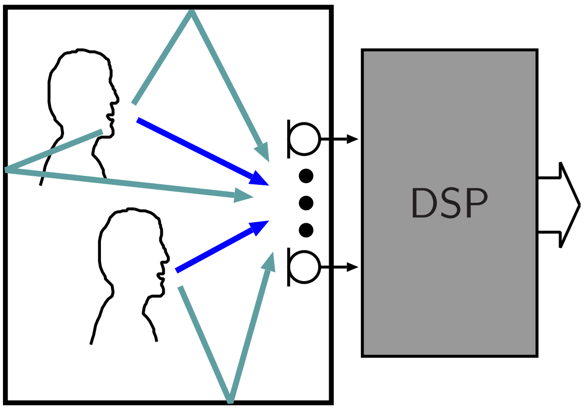
Blind Source Separation of Audio Signals

Herbert Buchner
Technische Universität Berlin
Lecture Machine Learning II

May 3, 2011

Introduction - Broadband Adaptive MIMO Filtering

Motivation: Signal aquisition by sensor arrays in convolutive environments
Example: Speech capture by microphone arrays in reverberant rooms

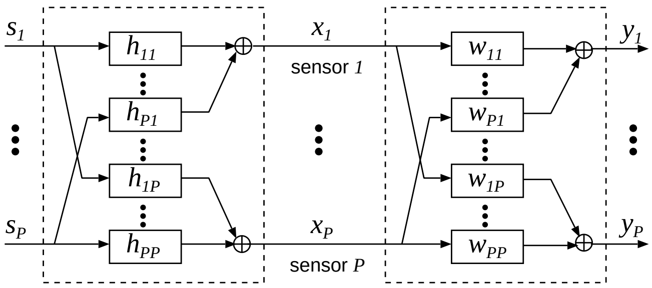


Applications: Teleconferencing, hands-free speech recognition, modern hearing aids,...

Introduction - Broadband Adaptive MIMO Filtering

Motivation: Signal aquisition by sensor arrays in convolutive environments

MIMO model



mixing system $\mathbf{H}$ demixing system $\mathbf{W}$

$$x_p(n) = \sum_{q=1}^{P} \sum_{\kappa=0}^{M-1} h_{qp}(\kappa) s_q(n-\kappa),$$

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

Adaptive Signal Processing Tasks:

- **Signal Separation**
  $\check{\check{\mathbf{y}}} = \check{\check{\mathbf{W}}} * \check{\check{\mathbf{H}}} * \check{\check{\mathbf{s}}} \stackrel{!}{=} \mathrm{diag}\{\check{\check{\mathbf{W}}} * \check{\check{\mathbf{H}}}\} * \check{\check{\mathbf{s}}}$
- **Deconvolution**:
  $\check{\check{\mathbf{y}}} = \check{\check{\mathbf{W}}} * \check{\check{\mathbf{H}}} * \check{\check{\mathbf{s}}} \stackrel{!}{=} \alpha \cdot \delta(n-n_0)\mathbf{I} * \check{\check{\mathbf{s}}}$
- **System Identification**:
  estimate system $\check{\check{\mathbf{H}}}$ using signals x_p

- **Blind Estimation:** Propagation paths **and** original source signals are unknown
- **Supervised Estimation:** (Some) source signals and/or side info on paths are known

Introduction - Broadband Adaptive MIMO Filtering

Classification of the linear adaptive filtering problems

	supervised adaptive filtering problems	blind adaptive filtering problems
"direct adaptive filtering problems"	system identification interference cancellation	blind system identification blind source separation/ blind interference cancellation
"inverse adaptive filtering problems"	inverse modeling/equalization linear prediction	blind (partial) deconvolution linear prediction

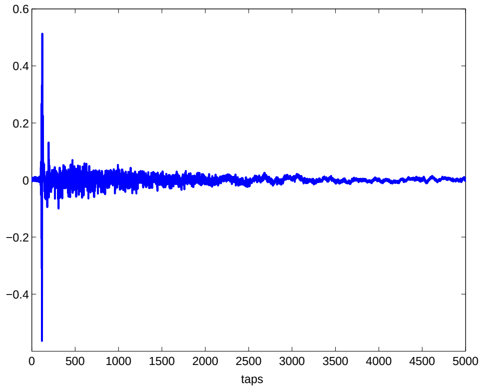
Semi-blind adaptive filtering:
prominent example: adaptive beamforming (=spatial filtering).
Can typically be implemented using (supervised) interference cancellation, but requires prior information on source locations for beamsteering.

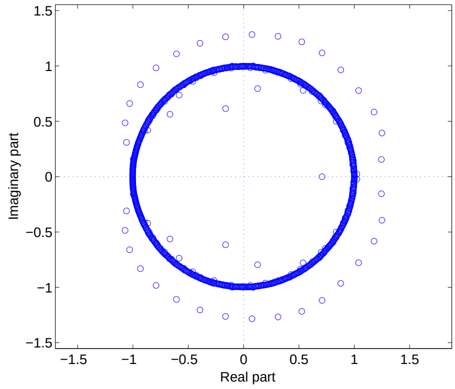
Acoustic Environment

- **Reverberation time T_{60}**
  (sound energy decayed by 60dB)
  - ▷ car $\approx$ 50ms
  - ▷ concert halls $\approx 1 \dots 2$s

- **FIR models**
  - ▷ typically $L_H = T_{60} \cdot f_s / 3$ coefficients
    $\Rightarrow$ hundreds...thousands of coeffs
  - ▷ nonminimum-phase
  - ▷ many zeros close to unit circle

Example: Office 5.5m $\times$ 3m $\times$ 2.8m, $T_{60} \approx$ 300msec, sampling frequency f_s = 12kHz.





- **Excitation by speech and audio signals:**
  $\Rightarrow$ **nonwhiteness, nonstationarity, nongaussianity**

Overview

- **TRINICON - A Generic Concept for Adaptive MIMO Filtering**
  - ▷ Optimization Criterion and Generic Adaptation Algorithms
  - ▷ Incorporation of Stochastic Source Models
- **Applications to Signal Processing Problems for Speech Capture**
  - ▷ Blind Source Separation (BSS) / Interference Suppression
  - ▷ Supervised Separation and System Identification / Echo Cancellation
  - ▷ Blind System Identification (BSI) / Localization of Multiple Sources
  - ▷ * Multich. Blind Deconvolution (MCBD) and
    Multich. Blind Partial Deconvolution (MCBPD) / Dereverberation
- **Concluding Remarks**

TRINICON - Optimization Criterion

As a generic framework, **TRINICON** [Buchner et al., 2003-] exploits

TRIple **N**, i.e., **N**onwhiteness, **N**onstationarity, and **N**ongaussianity of $s_q(n)$,
for **I**ndependent component analysis of **CON**volutive mixtures

Optimization Criterion

$$\mathcal{J}(m,\mathbf{W}) = -\sum_{i=0}^{\infty} \beta(i,m) \frac{1}{N} \sum_{j=iN_L}^{iN_L+N-1} \left\{ \log(\hat{p}_{s,PD}(\mathbf{y}(j))) - \log(\hat{p}_{y,PD}(\mathbf{y}(j))) \right\}$$

- **Nongaussianity** by minimizing Kullback-Leibler divergence (KLD) between the PD-variate probability density functions with data-dependent parameterizations
  - $\hat{p}_{s,PD}(\mathbf{y})$ for sources (*desired*)
  - $\hat{p}_{y,PD}(\mathbf{y})$ for outputs (*actual*)
- **Nonwhiteness** by simultaneous minimization for D time-lags (blocks of D output values per channel in vector $\mathbf{y}$)
- **Nonstationarity** by simultaneous minimization of N blocks (of D output samples per channel)

Windowing $\beta(i,m)$ defines online, block-online, or offline adaptation

TRINICON - MIMO Filtering of Broadband Signals

Latest D samples at **output** y_q

$2L \times D$ **Sylvester matrix** with D times **FIR filter** impulse response w_{pq} of length L

Latest $2L$ samples of **input** x_p

$$\underbrace{\begin{bmatrix} y_q(n) \\ \vdots \\ y_q(n-D+1) \end{bmatrix}}_{:=\,\mathbf{y}_q(n)} = \sum_{p=1}^{P} \underbrace{\begin{bmatrix} w_{pq,0} & 0 & \cdots & 0 \\ \vdots & w_{pq,0} & \ddots & \vdots \\ w_{pq,L-1} & \vdots & \ddots & 0 \\ 0 & w_{pq,L-1} & \ddots & w_{pq,0} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & w_{pq,L-1} \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}^T}_{:=\,\mathbf{W}_{pq}^T} \underbrace{\begin{bmatrix} x_p(n) \\ \vdots \\ x_p(n-2L+1) \end{bmatrix}}_{:=\,\mathbf{x}_p(n)}$$

Input/output of complete MIMO system:

$$\underbrace{\begin{bmatrix} \mathbf{y}_1(n) \\ \vdots \\ \mathbf{y}_P(n) \end{bmatrix}}_{\mathbf{y}(n)} = \underbrace{\begin{bmatrix} \mathbf{W}_{11} & \cdots & \mathbf{W}_{1P} \\ \vdots & \ddots & \vdots \\ \mathbf{W}_{P1} & \cdots & \mathbf{W}_{PP} \end{bmatrix}^T}_{\mathbf{W}^T} \cdot \underbrace{\begin{bmatrix} \mathbf{x}_1(n) \\ \vdots \\ \mathbf{x}_P(n) \end{bmatrix}}_{\mathbf{x}(n)}$$

TRINICON - Euclidean Gradient-Based Coefficient Update

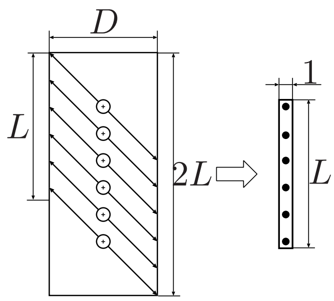
- Block-online update rule:
$$\begin{aligned}\check{\mathbf{W}}^0(m) &:= \check{\mathbf{W}}(m-1),\\ \check{\mathbf{W}}^\ell(m) &= \check{\mathbf{W}}^{\ell-1}(m) - \mu\Delta\check{\mathbf{W}}^\ell(m), \quad \ell = 1,\ldots,\ell_{\max},\\ \check{\mathbf{W}}(m) &:= \check{\mathbf{W}}^{\ell_{\max}}(m)\end{aligned}$$

- Coefficient updates: Euclidean gradient of $\mathcal{J}$ w.r.t. $\check{\mathbf{W}}$ yields

$$\Delta\check{\mathbf{W}}^\ell(m) = \frac{1}{N}\sum_{i=0}^{\infty}\beta(i,m)\,\mathcal{SC}\left\{\sum_{j=iL}^{iL+N-1}\left[\mathbf{x}(j)\mathbf{\Phi}_{s,PD}^T(\mathbf{y}(j)) - \left(\left(\mathbf{W}^{\ell-1}(m)\right)^T\right)^+\right]\right\}$$

Selection of practical algorithms for specific applications by

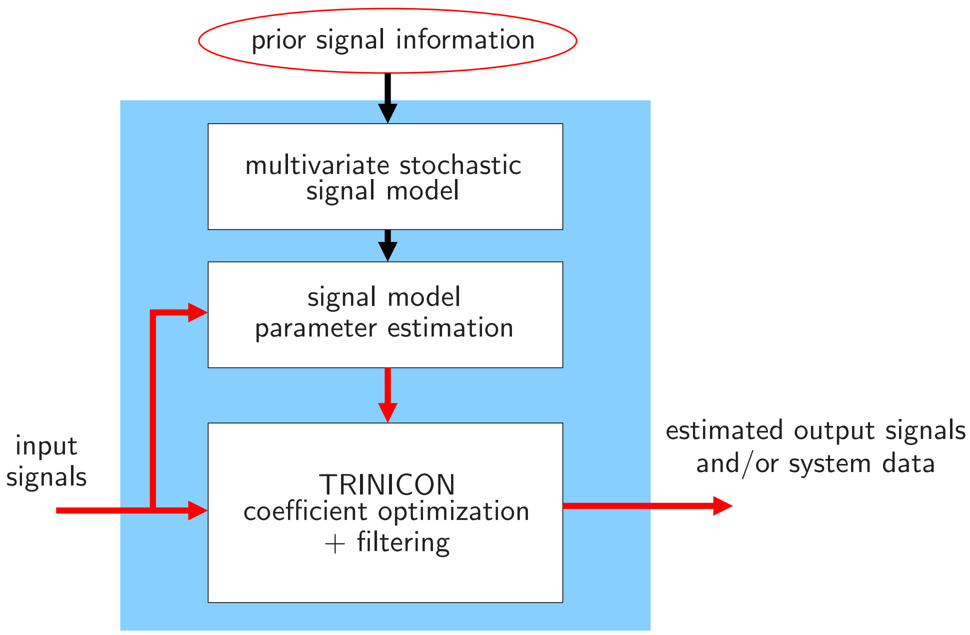
- *Sylvester constraint* $\mathcal{SC}\{\bullet\}$ linking $\mathbf{W}$ (in the cost function) and $\check{\mathbf{W}}$ (in the optimization procedure).
General realization: sums within Sylvester diagonals.



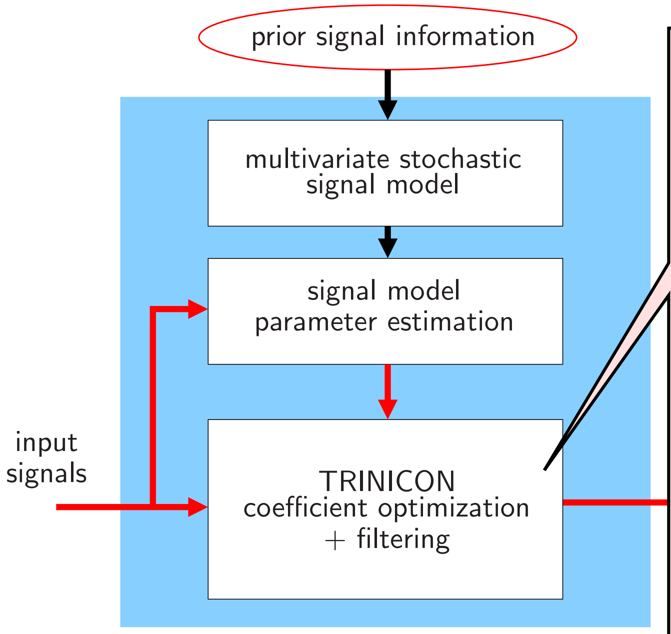
- *multivariate score function*

$$\mathbf{\Phi}_{s,PD}(\mathbf{y}) = -\frac{\partial \log \hat{p}_{s,PD}(\mathbf{y})}{\partial \mathbf{y}}$$

TRINICON-Based Algorithm Design



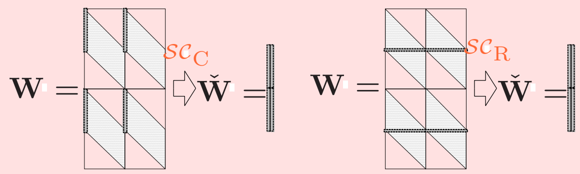
TRINICON-Based Algorithm Design



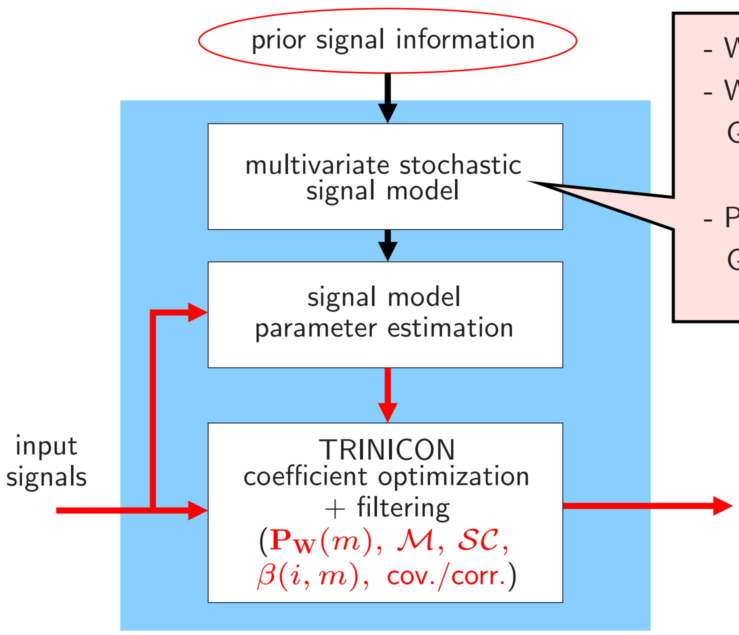
- general: Newton-type, Hessian $\mathbf{P}_{\mathbf{W}}(m)$
- general: on 'arbitrary' manifold $\mathcal{M}$, e.g.,
  - 'natural gradient adaptation' (MIMO)

$$\Delta \mathbf{W} \propto \mathbf{W}\mathbf{W}^T \frac{\partial \mathcal{J}}{\partial \mathbf{W}}$$

  - system structure/application: blind/supervised/semi-blind
  - ...
- realization of Sylvester Constraint $\mathcal{SC}$

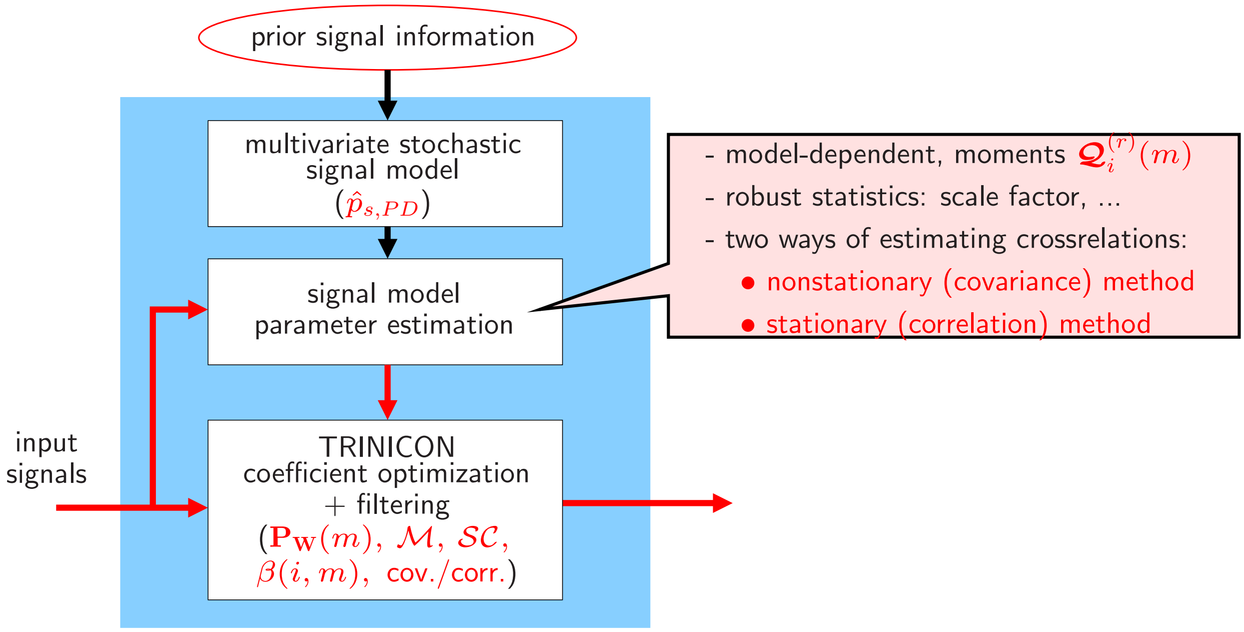


TRINICON-Based Algorithm Design



- Which application?
- Which priors $\hat{p}_{s,PD}$? Which moments?
  Generic concept: Maximum-Entropy principle
  based on variational calculus
- Possible: robust statistics
  Generic concept: Minimum Fisher information
  based on variational calculus

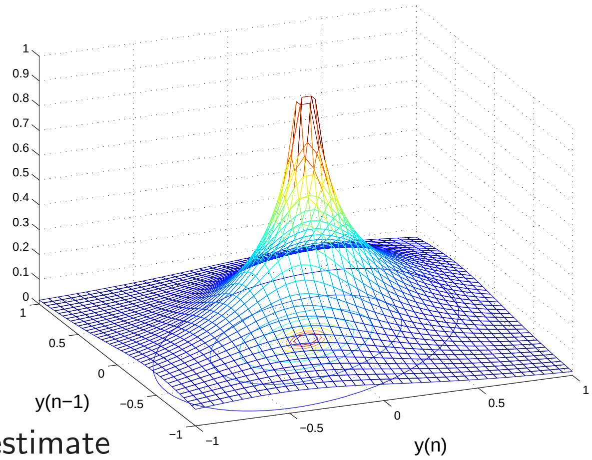
TRINICON-Based Algorithm Design



TRINICON - Exploiting HOS: SIRPs as Source Model

Spherically Invariant Random Processes (SIRPs) are described by multivariate pdfs of the form (with suitable function f_D):

$$\hat{p}_D(\mathbf{y}_p) = \frac{1}{\sqrt{\pi^D \det(\hat{\mathbf{R}}_{pp})}} f_D\left(\mathbf{y}_p^T \hat{\mathbf{R}}_{pp}^{-1} \mathbf{y}_p\right)$$



Several attractive properties:

- Good model for speech signals
- Reduced number of model parameters to estimate
- Multivariate pdfs can be derived analytically from corresponding univariate pdfs
- Incorporation into TRINICON leads to inherent stepsize normalization of the update equation: $\boldsymbol{\Phi}_{p,D}(\mathbf{y}_p(j)) = 2\,\phi_{y_p,D}\left(\mathbf{y}_p^T(j)\mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}^{-1}(i)\mathbf{y}_p(j)\right) \cdot \mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}^{-1}(i)\mathbf{y}_p(j),$

SIRP score: $\phi_{y_p,D}(u_p) = -\partial \log f_{p,D}(u_p)/\partial u_p \;\rightarrow$ Gauss: $\phi_{y_p,D} = 1/2$

Overview

- TRINICON - A Generic Concept for Adaptive MIMO Filtering
  - ▷ Optimization Criterion and Generic Adaptation Algorithms
  - ▷ Incorporation of Stochastic Source Models
- **Applications to Signal Processing Problems for Speech Capture**
  - ▷ Blind Source Separation (BSS) / Interference Suppression
  - ▷ Supervised Separation and System Identification / Echo Cancellation
  - ▷ Blind System Identification (BSI) / Localization of Multiple Sources
  - ▷ Multich. Blind Deconvolution (MCBD) and
    Multich. Blind Partial Deconvolution (MCBPD) / Dereverberation
- Concluding Remarks

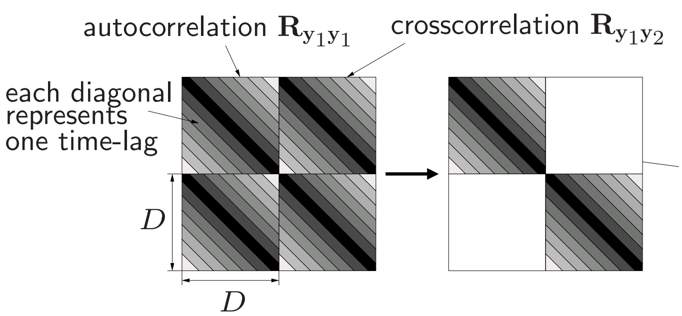
TRINICON for Blind Source Separation

- Desired score function for BSS: Independence between channels

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

- Illustration for SOS (Gaussian source model $\Rightarrow$ $\mathbf{\Phi}_{p,D}(\mathbf{y}_p(j)) = \mathbf{R}^{-1}_{\mathbf{y}_p\mathbf{y}_p}(i)\mathbf{y}_p(j)$):
  Natural gradient-based update

$$\Delta\check{\mathbf{W}}(m) \quad = \quad 2\sum_{i=0}^{\infty}\beta(i,m)\mathcal{SC}\left\{\mathbf{W}(i)\left\{\hat{\mathbf{R}}_{\mathbf{yy}} - \hat{\mathbf{R}}_{\mathbf{ss}}\right\}\hat{\mathbf{R}}_{\mathbf{ss}}^{-1}\right\}$$

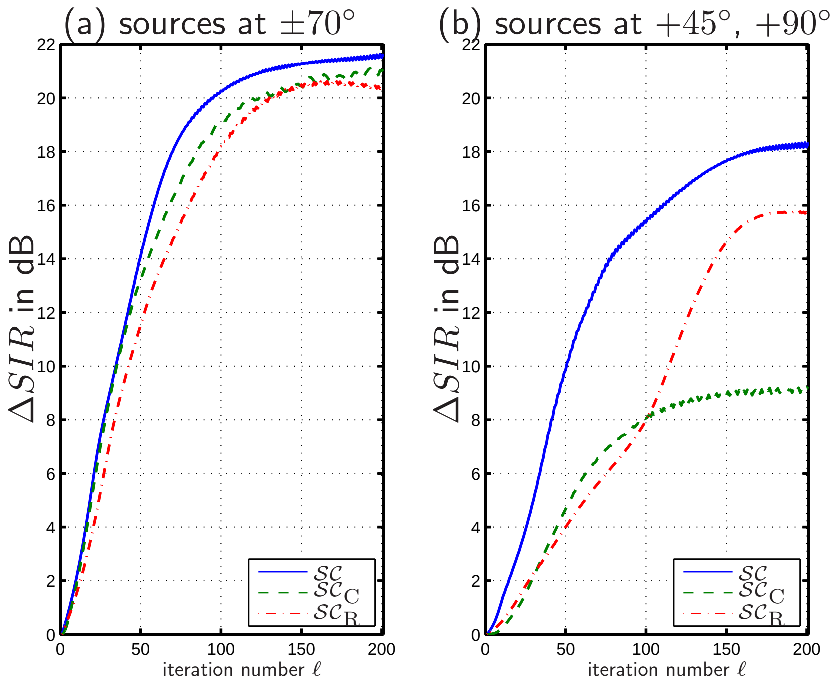


$$\hat{\mathbf{R}}_{\mathbf{ss}}(i) \stackrel{\text{(BSS)}}{=} \text{bdiag}_D\,\hat{\mathbf{R}}_{\mathbf{yy}}(i).$$

- Inherent normalization ($\rightarrow$ stepsize control) in each channel by $\text{bdiag}_D\hat{\mathbf{R}}_{\mathbf{yy}}^{-1}$

BSS - Results: Generic SOS + Sylvester Constraint

Offline BSS generic SOS natural gradient adaptation, cov. method, $L = 256$

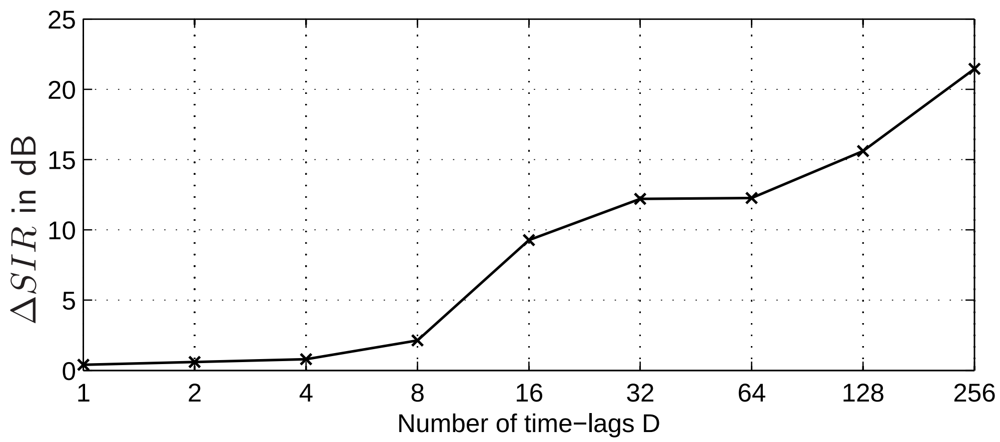


$\Rightarrow$Exact $\mathcal{SC}$ combines versatility ($\mathcal{SC}_\mathrm{R}$) and robust convergence ($\mathcal{SC}_\mathrm{C}$)

TRINICON for BSS - Exploitation of Nonwhiteness

Influence of the number D of simultaneously optimized time lags for exploiting nonwhiteness

Offline BSS generic SOS natural gradient adaptation, $\mathcal{SC}_\mathrm{R}$, cov. method, $L = 256$

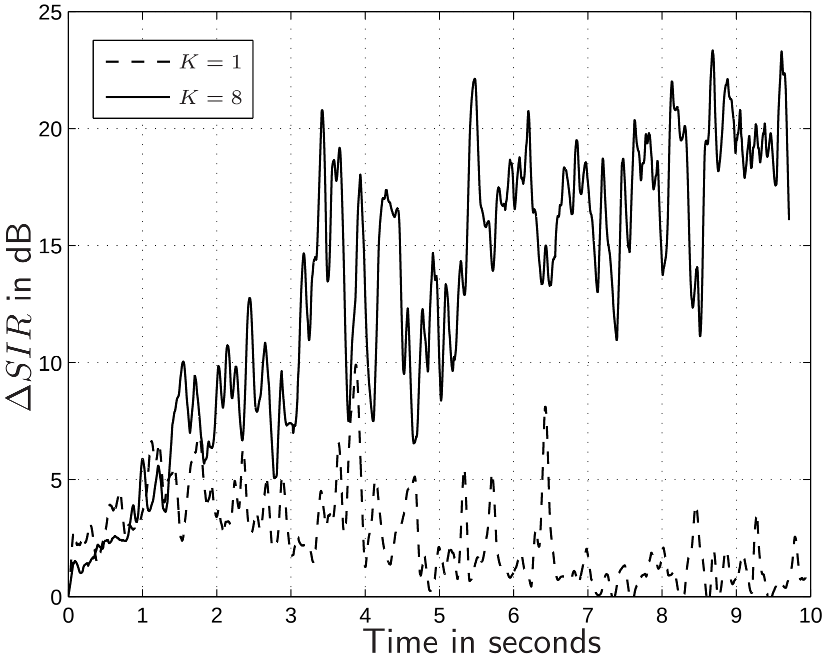


(no further improvement can be achieved for $D > L = 256$)

TRINICON for BSS - Exploitation of Nonstationarity

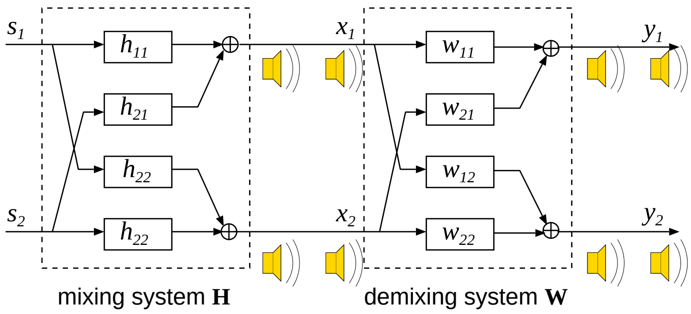
Influence of the number K of simultaneously optimized blocks

Block-offline BSS generic SOS natural gradient adaptation, $\mathcal{SC}_{\mathrm{R}}$, cov. method, $L = 256$



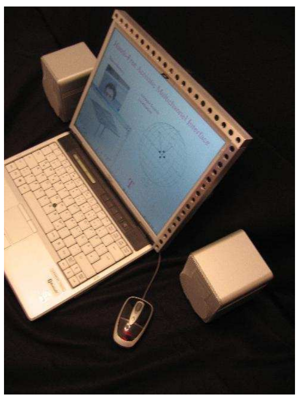
TRINICON for BSS - Results

- Example: Convergence behaviour
  - ▷ SOS-BSS with diagonal power normalization
  - ▷ $f_S = 16$kHz
  - ▷ $T_{60} = 50, 200$msec
  - ▷ $D = L = 1024$
  - ▷ block-online adaptation

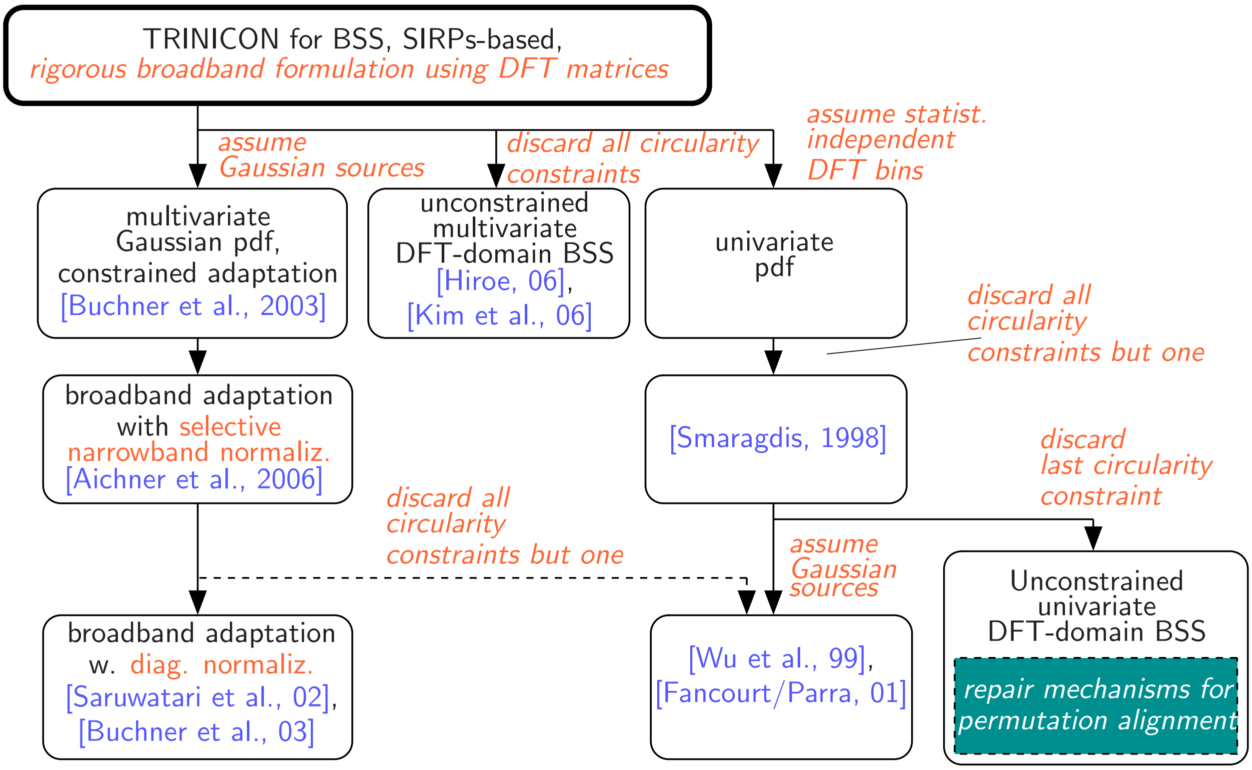


- Real-time implementation

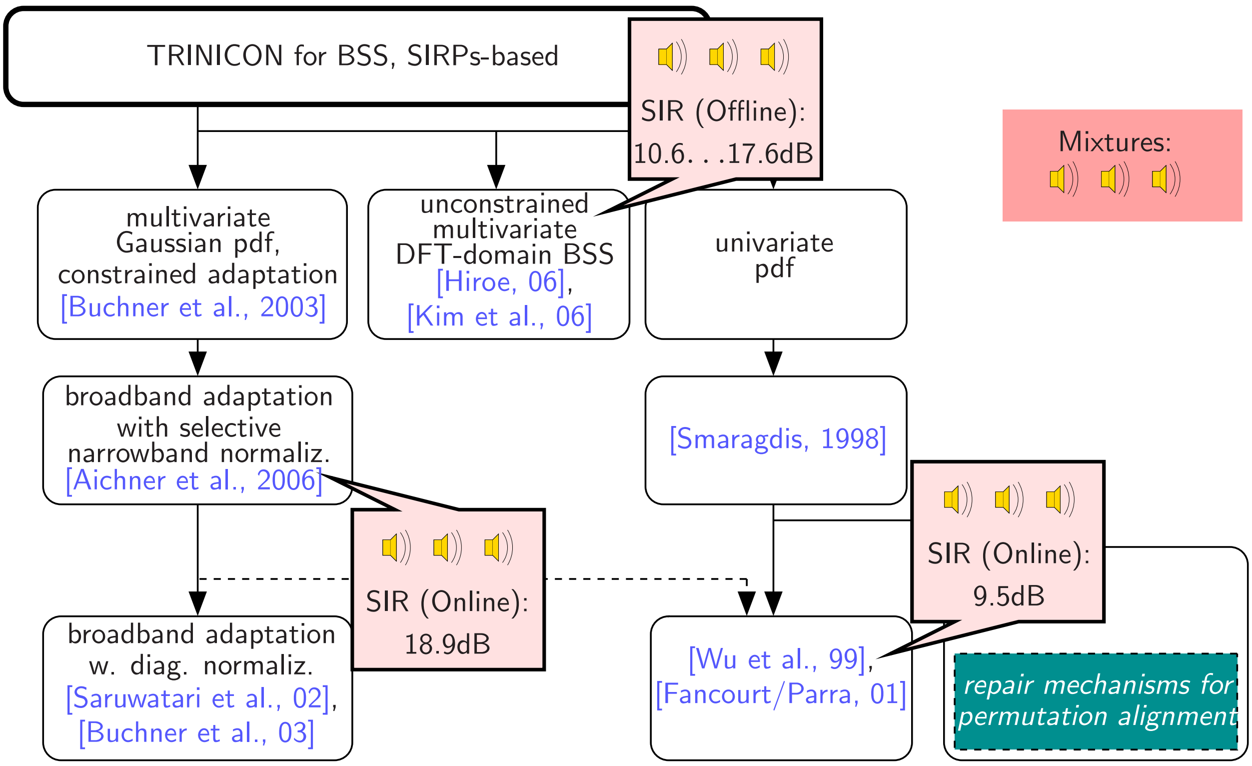
Implementation in the DFT domain equivalent to time-domain $\rightsquigarrow$ Broadband algorithm, no internal permutation or circular convolution



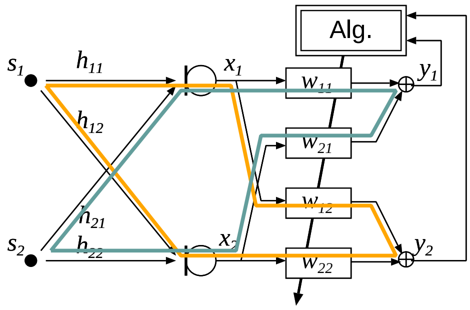
TRINICON for BSS - Efficient Frequency-Domain Realizations



TRINICON for BSS - Efficient Frequency-Domain Realizations



TRINICON for Blind System Identification



Equilibria of overall system $\mathbf{C} = \mathbf{HW}$ in Sylvester structure (Buchner et al., 2005):

$$\mathrm{boff}\{\mathbf{C}\} = \mathbf{0} \quad \text{iff } D = L$$

$$\Rightarrow \quad h_{11} * w_{12} = -h_{12} * w_{22}$$

$$h_{21} * w_{11} = -h_{22} * w_{21}$$

Ideal separation filters: $\begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix} = \begin{bmatrix} \alpha_1 \mathbf{H}_{22} & -\alpha_2 \mathbf{H}_{12} \\ -\alpha_1 \mathbf{H}_{21} & \alpha_2 \mathbf{H}_{11} \end{bmatrix}$

Unique solution up to scaling (α_1, α_2) iff

- No common zeros in $H_{11}(z)$, $H_{12}(z)$/no common zeros in $H_{21}(z)$, $H_{22}(z)$
- Demixing filter length $L \leq$ length M of mixing system

$\Rightarrow$ **Application to localization of simultaneously active sources**

TRINICON for Localization $\rightarrow$ TDOA estimation

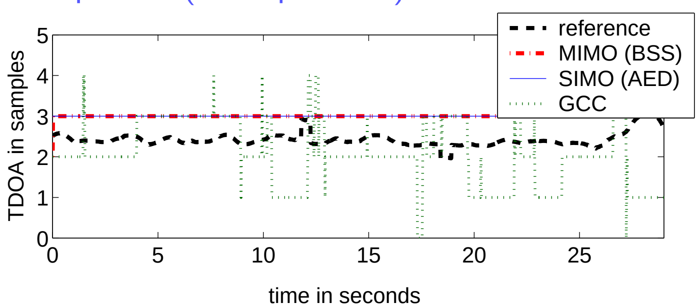
Simultaneous Localization of **Multiple** Sound Sources in **Reverberant** Environments

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

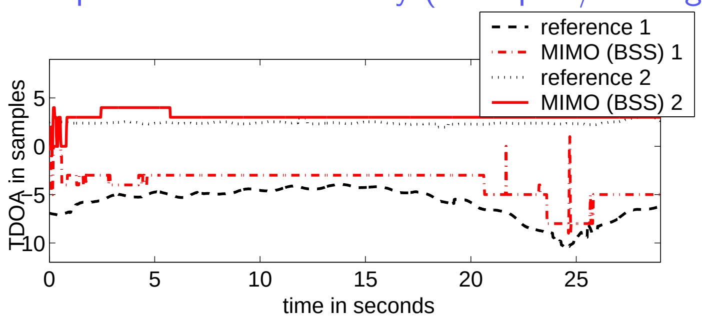
TDOA estimation:

- Generalized cross-correlation (GCC) with phase-transform (PHAT) weighting + VAD
- **SIMO-based BSI** + VAD (AED, Benesty et al., 1999)
- **MIMO-based BSI** (may be seen as a generalization of AED)

one speaker (fixed position)

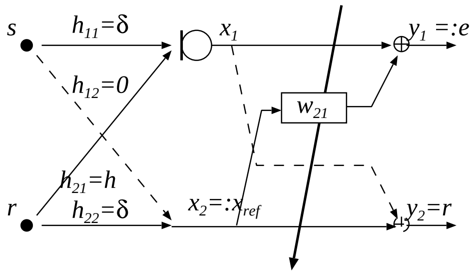


two speakers simultaneously (fixed pos./moving)



TRINICON for Supervised Adaptive Filtering Problems

Example: acoustic echo cancellation (AEC) problem from a specialized mixing model



Ideal filters:

$$\begin{bmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} \\ \mathbf{w}_{21} & \mathbf{w}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{22} & -\mathbf{h}_{12} \\ -\mathbf{h}_{21} & \mathbf{h}_{11} \end{bmatrix}$$

Here:

$$\begin{bmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} \\ \mathbf{w}_{21} & \mathbf{w}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{1}_1 & \mathbf{0} \\ -\mathbf{h} & \mathbf{1}_1 \end{bmatrix}$$

Lower left sub-matrix of simple gradient-based SIRPs TRINICON update:

$$\hat{\mathbf{h}}^{\ell}(m) = \hat{\mathbf{h}}^{\ell-1}(m) + \frac{\mu}{N}\sum_{i=0}^{\infty}\beta(i,m)\mathcal{SC}\left\{\sum_{j=iL}^{iL+N-1}\mathbf{x}_{\mathrm{ref}}(j)\mathbf{e}^T(j)\mathbf{R}_{\mathbf{ee}}^{-1}(i)\underbrace{\phi_{e,D}\left(\mathbf{e}^T(j)\mathbf{R}_{\mathbf{ee}}^{-1}(i)\mathbf{e}(j)\right)}_{\text{'SIRP score'}}\right\}$$

$\Rightarrow$ Generalization of Least-Mean-Squares (LMS): **inherent 'stepsize control'**

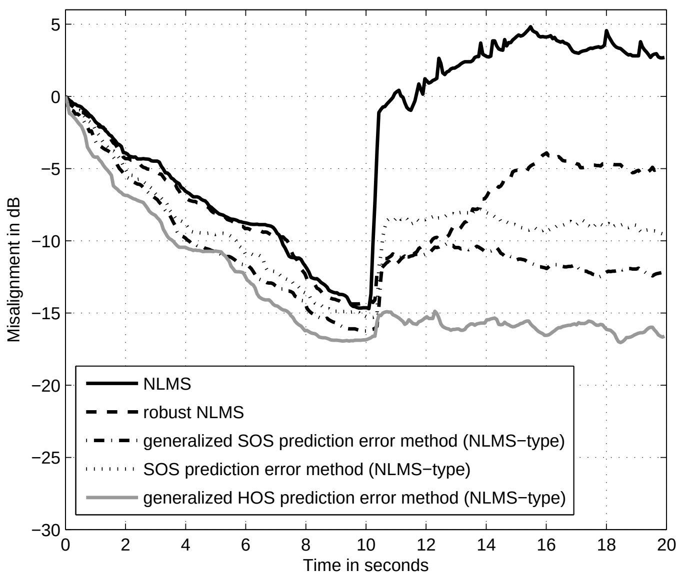
- **HOS case:** nongaussianity of local speech, **incorporation of 'robust statistics'**
- **Analogously:** generalization of other supervised algorithms (NLMS, Newton, RLS,...)

TRINICON for Supervised Adaptive Filtering Problems

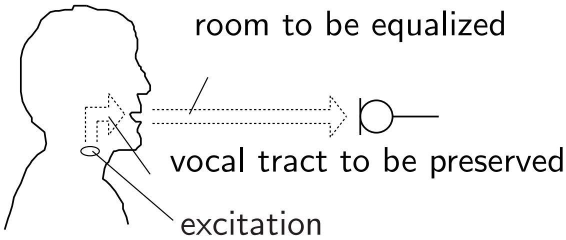
Misalignment convergence ($\|\mathbf{h}-\hat{\mathbf{h}}\|^2/\|\mathbf{h}\|^2$) for gradient-based updates with NLMS-type normalization

$L = 1024$, $f_S = 16kHz$, without double-talk detector

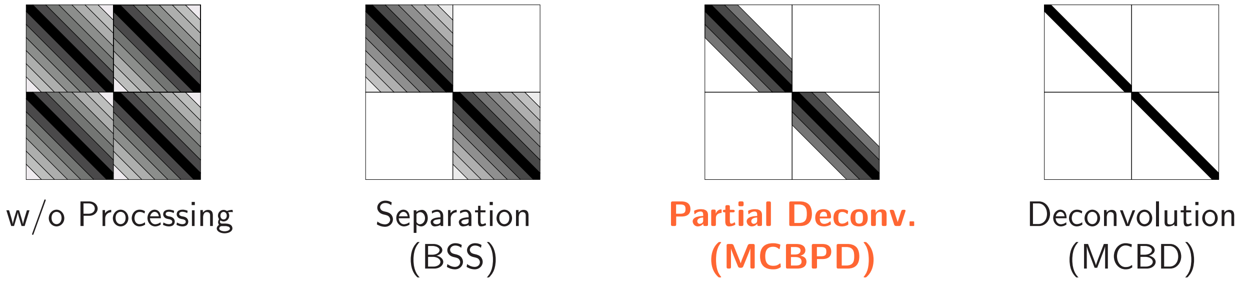
Single talk during the first 10sec, double talk starts after 10sec



TRINICON for Dereverberation = Partial Deconvolution

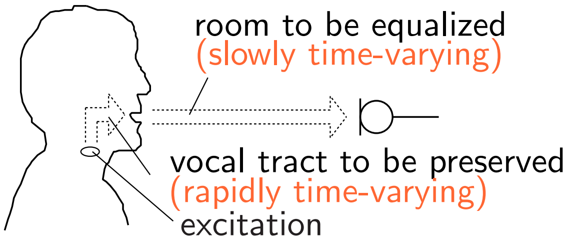


Desired output statistics - illustration for SOS:

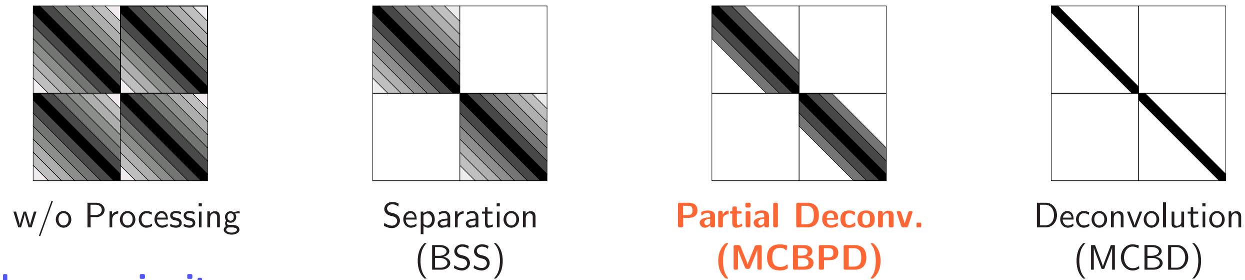


TRINICON for Dereverberation = Partial Deconvolution

Nonstationary:



Nonwhiteness:



Nongaussianity:

- Probability density of speech signal: supergaussian
- Room acoustics described by convolutional sum $\rightarrow$ mic signals closer to Gaussian
- Aim: maximize nongaussianity of demixing filter outputs $\rightarrow$ e.g., max. **kurtosis**

TRINICON - Nearly Gaussian Densities as Source Model

Expansions based on Chebyshev-Hermite polynomials $P_{\mathrm{H},n}(\cdot)$

Here: Gram-Charlier expansion

Univariate Example: fourth-oder approximation for a zero-mean process

$$\hat{p}_{y,1}(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2/2\sigma^2} \left(1 + \underbrace{\frac{\kappa_3}{3!\sigma^3} P_{\mathrm{H},3}\left(\frac{y}{\sigma}\right)}_{\kappa_3 = \text{skewness, negligible}} + \underbrace{\frac{\kappa_4}{4!\sigma^4} P_{\mathrm{H},4}\left(\frac{y}{\sigma}\right)}_{\kappa_4 = \text{kurtosis, nongaussianity}} \right)$$

TRINICON - Nearly Gaussian Densities as Source Model

Expansions based on Chebyshev-Hermite polynomials $P_{\mathrm{H},n}(\cdot)$

Here: Gram-Charlier expansion

Multivariate Case:

$$\hat{p}_{y_p,D}(\mathbf{y}_p(j)) = \frac{1}{\sqrt{(2\pi)^D \det \mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}(i)}} \mathrm{e}^{-\frac{1}{2}\mathbf{y}_p^T(j)\mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}^{-1}(i)\mathbf{y}_p(j)}$$

$$\cdot \sum_{n_1=0}^{\infty} \cdots \sum_{n_D=0}^{\infty} a_{n_1\cdots n_D,p}\, P_{\mathrm{H},n_1}\left(\left[\mathbf{L}_p^{-1}(i)\mathbf{y}_p(j)\right]_1\right) \cdot \cdots \cdot P_{\mathrm{H},n_D}\left(\left[\mathbf{L}_p^{-1}(i)\mathbf{y}_p(j)\right]_D\right)$$

For speech signals: we introduce the factorization

$$\mathbf{R}_{\mathbf{y}_p\mathbf{y}_p}^{-1}(i) = \mathbf{A}_p(i)\boldsymbol{\Sigma}_{\tilde{\mathbf{y}}_p\tilde{\mathbf{y}}_p}^{-1}(i)\mathbf{A}_p^T(i)$$

Unit lower triangular matrix $\mathbf{A}_p(i) \rightarrow$ interpreted as a whitening convolution matrix

- model $y_p(n)$ as an AR process of order $n_A = D - 1$
- shift prefilter matrix into the data terms:

$$\tilde{\mathbf{y}}_p := \mathbf{A}_p^T \mathbf{y}_p = [\tilde{y}_p(n), \tilde{y}_p(n-1), \ldots, \tilde{y}_p(n-D+1)]^T \tag{0}$$

TRINICON - Nearly Gaussian Densities as Source Model

⇒ **Multivariate speech model based on fourth-order approximation:**

$$\hat{p}_{y_p,D}(\mathbf{y}_p(j)) = \prod_{d=1}^{D} \frac{1}{\sqrt{2\pi\ \hat{\sigma}^2_{\tilde{y}_p}(j-d+1)}} \mathrm{e}^{-\frac{\tilde{y}_p^2(j-d+1)}{2\hat{\sigma}^2_{\tilde{y}_p}(j-d+1)}} \left(1 + \frac{\hat{\kappa}_{4,\tilde{y}_p}}{4!\sigma^4_{\tilde{y}_p}(j-d+1)} P_{\mathrm{H},n_d}\left(\frac{\tilde{y}_p(j-d+1)}{\hat{\sigma}_{\tilde{y}_p}(j-d+1)}\right)\right).$$

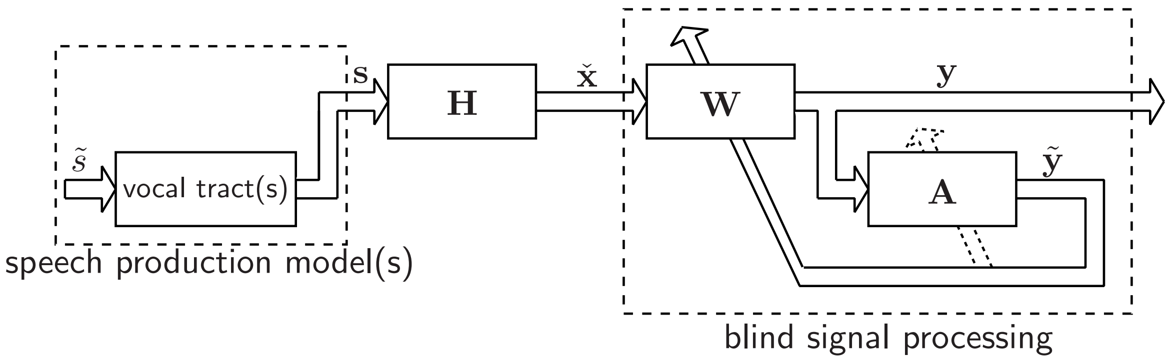
Resulting generalized multivariate score function:

$$\mathbf{\Phi}_{y,PD}(\mathbf{y}(j)) = \mathbf{A}(i)\left[\frac{\tilde{y}_p(j-d+1)}{\hat{\sigma}^2_{\tilde{y}_p}(j-d+1)} - \left(\frac{\sum_{j=iN_L}^{iN_L+N-1} \tilde{y}_p^4(j-d+1)}{3\left(\sum_{j=iN_L}^{iN_L+N-1} \tilde{y}_p^2(j-d+1)\right)^2} - 1\right) \cdot \left(\frac{\tilde{y}_p^3(j-d+1)}{\hat{\sigma}^4_{\tilde{y}_p}(j-d+1)} - \frac{\tilde{y}_p(j-d+1)\sum_{j=iN_L}^{iN_L+N-1}\tilde{y}_p^4(j-d+1)}{\hat{\sigma}^6_{\tilde{y}_p}(j-d+1)}\right)\right] \tag{0}$$

TRINICON - Nearly Gaussian Densities as Source Model

Filtered-x-type interpretation (filtered versions of microphone signals and output signals in the coeff. update):

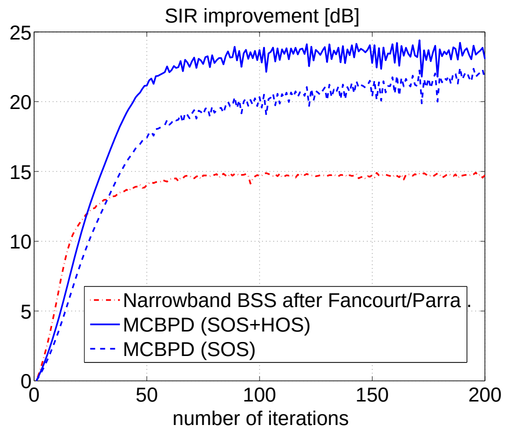
Inversion of the speech production models within the blind signal processing

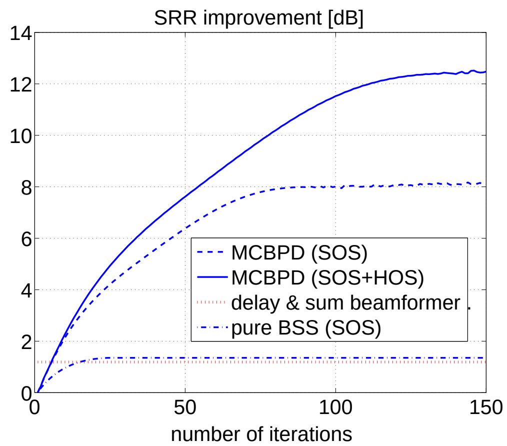


- The coefficients in $\mathbf{W}$ and $\mathbf{W}$ are estimated in an alternating way.
- Note: The filtered input vector is calculated using the LP of the output signals.

Experimental Results – TRINICON for Dereverberation

- MCBPD, 2 sources, 4 mics., $L = 3000$, offline adaptation; $T_{60} \approx 700$ms, $f_S = 16kHz$, $n_A = 32$





Sources: Mixture (mic 1): Outputs:

Summary

- In acoustic preprocessing for hands-free speech communication, **Blind MIMO signal processing** seems appropriate for practical separation and dereverberation tasks.
- **From a generic framework** for adaptive signal processing, we can establish links between various known algorithms, and, so far, it has led to various novel algorithms for
  - ▷ robust BSS real-time system
  - ▷ robust TDOA estimation for localization of multiple sources
  - ▷ dereverberation without artifacts
  - ▷ improved supervised adaptive filtering, e.g., for acoustic echo cancellation