Pattern Recognition

Part 3: Beamforming

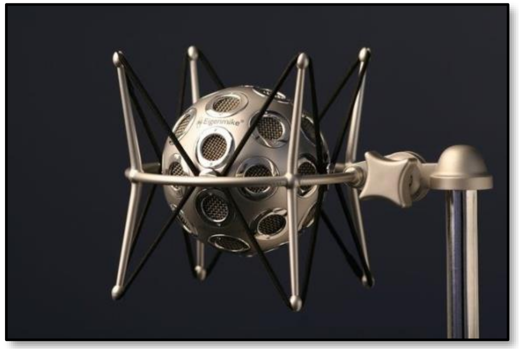
Gerhard Schmidt

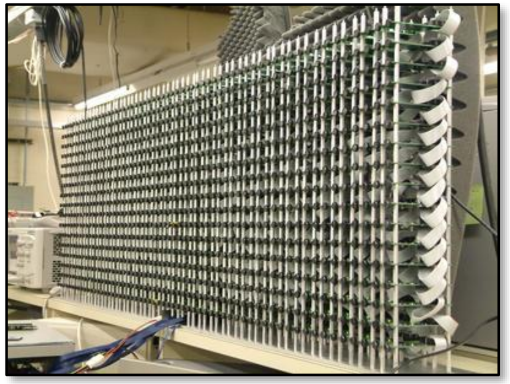
Christian-Albrechts-Universität zu Kiel
Faculty of Engineering
Institute of Electrical and Information Engineering
Digital Signal Processing and System Theory

Contents

- Introduction
- Characteristic of multi-microphone systems
- Delay-and-sum structures
- Filter-and-sum structures
- Interference compensation
- Audio examples and results
- Outlook on postfilter structures

Introduction – Part 1





Literature

Beamforming

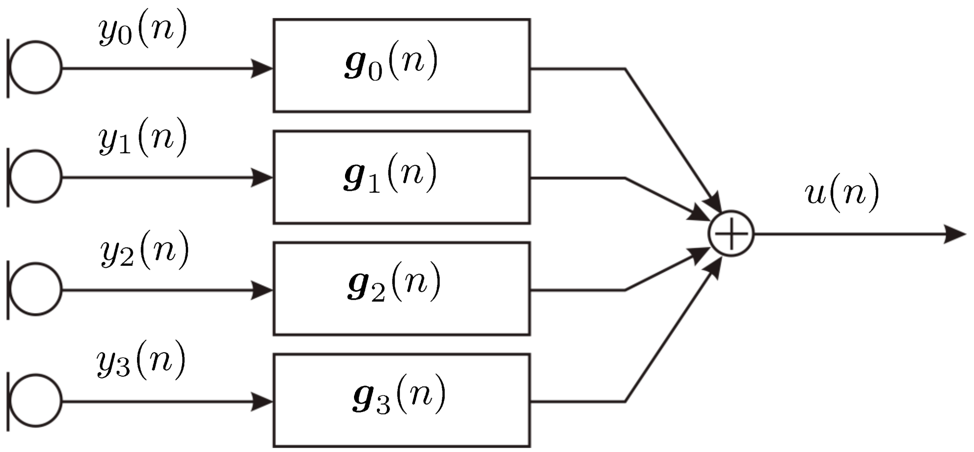
- E. Hänsler / G. Schmidt: ***Acoustic Echo and Noise Control – Chapater 11 (Beamforming)***, Wiley, 2004
- H. L. Van Trees: ***Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory***, Wiley, 2002
- W. Herbordt: ***Sound Capture for Human/Machine Interfaces: Practical Aspects of Microphone Array Signal Processing***, Springer, 2005

Postfiltering

- K. U. Simmer, J. Bitzer, C. Marro: ***Post-Filtering Techniques***, in M. Brandstein, D. Ward (editors), Microphone Arrays, Springer, 2001
- S. Gannot, I. Cohen: ***Adaptive Beamforming and Postfiltering***, in J. Benesty, M. M. Sondhi, Y. Huang (editors), Springer Handbook of Speech Processing, Springer, 2007

Introduction – Part 2

Basis structure:



Difference equation:

$$u(n) \;=\; \sum_{m=0}^{M-1} \sum_{i=0}^{N-1} y_m(n-i)\, g_{m,i}(n)$$

Introduction – Part 3

Difference equation in vector notation:

$$
\begin{aligned}
u(n) &= \sum_{m=0}^{M-1} \sum_{i=0}^{N-1} y_m(n-i)\, g_{m,i}(n) \\
&= \sum_{m=0}^{M-1} \boldsymbol{y}_m^{\mathrm{T}}(n)\, \boldsymbol{g}_m(n)
\end{aligned}
$$

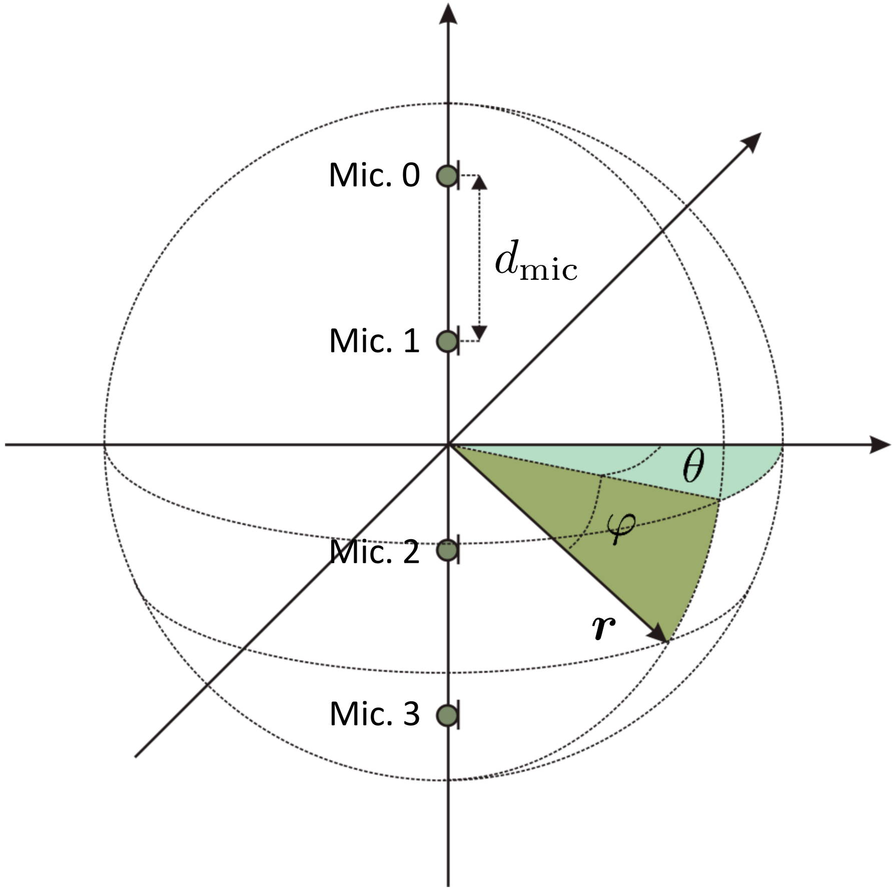
with

$$
\begin{aligned}
\boldsymbol{y}_m(n) &= \big[y_m(n),\, y_m(n-1),\, \ldots,\, y_m(n-N+1)\big]^{\mathrm{T}}, \\
\boldsymbol{g}_m(n) &= \big[g_{m,0}(n),\, g_{m,1}(n),\, \ldots,\, g_{m,N-1}(n)\big]^{\mathrm{T}}.
\end{aligned}
$$

For fixed (time-invariant) beamformers we get:

$$
u(n) = \sum_{m=0}^{M-1} \boldsymbol{y}_m^{\mathrm{T}}(n)\, \boldsymbol{g}_m \quad \Longleftrightarrow \quad U(e^{j\Omega}) = \sum_{m=0}^{M-1} Y_m(e^{j\Omega})\, G_m(e^{j\Omega})
$$

Microphone positions and coordinate systems:



- The ***origin of the coordinate system*** is often chosen as the sum of the vectors pointing at the individual microphones:

$$\sum_{m=0}^{M-1} \boldsymbol{r}_m = 0.$$

- The vector $\boldsymbol{r}$ points to the ***direction of the incoming sound*** and has a unit length:

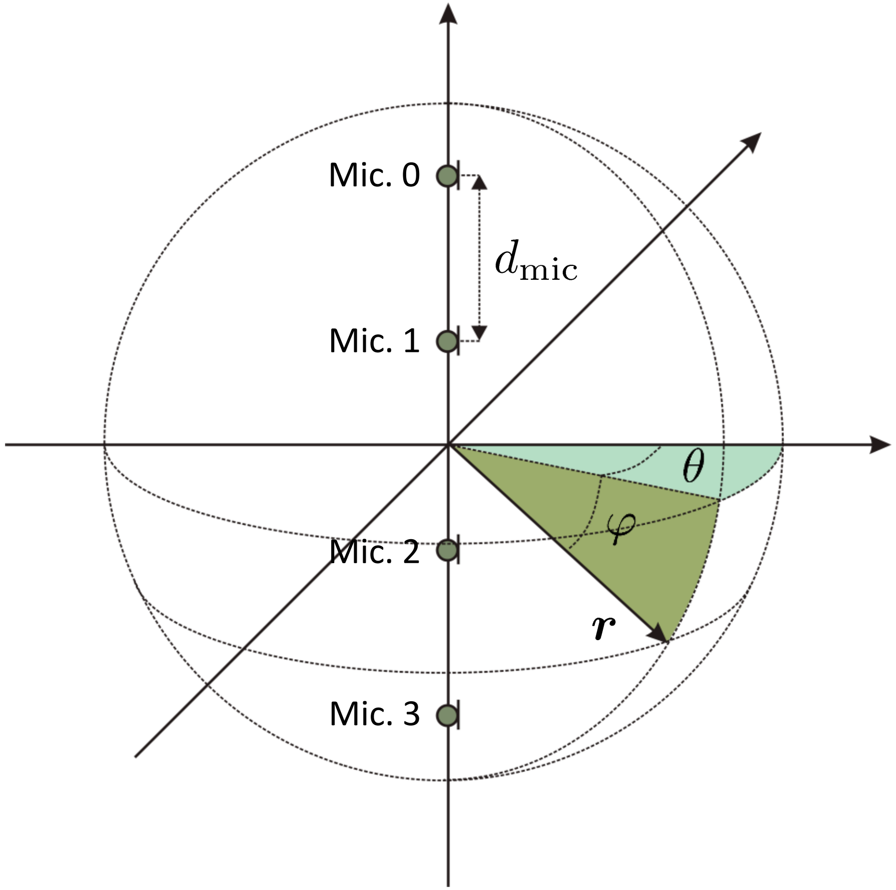
$$\|\boldsymbol{r}\| = 1.$$

- If we assume ***plain wave sound propagation*** (far-field approximation), we obtain a delay of

$$\tau_m = \frac{f_\mathrm{s}}{c} \boldsymbol{r}_m^\mathrm{T} \boldsymbol{r}.$$

for sound arriving from direction $\boldsymbol{r}$.

Directivity due to filtering and sensor characteristics:



- ***Directivity*** can be achieved either by ***spatial filtering*** of the microphone signals according to

$$u(n) = \sum_{m=0}^{M-1} \boldsymbol{y}_m^{\mathrm{T}}(n)\, \boldsymbol{g}_m$$

  or by the ***sensors themselves*** (e.g. due to cardioid characteristics).

- If we use spatial filtering a ***reference for the disturbing signal components*** can be estimated. This can be exploited by means of, e.g. a ***Wiener filter*** and leads to an additional directivity gain.

Quality Measures of Multi-Microphone Systems – Part 1

Assumptions for computing a „spatial frequency response”:

- The sound propagation is modeled as ***plane wave***:

$$S_m(e^{j\Omega}) \quad = \quad S(e^{j\Omega})\, e^{-j\Omega\tau_m}.$$

- Each microphone has got a ***receiving characteristic***, which can be described as

$$M_m(e^{j\Omega}, \boldsymbol{r}) \quad = \quad M_m(e^{j\Omega}, \varphi, \theta)\,.$$

For microphones with ***omnidirectional characteristic*** the following equation holds,

$$M_{m,\mathrm{omni}}(e^{j\Omega}, \varphi, \theta) \quad = \quad 1.$$

Microphones with ***cardioid characteristic*** can be described as

$$M_{m,\mathrm{card}}(e^{j\Omega}, \varphi, \theta) \quad = \quad \frac{1}{2}\Big[1 + \cos(\varphi)\Big].$$

Spatial frequency response

- With the above assumptions the ***desired signal component of the output spectrum of a single microphone*** can be written as

$$\begin{aligned} Y_m(e^{j\Omega}, \boldsymbol{r}) &= S_m(e^{j\Omega})\, M_m(e^{j\Omega}, \boldsymbol{r}) \\ &= S(e^{j\Omega})\; M_m(e^{j\Omega}, \boldsymbol{r})\, e^{-j\Omega\tau_m}. \end{aligned}$$

- The ***output spectrum of the beamformer*** can consequently be written as

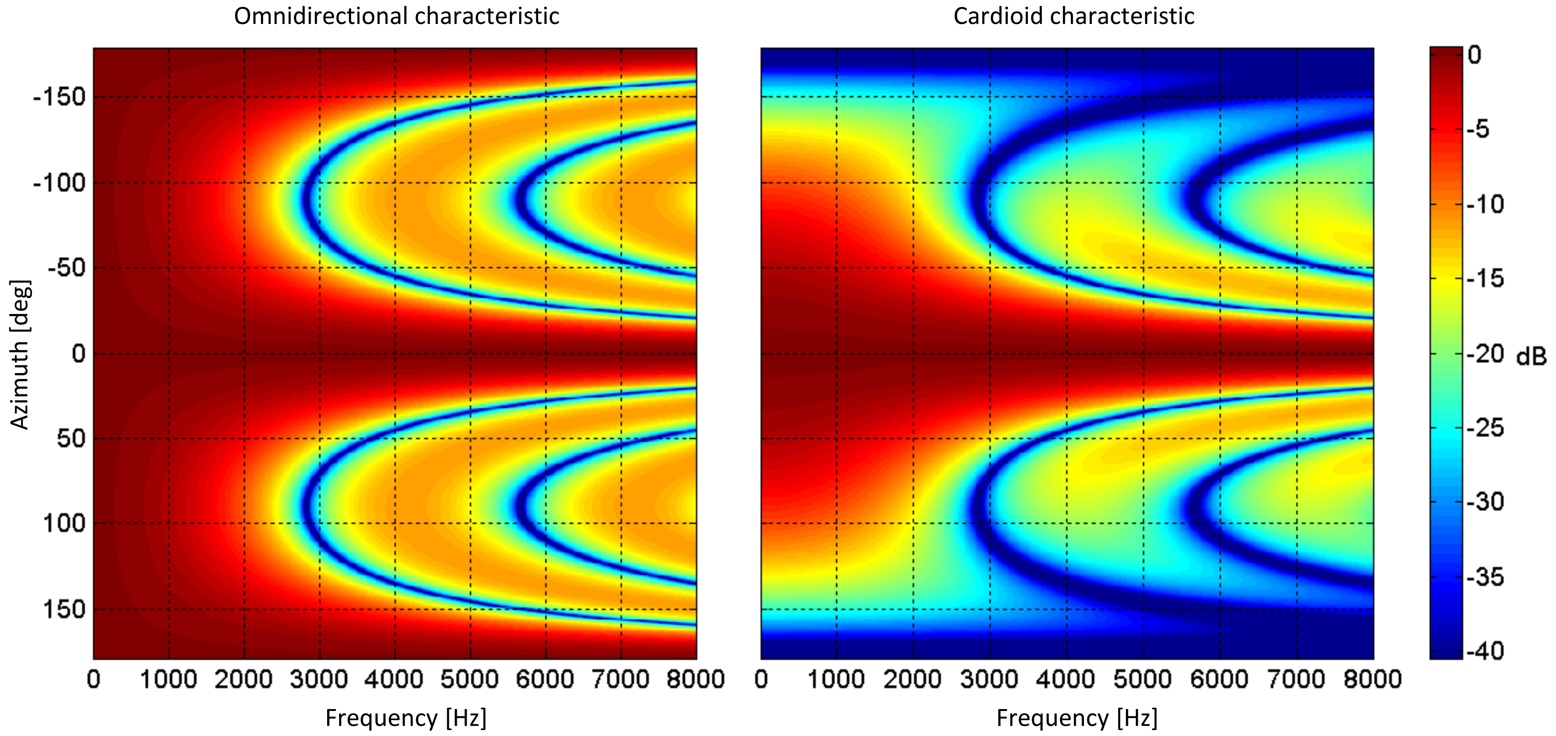
$$\begin{aligned} U(e^{j\Omega}, \boldsymbol{r}) &= \sum_{m=0}^{M-1} Y_m(e^{j\Omega}, \boldsymbol{r})\, G_m(e^{j\Omega}) \\ &= S(e^{j\Omega}) \sum_{m=0}^{M-1} M_m(e^{j\Omega}, \boldsymbol{r})\, G_m(e^{j\Omega})\, e^{-j\Omega\tau_m}. \end{aligned}$$

- Finally the ***spatial frequency response*** is defined as follows,

$$G_{\mathrm{BF}}(e^{j\Omega}, \boldsymbol{r}) = \frac{U(e^{j\Omega}, \boldsymbol{r})}{S(e^{j\Omega})} = \sum_{m=0}^{M-1} M_m(e^{j\Omega}, \boldsymbol{r})\, G_m(e^{j\Omega})\, e^{-j\Omega\tau_m}.$$

Quality Measures of Multi-Microphone Systems – Part 3

Examples of spatial frequency responses



- 4 microphones in a row in intervals of 3cm were used.
- The microphone signals were just added and weighted with ¼ .

Quality Measures of Multi-Microphone Systems – Part 4

Beampattern

- The squared absolute of the spatial frequency response is called ***beampattern***:

$$\Phi(\Omega, \boldsymbol{r}) \quad = \quad \left|G_{\mathrm{BF}}(e^{j\Omega}, \boldsymbol{r})\right|^2.$$

- If all microphones have the same ***beampattern***, the influences of the microphones and of the signal processing can be ***separated***:

$$\begin{aligned}\Phi(\Omega, \boldsymbol{r}) &= \left|\sum_{m=0}^{M-1} M_m(e^{j\Omega}, \boldsymbol{r})\, G_m(e^{j\Omega})\, e^{-j\Omega\tau_m}\right|^2 \\ &= \left|M(e^{j\Omega}, \boldsymbol{r})\right|^2 \left|\sum_{m=0}^{M-1} G_m(e^{j\Omega})\, e^{-j\Omega\tau_m}\right|^2 \\ &= \Phi_{\mathrm{mic}}(\Omega, \boldsymbol{r})\, \Phi_{\mathrm{sig}}(\Omega, \boldsymbol{r}).\end{aligned}$$

Quality Measures of Multi-Microphone Systems – Part 5

Array gain:

- If a characteristic number is needed, the so-called **array gain** can be used,

$$Q(\Omega, \boldsymbol{r}_\mathrm{s}) = \frac{\Phi(\Omega, \boldsymbol{r}_\mathrm{s})}{\frac{1}{A_\mathrm{sphere}} \int\limits_{\mathrm{surface}} \Phi(\Omega, \boldsymbol{r})\, dA}.$$

- The vector $\boldsymbol{r}_\mathrm{s}$ is pointing into the direction of the desired signal.

- The logarithmic array gain

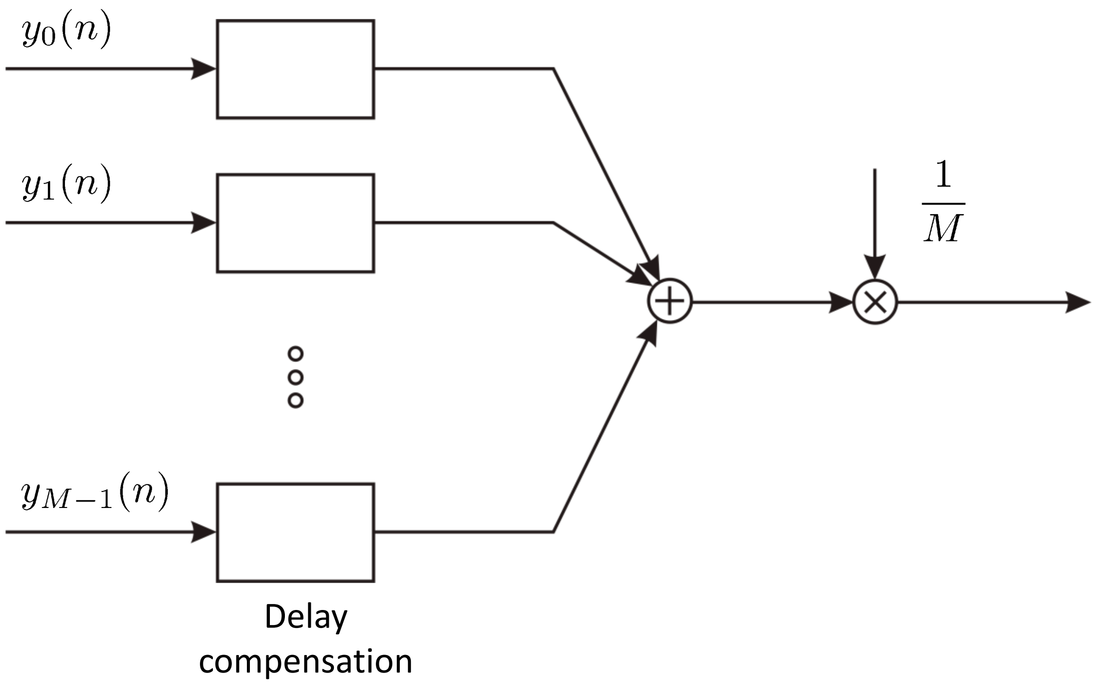
$$Q_\mathrm{log}(\Omega, \boldsymbol{r}_\mathrm{s}) = 10\,\log_{10}\left\{Q(\Omega, \boldsymbol{r}_\mathrm{s})\right\}$$

  is called ***directivity index***.

- Both quantities describe the gain compared to an onmidirectional sensor (e.g., a microphone with omnidirectional characteristic).

Delay-and-Sum Structure – Part 1

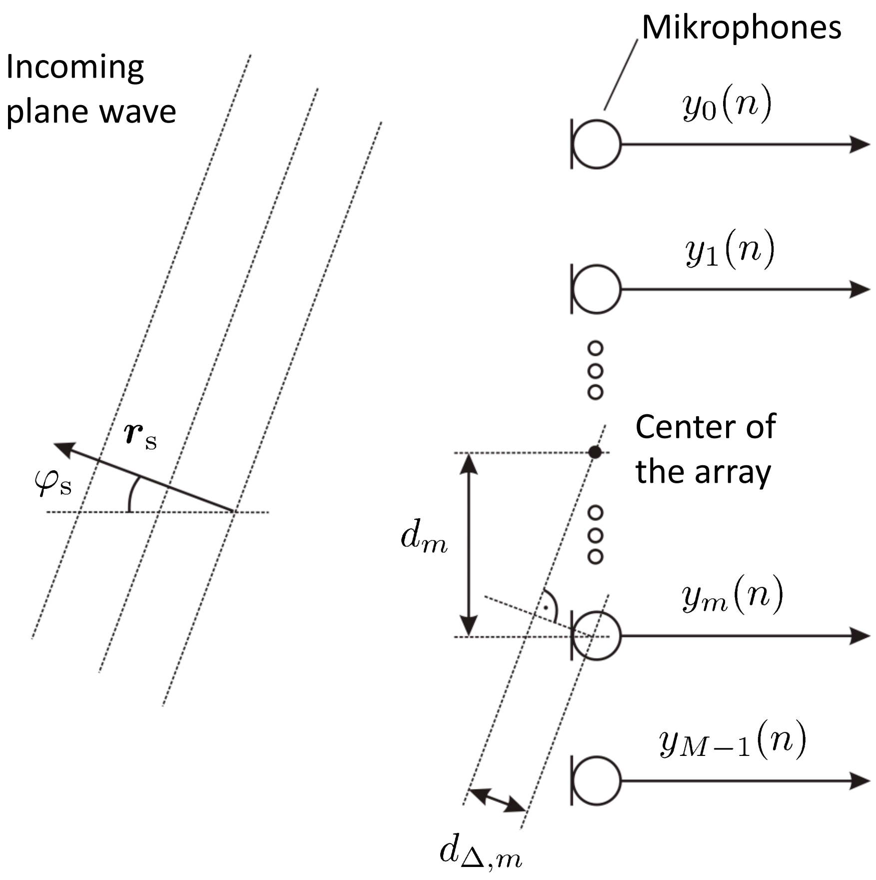
Basic structure



- The microphone signals are being delayed in such a way that all signals from a predefined preferred direction are synchronized after the delay compensation.
- In the next step, the signals are weighted and added in such a way that at the output, the signal power of the desired signal from the preferred direction is the same as at the input (but without reflections).
- ***Interferences*** which do not arrive from the preferred direction, will not be added in-phase and will therefore be attenuated.

Delay-and-Sum Structure – Part 2

Identify the necessary delays



- In the case of a linear array with constant microphone distance, the distance of the m^th microphone to the center of the array can be calculated as

$$d_m = \|\boldsymbol{r}_m\| = \left| m - \frac{M-1}{2} \right| d_{\text{mic}}.$$

- Based on this distance, we can calculate the ***time delay*** of the plane wave to arrive at the m^th microphone,

$$t_m = \frac{d_{\Delta,m}}{c} = \frac{d_m \, \sin(\varphi_\mathrm{s})}{c}.$$

- Using the sample rate, the time delay can be expressed in frames,

$$\tau_m = t_m \, f_\mathrm{s}.$$

Delay-and-Sum Structure – Part 3

Optimal solution

$$G_{\mathrm{opt},m}(e^{j\Omega}) = e^{-j\Omega\tau_m} \text{ with } -\pi < \Omega \leq \pi \quad \bullet\!\!-\!\!\!-\!\!\circ \quad g_{\mathrm{opt},\,m,\,i} = \frac{\sin\left(\pi(i-\tau_m)\right)}{\pi(i-\tau_m)}$$

Implementation in time domain (example)

- The optimal impulse response is delayed to make it causal, and is then „***windowed***“,

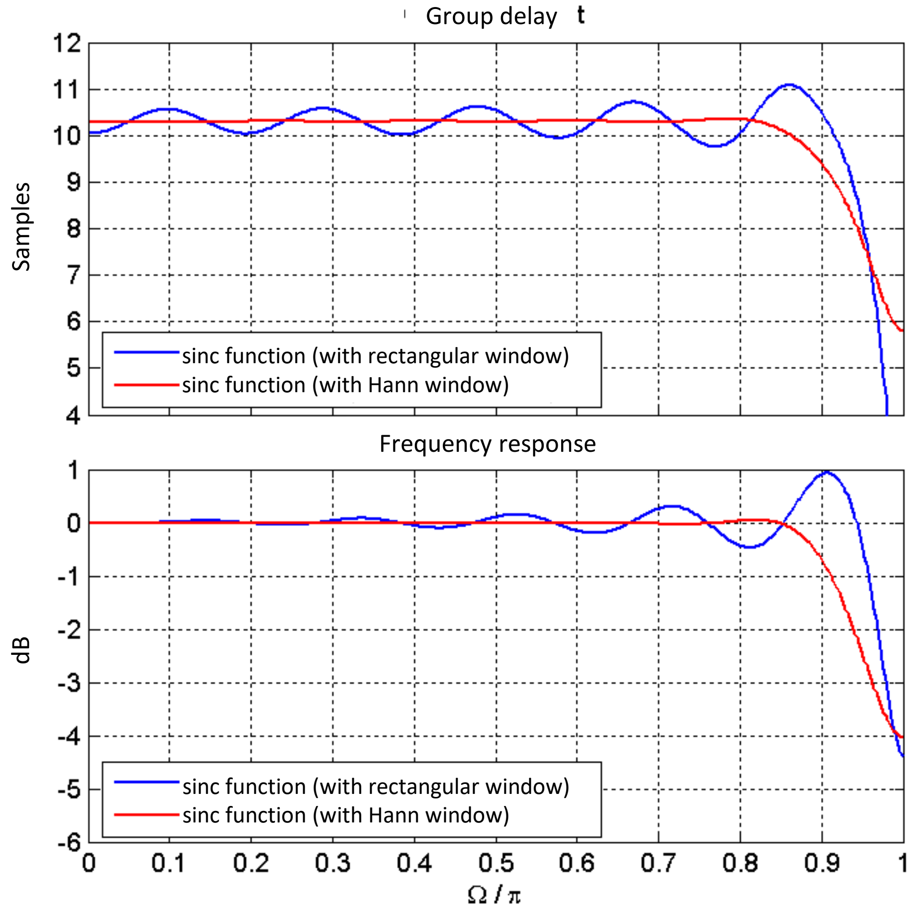
$$g_{m,i} = g_{\mathrm{opt},m,\,i\,-\,N/2}\, f_i.$$

- As window function, for example the Hann window can be chosen,

$$f_i = \begin{cases} K\left[1+\sin\left(\frac{2\pi}{N}(i-\frac{N}{2})\right)\right], & \text{for } i \in \{0,\,...,\,N-1\}, \\ 0, & \text{else.} \end{cases}$$

Delay-and-Sum Structure – Part 4

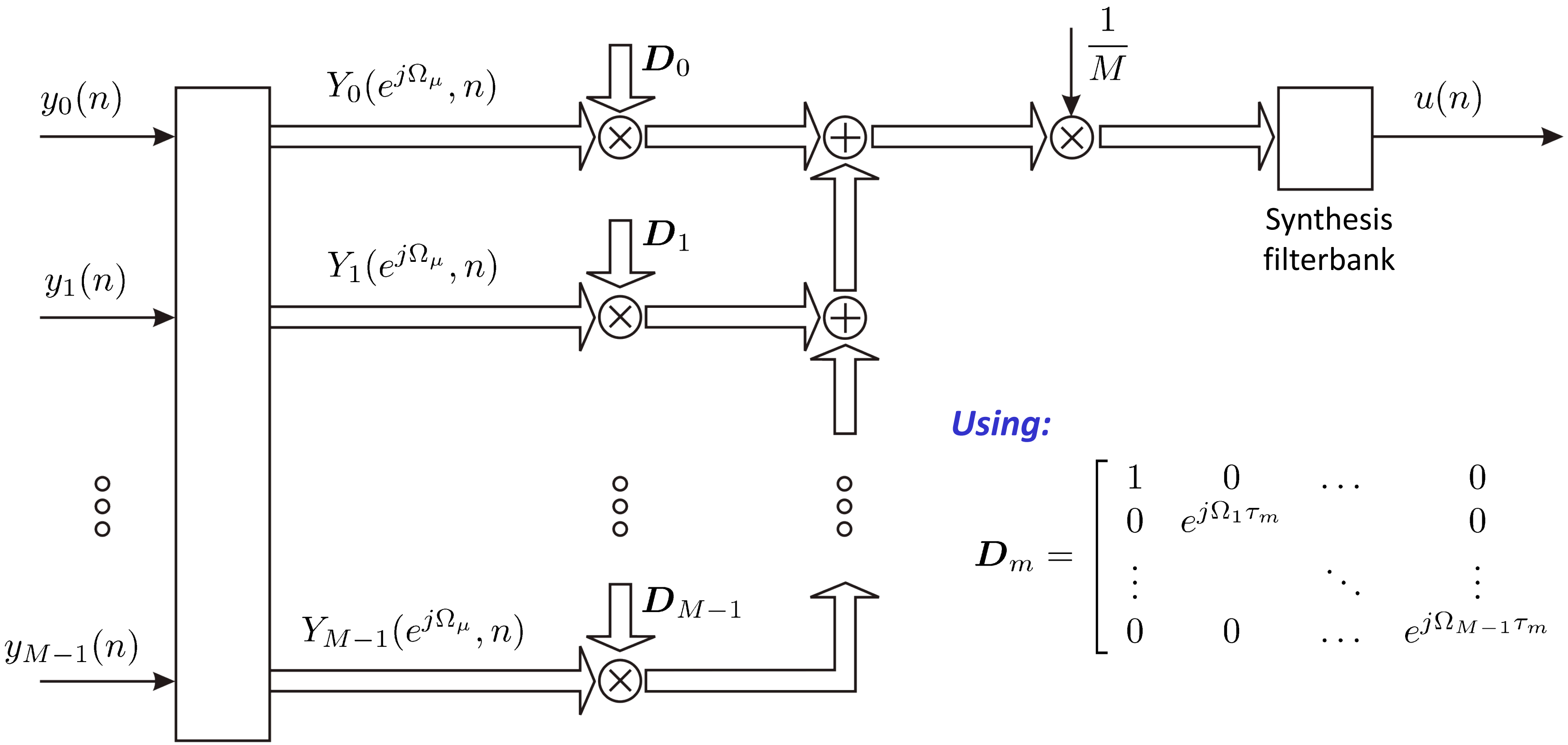
Implementation in time domain (example)



- Goal: Design a filter with group delay of 10.3 samples.
- Constraint: 21 filter coefficients may be used.

Delay-and-Sum Structure – Part 5

Implementation in the frequency domain

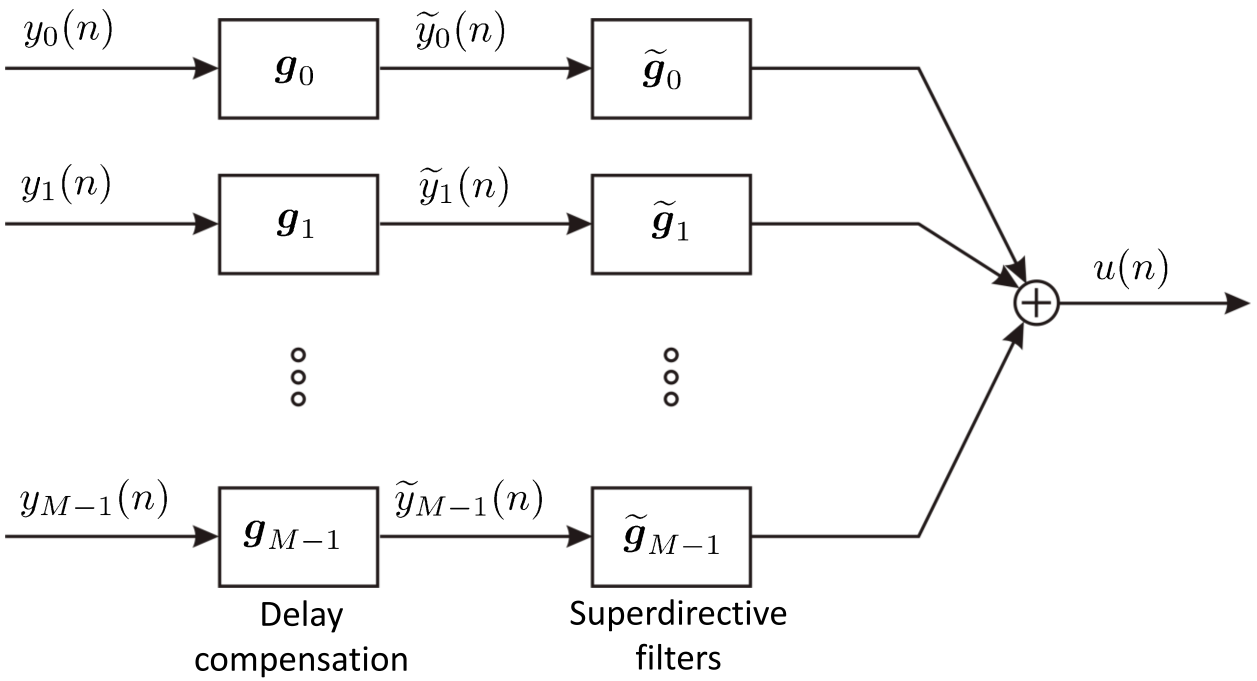


Using:

$$\boldsymbol{D}_m = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & e^{j\Omega_1 \tau_m} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & e^{j\Omega_{M-1}\tau_m} \end{bmatrix}$$

Filter-and-Sum Structure – Part 1

Basic principle



- In addition to the delay compensation, the array characteristic are to be improved using ***filters***.
- As soon as the beamformer properties are better than the delay-and-sum approach, the beamformer is called ***superdirective***.
- The introduced filters are designed to be optimal for the broadside direction as preferred direction.

Filter-and-Sum Structure – Part 2

Filter design

- Difference equation:

$$u(n) = \sum_{m=0}^{M-1} \sum_{i=0}^{\widetilde{N}-1} \widetilde{y}_m(n-i)\, \widetilde{g}_{m,i}$$

$$= \sum_{m=0}^{M-1} \widetilde{\boldsymbol{y}}_m^{\mathrm{T}}(n)\, \widetilde{\boldsymbol{g}}_m$$

- Optimization criterion:

$$\mathrm{E}\{u^2(n)\} \longrightarrow \min \qquad \text{with the constraint} \qquad \sum_{m=0}^{M-1} \widetilde{\boldsymbol{g}}_m = \boldsymbol{c}$$

Filter-and-Sum Structure – Part 3

Constraints of the filter design

$$\sum_{m=0}^{M-1} \widetilde{\boldsymbol{g}}_m = \boldsymbol{c}$$

$$\begin{array}{ccccccccc}
\widetilde{\boldsymbol{g}}_0 & = & \widetilde{g}_{0,0} & \dots & \widetilde{g}_{0,k-1} & \widetilde{g}_{0,k} & \widetilde{g}_{0,k+1} & \dots & \widetilde{g}_{0,\widetilde{N}-1} \\
\widetilde{\boldsymbol{g}}_1 & = & \widetilde{g}_{1,0} & \dots & \widetilde{g}_{1,k-1} & \widetilde{g}_{1,k} & \widetilde{g}_{1,k+1} & \dots & \widetilde{g}_{1,\widetilde{N}-1} \\
\dots & & \dots & & \dots & & \dots & \dots & \\
\widetilde{\boldsymbol{g}}_{M-1} & = & \widetilde{g}_{M-1,0} & \dots & \widetilde{g}_{M-1,k-1} & \widetilde{g}_{M-1,k} & \widetilde{g}_{M-1,k+1} & \dots & \widetilde{g}_{M-1,\widetilde{N}-1} \\
\hline
\boldsymbol{c} & = & 0 & \dots & 0 & 1 & 0 & \dots & 0
\end{array}$$

This means: Signals from the broadside direction can pass the filter network without any attenuation.

The „zero solution“ is excluded by introducing the constraint!

Filter-and-Sum Structure – Part 4

Filter design

- Introducing ***overall signal vectors*** and ***overall filter vectors***:

$$\begin{aligned}\tilde{\boldsymbol{g}} &= \left[\tilde{\boldsymbol{g}}_0^{\mathrm{T}}, \tilde{\boldsymbol{g}}_1^{\mathrm{T}}, \ldots, \tilde{\boldsymbol{g}}_{M-1}^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \tilde{\boldsymbol{y}}(n) &= \left[\tilde{\boldsymbol{y}}_0^{\mathrm{T}}(n), \tilde{\boldsymbol{y}}_1^{\mathrm{T}}(n), \ldots, \tilde{\boldsymbol{y}}_{M-1}^{\mathrm{T}}(n)\right]^{\mathrm{T}}.\end{aligned}$$

- Subsequently, the ***beamformer output signal*** can be written as follows:

$$u(n) = \tilde{\boldsymbol{g}}^{\mathrm{T}}\,\tilde{\boldsymbol{y}}(n) = \tilde{\boldsymbol{y}}^{\mathrm{T}}(n)\,\tilde{\boldsymbol{g}}.$$

- The ***mean output signal power*** results in:

$$\begin{aligned}\mathrm{E}\{u^2(n)\} &= \mathrm{E}\{\tilde{\boldsymbol{g}}^{\mathrm{T}}\,\tilde{\boldsymbol{y}}(n)\,\tilde{\boldsymbol{y}}^{\mathrm{T}}(n)\,\tilde{\boldsymbol{g}}\} \\ &= \tilde{\boldsymbol{g}}^{\mathrm{T}}\,\mathrm{E}\{\tilde{\boldsymbol{y}}(n)\,\tilde{\boldsymbol{y}}^{\mathrm{T}}(n)\}\,\tilde{\boldsymbol{g}} \\ &= \tilde{\boldsymbol{g}}^{\mathrm{T}}\,\boldsymbol{S}_{\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}}\,\tilde{\boldsymbol{g}}.\end{aligned}$$

Filter-and-Sum Structure – Part 5

Filter design

- The ***constraint*** can be rewritten as follows:

$$\begin{aligned} \boldsymbol{c} &= \sum_{m=0}^{M-1} \widetilde{\boldsymbol{g}}_m \\ &= \left[\boldsymbol{I}_{NxN},\, \boldsymbol{I}_{NxN},\, \ldots,\, \boldsymbol{I}_{NxN}\right] \widetilde{\boldsymbol{g}} \\ &= \boldsymbol{C}\,\widetilde{\boldsymbol{g}}. \end{aligned}$$

- Then, using a ***Lagrange approach*** the following function can be minimized:

$$F(\widetilde{\boldsymbol{g}}) = \frac{1}{2}\,\widetilde{\boldsymbol{g}}^{\mathrm{T}}\boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}\,\widetilde{\boldsymbol{g}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{C}\,\widetilde{\boldsymbol{g}} - \boldsymbol{c}).$$

- Calculating the ***gradient*** with respect to $\widetilde{\boldsymbol{g}}$ results in:

$$\nabla_{\widetilde{\boldsymbol{g}}} F(\widetilde{\boldsymbol{g}}) = \boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}\,\widetilde{\boldsymbol{g}} + \boldsymbol{C}^{\mathrm{T}}\,\boldsymbol{\lambda}.$$

Filter-and-Sum Structure – Part 6

Filter design

- ***Setting the gradient to zero*** results in:

$$\widetilde{\boldsymbol{g}}_{\mathrm{opt}} = -\boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}^{-1}\,\boldsymbol{C}^{\mathrm{T}}\,\boldsymbol{\lambda}.$$

- Inserting this result into the ***constraint*** we get:

$$\boldsymbol{c} = \boldsymbol{C}\left(-\,\boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}^{-1}\,\boldsymbol{C}^{\mathrm{T}}\,\boldsymbol{\lambda}\right).$$

- Resolving this equation to the ***Lagrange multiplication vector*** results in:

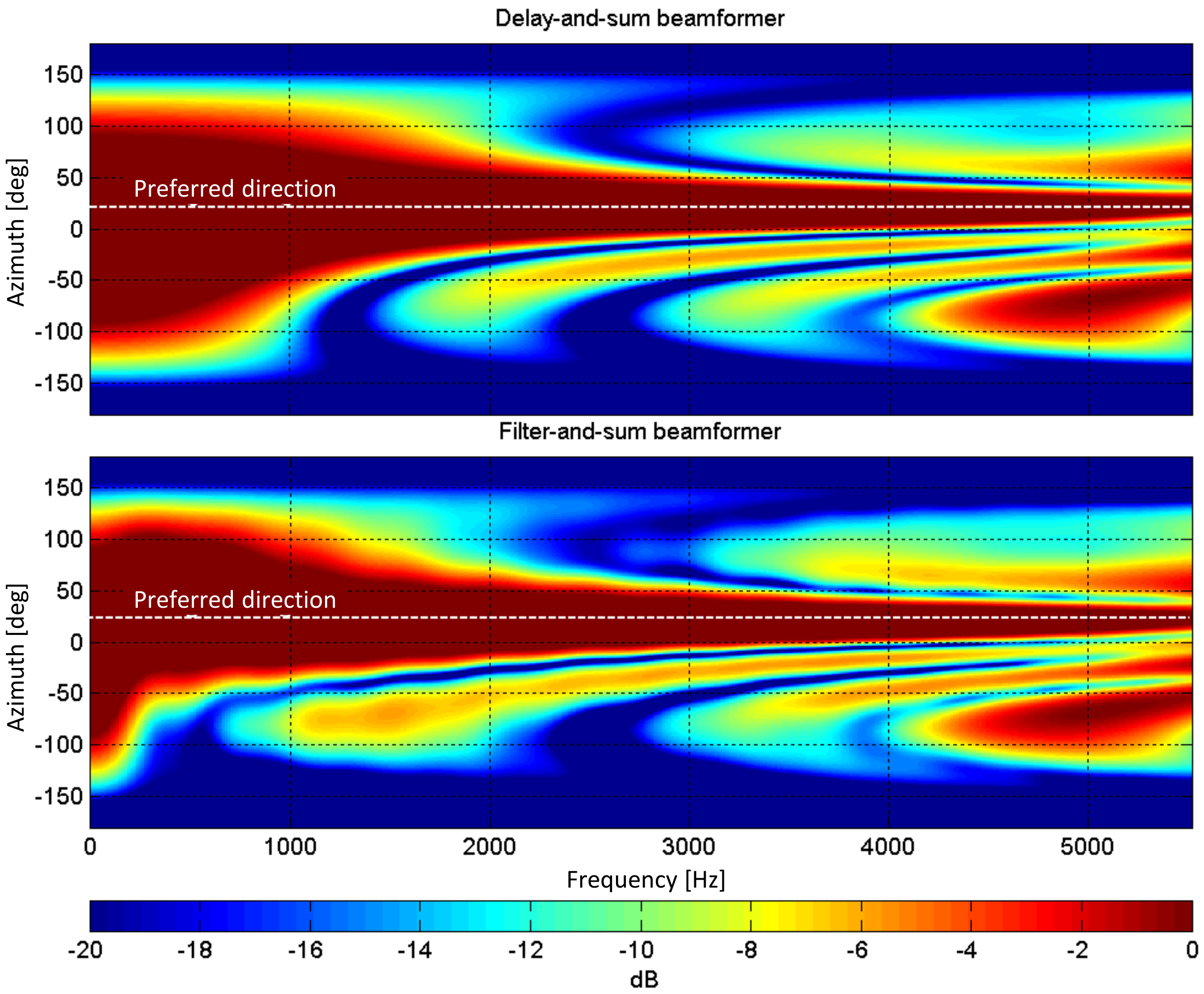
$$\boldsymbol{\lambda} = \left(-\,\boldsymbol{C}\,\boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}^{-1}\,\boldsymbol{C}^{\mathrm{T}}\right)^{-1}\boldsymbol{c}.$$

- Finally, we get:

$$\widetilde{\boldsymbol{g}}_{\mathrm{opt}} = \boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}^{-1}\,\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}\,\boldsymbol{S}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}}^{-1}\,\boldsymbol{C}^{\mathrm{T}}\right)^{-1}\boldsymbol{c}.$$

The filter coefficients are defined by the auto correlation matrix of the interference sound field!

Filter-and-Sum Structure – Part 7



- Goal: Design filters for a microphone array consisting of 4 microphones.
- The microphone distance is 4 cm.

Interference Cancellation

Basic principle

- Up to now, we had to make assumptions about the properties of the sound field. If this is not possible, we should use an ***adaptive error power minimization*** instead.

- A direct application of adaptive algorithms would lead to the so-called „zero solution“ (all filter coefficients are zero). So as before, we need to introduce a ***constraint***.

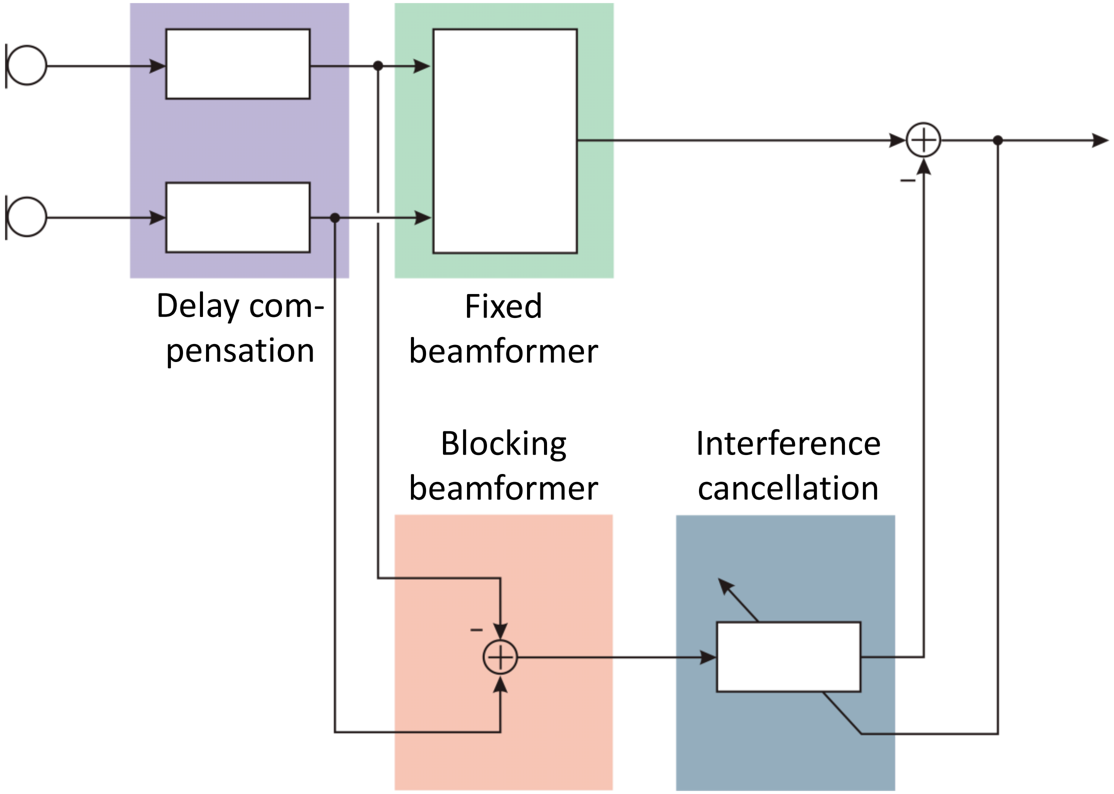
- This constraint can either be taken care of when calculating the gradient (e.g., using the Frost approach), or implemented in the filter structure using a ***desired signal blocking***. The latter is much more efficient.

- The desired signal blocking has the task to block the desired signal completely but to let pass all interferences. Using this output signal, a ***minimization of the error power without constraints*** can be applied.

C | A | U
Christian-Albrechts-Universität zu Kiel

Interference Cancellation – Blocking the desired signal (part 1)

Subtraction of delay-compensated microphone signals



Interference Cancellation – Blocking the Desired Signal (Part 2)

Subtracting the delay-compensated microphone signals

Advantages:

- Very simple and ***computationally efficient structure***.
- Besides just to subtract the signals, also the principles of filter design may be applied. Hereby, the ***width of the blocking*** can be controlled.

Drawbacks:

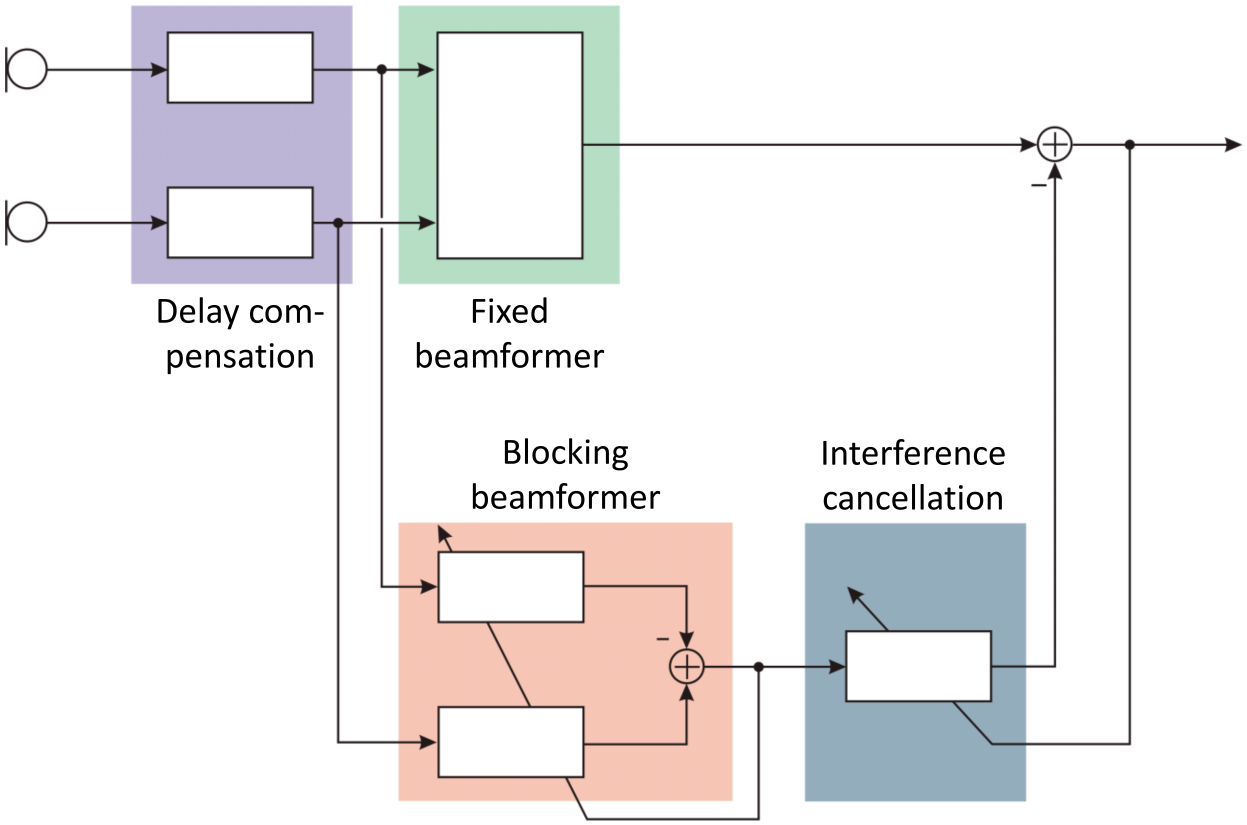
- In the case of errors in the delay compensation, or if different sensors are used, the desired signal may pass the blocking structure and may be ***compensated*** unintentionally.
- Echo components of the desired signal may pass the blocking structure, which may equally lead to a ***compensation of the desired signal***.

Conclusion:

- This blocking structure is usually used to classify the current situation (e.g., „desired signal active“, „interference active“, etc.). Based on this classification, further and more sophisticated approaches may be regulated.

Interference Cancellation – Blocking the Desired Signal (Part 3)

Adaptive subtraction of delay-compensated microphone signals



Interference Cancellation – Blocking the Desired Signal (Part 4)

Adaptive subtraction of delay-compensated microphone signals

Advantages:

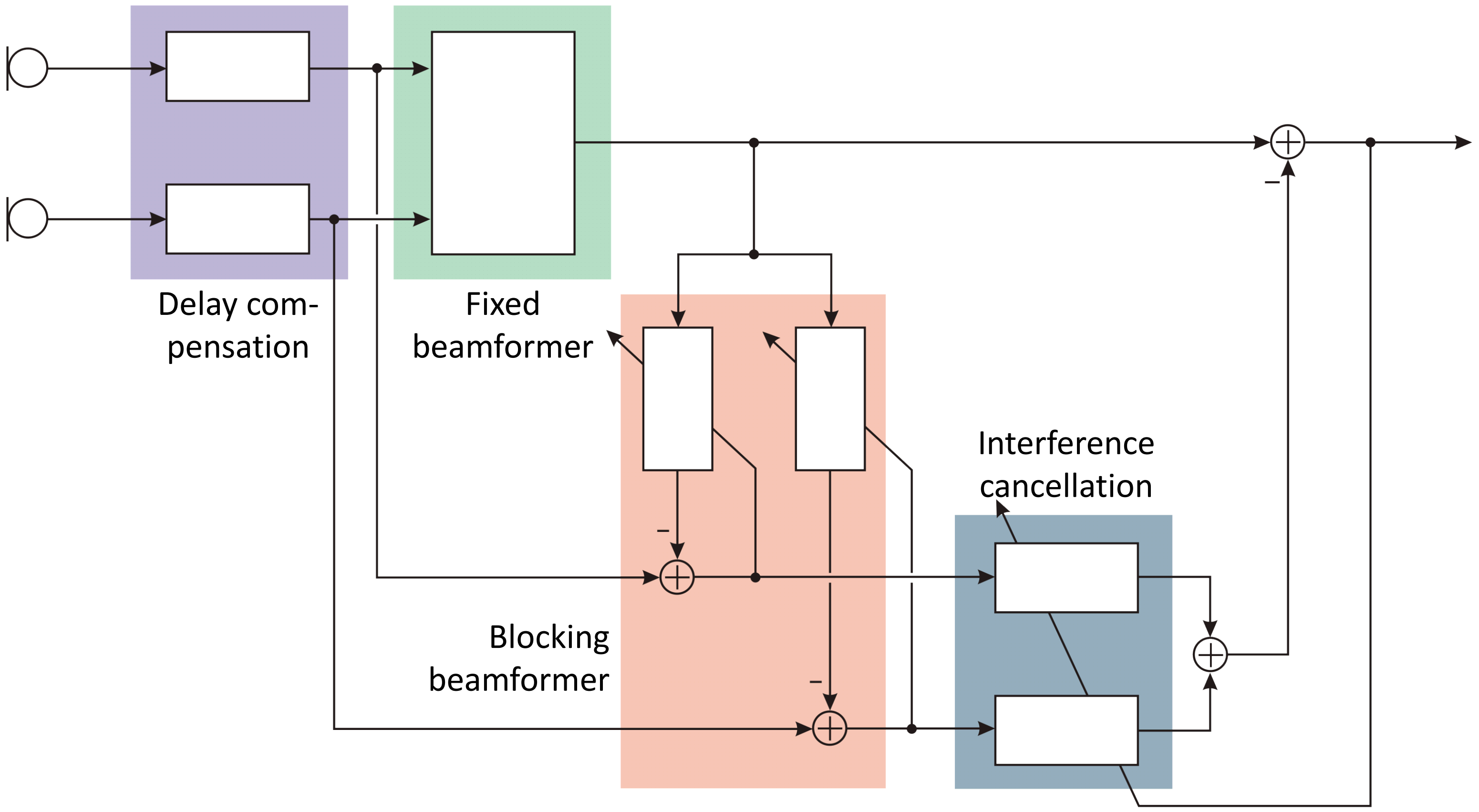
- ***Errors*** in the delay compensation ***may be compensated*** (provided that the situation was classified correctly).
- ***Echo components*** can be (partly) removed.
- The structure can be used to ***localize*** the desired speaker (topic for a talk...)

Drawbacks:

- In the adaption, a ***constraint*** has to be fulfilled (e.g., the sum of the norms of the filters has to be constant).
- A ***robust control*** of the filter adaption is necessary.

Interference Cancellation – Blocking the Desired Signal (Part 5)

Adaptive subtraction of delay-compensated microphone signals and beamformer output



Interference Cancellation – Blocking the Desired Signal (Part 6)

Adaptive subtraction of delay-compensated microphone signals and beamformer output

Advantages:

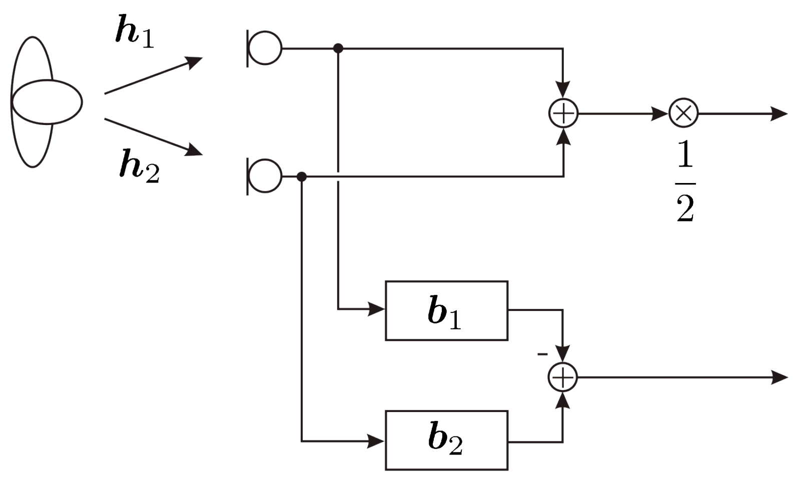
- ***Echo components*** can be (party) removed.
- The reference signal of the desired speaker (beamformer output) has a ***better signal-to-noise ratio*** than using the adaptive microphone signal filtering.
- Only one signal has to be kept in memory (***less memory requirements*** than the structure before).

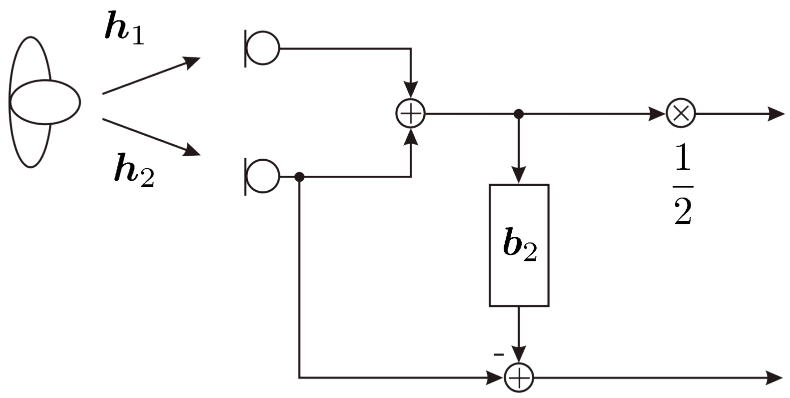
Drawbacks:

- To approximate the inverse room transfer function, usually ***more parameters*** are necessary (compared to direct approximation).
- A ***robust control*** of the filter adaption is necessary.

Interference Cancellation – Blocking the Desired Signal (Part 7)

Differences between the blocking structures:





$$H_1(e^{j\Omega})\,B_1(e^{j\Omega}) = H_2(e^{j\Omega})\,B_2(e^{j\Omega})$$

$$\Longrightarrow \quad \begin{aligned} B_1(e^{j\Omega}) &= H_2(e^{j\Omega})\,C(e^{j\Omega}) \\ B_2(e^{j\Omega}) &= H_1(e^{j\Omega})\,C(e^{j\Omega}) \end{aligned}$$

$$H_2(e^{j\Omega}) = \left[H_1(e^{j\Omega}) + H_2(e^{j\Omega})\right]\,B_2(e^{j\Omega})$$

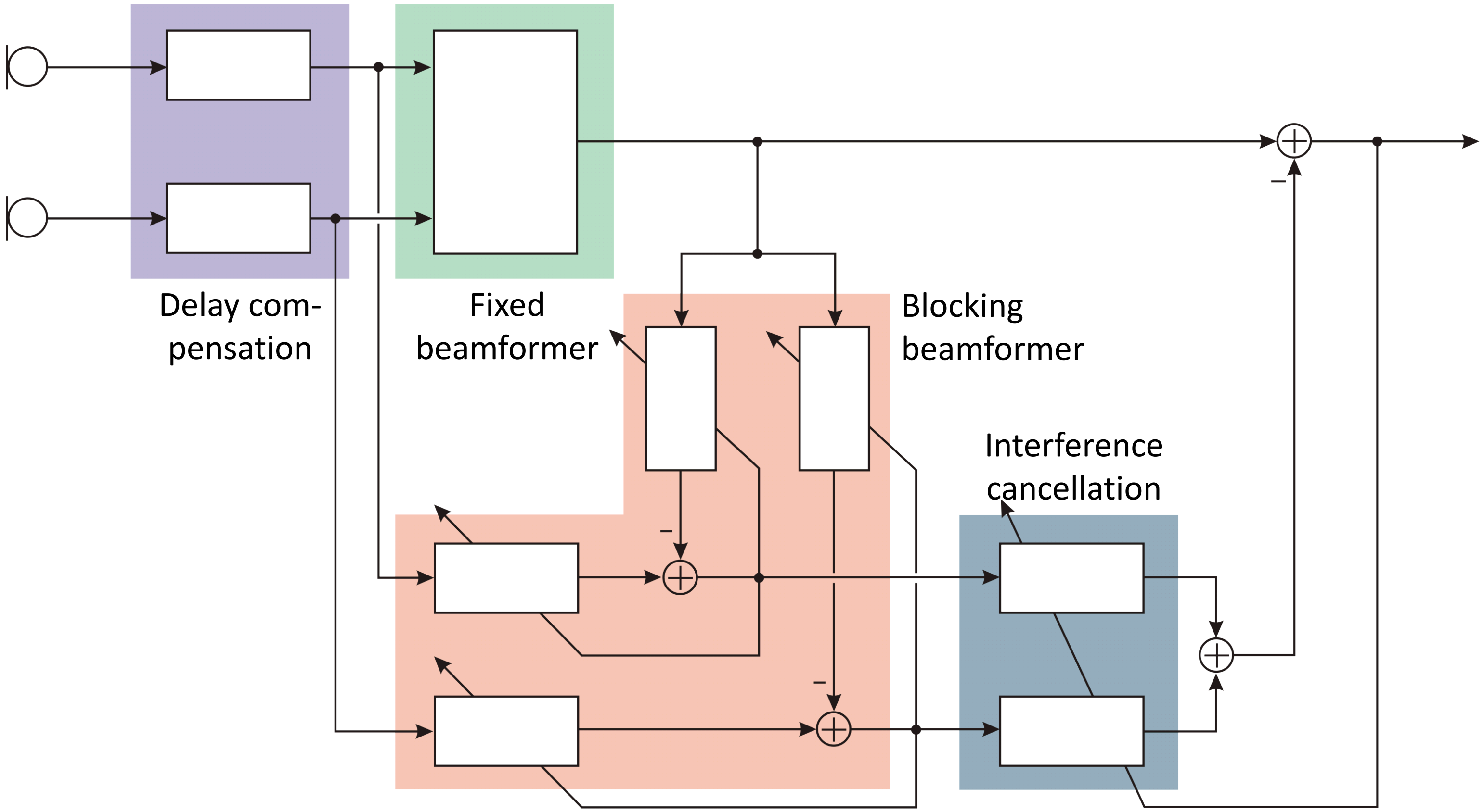
$$\Longrightarrow \quad B_2(e^{j\Omega}) = \frac{H_2(e^{j\Omega})}{H_1(e^{j\Omega}) + H_2(e^{j\Omega})}$$

The approximation of inverse impulse responses is necessary (zeros-only model)!

C | A | U
Christian-Albrechts-Universität zu Kiel

Interference Cancellation – Blocking the Desired Signal (Part 8)

Double-adaptive subtraction of microphone signals and beamformer output



Interference Cancellation – Blocking the Desired Signal (Part 9)

Double-adaptive subtraction of microphone signals and beamformer output

Advantages:

- ***Echo components*** can be (partly) removed.
- The reference signal of the desired speaker (***beamformer output***) has a ***better signal-to-noise ratio*** than using the adaptive microphone signal filtering.
- The approximation of inverted transfer functions is not necessary.

Drawbacks:

- A ***robust control*** of the filter adaption is necessary.
- Again, we need to ***normalize*** (at least one) filter norm.

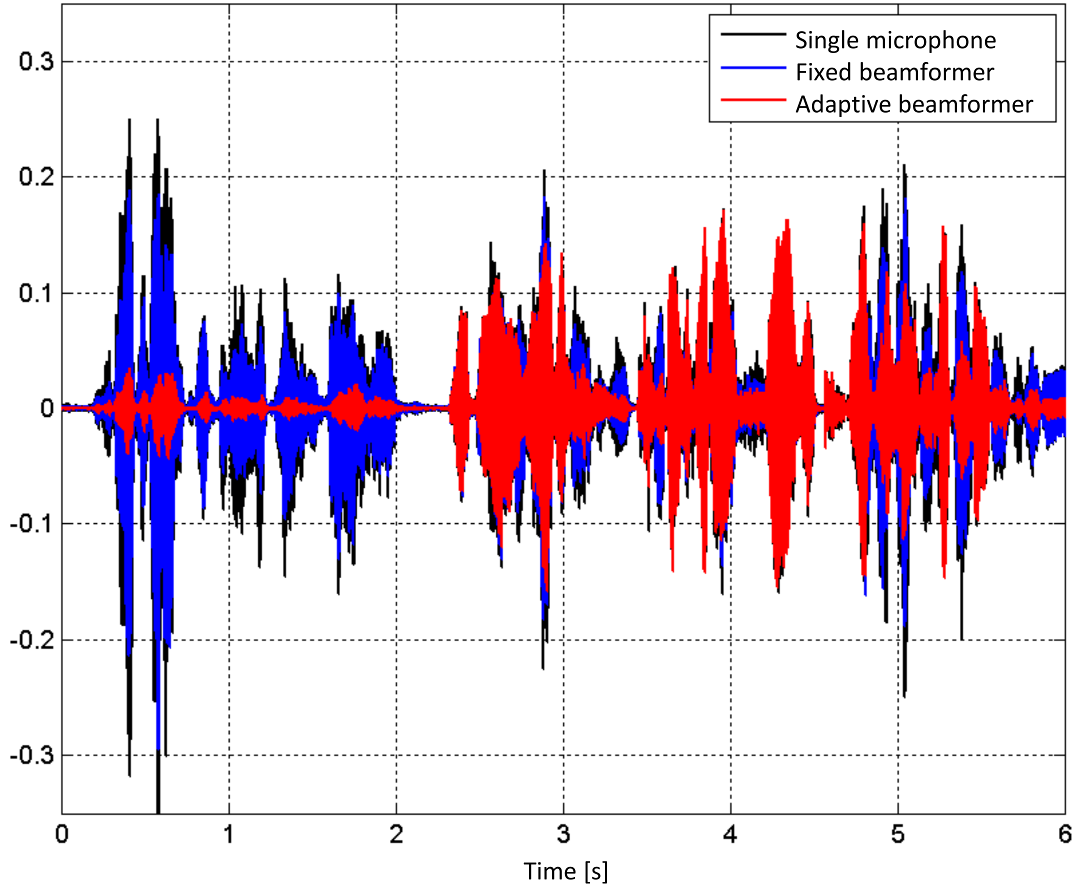
Audio Examples and Results – Part 1

- 4-channel microphone array
- Directional noise source (loudspeaker of the vehicle)
- Noise suppression > 15 dB by adaptive filtering of the microphone signals

Single microphone

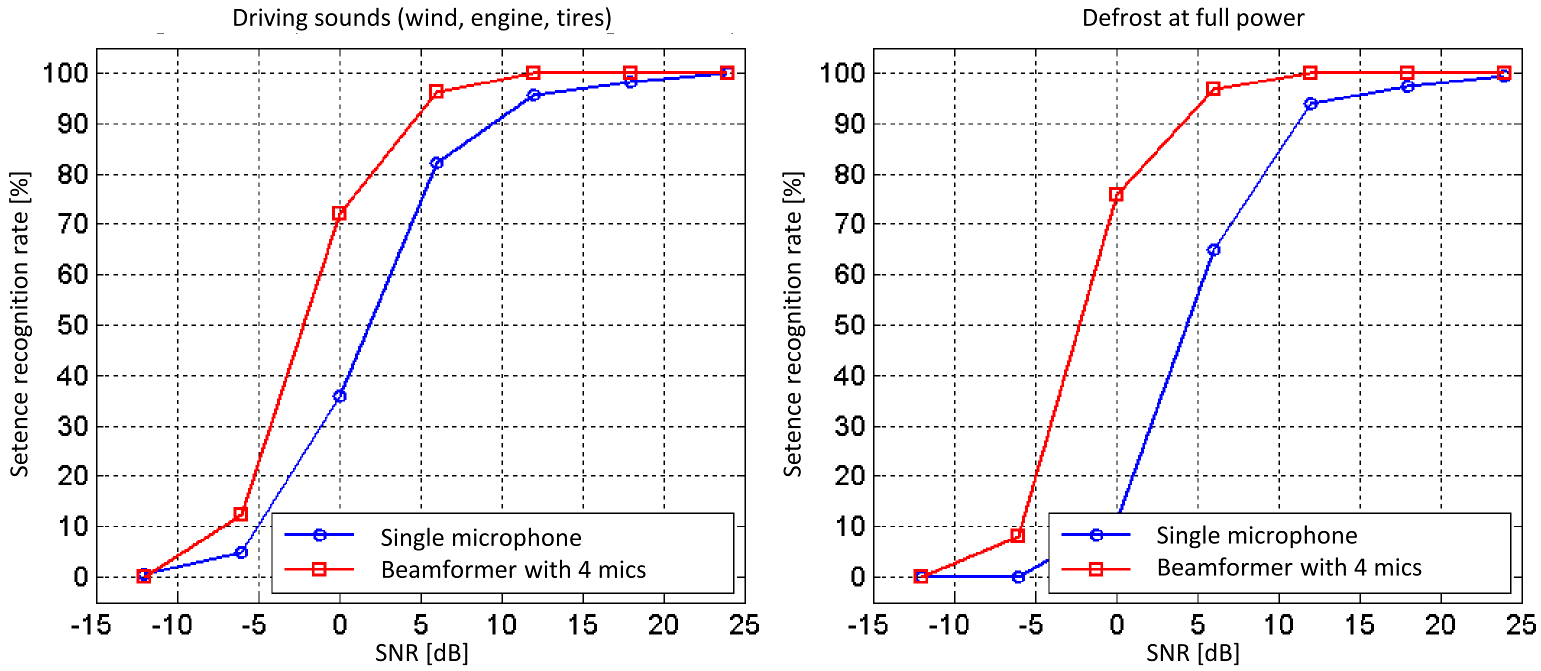
Fixed beamformer

Adaptive beamformer



Audio Examples and Results – Part 2

Recognition rates of a dialog system

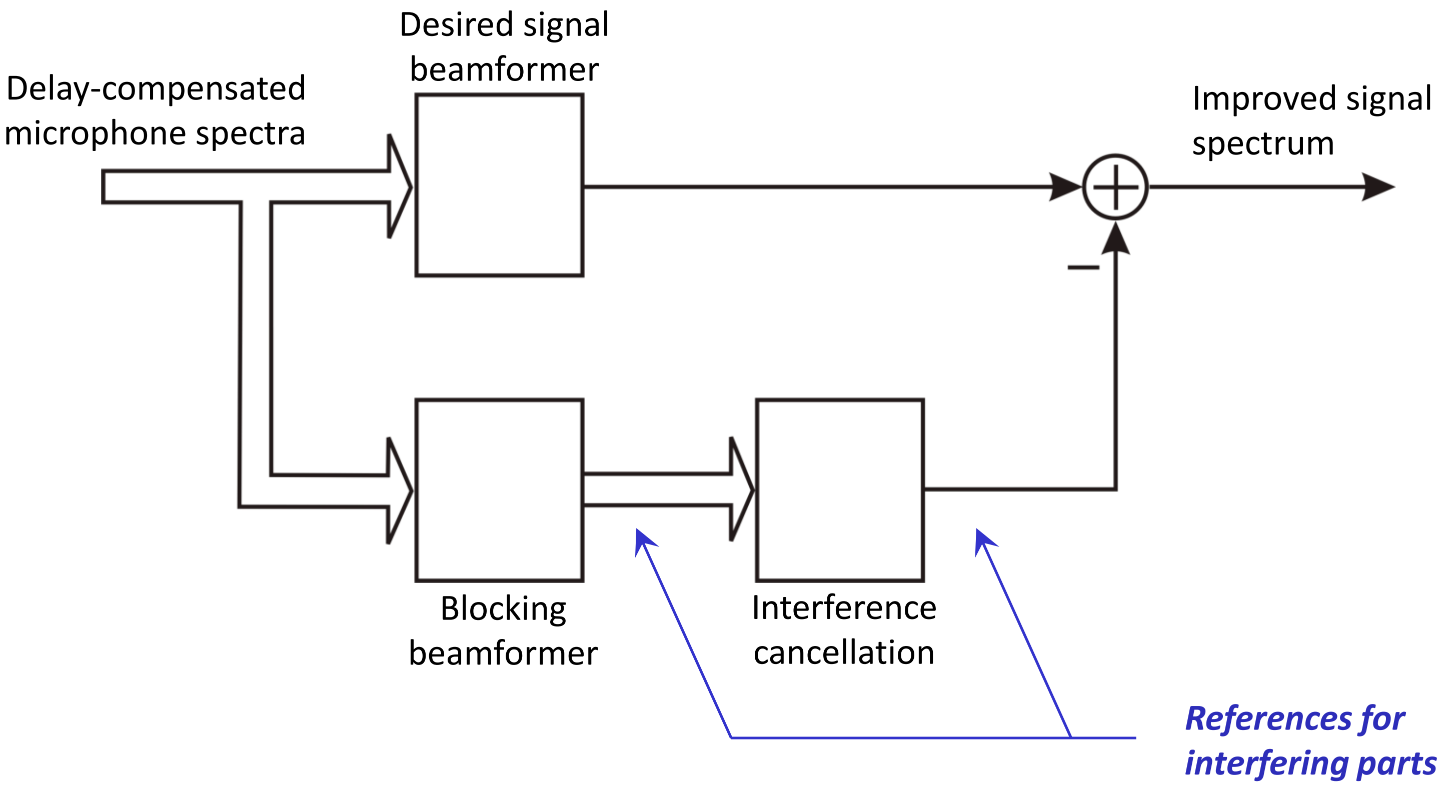


From E. Hänsler, G. Schmidt: *Acoustic Echo and Noise Control*, Wiley, 2004, with permission.

- Noise and speech have been added with different weights
- Speech model with 40 command words for radio and telephone applications
- 16 speakers (9 male, 7 female)

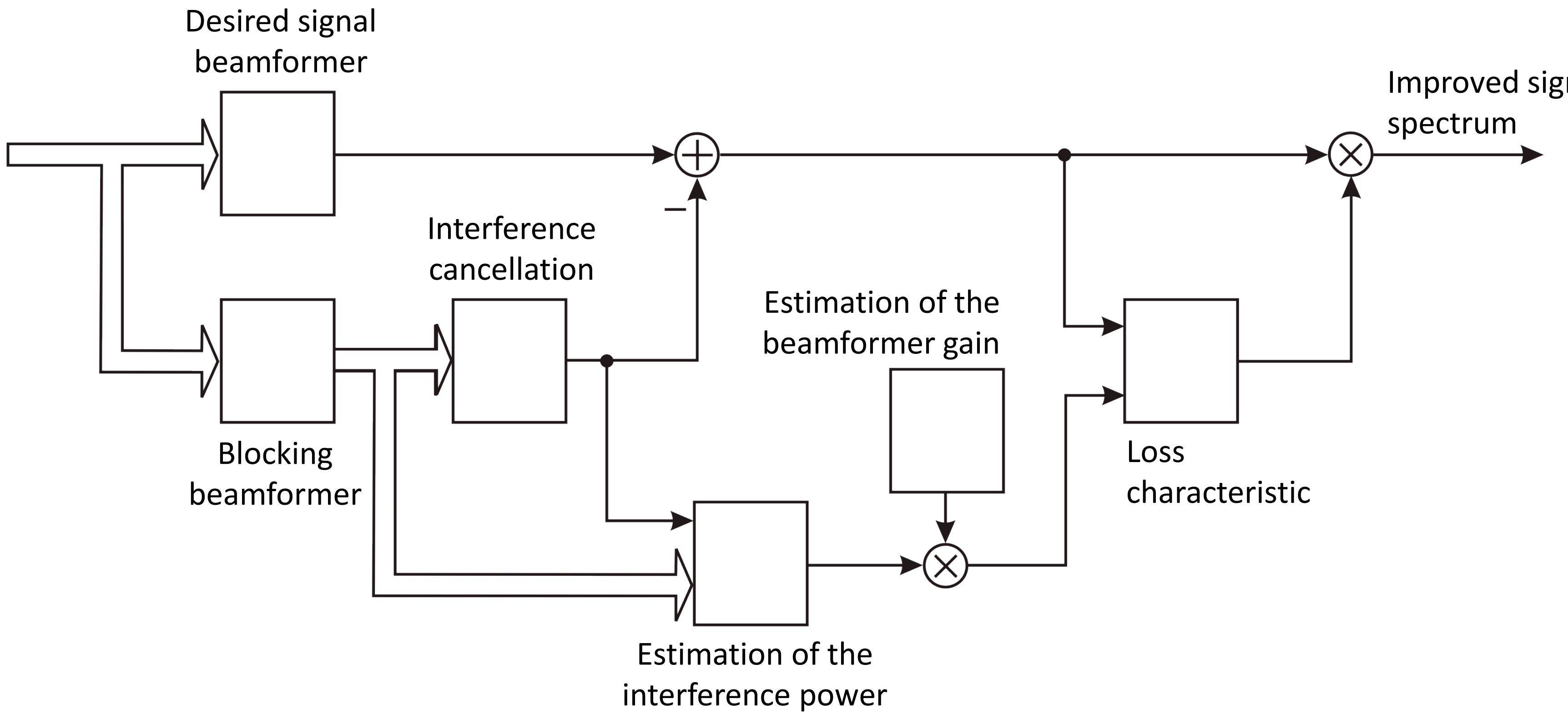
Postfiltering – Part 1

Previous structure (excerpt in subband domain)

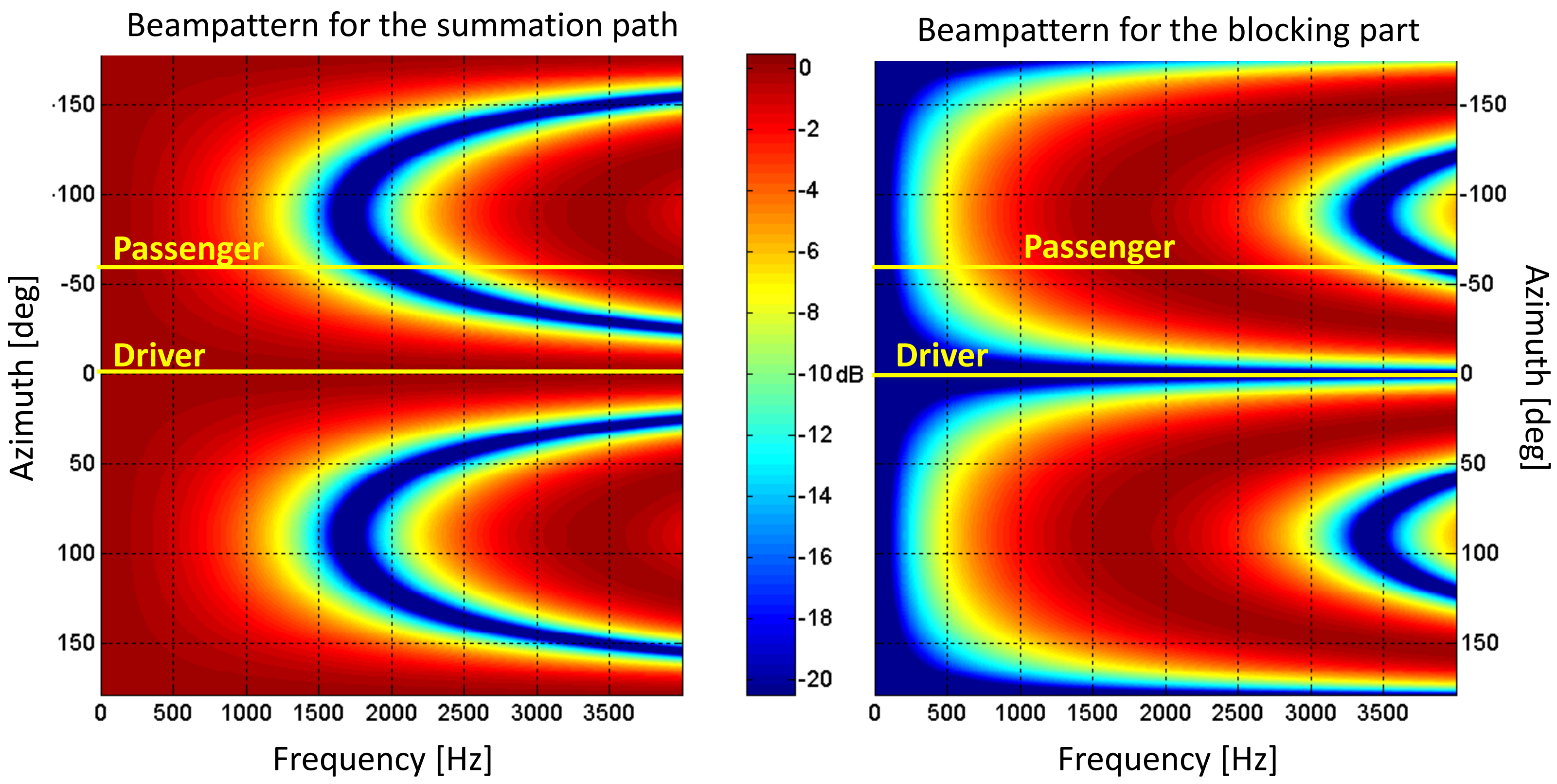


Postfiltering – Part 2

Extended structure (excerpt in subband domain)



Postfiltering – Part 3



Boundary conditions:

- Two (ideal) omnidirectional microphones
- Microphone distance 10 cm

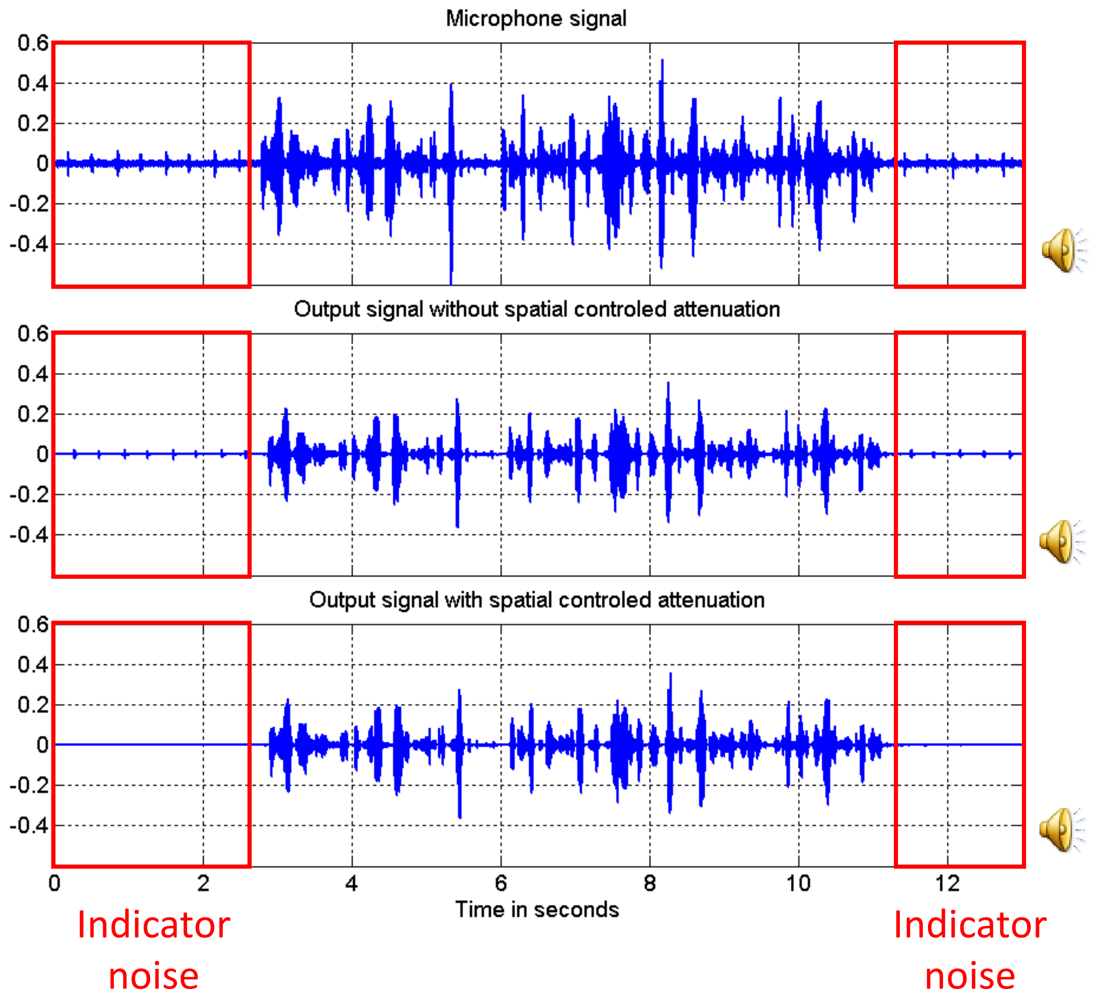
Postfiltering – Part 4

Boundary conditions

- Microphone array consisting of 4 microphones.
- While the recording, the direction indicator is active

Results

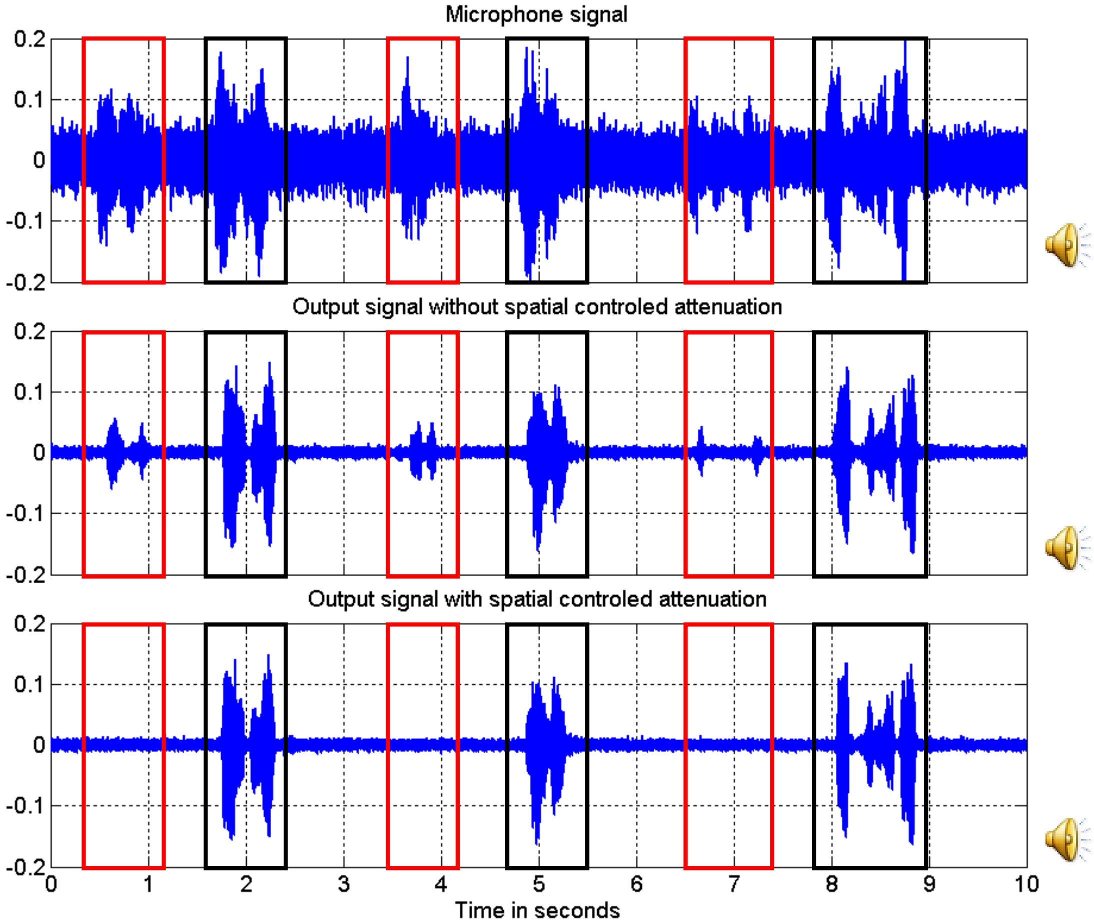
- The sound of the direction indicator can be removed during speech pauses.
- During speech activity, the indicator sound can be removed only partly.



Postfiltering – Part 5

Boundary conditions

- Microphone array consisting of 4 microphones.
- The passenger says the name of a city, where after the driver repeats the name of the city.



Passenger Driver Passenger Driver Passenger Driver

Summery and Outlook

Summary:

- Introduction
- Quality measures for multi-microphone systems
- Delay-and-sum schemes
- Filter-and-sum schemes
- Interference cancellation
- Audio examples and results
- Post-filter schemes

Next part:

- Feature extraction