is exploited by simultaneous minimization of output cross-relations at different time-instants. We assume ergodicity within blocks of length N so that the ensemble average is replaced by time averages over these blocks.

Throughout this section, we present the framework for $Q = P$ without loss of generality. In practice, the current number of simultaneously active sources is allowed to vary throughout the application and only the condition $Q \le P$ has to be fulfilled.

2.1 Optimization Criterion

To introduce an algorithm for broadband processing of convolutive mixtures, we first formulate the convolution of the FIR demixing system of length L in the following matrix form [8]:

$$\mathbf{y}(n) = \mathbf{W}^{\mathrm{T}}\mathbf{x}(n), \tag{1}$$

where n denotes the time index, and

$$\begin{aligned} \mathbf{x}(n) &= [\mathbf{x}_1^{\mathrm{T}}(n), \ldots, \mathbf{x}_P^{\mathrm{T}}(n)]^{\mathrm{T}}, & (2)\\ \mathbf{y}(n) &= [\mathbf{y}_1^{\mathrm{T}}(n), \ldots, \mathbf{y}_P^{\mathrm{T}}(n)]^{\mathrm{T}}, & (3)\\ \mathbf{x}_p(n) &= [x_p(n), \ldots, x_p(n-2L+1)]^{\mathrm{T}}, & (4)\\ \mathbf{y}_q(n) &= [y_q(n), \ldots, y_q(n-D+1)]^{\mathrm{T}}. & (5) \end{aligned}$$

The parameter D in (5), $1 \leq D < L$, denotes the number of time lags taken into account to exploit the nonwhiteness of the source signals as shown below. $\mathbf{W}_{pq}$, $p = 1, \ldots, P$, $q = 1, \ldots, P$ denote $2L \times D$ *Sylvester matrices* that contain all coefficients of the respective filters in each column by successive shifting, i.e., the first column reads $\left[\mathbf{w}_{pq}^{\mathrm{T}}, 0, \ldots, 0\right]^{\mathrm{T}}$, the second column $\left[0, \mathbf{w}_{pq}^{\mathrm{T}}, 0, \ldots, 0\right]^{\mathrm{T}}$, etc. Finally, the $2PL \times PD$ matrix $\mathbf{W}$ combines all Sylvester matrices $\mathbf{W}_{pq}$.

Based on the KLD, the following cost function was introduced in [7] taking into account all three fundamental signal properties (i)-(iii):

$$\begin{aligned} \mathcal{J}(m, \mathbf{W}) &= -\sum_{i=0}^{\infty} \beta(i,m) \frac{1}{N} \\ &\cdot \sum_{j=iL}^{iL+N-1} \left\{ \log(\hat{p}_{s,PD}(\mathbf{y}_p(j))) - \log(\hat{p}_{y,PD}(\mathbf{y}(j))) \right\}, \end{aligned} \tag{6}$$

where $\hat{p}_{s,PD}(\cdot)$ and $\hat{p}_{y,PD}(\cdot)$ are assumed or estimated PD-variate source model (i.e., desired) pdf and output pdf, respectively. The index m denotes the block time index for a block of N output samples shifted by L samples relatively to the previous block. Furthermore, β is a window function allowing for online, offline, or block-online algorithms [6].

2.2 Gradient-Based Coefficient Update

For brevity and simplicity we concentrate in this subsection on iterative Euclidean gradient-based block-online coefficient updates (TRINICON-based Newton-type algorithms have been developed in an analogous way but they are omitted here) which can be written in the general form

$$\begin{aligned} \check{\mathbf{W}}^0(m) &:= \check{\mathbf{W}}(m-1), & (7a)\\ \check{\mathbf{W}}^{\ell}(m) &= \check{\mathbf{W}}^{\ell-1}(m) - \mu\Delta\check{\mathbf{W}}^{\ell}(m),\ \ell = 1, \ldots, \ell_{\max}, & (7b)\\ \check{\mathbf{W}}(m) &:= \check{\mathbf{W}}^{\ell_{\max}}(m), & (7c) \end{aligned}$$

where μ is a stepsize parameter, and the superscript index ℓ denotes an iteration parameter to allow for multiple iterations ($\ell = 1, \ldots, \ell_{\max}$) within each block m. The downwards pointing hat symbol on top of $\mathbf{W}$ in (7) serves to distinguish the *condensed* $PL \times Q$ demixing coefficient matrix $\check{\mathbf{W}}$ to be optimized, from the corresponding larger Sylvester matrix $\mathbf{W}$ in the cost function. The matrix $\check{\mathbf{W}}$ consists of the first column of each submatrix $\mathbf{W}_{pq}$ without the L zeros.

Obviously, when calculating the gradient of $\mathcal{J}(m, \mathbf{W})$ w.r.t. $\check{\mathbf{W}}$ explicitly, we are confronted with the problem of the different matrix formulations $\mathbf{W}$ and $\check{\mathbf{W}}$. The larger dimensions of $\mathbf{W}$ are a direct consequence of taking into account the nonwhiteness signal property by choosing $D > 1$. The rigorous distinction between these different matrix structures is also an essential aspect of the general TRINICON framework and leads to an important building block whose actual implementation is fundamental to the properties of the resulting algorithm, the so-called *Sylvester constraint* ($\mathcal{SC}$) on the coefficient update, formally introduced in [6, 8]. Using the Sylvester constraint operator the gradient descent update can be written as

$$\Delta\check{\mathbf{W}}^{\ell}(m) = \mathcal{SC}\left\{\nabla_{\mathbf{W}}\mathcal{J}(m, \mathbf{W})\right\}\big|_{\mathbf{W}=\mathbf{W}^{\ell}(m)}. \tag{8}$$

Depending on the particular realization of ($\mathcal{SC}$), we are able to select both, well known and also novel improved adaptation algorithms [10] as we will demonstrate in Sect. 5. In [1] an explicit formulation of a *generic* Sylvester constraint was derived to further formalize and clarify this concept. It turns out that the generic Sylvester constraint corresponds – up to the constant D denoting the width of the submatrices – to a *channel-wise arithmetic averaging* of elements according to Fig. 2.

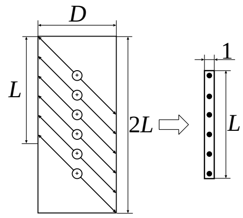


Figure 2: Illustration of the generic Sylvester constraint ($\mathcal{SC}$) after [1] for one channel.

2.3 Natural Gradient-Based Coefficient Update

It can be shown (after a somewhat tedious but straightforward derivation) that by taking the more efficient *natural gradient* [3] of $\mathcal{J}(m)$ with respect to the demixing filter matrix $\mathbf{W}(m)$ [8],

$$\Delta\check{\mathbf{W}} \propto \mathcal{SC}\left\{\mathbf{W}\mathbf{W}^{\mathrm{T}}\frac{\partial\mathcal{J}}{\partial\mathbf{W}}\right\}, \tag{9}$$

we obtain a generic TRINICON-based update rule which may be written in the following compact form:

$$\begin{aligned} \Delta\check{\mathbf{W}}(m) &= \sum_{i=0}^{\infty} \beta(i,m) \\ &\cdot \mathcal{SC}\left\{\mathbf{W}(i)\left[\frac{1}{N}\sum_{j=iL}^{iL+N-1}\mathbf{y}(j)\Phi_{s,PD}^{\mathrm{T}}(\mathbf{y}(j)) - \mathbf{I}\right]\right\}, \end{aligned} \tag{10a}$$

with the score function

$$\Phi_{s,PD}(\mathbf{y}(j)) = -\frac{\partial \log \hat{p}_{s,PD}(\mathbf{y}(j))}{\partial \mathbf{y}(j)} \tag{10b}$$

resulting from the hypothesized source model $\hat{p}_{s,PD}(\cdot)$. The choice of $\hat{p}_{s,PD}(\cdot)$ in (10b) depends on the class of signal processing problem to be solved and on the type of source signals, as detailed in the next two sections.

3 Incorporation of Stochastic Source Models

The general update equations (7),(10) provide the possibility to take into account all available information on the statistical properties of the desired source signals. To apply the general approach in a real-world scenario, appropriate multivariate score functions in (10) have to be determined.

An efficient solution to this problem, proposed in [5, 6], is obtained by assuming so-called *spherically invariant random processes* (SIRPs), e.g., [11]. The general form of correlated SIRPs of D-th order is given with a properly chosen function $f_D(\cdot)$ by

$$\hat{p}_{y_p,D}(\mathbf{y}_p(j)) = \frac{1}{\sqrt{\pi^D \det(\mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}(i))}} f_D\left(\mathbf{y}_p^{\mathrm{T}}(j)\mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}^{-1}(i)\mathbf{y}_p(j)\right) \tag{11}$$

for the p-th channel, where $\mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}$ denotes the corresponding auto-correlation matrix with the corresponding number of lags. These models are representative for a wide class of stochastic processes. Speech signals in particular can very accurately be represented by SIRPs [11]. A great advantage arising from the SIRP model is that multivariate pdfs can be derived analytically from the corresponding univariate pdf together with the (lagged) correlation matrices. The function $f_D(\cdot)$ can thus be calculated from the well-known univariate models for speech, e.g., the Laplacian density. Using the chain rule, the corresponding score function (10b) can be derived from (11), as detailed in [5, 6].

Note that the *multivariate Gaussian* is a special case of a SIRP and thus, most of the popular algorithms based on second-order statistics (SOS) represent special cases of the corresponding algorithms based on SIRPs [5, 6]. In both cases, by transforming the model into the DFT domain, various links to novel and existing popular frequency-domain algorithms can be established [1, 6].

Another very powerful family of stochastic models are given by the theory of multivariate *robust statistics* which was recently introduced in TRINICON and related to the SIRP model in [12].

Finally, it should be noted that in addition to the model selection the choice of estimation procedure for the corresponding *stochastic model parameters* (e.g., correlation matrices in (11), scaling parameter in [12], etc.) is another important design consideration. Similar to the estimation of correlation matrices in linear prediction problems [13] we have to distinguish in actual implementations between the more accurate so-called *covariance method* and the approximative *correlation method* leading to a lower complexity, e.g., [10].

A natural but in practice somewhat more sophisticated extension would be to incorporate more macroscopic temporal dependencies of the stochastic model parameters by a codebook approach or another stochastic process leading to a hidden-markov-like structure (e.g., [9]).

4 Applications to Signal Processing Problems for Speech Capture

The various classes of array processing problems mentioned in Sect. 1 are also closely related to the choice of stochastic *signal* model [7] as we illustrate in this section. Another differentiating factor mentioned in Sect. 1 is the degree of "blindness" which is defined by the amount of prior knowledge on the mixing *system*, and thus, via the relation given by the ideal solution, on the demixing system [12]. In this way, all major classes of MIMO adaptation problems having a unique solution may be addressed systematically by the TRINICON framework.

4.1 Direct adaptive MIMO filtering problems

4.1.1 Blind Source Separation (BSS)

The separation variant of the generic natural gradient update (10a) follows immediately by setting

$$\hat{p}_{s,PD}(\mathbf{y}(j)) \overset{\text{(BSS)}}{=} \prod_{q=1}^{P} \hat{p}_{y_q,D}(\mathbf{y}_q(j)), \qquad (12)$$

in (10b). i.e., the original source signals may be colored, but are assumed to be mutually stochastically independent to each other. Figure 3(a) illustrates this model of desired signal statistics for the special case of Gaussian sources, i.e., second-order statistics, in terms of the desired $PD \times PD$ correlation matrix $\hat{\mathbf{R}}_{\mathbf{ss}} = \text{bdiag}\,\hat{\mathbf{R}}_{\mathbf{yy}}$. Thus, by minimizing $\mathcal{J}(m,\mathbf{W})$, all cross-correlations for D time-lags are reduced and will ideally vanish, while the auto-correlations are untouched to preserve the structure of the individual signals. This general class of broadband BSS algorithms leads to very robust practical solutions even for a large number of filter taps due to an inherent normalization of the coefficient update by the auto-correlation matrices [8, 12]. Note that there are also various efficient approximations of this algorithm, e.g, [10, 14] still preserving its broadband nature (and thus avoiding the problem of bin-wise permutations in narrowband approaches [8]) but with a reduced computational complexity that have allowed some of the first real-time systems of this kind on regular PC platforms. Moreover, a close link has been established to various popular frequency-domain algorithms [6, 8].

4.1.2 Blind System Identification (BSI) and Acoustic Source Localization for Multiple Sources in Reverberant Environments

The efficient broadband BSS algorithms also form a powerful basis for blind system identification and acoustic source localization. Under certain conditions [1], such as a suitable choice of filter length L, the broadband separation problem can be turned into a MIMO system identification problem, and vice versa. This important link has led to practical localization systems which are not only suitable for multiple simultaneously active sources, but also generalize powerful single-source approaches for reverberant environments, e.g., [15], to this case [1, 16].

4.1.3 Supervised and Semi-Blind System Identification and Interference Cancellation

The field of supervised adaptive filtering (see Sect. 1) has a much longer history and has been treated largely independently from the blind case so far [4]. The general broadband approach of TRINICON allows to connect both theories in a systematic way by introducing the respective prior knowledge on the mixing system, as recently demonstrated in [12] for the important application of acoustic echo cancellation. Analogous considerations are possible for other applications, e.g., adaptive beamforming. It was shown that this connection opens up a great potential for various synergy effects and improved algorithms.

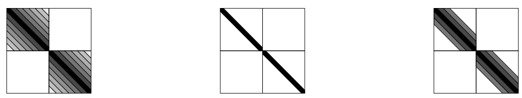
4.2 Inverse adaptive MIMO filtering problems

4.2.1 Multichannel Blind Deconvolution (MCBD)

Traditionally, ICA-based MCBD algorithms assume i.i.d. source models, e.g., [17]. This corresponds to a complete factorization of the hypothesized source model $\hat{p}_{s,PD}(\cdot)$ in (10b), i.e.,

$$\hat{p}_{s,PD}(\mathbf{y}(j)) \overset{\text{(MCBD)}}{=} \prod_{q=1}^{P}\prod_{d=1}^{D} \hat{p}_{y_q,1}(\mathbf{y}_q(j-d)). \qquad (13)$$

In the SOS case, this translates to a complete whitening of the output signals by not only applying a joint de-cross-correlation, but also a joint de-auto-correlation, i.e., $\hat{\mathbf{R}}_{\mathbf{ss}} = \text{diag}\,\hat{\mathbf{R}}_{\mathbf{yy}}$ over multiple time-instants, as illustrated in Fig. 3 (b).



(a) BSS (b) MCBD (c) MCBPD

Figure 3: Desired correlation matrices $\hat{\mathbf{R}}_{\mathbf{ss}}$ for BSS, MCBD, and MCBPD with TRINICON in the SOS case.

4.2.2 Multichannel Blind Partial Deconvolution (MCBPD)

Signal sources which are non i.i.d. should not become i.i.d. at the output of the blind adaptive filtering stage. Therefore, their statistical dependencies should be preserved, i.e., the adaptation algorithm has to distinguish between the statistical dependencies within the source signals (e.g., by the vocal tract for speech), and the statistical dependencies introduced by the mixing system $\mathbf{H}$ (e.g., a reverberant room). We denote the corresponding generalization of the traditional MCBD technique as *MultiChannel Blind Partial Deconvolution* (MCBPD) [7]. Equations (10) inherently contain a statistical source model (signal properties (i)-(iii) in Sect. 2), expressed by the multivariate densities, and thus provide all necessary requirements for the MCBPD approach.

Ideally, only the influence of the room acoustics should be minimized. In the typical example of speech dereverberation the auto-correlation structure of the speech signals can be taken into account, as shown in Fig. 3 (c). While the room acoustics influences all off-diagonals, the effect of the vocal tract is concentrated in the first few off-diagonals around the main diagonal. These first off-diagonals of $\hat{\mathbf{R}}_{\mathbf{yy}}$ are now taken over into $\hat{\mathbf{R}}_{\mathbf{ss}}$, as shown in Fig. 3 (c). Note that there is a close link to linear prediction techniques [13] which gives guidelines for the number of lags to be preserved.

5 Discussion of Different Sylvester Constraint Realizations

As mentioned in Sect. 2.2 the actual realization of the Sylvester constraint is fundamental to the properties of the resulting algorithm. So far there are two particularly popular and simple realizations of ($\mathcal{SC}$) leading to two different classes of algorithms [10]:

(1) Computing only the *first column* of each channel of the update matrix to obtain the new coefficient matrix $\check{\mathbf{W}}$. We denote this method as ($\mathcal{SC}_\mathrm{C}$).

(2) Computing only the *L-th row* of each channel of the update matrix to obtain the new coefficient matrix $\check{\mathbf{W}}$. We denote this method as ($\mathcal{SC}_\mathrm{R}$).

From Fig. 2 in Sect. 2.2 we now see that these realizations represent certain approximations by neglecting some of the elements within the summation process in ($\mathcal{SC}$).

In [10, 18] it was shown that with both of the above approximations the update process is significantly simplified. Moreover, based on $\mathcal{SC}_\mathrm{C}$ and $\mathcal{SC}_\mathrm{R}$ we established links for the special case of MCBD to various existing natural gradient-based algorithms in [18] by additionally distinguishing between the covariance method and correlation method as mentioned in Sect. 3.

However, in general, both $\mathcal{SC}_\mathrm{C}$ and $\mathcal{SC}_\mathrm{R}$ require some tradeoff in the algorithm performance. While $\mathcal{SC}_\mathrm{C}$ may provide a potentially more robust convergence behaviour, it will not work for arbitrary source positions (e.g., in the case of two sources, they are required to be located in different half-planes w.r.t. the orientation of the microphone array), which is in contrast to the more versatile $\mathcal{SC}_\mathrm{R}$ [10].

In this section we compare the performance of the different Sylvester constraint realizations by using the generic SOS natural gradient algorithm for BSS. For this comparison the coefficient adaptation is performed in an offline fashion and the correlation matrices are estimated by the more accurate covariance method. Two different setups in a moderately reverberant room for $L = 256$ are examined: (a) two sources positioned in different half planes (at $\pm 70°$) and (b) two sources positioned in the same half plane (at $+45°$, $+90°$).

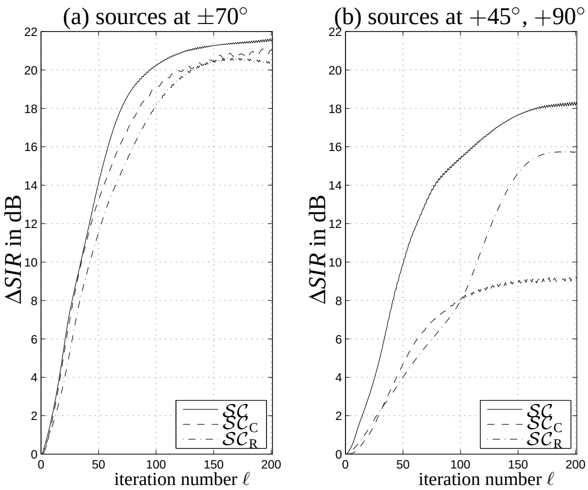


Figure 4: Comparison of different Sylvester constraint implementations with the offline generic SOS natural gradient algorithm.

In the results for the two-sided setup in Fig. 4(a) it can be seen that the original Sylvester constraint operator $\mathcal{SC}$ achieves the highest separation performance in terms of signal-to-interference ratio improvement ΔSIR. Approximating $\mathcal{SC}$ by the row Sylvester constraint $\mathcal{SC}_\mathrm{R}$ or the column Sylvester constraint $\mathcal{SC}_\mathrm{C}$ only leads to a slight degradation of the separation performance. In the one-sided scenario (b) the original Sylvester constraint operator $\mathcal{SC}$ exhibits again the highest separation performance. The approximation by $\mathcal{SC}_\mathrm{R}$ still achieves reasonable separation. The application of the column Sylvester constraint $\mathcal{SC}_\mathrm{C}$ is not recommendable for one-sided setups or for the case $P > 2$ as mentioned above.

6 Conclusions

TRINICON provides a versatile tool to the design of adaptive systems. The "top-down approach" of this framework has led to various recent advances in the field, but also shows opportunities for many new developments in the future. The Sylvester constraint represents one of the generic elements of the framework which unify previous algorithmic results from the existing literature and facilitate the development of new efficient algorithms.

References

[1] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON-based blind system identification with application to multiple-source localization and separation," in S. Makino, T.-W. Lee, and S. Sawada (eds.), *Blind Speech Separation*, Springer-Verlag, Berlin/Heidelberg, pp. 101-147, Sept. 2007.

[2] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol 36, no. 2, pp. 145-152, Feb. 1988.

[3] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley & Sons, Inc., New York, 2001.

[4] S. Haykin, *Adaptive Filter Theory*, 4th ed., Prentice-Hall, Englewood Cliffs, NJ, 2002.

[5] H. Buchner, R. Aichner, and W. Kellermann, "Blind source separation for convolutive mixtures exploiting nongaussianity, nonwhiteness, and nonstationarity," in *Proc. Int. Workshop Acoustic Echo and Noise Control (IWAENC)*, Kyoto, Japan, pp. 223-226, Sept. 2003.

[6] H. Buchner, R. Aichner, and W. Kellermann, "Blind source separation for convolutive mixtures: A unified treatment," in Y. Huang and J. Benesty (eds.), *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Kluwer Academic Publ., Boston, pp. 255-293, Feb. 2004.

[7] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON: A versatile framework for multichannel blind signal processing," in *Proc. IEEE ICASSP*, Montreal, Canada, vol. 3, pp. 889-892, May 2004.

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

[9] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley & Sons, New York, 1991.

[10] 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.

[11] H. Brehm and W. Stammler, "Description and generation of spherically invariant speech-model signals," *Signal Processing*, vol. 12, pp. 119-141, 1987.

[12] H. Buchner and W. Kellermann, "A fundamental relation between blind and supervised adaptive filtering illustrated for blind source separation and acoustic echo cancellation," *Conf. Rec. HSCMA* , Trento, Italy, May 2008.

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

[14] R. Aichner, H. Buchner, and W. Kellermann, "Exploiting narrowband efficiency for broadband convolutive blind source separation," *EURASIP Journal on Applied Signal Processing*, vol. 2007, pp. 1-9, Sept. 2006.

[15] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *J. Acoust. Soc. Am.*, vol. 107, pp. 384-391, Jan. 2000.

[16] H. Buchner, R. Aichner, J. Stenglein, H. Teutsch, and W. Kellermann, "Simultaneous localization of multiple sound sources using blind adaptive MIMO filtering," in *Proc. IEEE ICASSP*, Philadelphia, PA, USA, Mar. 2005.

[17] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing*, Wiley & Sons, Chichester, UK, 2002.

[18] R. Aichner, H. Buchner, and W. Kellerman, "On the causality problem in time-domain blind source separation and deconvolution algorithms," in *Proc. IEEE ICASSP*, Philadelphia, PA, Mar. 2005, pp. 181-184.