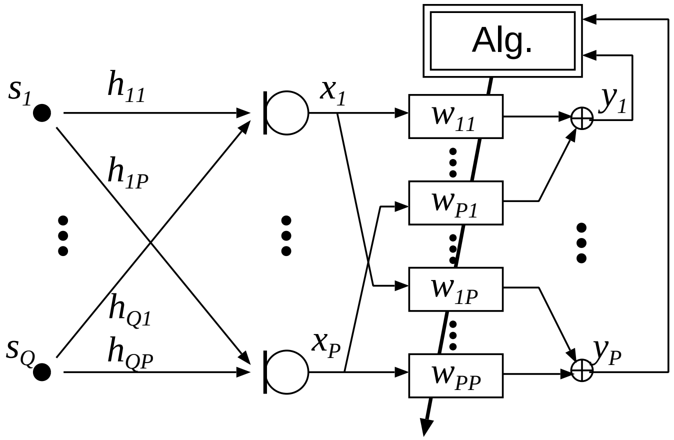
Relation Between Blind System Identification and Convolutive Blind Source Separation

Herbert Buchner, Robert Aichner, Walter Kellermann
{buchner,aichner,wk}@LNT.de

March 18, 2005

Multimedia Communications and Signal Processing
Telecommunications Laboratory
University of Erlangen-Nuremberg

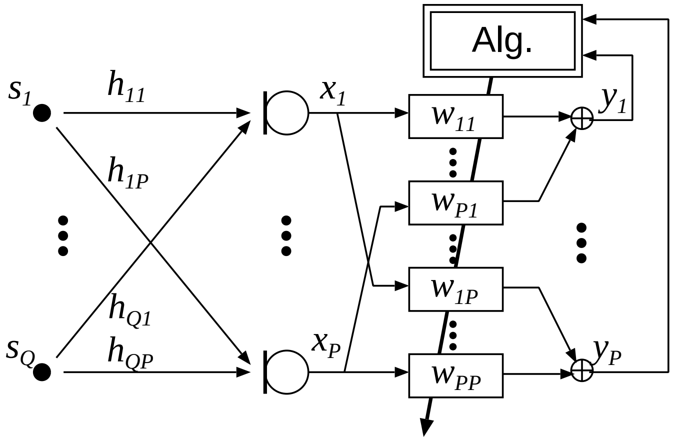
MIMO FIR model $(Q \leq P)$



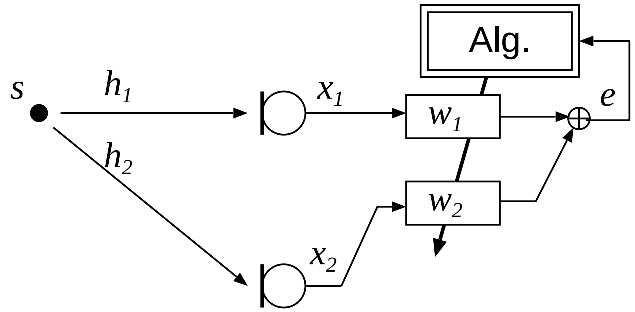
Blind signal processing problems:

- Blind source separation (BSS) for convolutive mixtures
- Multichannel blind deconvolution and dereverberation (MCB[P]D)

MIMO FIR model $(Q \leq P)$



SIMO FIR model $(Q = 1)$

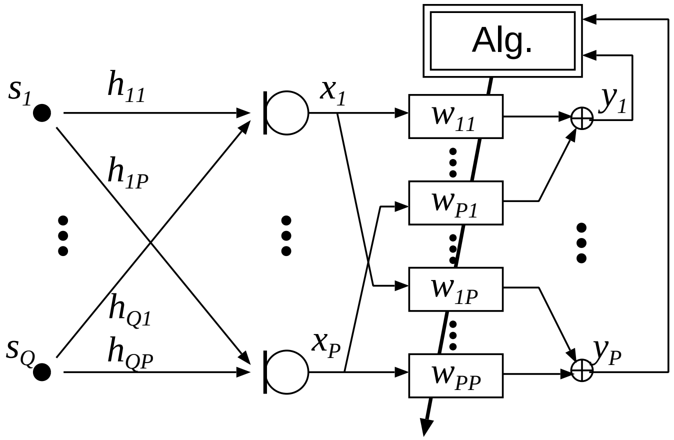


Blind signal processing problems:

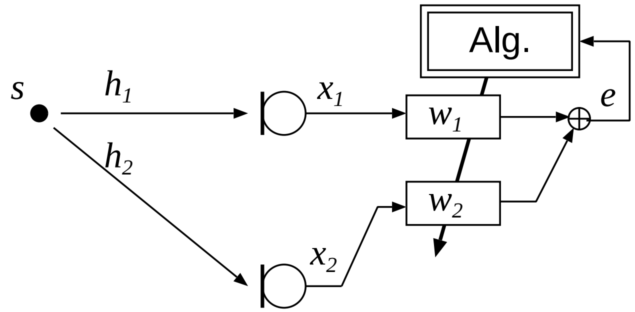
- Blind source separation (BSS) for convolutive mixtures
- Multichannel blind deconvolution and dereverberation (MCB[P]D)
- Blind system identification (BSI)

Introduction

MIMO FIR model $(Q \leq P)$



SIMO FIR model $(Q = 1)$



Blind signal processing problems:

- Blind source separation (BSS) for convolutive mixtures
- Multichannel blind deconvolution and dereverberation (MCB[P]D)
- Blind system identification (BSI)

Here we consider:

- What is the relation between BSS and the SIMO-based BSI approach?
- What is the optimum convolutive BSS solution? Ambiguities?

Contents

Known Ambiguities of Blind Adaptive MIMO Filtering

BSS: separation by forcing the outputs to be mutually independent

- **instantaneous BSS:**
  permutation and scaling

Known Ambiguities of Blind Adaptive MIMO Filtering

BSS: separation by forcing the outputs to be mutually independent

- **instantaneous BSS:**
  permutation and scaling

- **narrowband (DFT-domain) convolutive BSS:**
  internal bin-wise permutations and scaling in each freq. bin $\Rightarrow$ filtering
  $\longrightarrow$ *additional measures*: bin-wise reordering, e.g., [Sawada et al., 2004] and minimum distortion principle [Matsuoka et al., 2001]

Known Ambiguities of Blind Adaptive MIMO Filtering

BSS: separation by forcing the outputs to be mutually independent

- **instantaneous BSS:**
  permutation and scaling

- **narrowband (DFT-domain) convolutive BSS:**
  internal bin-wise permutations and scaling in each freq. bin $\Rightarrow$ filtering
  $\longrightarrow$ *additional measures*: bin-wise reordering, e.g., [Sawada et al., 2004] and minimum distortion principle [Matsuoka et al., 2001]

- **convolutive BSS via (time-domain) MCBD:**
  permutation and whitening of output
  $\longrightarrow$ *additional measures*: minimum distortion principle [Matsuoka et al., 2001] or linear prediction [Douglas et al., 2001]

Known Ambiguities of Blind Adaptive MIMO Filtering

BSS: separation by forcing the outputs to be mutually independent

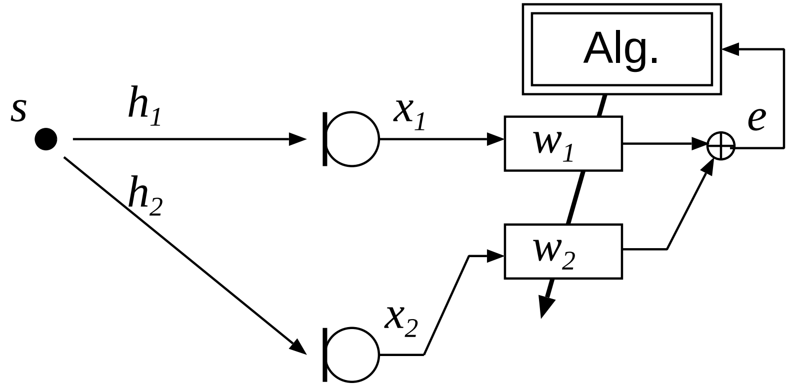
- **instantaneous BSS:**
  permutation and scaling

- **narrowband (DFT-domain) convolutive BSS:**
  internal bin-wise permutations and scaling in each freq. bin ⇒ filtering
  → *additional measures*: bin-wise reordering, e.g., [Sawada et al., 2004] and minimum distortion principle [Matsuoka et al., 2001]

- **convolutive BSS via (time-domain) MCBD:**
  permutation and whitening of output
  → *additional measures*: minimum distortion principle [Matsuoka et al., 2001] or linear prediction [Douglas et al., 2001]

- **broadband convolutive BSS:**
  permutation and scaling/filtering ??? ⇒ addressed here

SIMO-based BSI

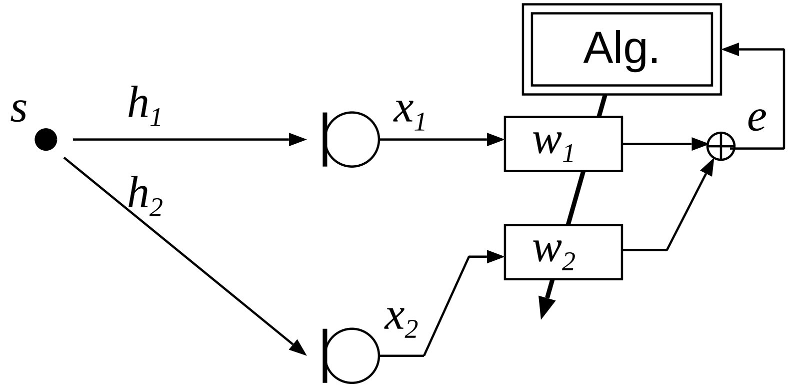


Approach: minimize $E\{e^2(n)\}$
with suitable coefficient initialization

Ideal solution for $e(n) = 0$:

$$h_1(n) * w_1(n) = -h_2(n) * w_2(n)$$

SIMO-based BSI



Approach: minimize $E\{e^2(n)\}$
with suitable coefficient initialization

Ideal solution for $e(n) = 0$:

$$h_1(n) * w_1(n) = -h_2(n) * w_2(n)$$

...in z-domain with FIR model structure (zeros $z_{0H_i,\nu}$, $z_{0W_i,\mu}$, gains A_{H_i}, A_{W_i}):

$$A_{H_1} \prod_{\nu=1}^{M-1} (z - z_{0H_1,\nu}) A_{W_1} \prod_{\mu=1}^{L-1} (z - z_{0W_1,\mu}) = -A_{H_2} \prod_{\nu=1}^{M-1} (z - z_{0H_2,\nu}) A_{W_2} \prod_{\mu=1}^{L-1} (z - z_{0W_2,\mu}) \tag{1}$$

Assumption: $H_1(z)$ and $H_2(z)$ have no common zeros.
$\Rightarrow$ optimum filters: $W_1(z) = \alpha H_2(z)$ and $W_2(z) = -\alpha H_1(z)$

- for $L \leq M$: arbitrary *scaling* $\alpha = \frac{A_{W_1}}{A_{H_2}} = \frac{A_{W_2}}{A_{H_1}}$
- for $L > M$: arbitrary *filtering*

Convolutive Broadband BSS

For the MIMO case, a different optimization criterion has to be used leading to, e.g., generic SOS-based coefficient update (see also → keynote talk, Kellermann et al.):

$$\Delta \mathbf{W}(m) = 2 \sum_{i=0}^{\infty} \beta(i,m) \mathbf{W} \left\{ \mathbf{R}_{\mathbf{yy}} - \mathrm{bdiag}_D \, \mathbf{R}_{\mathbf{yy}} \right\} \mathrm{bdiag}_D^{-1} \, \mathbf{R}_{\mathbf{yy}}$$

($\mathbf{W}$: matrix of adaptive filter coefficients in Sylvester structure)

Convolutive Broadband BSS

For the MIMO case, a different optimization criterion has to be used leading to, e.g., generic SOS-based coefficient update (see also → keynote talk, Kellermann et al.):

$$\Delta \mathbf{W}(m) = 2 \sum_{i=0}^{\infty} \beta(i,m) \mathbf{W} \left\{ \mathbf{R}_{\mathbf{yy}} - \mathrm{bdiag}_D \, \mathbf{R}_{\mathbf{yy}} \right\} \mathrm{bdiag}_D^{-1} \, \mathbf{R}_{\mathbf{yy}}$$

($\mathbf{W}$: matrix of adaptive filter coefficients in Sylvester structure)

Overall system in Sylvester structure: $\mathbf{C} = \mathbf{H}\mathbf{W}$
With $\mathbf{R}_{\mathbf{yy}} = \mathbf{C}^T \mathbf{R}_{\mathbf{ss}} \mathbf{C}$ follows:

$$\Delta \mathbf{C}(m) = 2 \sum_{i=0}^{\infty} \beta(i,m) \mathbf{C} \left\{ \mathbf{C}^T \mathbf{R}_{\mathbf{ss}} \mathbf{C} \, \mathrm{bdiag}_D^{-1} \left\{ \mathbf{C}^T \mathbf{R}_{\mathbf{ss}} \mathbf{C} \right\} - \mathbf{I} \right\}$$

Convolutive Broadband BSS

For the MIMO case, a different optimization criterion has to be used leading to, e.g., generic SOS-based coefficient update (see also → keynote talk, Kellermann et al.):

$$\Delta \mathbf{W}(m) = 2 \sum_{i=0}^{\infty} \beta(i,m) \mathbf{W} \left\{ \mathbf{R}_{\mathbf{yy}} - \mathrm{bdiag}_D \, \mathbf{R}_{\mathbf{yy}} \right\} \mathrm{bdiag}_D^{-1} \, \mathbf{R}_{\mathbf{yy}}$$

($\mathbf{W}$: matrix of adaptive filter coefficients in Sylvester structure)

Overall system in Sylvester structure: $\mathbf{C} = \mathbf{HW}$

With $\mathbf{R}_{\mathbf{yy}} = \mathbf{C}^T \mathbf{R}_{\mathbf{ss}} \mathbf{C}$ follows:

$$\Delta \mathbf{C}(m) = 2 \sum_{i=0}^{\infty} \beta(i,m) \mathbf{C} \left\{ \mathbf{C}^T \mathbf{R}_{\mathbf{ss}} \mathbf{C} \, \mathrm{bdiag}_D^{-1} \left\{ \mathbf{C}^T \mathbf{R}_{\mathbf{ss}} \mathbf{C} \right\} - \mathbf{I} \right\}$$

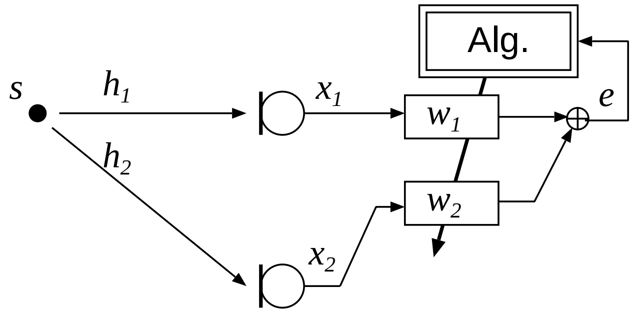
Equilibria: It can be shown that $\Delta \mathbf{C} = 0$ leads for $D = L$ time lags to the desired solution [Buchner et al., TR-SAP 2005]

$$\mathrm{boff}\{\mathbf{C}\} = \mathbf{0}, \tag{2}$$

i.e., ideally, all cross-channels of the overall system will be zero.

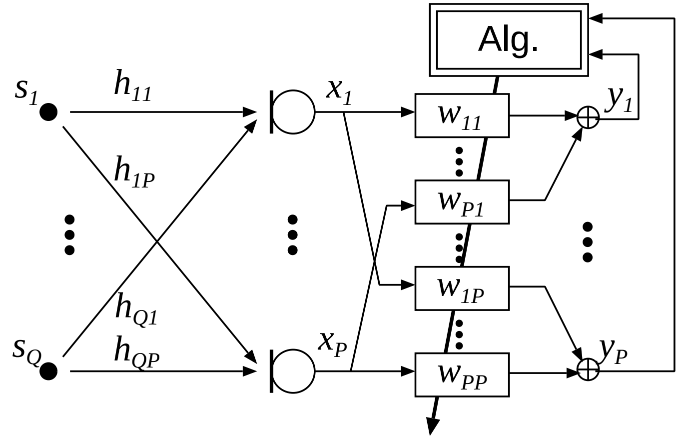
Comparison of Ideal Solutions

SIMO FIR model



$$h_1 * w_1 = -h_2 * w_2 \qquad (3)$$

MIMO FIR model



for two sources (from $\text{boff}\{\mathbf{C}\} = \mathbf{0}$):

$$h_{11} * w_{12} = -h_{12} * w_{22} \qquad (4)$$

$$h_{21} * w_{11} = -h_{22} * w_{21} \qquad (5)$$

- (4) and (5) are the generalization of (3) for two sources
- broadband BSS performs blind MIMO system identification
- no ambiguity in filtering with broadband algorithms for optimum filter length L

Optimum BSS Solution

As in the SIMO case, we obtain for two sources and $L = L_{\mathrm{opt,BSS}} = M$:

$$\mathbf{W} = \begin{bmatrix} \alpha_1 \mathbf{H}_{22} & -\alpha_2 \mathbf{H}_{12} \\ -\alpha_1 \mathbf{H}_{21} & \alpha_2 \mathbf{H}_{11} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{22} & -\mathbf{H}_{12} \\ -\mathbf{H}_{21} & \mathbf{H}_{11} \end{bmatrix} \begin{bmatrix} \alpha_1 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha_2 \mathbf{I} \end{bmatrix}$$

Optimum BSS Solution

As in the SIMO case, we obtain for two sources and $L = L_{\mathrm{opt,BSS}} = M$:

$$\mathbf{W} = \begin{bmatrix} \alpha_1 \mathbf{H}_{22} & -\alpha_2 \mathbf{H}_{12} \\ -\alpha_1 \mathbf{H}_{21} & \alpha_2 \mathbf{H}_{11} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{22} & -\mathbf{H}_{12} \\ -\mathbf{H}_{21} & \mathbf{H}_{11} \end{bmatrix} \begin{bmatrix} \alpha_1 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha_2 \mathbf{I} \end{bmatrix}$$

In general (for $L = L_{\mathrm{opt,BSS}}$):

$$\mathbf{W} \propto \mathrm{badj}\left\{\mathbf{H}\right\} \tag{6}$$

blind source separation

- "intra-channel **identification**"
- "inter-channel inversion"

Optimum BSS Solution

As in the SIMO case, we obtain for two sources and $L = L_{\mathrm{opt,BSS}} = M$:

$$\mathbf{W} = \begin{bmatrix} \alpha_1 \mathbf{H}_{22} & -\alpha_2 \mathbf{H}_{12} \\ -\alpha_1 \mathbf{H}_{21} & \alpha_2 \mathbf{H}_{11} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{22} & -\mathbf{H}_{12} \\ -\mathbf{H}_{21} & \mathbf{H}_{11} \end{bmatrix} \begin{bmatrix} \alpha_1 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha_2 \mathbf{I} \end{bmatrix}$$

In general (for $L = L_{\mathrm{opt,BSS}}$):

$$\mathbf{W} \propto \mathrm{badj}\left\{\mathbf{H}\right\} \quad (6)$$

blind source separation

- "intra-channel **identification**"
- "inter-channel inversion"

blind deconvolution

- "intra-channel inversion"
- "inter-channel inversion"

Optimum BSS Solution

As in the SIMO case, we obtain for two sources and $L = L_{\text{opt,BSS}} = M$:

$$\mathbf{W} = \begin{bmatrix} \alpha_1 \mathbf{H}_{22} & -\alpha_2 \mathbf{H}_{12} \\ -\alpha_1 \mathbf{H}_{21} & \alpha_2 \mathbf{H}_{11} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{22} & -\mathbf{H}_{12} \\ -\mathbf{H}_{21} & \mathbf{H}_{11} \end{bmatrix} \begin{bmatrix} \alpha_1 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha_2 \mathbf{I} \end{bmatrix}$$

In general (for $L = L_{\text{opt,BSS}}$):

$$\mathbf{W} \propto \text{badj}\{\mathbf{H}\} \tag{6}$$

blind source separation

- "intra-channel **identification**"
- "inter-channel inversion"

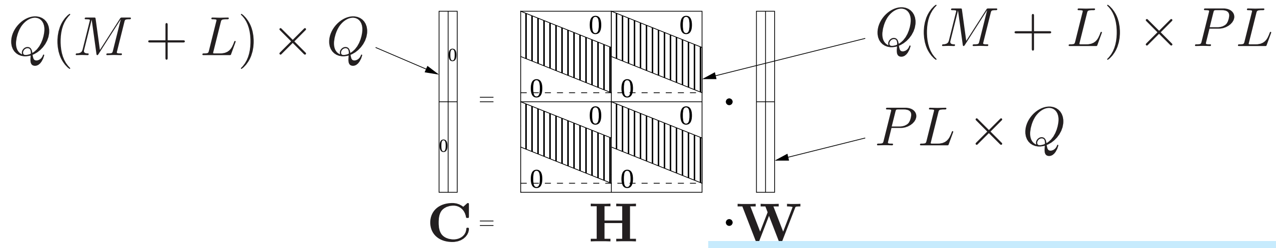
blind deconvolution

- "intra-channel inversion"
- "inter-channel inversion"

Remaining question: $L_{\text{opt,BSS}}$ for more than two channels?

Optimum Filter Length

... by considering matrix dimensions



For deconvolution: MINT [Miyoshi, 1988] $\Rightarrow$ $L_{\text{opt,MCBD}} = \frac{Q}{P-Q}(M-1)$.

For BSS: using $\text{boff}\{\mathbf{C}\} = \mathbf{0}$ (see paper) $\Rightarrow$ $L_{\text{opt,BSS}} = \frac{Q-1}{P-Q+1}M$.

- ideally for MCBD: $P > Q$ / for BSS: $P \geq Q$
- in practice: $L_{\text{opt,BSS}} < L_{\text{opt,MCBD}}$
- **for $P = Q = 2$:** $L_{\text{opt,BSS}} = M$ as in the SIMO case

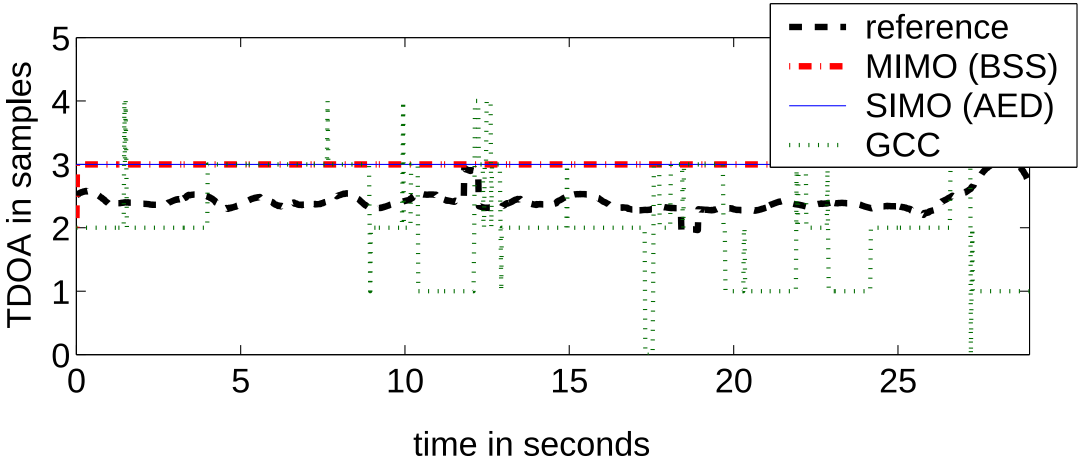
Simultaneous Localization of **Multiple** Sound Sources in **Reverberant** Environments

Setup:
Two speakers
recorded in a TV studio
$T_{60} \approx 700\text{ms},\ f_s = 48\text{kHz}$

Algorithms for estimation of time difference(s) of arrival (TDOA):

- Generalized cross-correlation (GCC) with phase-transform (PHAT) weighting with VAD
- SIMO-based BSI + VAD
- Blind MIMO identification based on BSS (second-order version)

one speaker (fixed position)



Application: Acoustic Source Localization

Simultaneous Localization of **Multiple** Sound Sources in **Reverberant** Environments

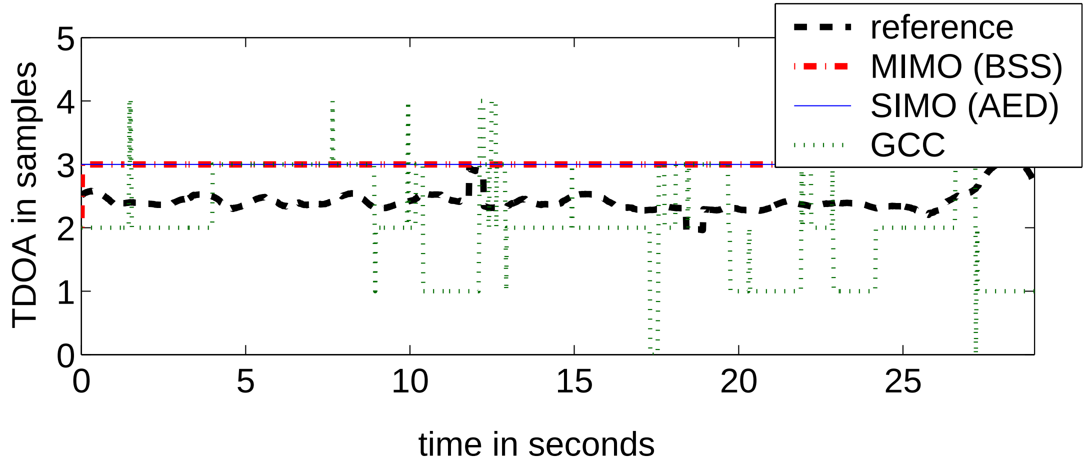
Setup:

Two speakers
recorded in a TV studio
$T_{60} \approx 700\text{ms}, \ f_s = 48\text{kHz}$

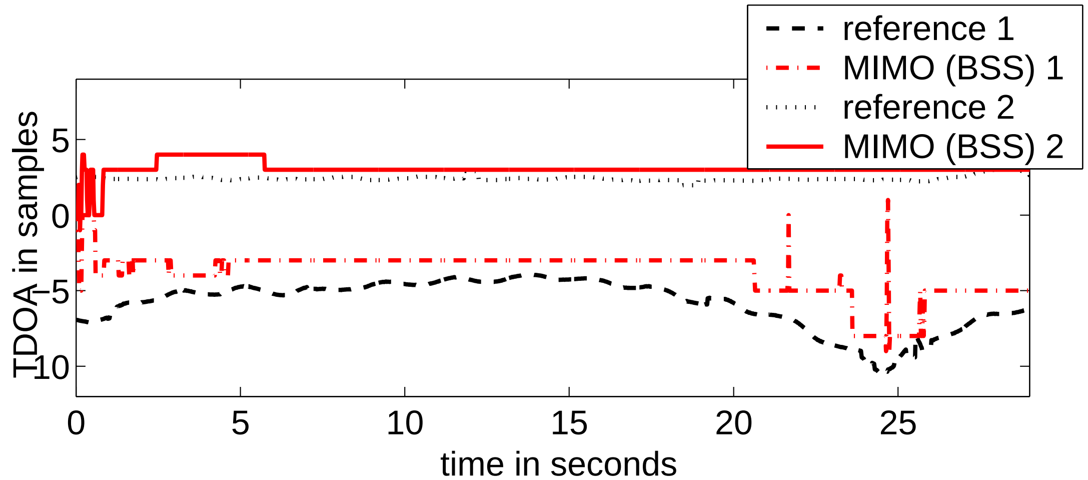
Algorithms for estimation of time difference(s) of arrival (TDOA):

- Generalized cross-correlation (GCC) with phase-transform (PHAT) weighting with VAD
- SIMO-based BSI + VAD
- Blind MIMO identification based on BSS (second-order version)

one speaker (fixed position)



two speakers simultaneously (fixed pos./moving)



Conclusions

Convolutive broadband BSS corresponds to a channel-wise blind MIMO system identification.

Consequences for broadband convolutive BSS:

- Ambiguity of filtering avoidable
- Whitening problem avoidable

Consequence for broadband convolutive BSI:

- Blind adaptive MIMO filtering as known from BSS also allows for new applications, such as simultaneous localization of multiple sound sources in reverberant environments (see also → ICASSP 2005)