Wave-Domain Adaptive Filtering for Acoustic Human-Machine Interfaces based on Wavefield Analysis and Synthesis

Herbert Buchner, Sascha Spors, Wolfgang Herbordt, Walter Kellermann
{buchner,spors,herbordt,wk}@LNT.de

September 9, 2004

Multimedia Communications and Signal Processing
University of Erlangen-Nuremberg

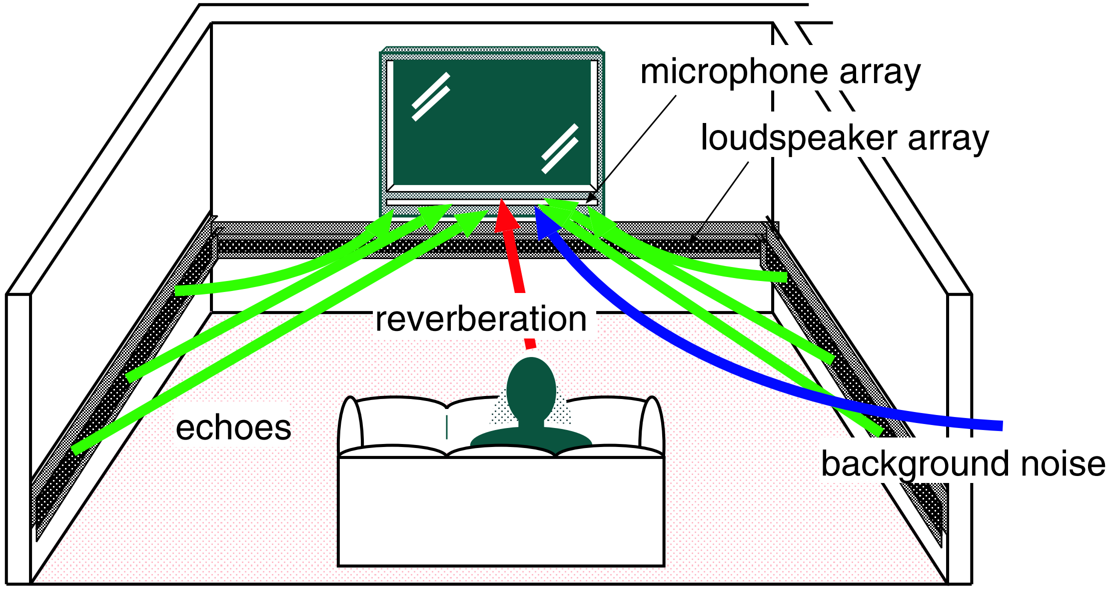
Contents

- Motivation: Hands-free full-duplex communication with 3D audio
- Background: Conventional point-to-point adaptive multichannel processing
- Novel approach: Wave-Domain Adaptive Filtering (WDAF)
- Evaluation of WDAF for acoustic echo cancellation
- Summary and outlook

Motivation (1): Hands-Free Full-Duplex Communication

- **Multichannel reproduction:**
  ⇒ wave field synthesis using a loudspeaker array

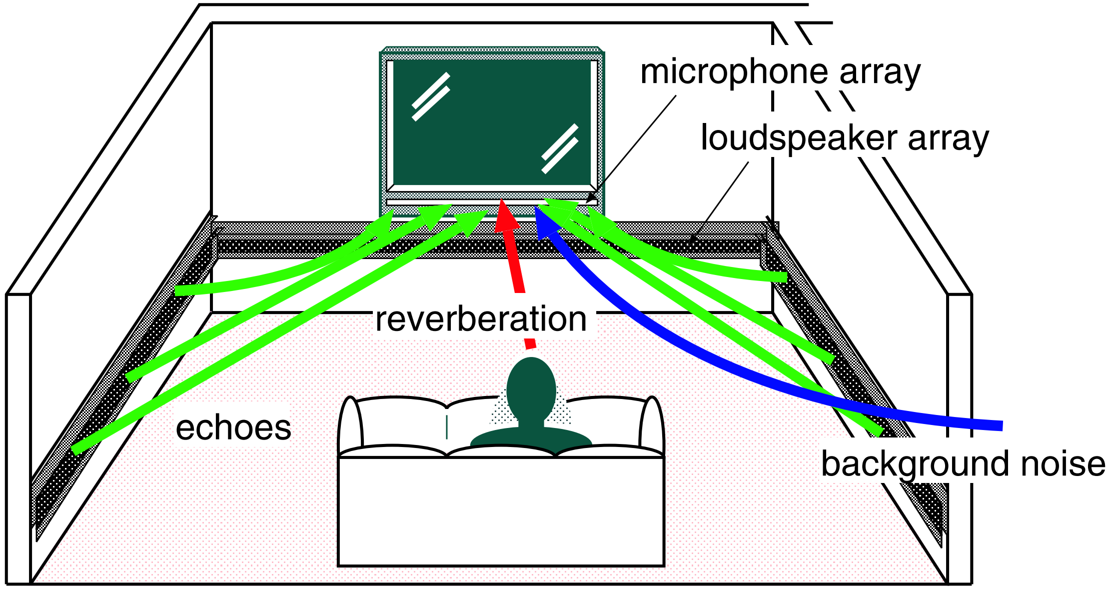
- **Multichannel recording:**
  ⇒ microphone array processing



Motivation (1): Hands-Free Full-Duplex Communication

- **Multichannel reproduction:**
  ⇒ wave field synthesis using a loudspeaker array

- **Multichannel recording:**
  ⇒ microphone array processing

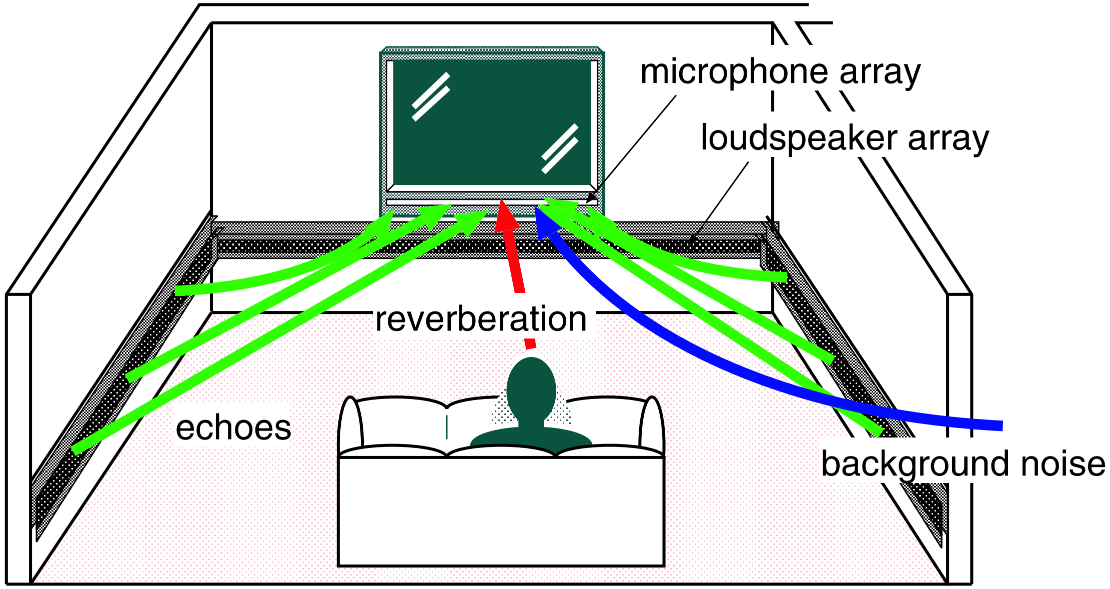


Major challenge: adaptive signal processing for massive multichannel systems

Motivation (1): Hands-Free Full-Duplex Communication

- **Multichannel reproduction:**
  ⇒ wave field synthesis using a loudspeaker array

- **Multichannel recording:**
  ⇒ microphone array processing



Major challenge: adaptive signal processing for massive multichannel systems

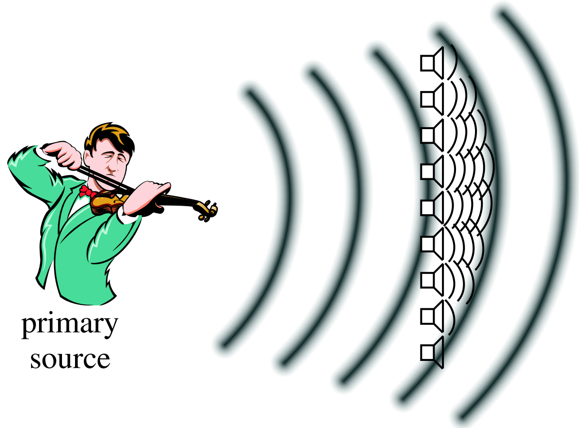
We propose:
framework for efficient spatio-temporal transform-domain adaptive filtering exploiting foundations of wave physics:

Wave-Domain Adaptive Filtering

Motivation (2): Wave-Field Synthesis and Wave-Field Analysis

Huygens (1690):

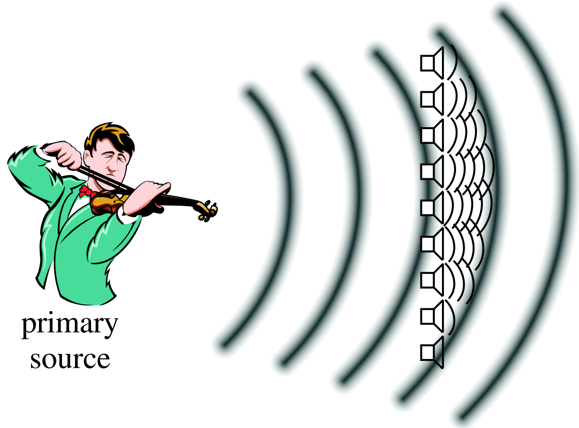
Any point on a propagating wavefront can be taken as a point source for the production of spherical secondary waves.



Motivation (2): Wave-Field Synthesis and Wave-Field Analysis

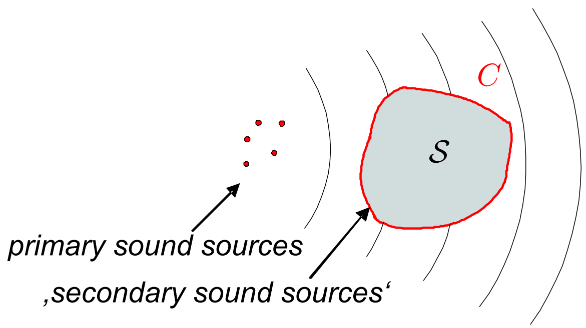
Huygens (1690):

Any point on a propagating wavefront can be taken as a point source for the production of spherical secondary waves.



Mathematical formulation by the Kirchhoff-Helmholtz integrals:

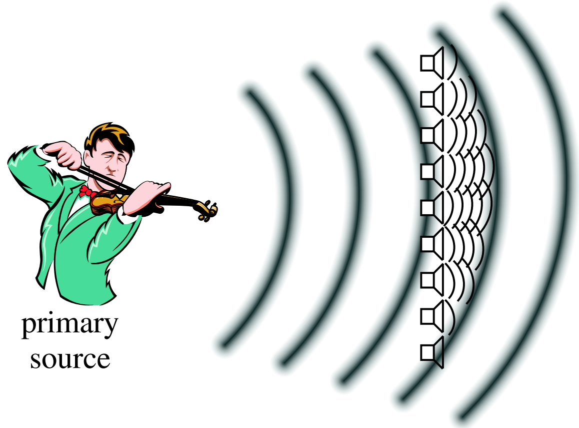
At any listening point within the source-free listening area $\mathcal{S}$, the sound pressure field $p(\mathbf{r}, t)$ can be calculated if both, the *sound pressure and its gradient* are *known on the contour* $\mathcal{C}$ enclosing this area.



Motivation (2): Wave-Field Synthesis and Wave-Field Analysis

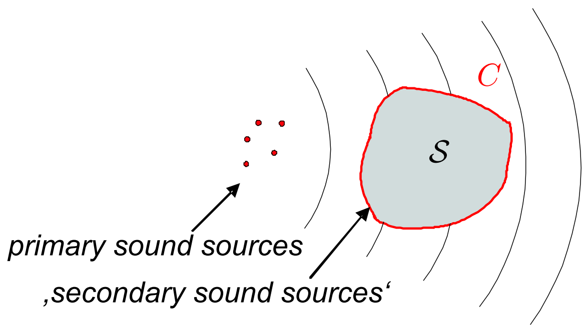
Huygens (1690):

Any point on a propagating wavefront can be taken as a point source for the production of spherical secondary waves.



Mathematical formulation by the Kirchhoff-Helmholtz integrals:

At any listening point within the source-free listening area $\mathcal{S}$, the sound pressure field $p(\mathbf{r}, t)$ can be calculated if both, the *sound pressure and its gradient* are *known on the contour* $\mathcal{C}$ enclosing this area.

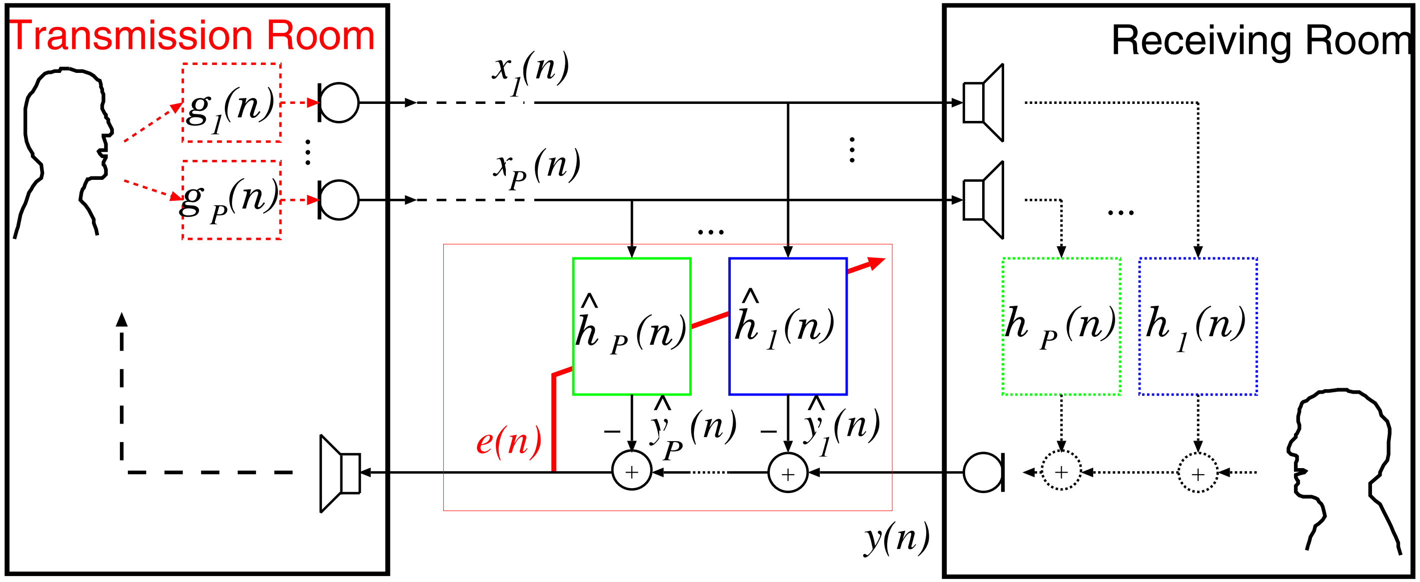


Practical realization:

WFS and WFA by spatial sampling using loudspeaker arrays and microphone arrays, resp. $\Rightarrow$ **large number of loudspeaker and microphone channels**

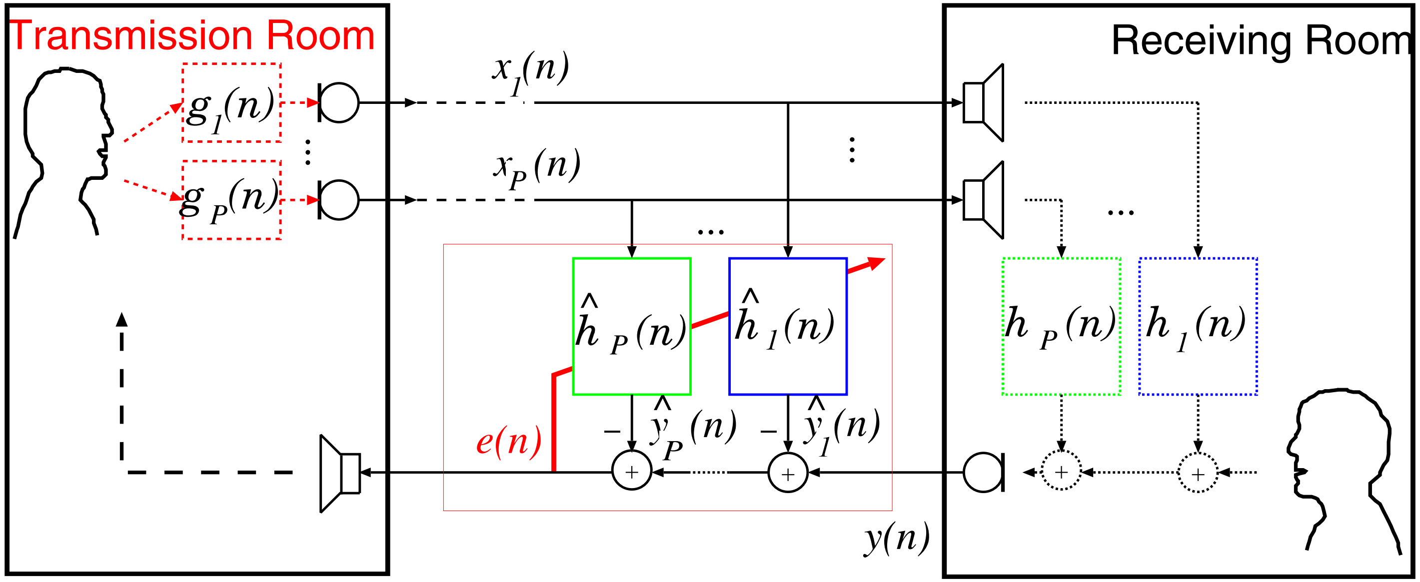
Background: Point-to-Point Adaptive Multichannel Processing (1)

Prominent example: multichannel acoustic echo cancellation (AEC)



Background: Point-to-Point Adaptive Multichannel Processing (1)

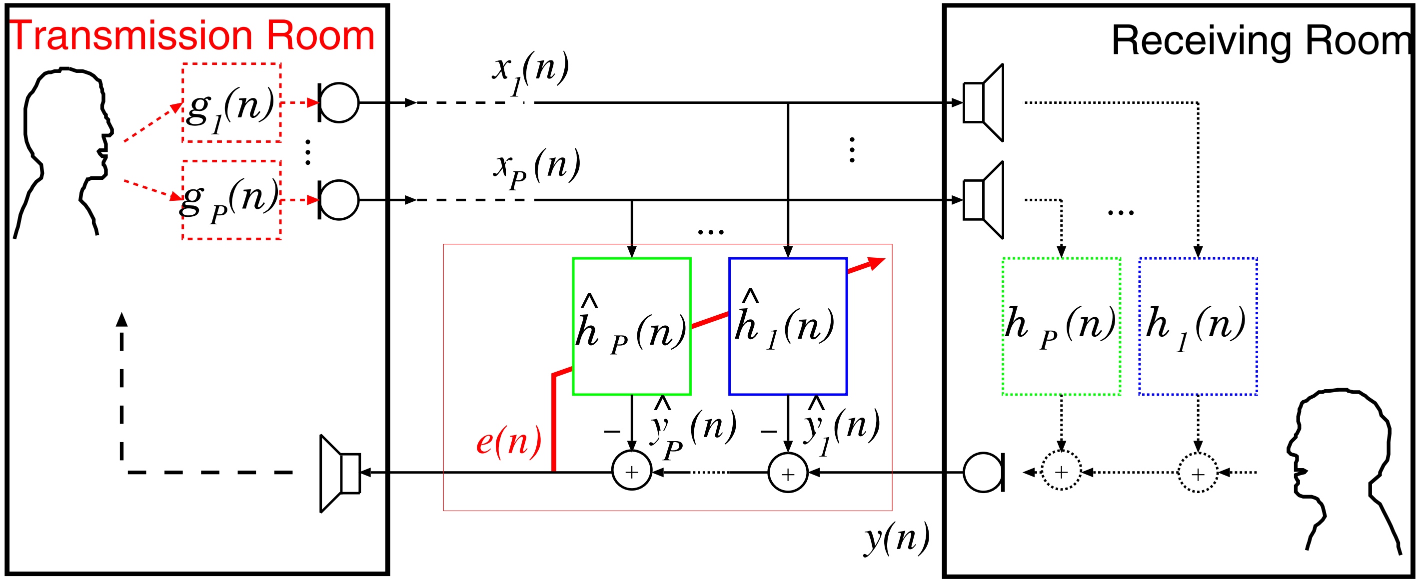
Prominent example: multichannel acoustic echo cancellation (AEC)



⇒ adaptive system identification

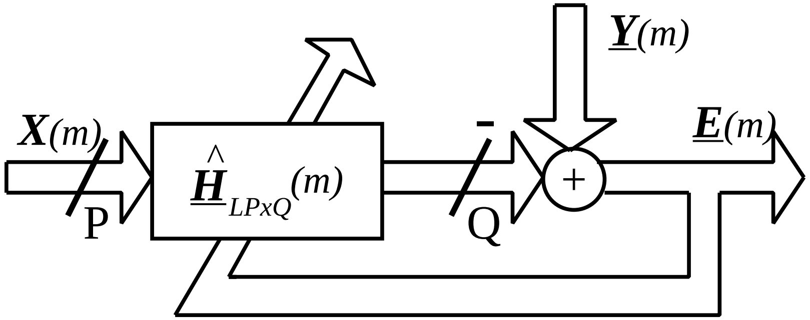
Background: Point-to-Point Adaptive Multichannel Processing (1)

Prominent example: multichannel acoustic echo cancellation (AEC)



$\Rightarrow$ adaptive system identification

Generalization: adaptive FIR MIMO filter with P inputs and Q outputs

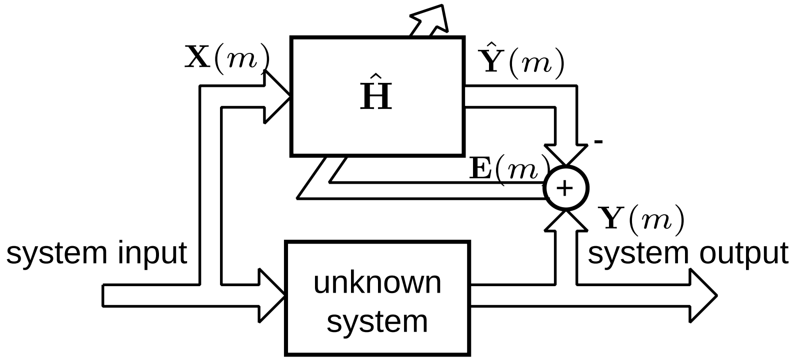


Background: Point-to-Point Adaptive Multichannel Processing (2)

Adaptive Filtering Constellations (multichannel version of [Haykin]):

Adaptive Filtering Constellations (multichannel version of [Haykin]):

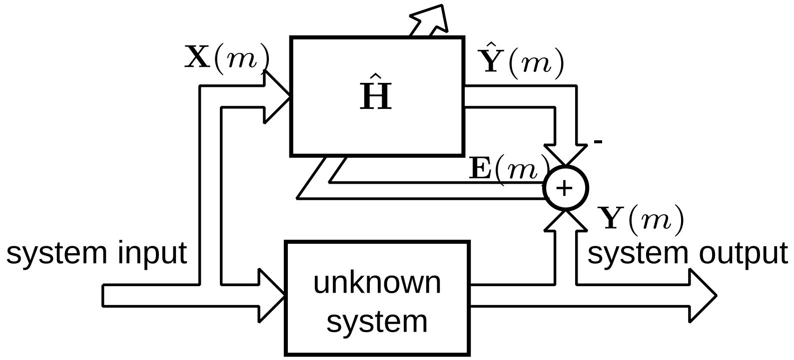
- **System identification**
  (e.g., acoustic echo cancellation)



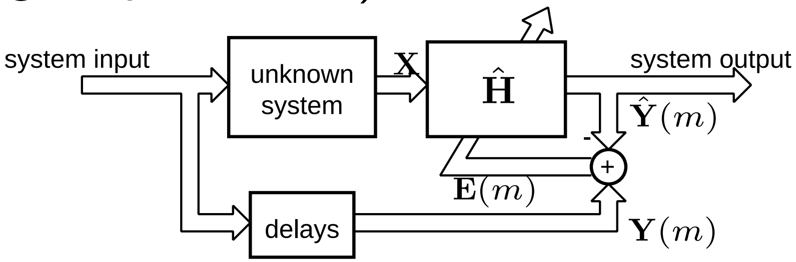
Background: Point-to-Point Adaptive Multichannel Processing (2)

Adaptive Filtering Constellations (multichannel version of [Haykin]):

- **System identification**
  (e.g., acoustic echo cancellation)



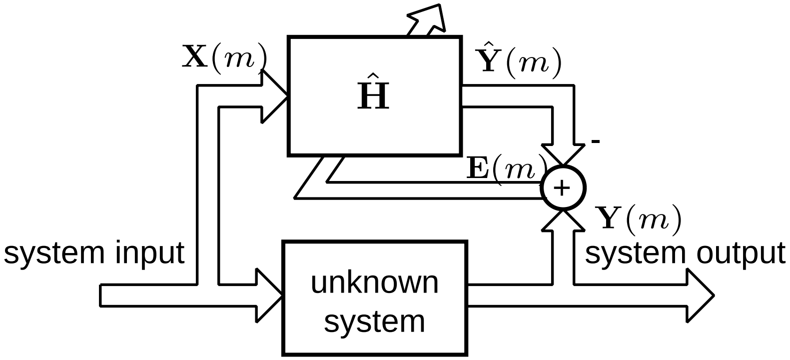
- **Inverse modeling**
  (e.g., equalization)



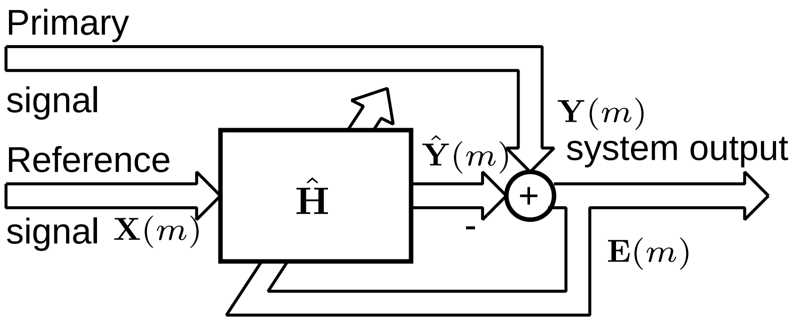
Background: Point-to-Point Adaptive Multichannel Processing (2)

Adaptive Filtering Constellations (multichannel version of [Haykin]):

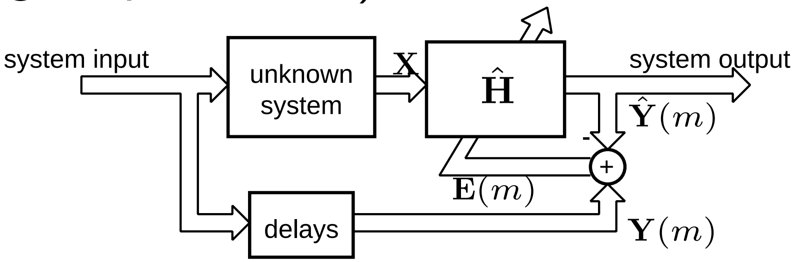
- **System identification**
  (e.g., acoustic echo cancellation)



- **Interference cancellation**
  (e.g., adaptive beamforming, noise reduction)



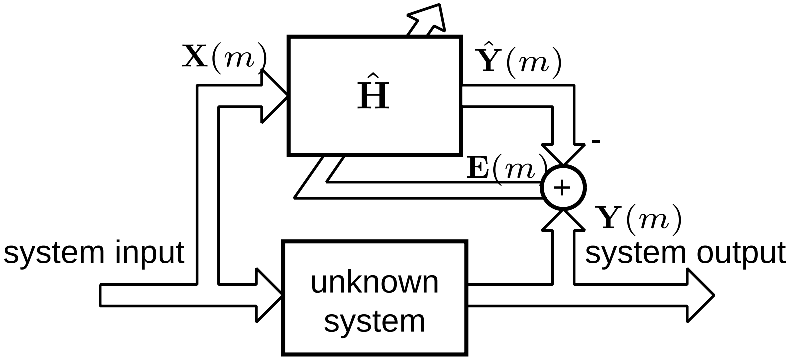
- **Inverse modeling**
  (e.g., equalization)



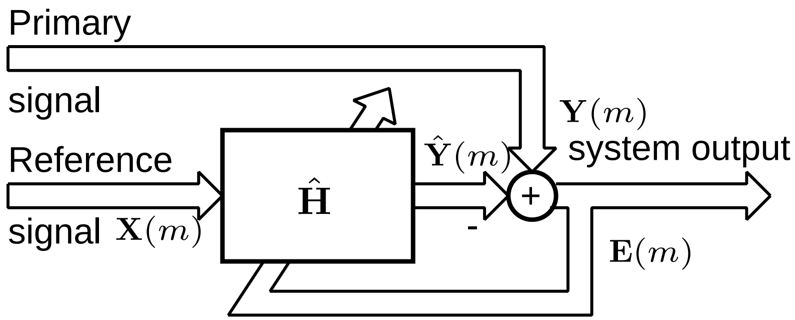
Background: Point-to-Point Adaptive Multichannel Processing (2)

Adaptive Filtering Constellations (multichannel version of [Haykin]):

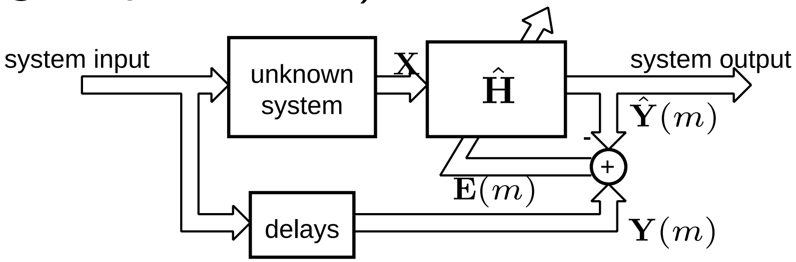
- **System identification**
  (e.g., acoustic echo cancellation)



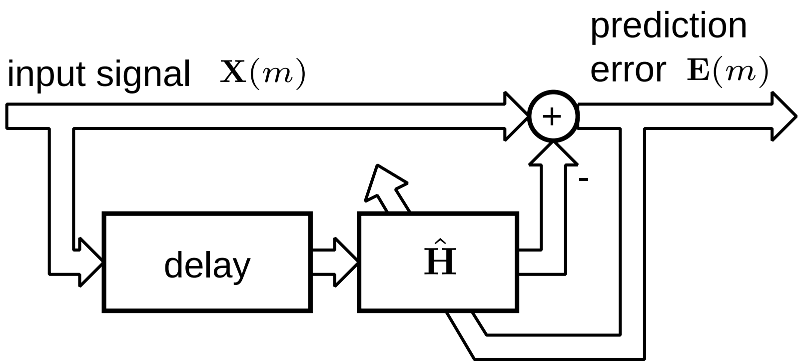
- **Interference cancellation**
  (e.g., adaptive beamforming, noise reduction)



- **Inverse modeling**
  (e.g., equalization)



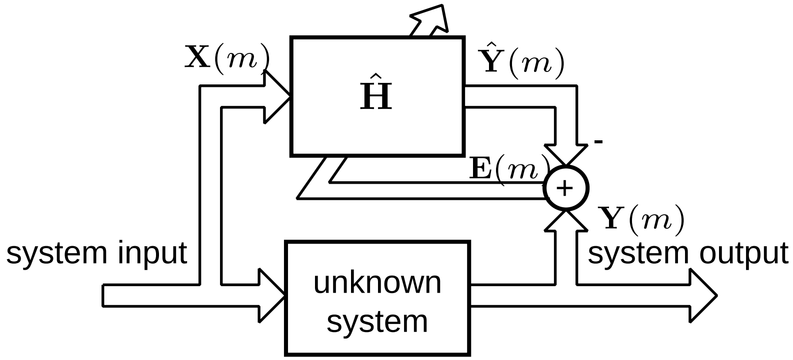
- **Prediction**



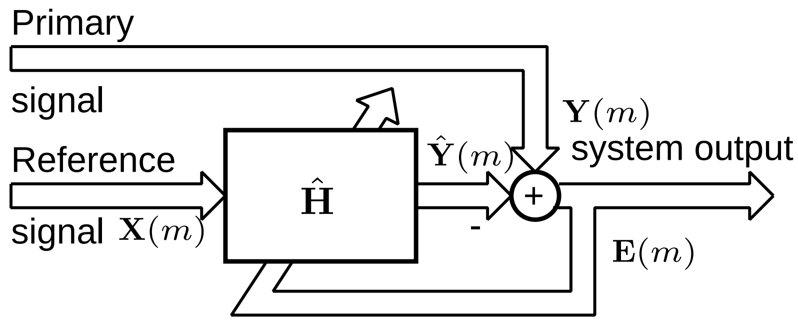
Background: Point-to-Point Adaptive Multichannel Processing (2)

Adaptive Filtering Constellations (multichannel version of [Haykin]):

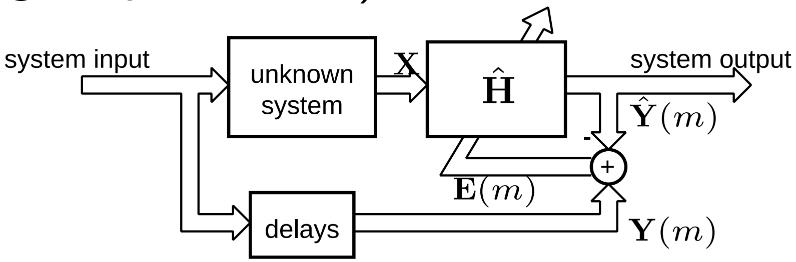
- **System identification**
  (e.g., acoustic echo cancellation)



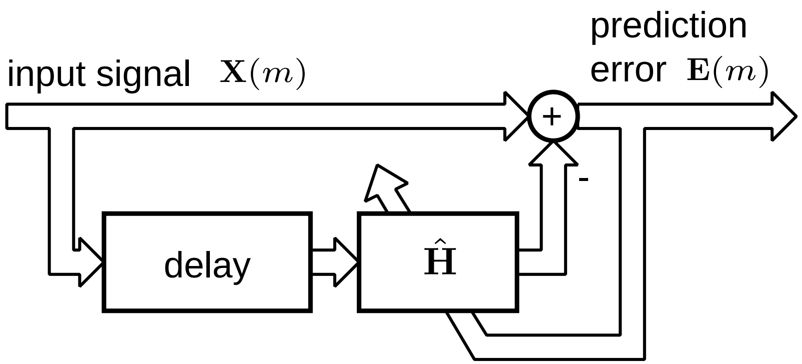
- **Interference cancellation**
  (e.g., adaptive beamforming, noise reduction)



- **Inverse modeling**
  (e.g., equalization)



- **Prediction**



$\Rightarrow$ **Least-Squares problems,** adaptive solutions with continuous updates for tracking

Multichannel Adaptive Filtering for Solving the LS Problem

Most known adaptation algorithms (e.g., LMS/NLMS) exhibit slow convergence for highly correlated input signals.

Multichannel Adaptive Filtering for Solving the LS Problem

Most known adaptation algorithms (e.g., LMS/NLMS) exhibit slow convergence for highly correlated input signals.

Theoretically Optimum Solution: **Multichannel Recursive Least-Squares (MC RLS)**

$$\mathbf{e}(n) = \mathbf{y}(n) - \hat{\mathbf{H}}^T(n-1)\mathbf{x}(n)$$
$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n)\mathbf{e}^T(n)$$

with input correlation matrix

$$\mathbf{R}_{\mathbf{xx}}(n) = \begin{bmatrix} \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}(n) & \cdots & \mathbf{R}_{\mathbf{x}_1\mathbf{x}_P}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{x}_P\mathbf{x}_1}(n) & \cdots & \mathbf{R}_{\mathbf{x}_P\mathbf{x}_P}(n) \end{bmatrix}$$

- all correlations taken into account by (large) matrix $\mathbf{R}_{\mathbf{xx}}^{-1}(n) \Rightarrow$ rapid convergence
- computationally very expensive
- in practice often numerical problems

Multichannel Adaptive Filtering for Solving the LS Problem

Most known adaptation algorithms (e.g., LMS/NLMS) exhibit slow convergence for highly correlated input signals.

Theoretically Optimum Solution: **Multichannel Recursive Least-Squares (MC RLS)**

$$\mathbf{e}(n) = \mathbf{y}(n) - \hat{\mathbf{H}}^T(n-1)\mathbf{x}(n)$$
$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n)\mathbf{e}^T(n)$$

with input correlation matrix

$$\mathbf{R}_{\mathbf{xx}}(n) = \begin{bmatrix} \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}(n) & \cdots & \mathbf{R}_{\mathbf{x}_1\mathbf{x}_P}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{x}_P\mathbf{x}_1}(n) & \cdots & \mathbf{R}_{\mathbf{x}_P\mathbf{x}_P}(n) \end{bmatrix}$$

Multichannel Frequency-Domain Adaptive Filtering (MC FDAF)

Multichannel Adaptive Filtering for Solving the LS Problem

Most known adaptation algorithms (e.g., LMS/NLMS) exhibit slow convergence for highly correlated input signals.

Theoretically Optimum Solution: **Multichannel Recursive Least-Squares (MC RLS)**

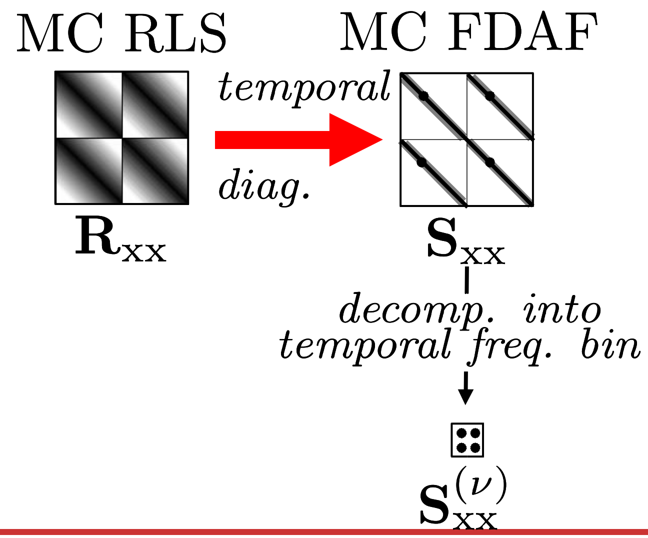
$$\mathbf{e}(n) = \mathbf{y}(n) - \hat{\mathbf{H}}^T(n-1)\mathbf{x}(n)$$
$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{R}_{\mathrm{xx}}^{-1}(n)\mathbf{x}(n)\mathbf{e}^T(n)$$

with input correlation matrix

$$\mathbf{R}_{\mathrm{xx}}(n) = \begin{bmatrix} \mathbf{R}_{\mathrm{x}_1\mathrm{x}_1}(n) & \cdots & \mathbf{R}_{\mathrm{x}_1\mathrm{x}_P}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathrm{x}_P\mathrm{x}_1}(n) & \cdots & \mathbf{R}_{\mathrm{x}_P\mathrm{x}_P}(n) \end{bmatrix}$$

Multichannel Frequency-Domain Adaptive Filtering (MC FDAF)

- approximate diagonalization of the correlation matrices $\mathbf{R}_{\mathrm{xx}}$ by DFT
- very efficient use of the FFT
  $\Rightarrow$ gains for both, adaptation and filtering
- there are algorithms fully taking into account all cross-correlations using $\mathbf{S}_{\mathrm{xx}}^{(\nu)}$



Massive Multichannel Systems

For massive WFS-based multichannel applications:

- High complexity
  Example (Echo cancellation for typical 48-channel WFS-based system):
  $P \cdot Q \cdot L = 48 \cdot 48 \cdot 1024 = 2304 \cdot 1024 = 2359296$ filter taps to optimize

- Even bin-wise matrices $\mathbf{S}_{\mathbf{xx}}^{(\nu)}$ become large and ill-conditioned

$\Rightarrow$ Current algorithms cannot be used

$\Rightarrow$ **Is there a more "global" point of view?**

Novel Approach: Wave-Domain Adaptive Filtering (WDAF)

Basic concept:

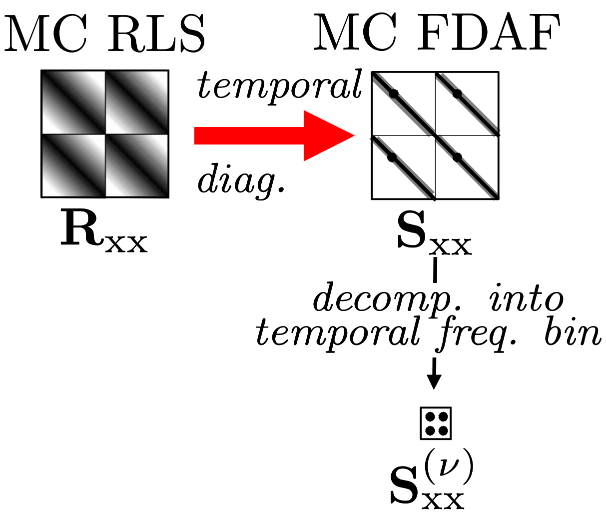
- give up point-to-point model ($\rightarrow$ ordinary difference equation) in favour of a more detailed spatial consideration exploiting wave-physics foundations

Novel Approach: Wave-Domain Adaptive Filtering (WDAF)

Basic concept:

- give up point-to-point model (→ ordinary difference equation) in favour of a more detailed spatial consideration exploiting wave-physics foundations

- extend conventional MC FDAF approach by a suitable spatio-temporal transform domain for efficiency

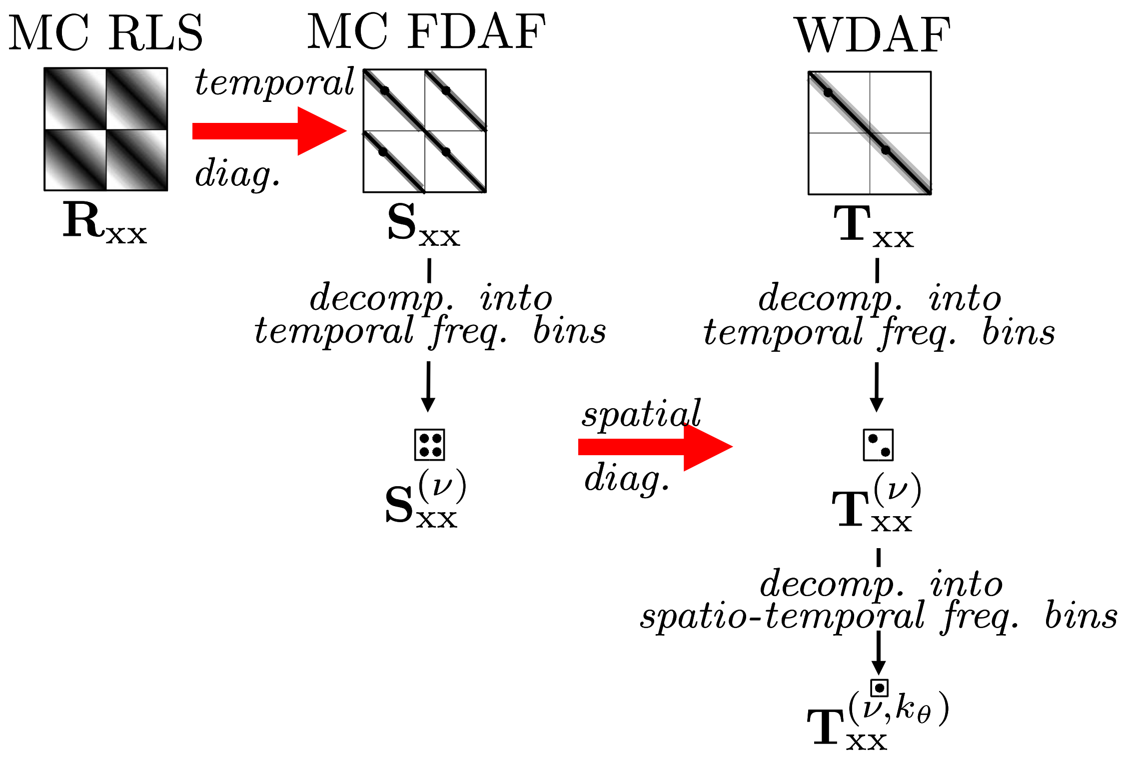


Novel Approach: Wave-Domain Adaptive Filtering (WDAF)

Basic concept:

- give up point-to-point model (→ ordinary difference equation) in favour of a more detailed spatial consideration exploiting wave-physics foundations

- extend conventional MC FDAF approach by a suitable spatio-temporal transform domain for efficiency

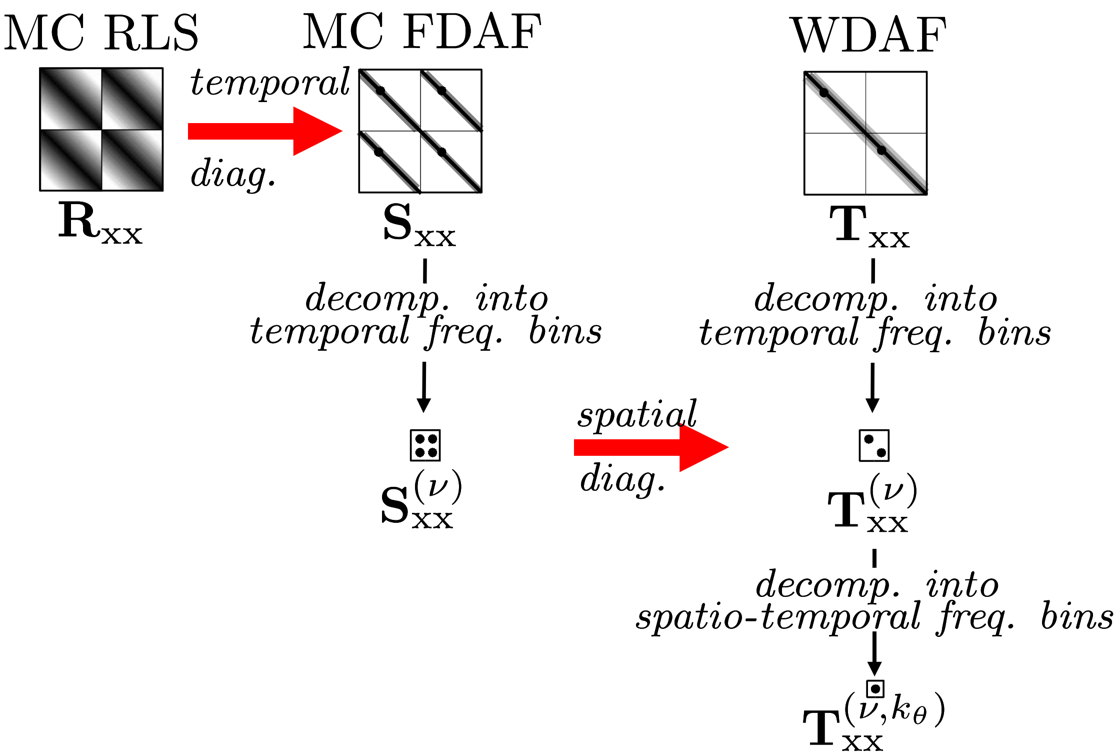


Novel Approach: Wave-Domain Adaptive Filtering (WDAF)

Basic concept:

- give up point-to-point model ($\rightarrow$ ordinary difference equation) in favour of a more detailed spatial consideration exploiting wave-physics foundations

- extend conventional MC FDAF approach by a suitable spatio-temporal transform domain for efficiency



Desirable:

- orthogonal basis functions
- approximate decomposition among temporal frequencies as in MC FDAF
- approximate spatial decomposition
- spatio-temporal basis functions must fulfill (acoustic) wave equation

Wave-Domain Adaptive Filtering (2)

Decoupling due to spatio-temporal transformation - Advantages

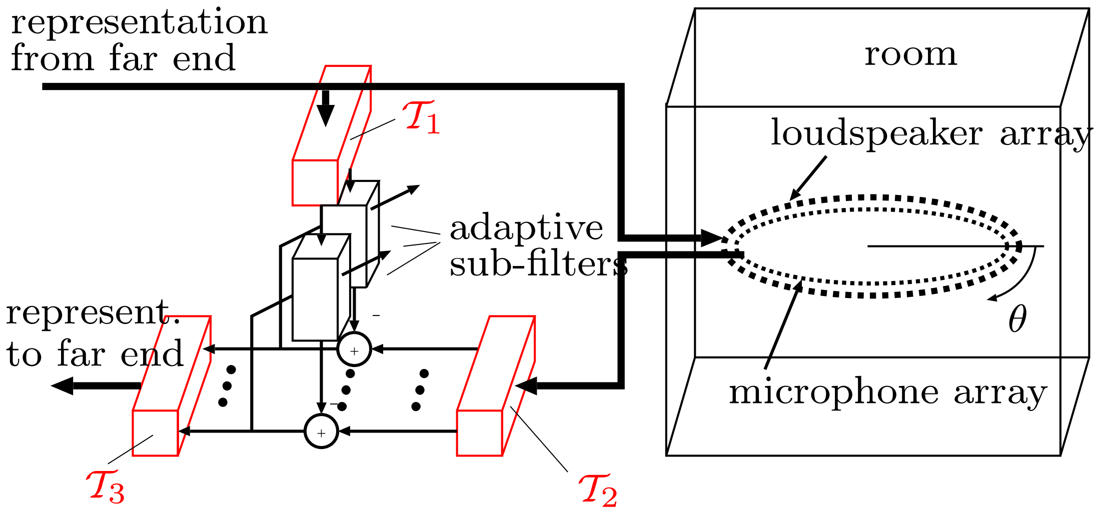
- improved convergence and very stable adaptation
- significant complexity reduction
  Example: $P = Q = 48 \Rightarrow$ only, e.g., 70 filters instead of $48 \cdot 48 = 2304$

Wave-Domain Adaptive Filtering (2)

Decoupling due to spatio-temporal transformation - Advantages

- improved convergence and very stable adaptation
- significant complexity reduction
  Example: $P = Q = 48 \Rightarrow$ only, e.g., 70 filters instead of $48 \cdot 48 = 2304$

Example: acoustic echo cancellation



- all microphone signals simultaneously taken into account for adaptive processing
- most cross-channels in the transform domain completely negligible

Wave-Domain Adaptive Filtering (3)

Transformations $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$ for real implementations

First approach: straightforward spatial Fourier transform ($\rightarrow$ plane waves)
$\Rightarrow$ Problem: transducers at each point of the listening room would be needed!

Wave-Domain Adaptive Filtering (3)

Transformations $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_3$ for real implementations

First approach: straightforward spatial Fourier transform ($\rightarrow$ plane waves)
$\Rightarrow$ Problem: transducers at each point of the listening room would be needed!

Decomposition taking into account the Kirchhoff-Helmholtz Integrals

- transformations depend on array geometries
- circular arrays known to perform well in wave field analysis [Hulsebos 2001] and lead to efficient WDAF solution in cylindrical coordinates:

Example: Transformation $\mathcal{T}_2$

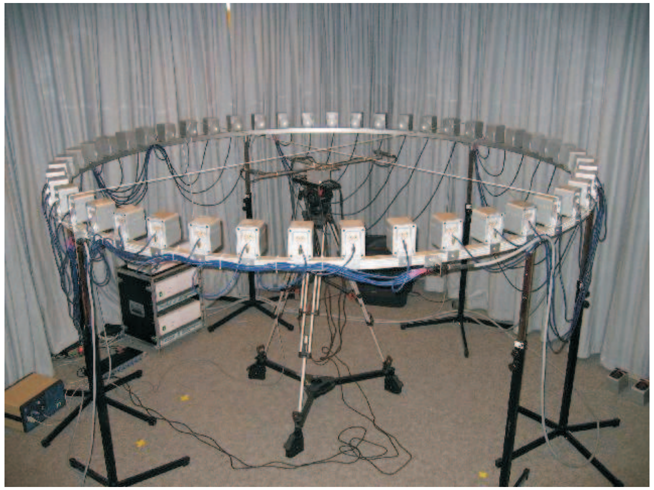
$$
\begin{aligned}
\underline{\tilde{y}}^{(1)}(k_\theta,\omega) &= \frac{j^{1-k_\theta}}{D_R(k_\theta,\omega)} \left\{ H_{k_\theta}^{(2)'}(kR)\underline{\tilde{p}}_{\mathrm{y}}(k_\theta,\omega) - H_{k_\theta}^{(2)}(kR) j\rho c \underline{\tilde{v}}_{\mathrm{y,n}}(k_\theta,\omega) \right\}, \\
\underline{\tilde{y}}^{(2)}(k_\theta,\omega) &= \frac{-j^{1+k_\theta}}{D_R(k_\theta,\omega)} \left\{ H_{k_\theta}^{(1)'}(kR)\underline{\tilde{p}}_{\mathrm{y}}(k_\theta,\omega) - H_{k_\theta}^{(1)}(kR) j\rho c \underline{\tilde{v}}_{\mathrm{y,n}}(k_\theta,\omega) \right\}
\end{aligned}
$$

(For AEC: recording signals using pressure and pressure gradient microphone elements)

Evaluation of WDAF for Acoustic Echo Cancellation (1)

Example setup:

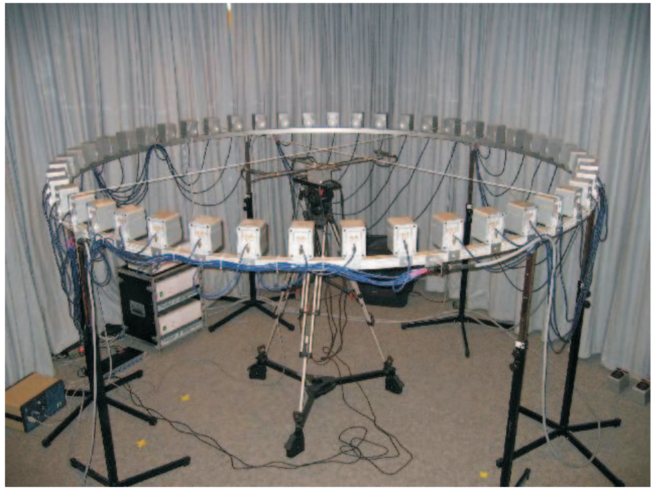
- Measured data, $T_{60} \approx 500$ ms
- Two concentric circular arrays
- 48 loudsp., $R_{loudsp} = 142$ cm, spacing 19 cm
- 48 mics, $R_{mic} = 75$ cm, spacing 9.8 cm
- Adaptation: wave-number selective FDAFs, $L = 1024$, overlap factor 256



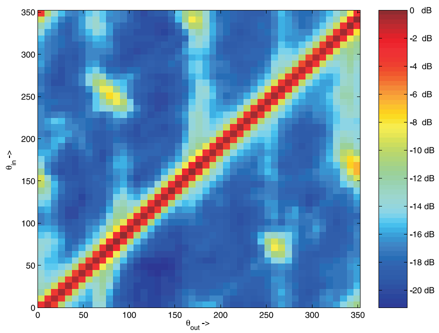
Evaluation of WDAF for Acoustic Echo Cancellation (1)

Example setup:

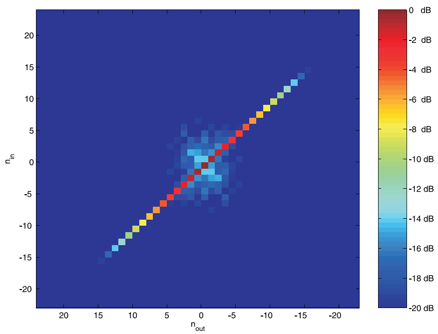
- Measured data, $T_{60} \approx 500$ ms
- Two concentric circular arrays
- 48 loudsp., $R_{loudsp} = 142$ cm, spacing 19 cm
- 48 mics, $R_{mic} = 75$ cm, spacing 9.8 cm
- Adaptation: wave-number selective FDAFs, $L = 1024$, overlap factor 256



Loudspeaker-enclosure-microphone system in the transform domain:



spatial domain (angles)



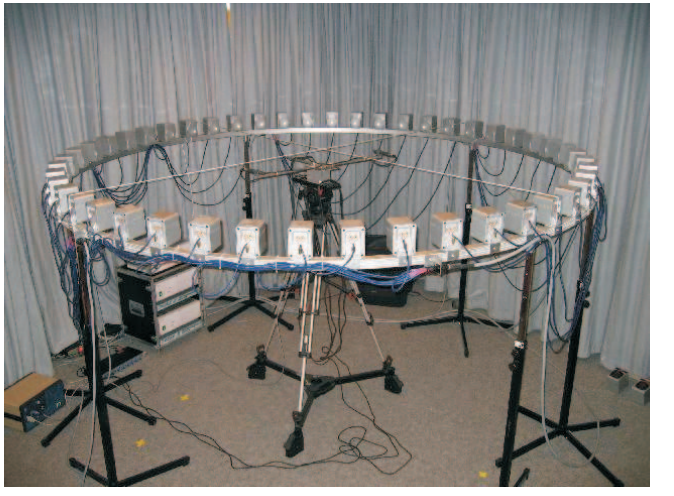
angular wave number domain

$\Rightarrow$ 70 (ideally 48) filters instead of $48 \cdot 48 = 2304$ filters sufficient

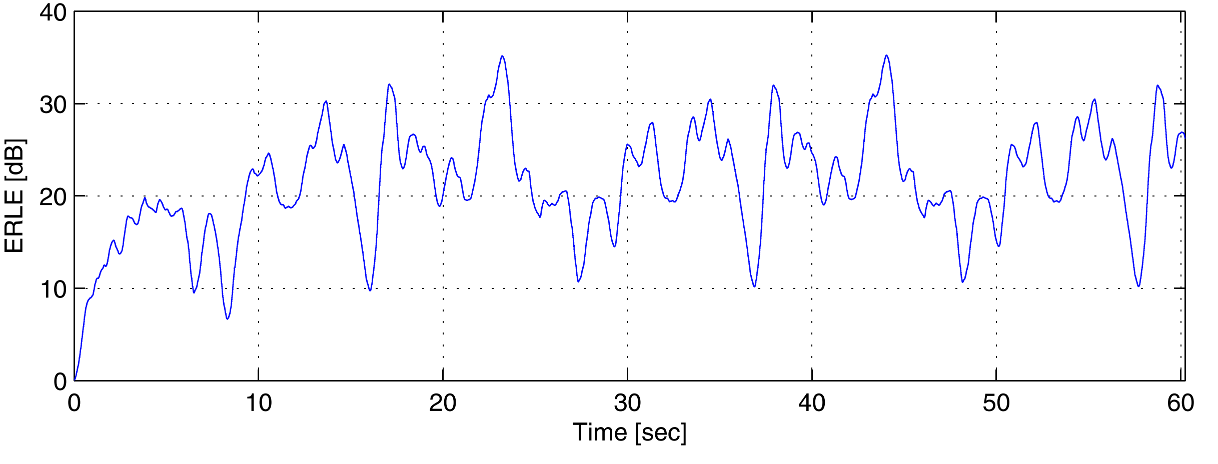
Evaluation of WDAF for Acoustic Echo Cancellation (2)

Example setup:

- Measured data, $T_{60} \approx 500$ ms
- Two concentric circular arrays
- 48 loudsp, $R_{loudsp} = 142$ cm, spacing 19 cm
- 48 mics, $R_{mic} = 75$ cm, spacing 9.8 cm
- Adaptation: wave-number selective FDAFs, $L = 1024$, overlap factor 256



Echo attenuation for music:



Summary and Outlook

Wave-Domain Adaptive Filtering: a novel concept for efficient adaptive multichannel systems

- Spatio-temporal orthogonalization of large MIMO systems
- Very robust and fast convergence
- Low computational complexity
- Verified here for acoustic echo cancellation in a WFS-based system
- Allows integrated solutions for human-machine interfaces with wave field synthesis