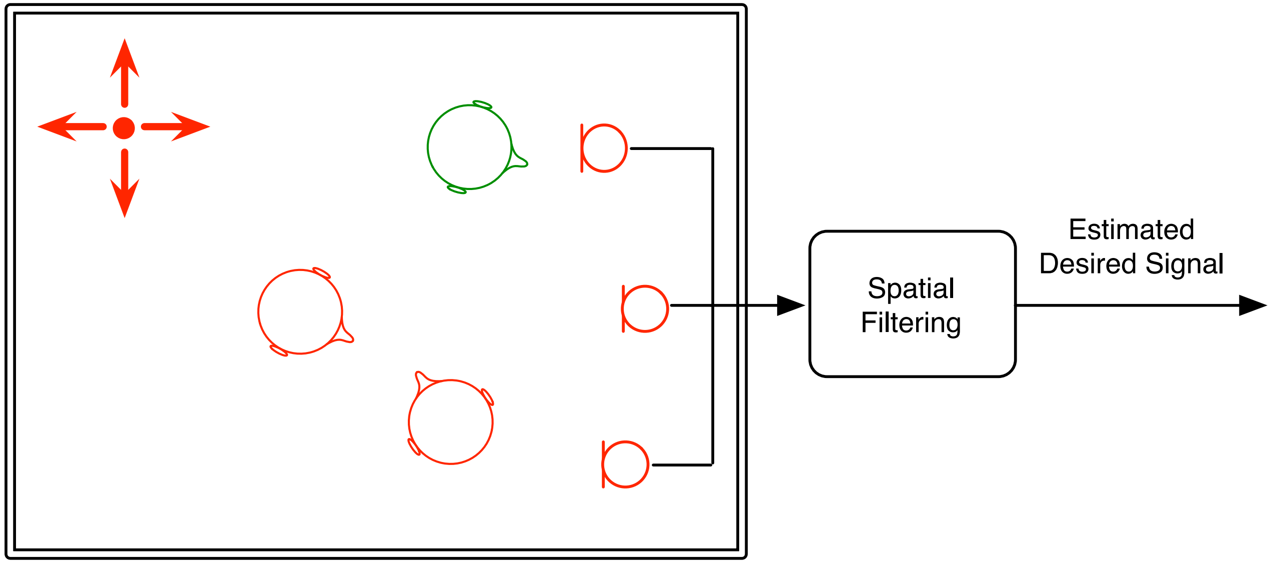
INTERNATIONAL AUDIO LABORATORIES ERLANGEN
A joint institution of Fraunhofer IIS and Universität Erlangen-Nürnberg

Recent Advances in Acoustic Signal Extraction and Dereverberation

Emanuël Habets

Erlangen Colloquium 2016

Scenario



Undesired sound components:

- Sensor noise
- Competing speakers
- Ambient sounds (e.g., generated by an air conditioner, fan, or babble)
- Reverberation (due to wall reflections, etc.)

Applications

- Hands-free Communication
- Human-Machine Interfaces
- Hearing Aids
- Music Recording and Post-Production




Outline

- Acoustic Signal Extraction
- Dereverberation
    - Reverberation Cancellation
    - Reverberation Suppression
- Conclusions and Future Challenges

Outline

- Acoustic Signal Extraction
- Dereverberation
  - Reverberation Cancellation
  - Reverberation Suppression
- Conclusions and Future Challenges

Acoustic Signal Extraction

Goal

- Extract the desired signal while reducing undesired signals from one or more microphone signals

Solutions

- Single-channel filters
- Data-independent beamformers
- Data-dependent beamformers

Challenges

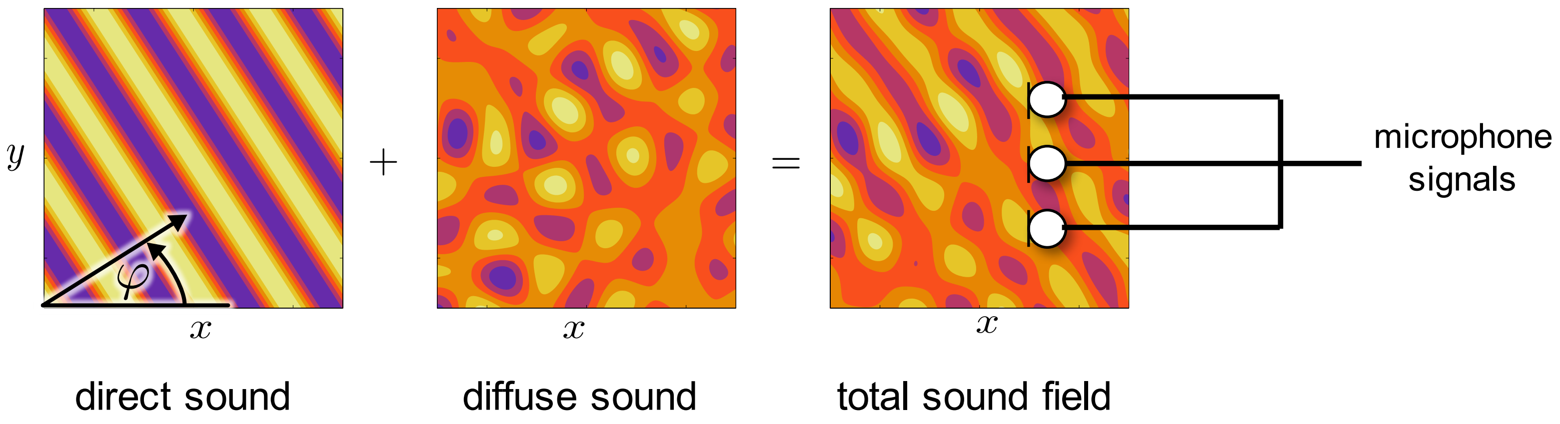
- Defining the desired signal
- Estimating the spatio-temporal statistics in non-stationary scenarios

Acoustic Signal Extraction

- We developed two approaches that use nearly **instantaneous information about the acoustic scene** to overcome the challenges of estimating the spatio-temporal statistics
- We refer to these as **informed spatial filtering approaches**
- The main difference lies in the way the spatial information is used
- The **direct approach** uses the spatial information to control the filters
- The **indirect approach** uses the spatial information to distinguish between desired and undesired sounds

Direct Informed Spatial Filtering
Parametric Sound Field Model

- The total sound field is formed as a superposition of the direct sound field and diffuse sound field



- In practice, the DOA of the direct sound can vary quickly, for instance, when multiple talkers are active at the same time

Direct Informed Spatial Filtering
Parametric Sound Field Model

- In the TF domain the microphone signals can be expressed as:

$$\mathbf{y}(n,k) = \sum_{l=1}^{L} \overbrace{\boxed{\mathbf{a}(k,\theta_l)}}^{\text{relative TFs}} \underbrace{\boxed{X_l(n,k)}}_{l\text{-th plane wave}} + \overbrace{\boxed{\mathbf{d}(n,k)}}^{\text{diffuse sound}} + \underbrace{\boxed{\mathbf{v}(n,k)}}_{\substack{\text{stationary noise} \\ \text{(e.g., fan, sensor-noise)}}}$$

- The diffuse sound power varies quickly across time. The PSD matrix of the diffuse sound component can be expressed as

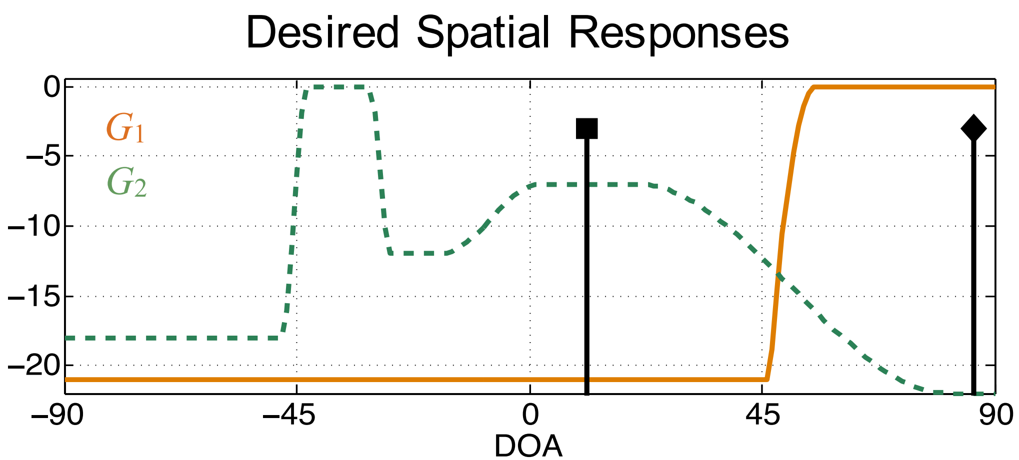
$$\mathbf{\Phi}_{\mathbf{d}}(n,k) = \mathrm{E}\left\{\mathbf{d}(n,k)\mathbf{d}^{\mathrm{H}}(n,k)\right\} = \underbrace{\boxed{\phi_D(n,k)}}_{\text{diffuse sound power}} \overbrace{\boxed{\mathbf{\Gamma}(k)}}^{\substack{\text{coherence matrix:} \\ \text{time-invariant and known}}}$$

- The statistics of the background noise vary slowly across time and can be estimated from the microphone signals

Direct Informed Spatial Filtering
Desired Signal

- Our objective is to capture *L* plane waves with desired gains while suppressing diffuse sound and noise (Thiergart et al., 2014)

- The desired signal is given by $$Z(n,k) = \sum_{l=1}^{L} \underbrace{\boxed{G(k,\theta_l)}}_{\text{desired spatial response for the } l\text{-th plane wave}} X_l(n,k)$$



Direct Informed Spatial Filtering
Estimation of the Desired Signal

- The desired signal is estimated using a spatial filter

$$\widehat{Z}(n,k) = \mathbf{w}^{\mathrm{H}}(n,k)\,\mathbf{y}(n,k)$$

- The informed LCMV filter is for example given by

$$\mathbf{w}_{\mathrm{iLCMV}}(n,k) = \arg\min_{\mathbf{w}} \overbrace{\boxed{\mathbf{w}^{\mathrm{H}}\,\left[\mathbf{\Phi}_{\mathbf{d}}(n,k) + \mathbf{\Phi}_{\mathbf{v}}(n,k)\right]\,\mathbf{w}}}^{\text{Residual Noise plus Reverberation}}$$

$$\text{subject to} \quad \mathbf{w}^{\mathrm{H}}(n,k)\,\mathbf{a}(k,\theta_l) = G(k,\theta_l), \quad l \in \{1,2,\ldots,L\}$$

- The required narrowband DOAs of the plane waves can be estimated using ESPRIT or root-MUSIC

- An estimator for the DNR was proposed in (Thiergart et al., 2014)

Direct Informed Spatial Filtering
Estimation of the Desired Signal

- The desired signal is estimated using a spatial filter

$$\widehat{Z}(n,k) = \mathbf{w}^{\mathrm{H}}(n,k)\,\mathbf{y}(n,k)$$

- The informed LCMV filter is for example given by

$$\mathbf{w}_{\mathrm{iLCMV}}(n,k) = \arg\min_{\mathbf{w}} \; \mathbf{w}^{\mathrm{H}} \left[\underbrace{\boxed{\Psi(n,k)}}_{\text{Diffuse-to-Noise Ratio}}\boldsymbol{\Gamma}(k) + \mathbf{I}\right] \mathbf{w}$$

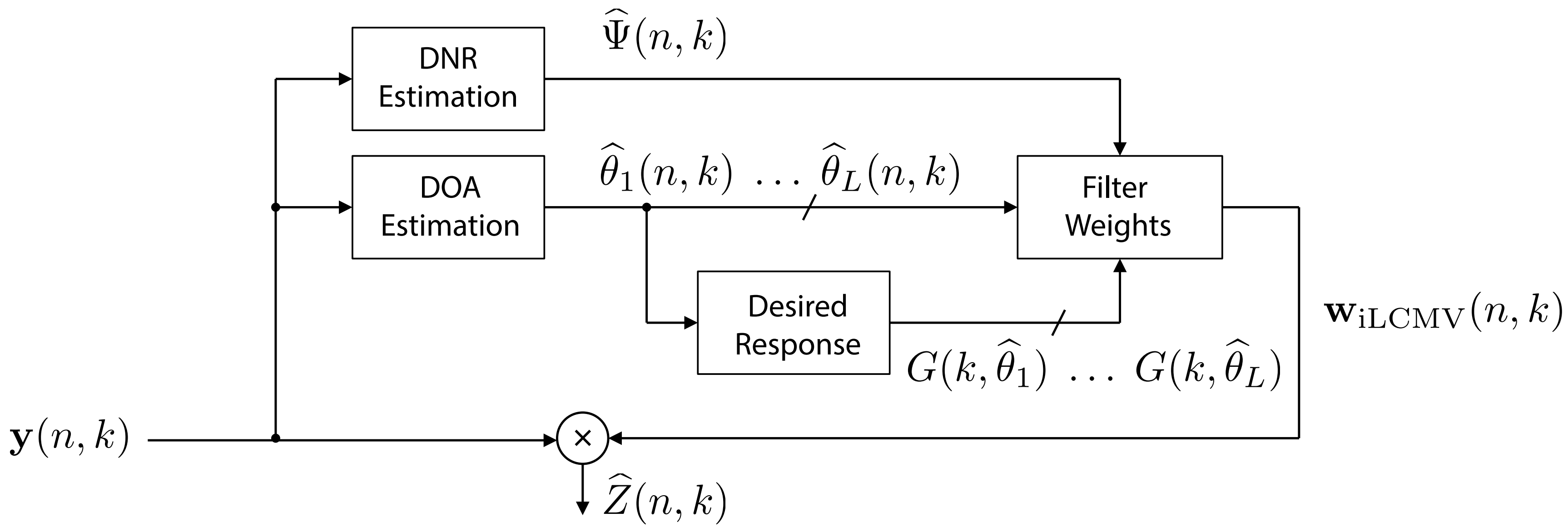
$$\text{subject to} \quad \mathbf{w}^{\mathrm{H}}(n,k)\,\mathbf{a}(k,\theta_l) = G(k,\theta_l), \quad l \in \{1,2,\ldots,L\}$$

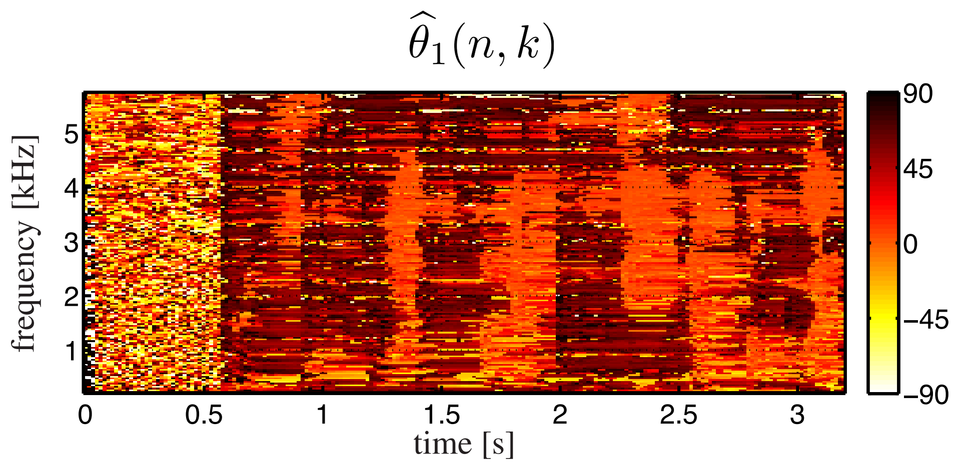
- The required narrowband DOAs of the plane waves can be estimated using ESPRIT or root-MUSIC

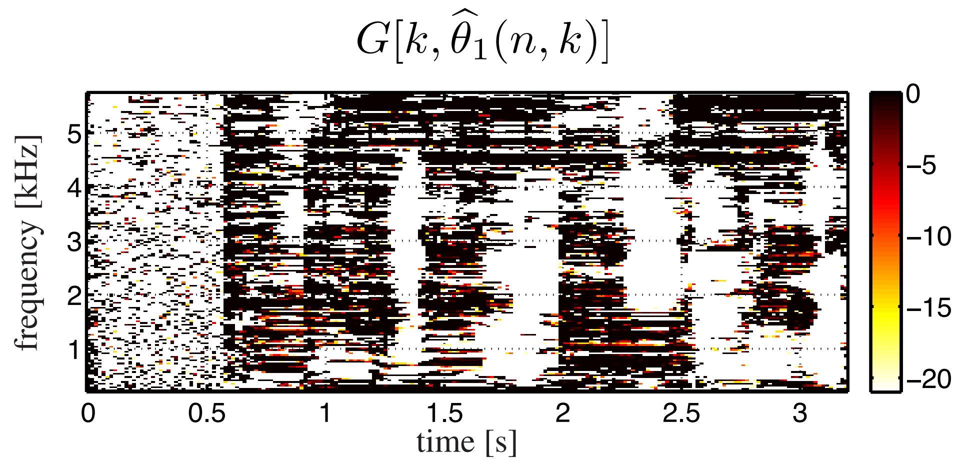
- An estimator for the DNR was proposed in (Thiergart et al., 2014)

Direct Informed Spatial Filtering

Proposed System

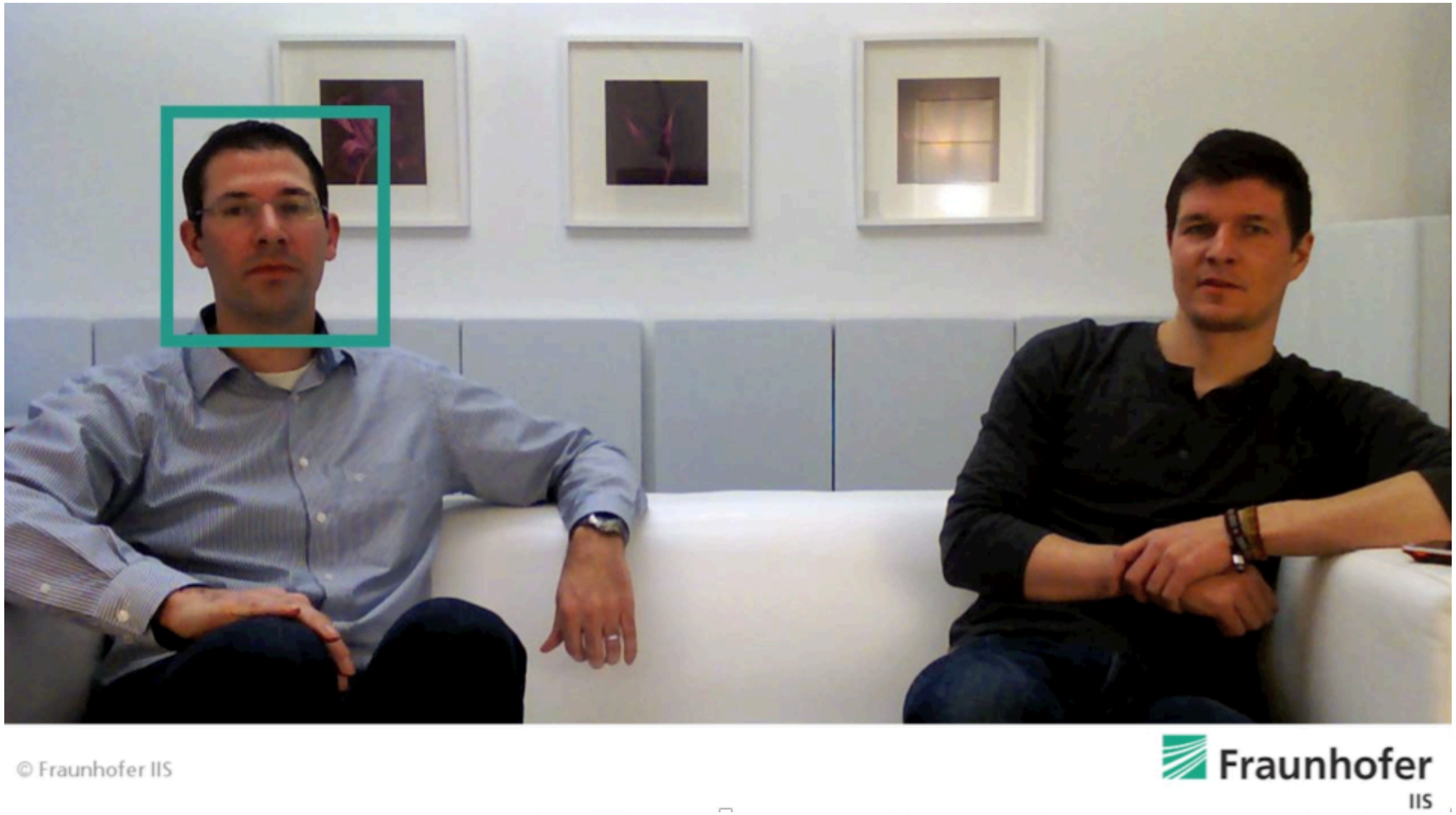






Direct Informed Spatial Filtering

Example with Face Tracking



The demo can be found at
https://www.audiolabs-erlangen.de/resources/2015-MCSE

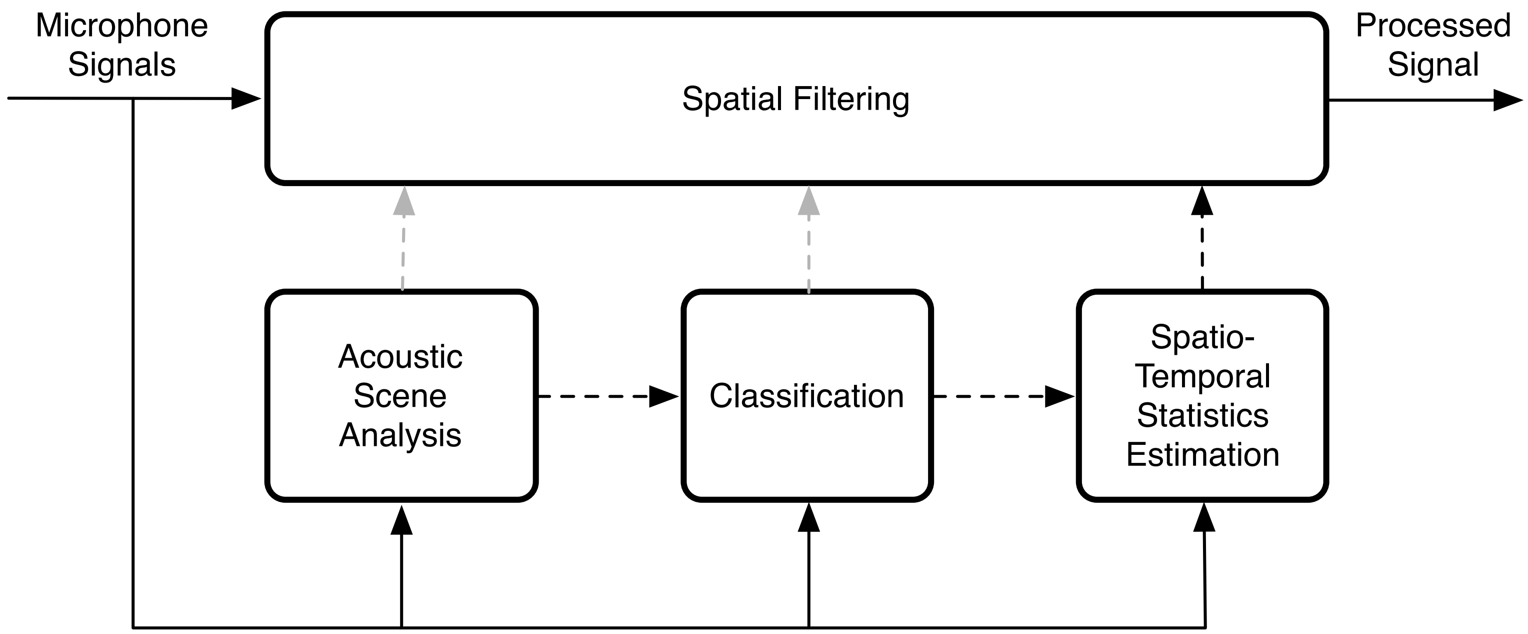
Direct Informed Spatial Filtering
Conclusions and Current Work

- A flexible spatial filtering approach that can be used to realize different audio applications independent of the microphone setup
- The approach offer **a high robustness in quickly changing acoustic scenarios**
- The achievable spatial selectively depends on the DOA accuracy

- Current Work
  - Developing even more robust parameter estimators
  - Incorporating DOA uncertainties
  - Applying this approach to binaural hearing aids

Indirect Informed Spatial Filtering

- An alternative approach was developed that aims in particular at acoustic signal extraction
- In this case instantaneous information about the acoustic scene is used to **classify each TF instance as desired or undesired**



Indirect Informed Spatial Filtering
Signal Model

- In the TF domain the *m*-th microphone signal can be expressed as

$$Y_m(n,k) = X_m(n,k) + \underbrace{B_m(n,k) + V_m(n,k)}_{U_m(n,k)}$$

- Input vector:

$$\mathbf{y}(n,k) = \left[Y_1(n,k), Y_2(n,k), \dots, Y_M(n,k)\right]^{\mathrm{T}}$$

- Desired PSD matrix:

$$\boldsymbol{\Phi}_{\mathbf{x}}(n,k) = \mathrm{E}\{\mathbf{x}(n,k)\mathbf{x}^{\mathrm{H}}(n,k)\}$$

- Undesired PSD matrix:

$$\boldsymbol{\Phi}_{\mathbf{u}}(n,k) = \mathrm{E}\{\mathbf{b}(n,k)\mathbf{b}^{\mathrm{H}}(n,k)\} + \mathrm{E}\{\mathbf{v}(n,k)\mathbf{v}^{\mathrm{H}}(n,k)\}$$

Indirect Informed Spatial Filtering
Power Spectral Density Estimation

- Using the estimated PSD matrices optimal filters can be computed
  - Minimum variance distortionless response filter
  - Multi-channel Wiener filter
- Estimate the PSD matrix (for each frequency bin *k)*

$$\widehat{\mathbf{\Phi}}_{\mathbf{x}+\mathbf{v}}(n) = \alpha_{\mathrm{d}}(n)\,\widehat{\mathbf{\Phi}}_{\mathbf{x}+\mathbf{v}}(n-1) + (1-\alpha_{\mathrm{d}}(n))\,\mathbf{y}(n)\mathbf{y}^{\mathrm{H}}(n)$$

$$\widehat{\mathbf{\Phi}}_{\mathbf{u}}(n) = \alpha_{\mathrm{u}}(n)\,\widehat{\mathbf{\Phi}}_{\mathbf{u}}(n-1) + (1-\alpha_{\mathrm{u}}(n))\,\mathbf{y}(n)\mathbf{y}^{\mathrm{H}}(n)$$

- Each TF instance can be classified as

Desired signal present	$\mathcal{H}_x$
Desired signal absent	$\mathcal{H}_u = \mathcal{H}_b \cup \mathcal{H}_v$

- The smoothing constants depend on the classification

Indirect Informed Spatial Filtering
Minimum Bayes-risk Detector

- We propose to classify each TF instance using **spatial features** and a minimum Bayes-risk decision rule

$$\mathcal{H}_x = \begin{cases} 1 & \text{if } \frac{p(\mathcal{H}_x|\hat{\Omega})}{p(\mathcal{H}_u|\hat{\Omega})} > \frac{\text{cost of a false alarm}}{\text{cost of a miss}} \\ 0 & \text{otherwise} \end{cases}$$

Probabilities of the hypotheses given the estimated features

- Bayes costs control the tradeoff between **speech distortion** and **interference reduction**

- How to obtain the posterior probabilities of the hypotheses?

Indirect Informed Spatial Filtering

DOA-based Minimum Bayes-risk Detector

- Let us assume a single microphone array and **known target direction**
- We propose to classify each TF instance using **narrowband direction-of-arrivals** (DOA) and **signal-to-diffuse ratio** (SDR) estimates
- Mixture model for the estimated DOAs

$$f(\hat{\theta}) = \beta_x\, f(\hat{\theta}|\mathcal{H}_x) + \beta_b\, f(\hat{\theta}|\mathcal{H}_b) + \beta_v\, \underset{\substack{\uparrow \\ \text{uniformly distributed}}}{f(\hat{\theta}|\mathcal{H}_v)}$$

- Posterior probability of the hypothesis

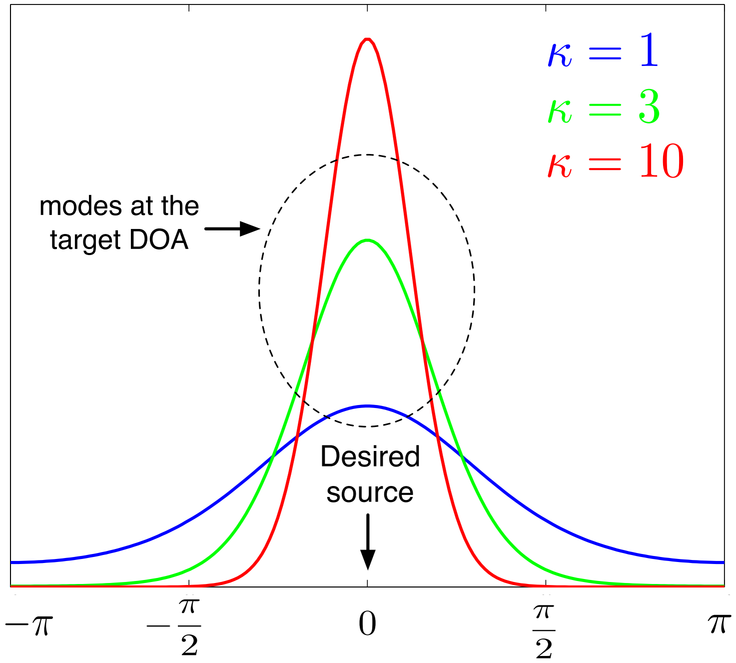
$$p(\mathcal{H}_x|\hat{\theta}) = \frac{\beta_x\, f(\hat{\theta}|\mathcal{H}_x)}{f(\hat{\theta})}$$

Source: (Taseska and Habets, 2015)

Indirect Informed Spatial Filtering

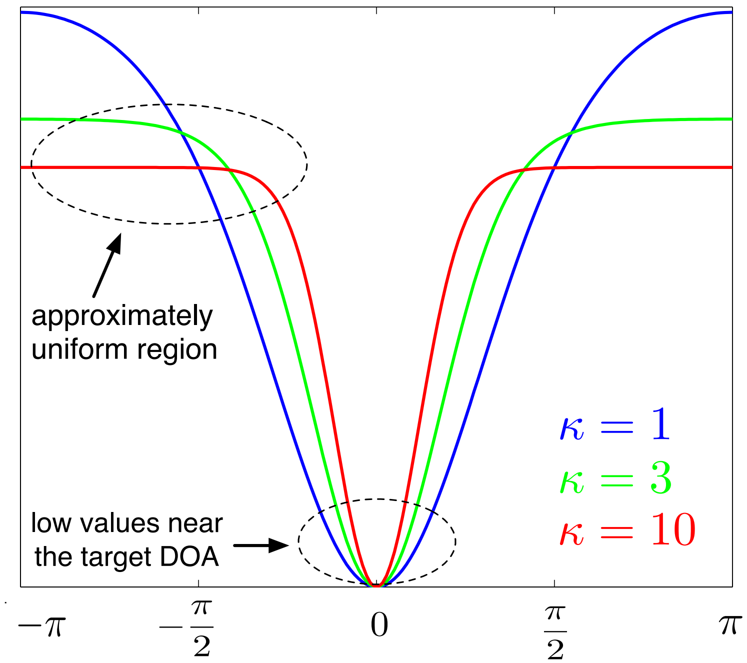
DOA-based Minimum Bayes-risk Detector

Von Mises distribution



$f(\hat{\theta}|\mathcal{H}_x)$

"Notched" distribution

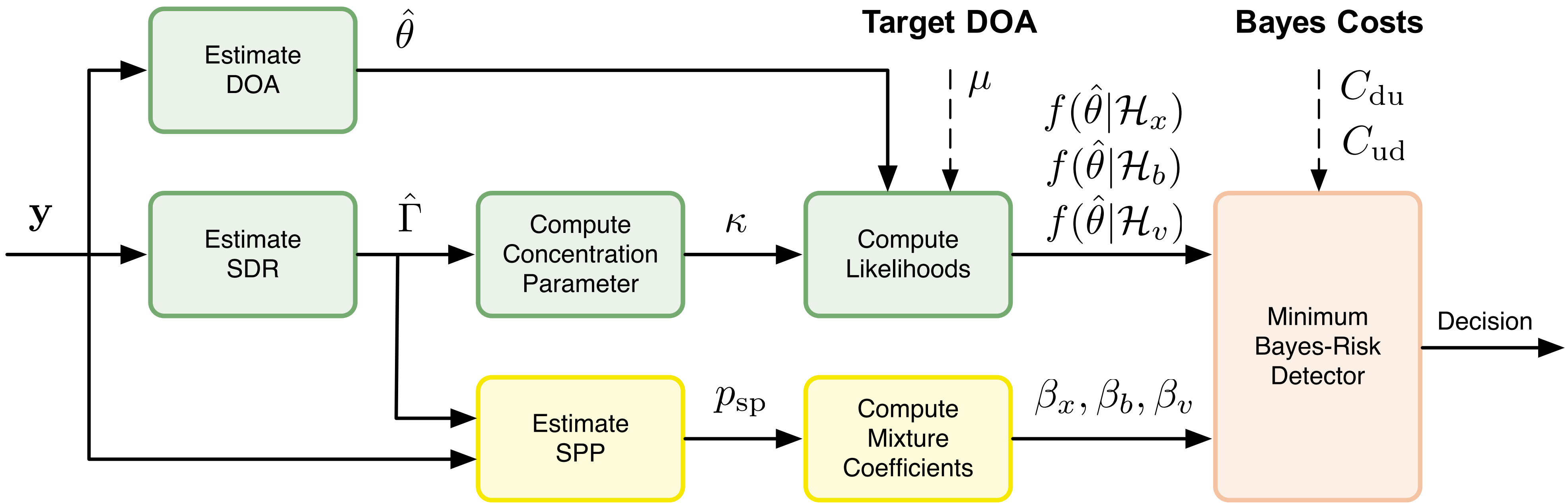


$f(\hat{\theta}|\mathcal{H}_b)$

- Mode μ corresponds to the target DOA
- The concentration parameter κ reflects the DOA estimator uncertainty

Indirect Informed Spatial Filtering

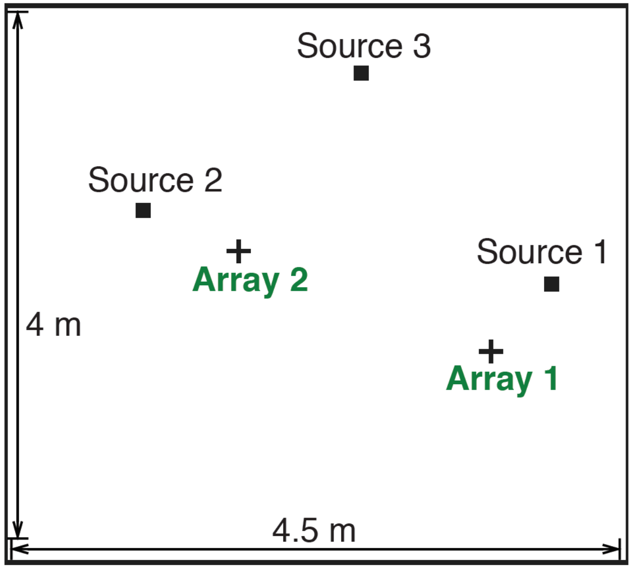
DOA-based Minimum Bayes-risk Detector



- The **mixture coefficients** (i.e., prior probabilities) are computed using the speech presence probability

Indirect Informed Spatial Filtering
Example Using DOA Model-based Detector

- Setup
  - Sampling frequency 16 kHz
  - STFT frames 64 ms, 50% overlap
  - Circular array (3 DPAs, 1.5 cm radius)
  - Reverberation time 0.18 s
  - Sensor and diffuse noise
  - Signal-to-interference ratio approx. 3 dB



- Acoustic Signal Extraction (Taseska and Habets, 2015)
  - Minimum variance distortionless response (MVDR) beamformer
  - PSD matrices are estimated using three different detectors

Indirect Informed Spatial Filtering

Example Using DOA Model-based Detector

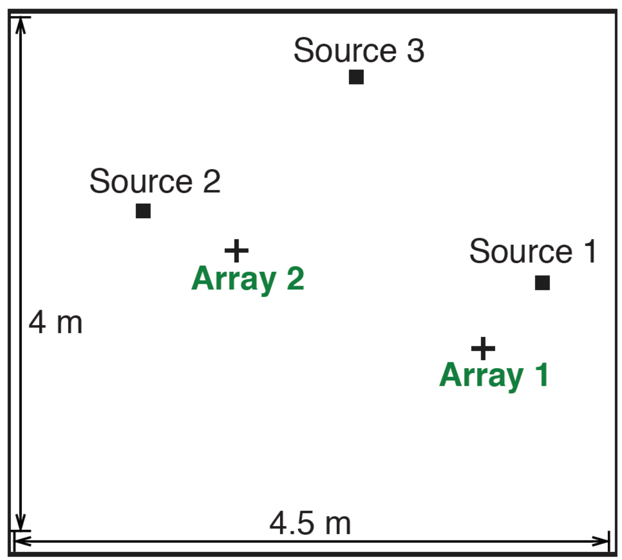
Source1 Source 2

Reference microphone

Ideal detector

Signal model-based detector (Jarrett et al., 2014)

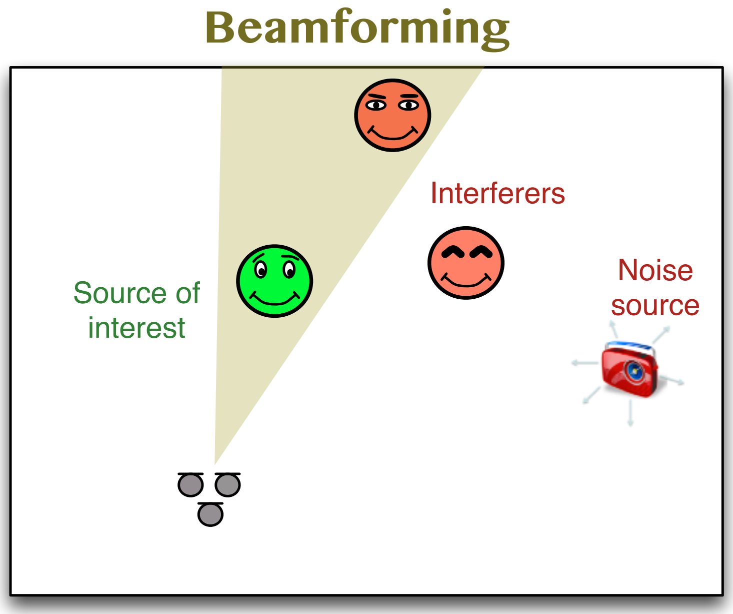
DOA model–based detector (Taseska and Habets, 2015)

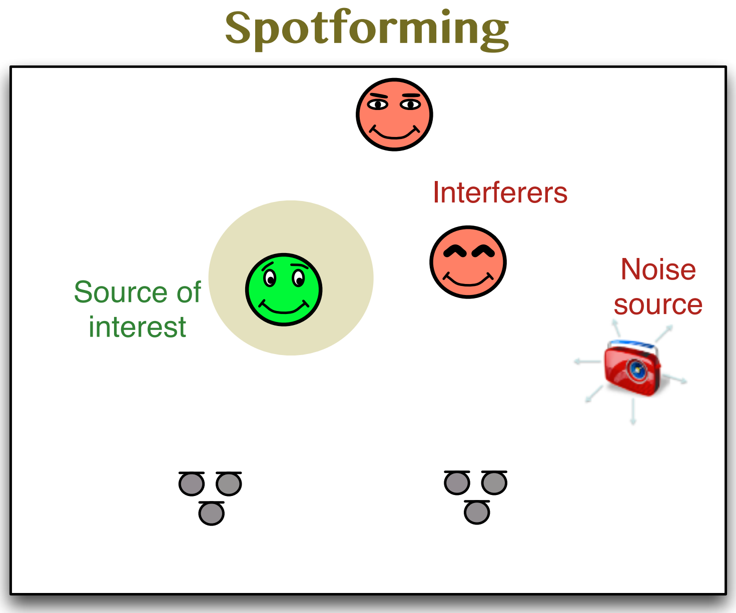


Some audio demos can be found at https://www.audiolabs-erlangen.de/resources/2015-ICASSP-DOAdet

Indirect Informed Spatial Filtering

Spotforming (Taseska and Habets, 2013)





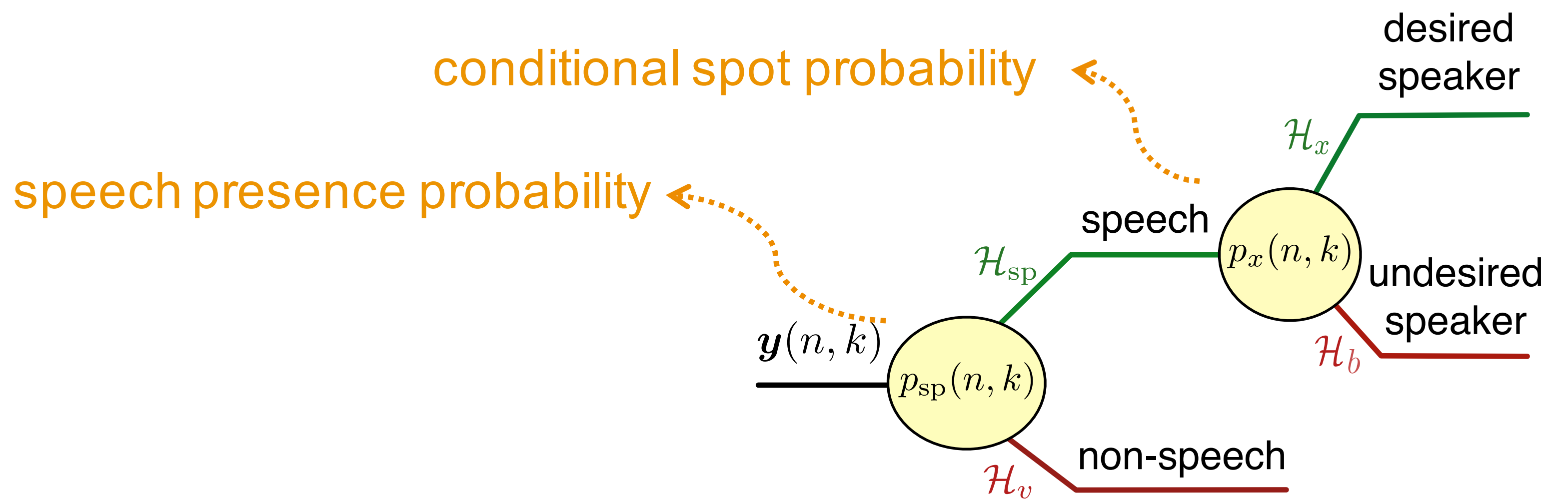
$$Y_m(n,k) = \underbrace{\int_{\mathbf{r}\in\mathcal{S}} H_{\mathbf{r},m}(k) S_{\mathbf{r}}(n,k)\, \mathrm{d}\mathbf{r}}_{X_m(n,k)} + \underbrace{B_m(n,k) + V_m(n,k)}_{U_m(n,k)}$$

Indirect Informed Spatial Filtering
Position-based Minimum Bayes-risk Detector

- Using distributed arrays narrowband position estimates can be obtained
- These positions are used as a spatial feature for the classification

$$\frac{p(\mathcal{H}_x|\hat{\mathbf{r}})}{1 - p(\mathcal{H}_x|\hat{\mathbf{r}})} > \frac{\text{cost of a false alarm}}{\text{cost of a miss}}$$

$$p(\mathcal{H}_x|\hat{\mathbf{r}}) = p(\mathcal{H}_x, \mathcal{H}_{\text{sp}}|\hat{\mathbf{r}}) = \boxed{p(\mathcal{H}_x|\mathcal{H}_{\text{sp}}, \hat{\mathbf{r}})}\, p(\mathcal{H}_{\text{sp}})$$



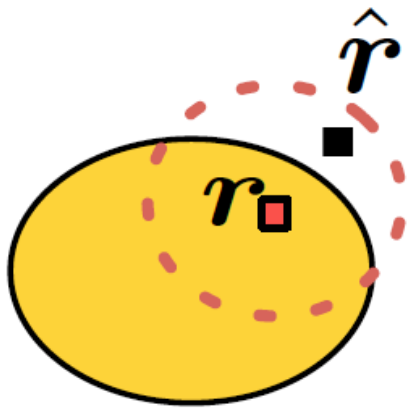
Indirect Informed Spatial Filtering
Position-based Minimum Bayes-risk Detector

- The conditional spot probability is given by

$$p(\mathcal{H}_x|\mathcal{H}_{\mathrm{sp}},\hat{\mathbf{r}}) = \int_{\mathbf{r}\in\mathcal{S}} f(\mathbf{r}|\mathcal{H}_{\mathrm{sp}},\hat{\mathbf{r}})\,\mathrm{d}\mathbf{r}$$

$$\propto$$

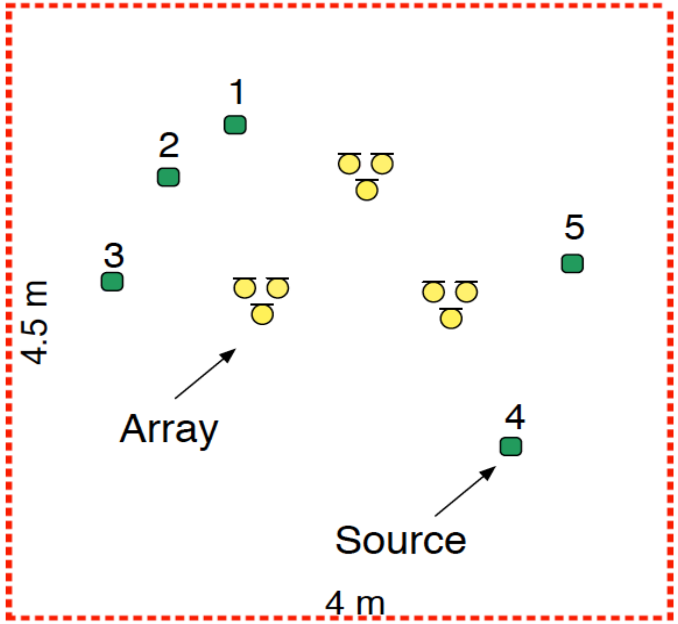
$$\underbrace{f(\hat{\boldsymbol{r}}\,|\,\mathcal{H}_{\mathrm{sp}},\boldsymbol{r})}_{\text{Gaussian likelihood model}}\,\underbrace{f(\boldsymbol{r}\,|\,\mathcal{H}_{\mathrm{sp}})}_{\text{Uniform prior}}$$



Indirect Informed Spatial Filtering

Example Using Position Model-based Detector

- Reverberation time approx. 0.18 s
- Three circular arrays, three DPA microphones per array, 3 cm diameter
- 16 kHz sampling rate, STFT frame 64 ms with 50 % overlap
- Spotformer using an MVDR filter



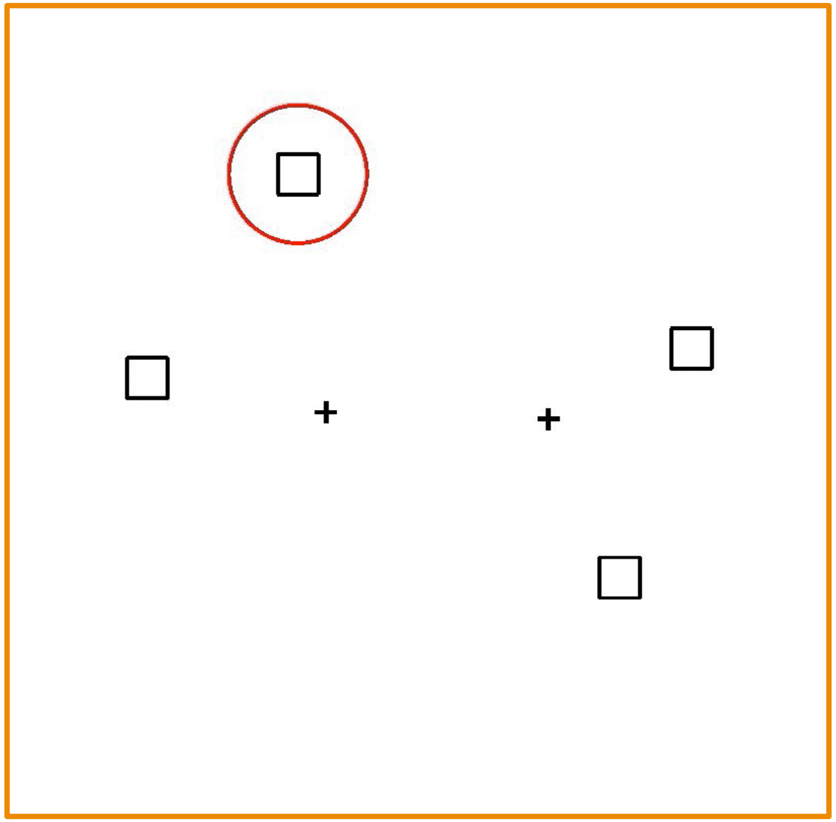
- Comparison using a oracle fixed spotformer where the PSD matrices were optimal during the first 10 seconds

Some audio demos can be found at https://www.audiolabs-erlangen.de/resources/2015-Spotformer

Indirect Informed Spatial Filtering

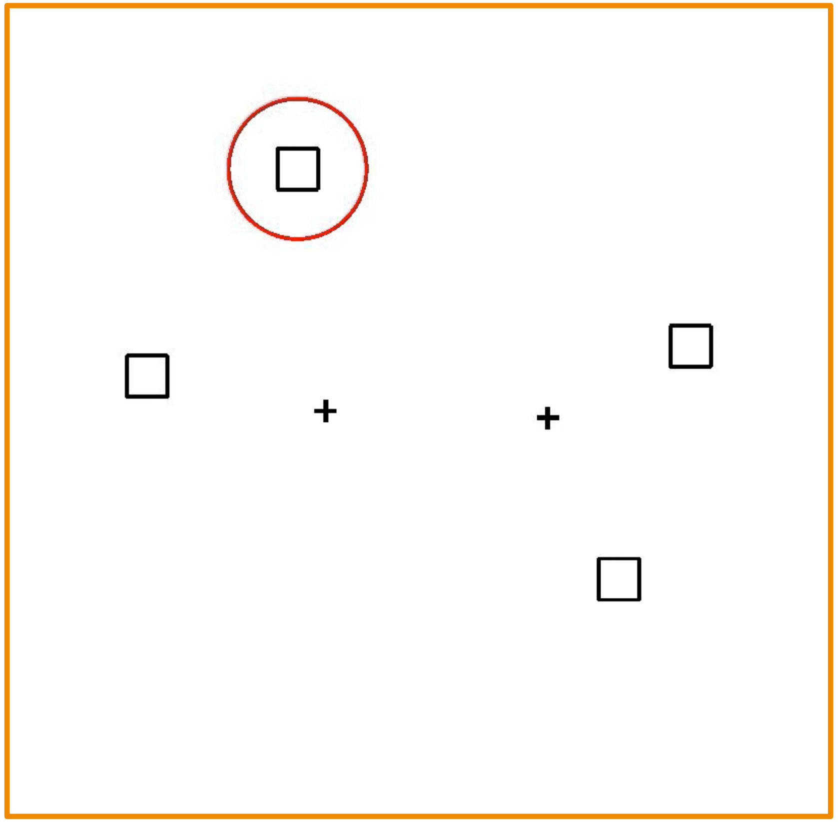
Example Using Position Model-based Detector

Reference Microphone



Indirect Informed Spatial Filtering
Example Using Position Model-based Detector

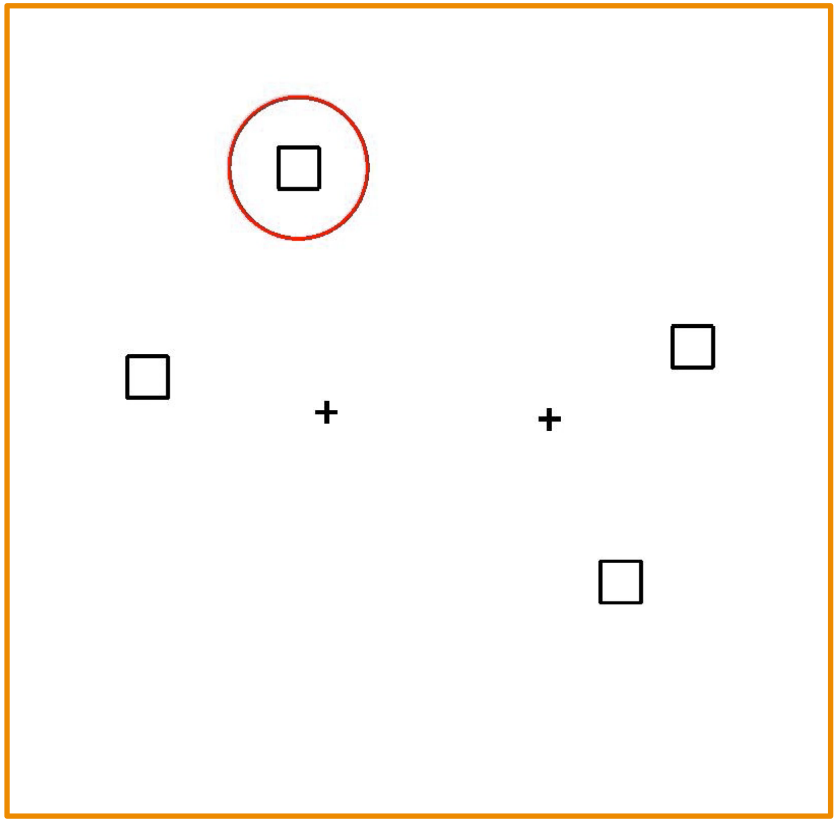
Oracle Fixed Spotformer



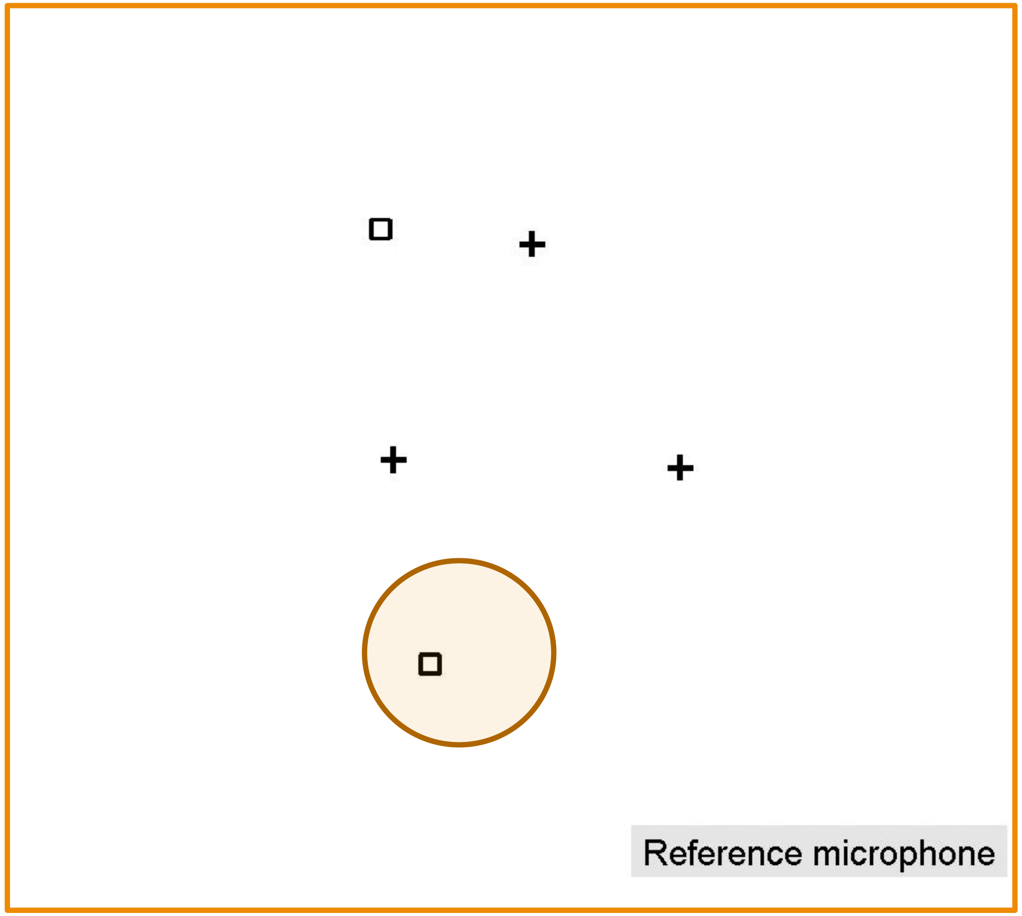
Indirect Informed Spatial Filtering

Example Using Position Model-based Detector

Proposed Data-Dependent Spotformer



Indirect Informed Spatial Filtering
Example Using Position Model-based Detector



- Scenario: moving sources
- Reverberation time: 0.3 s

- Scene analysis: 9 mics
- Spatial filtering: **3 mics**

- SIR Input 0 dB
- SIR Output 12.1 dB

- SNR Input 0 dB
- SNR Output 6.8 dB

Indirect Informed Spatial Filtering
Conclusions and Current Work

- The indirect ISF approach can be used to extract acoustic signals arriving from a **specific direction or location**
- Provides **low speech distortion** and **high interference reduction**

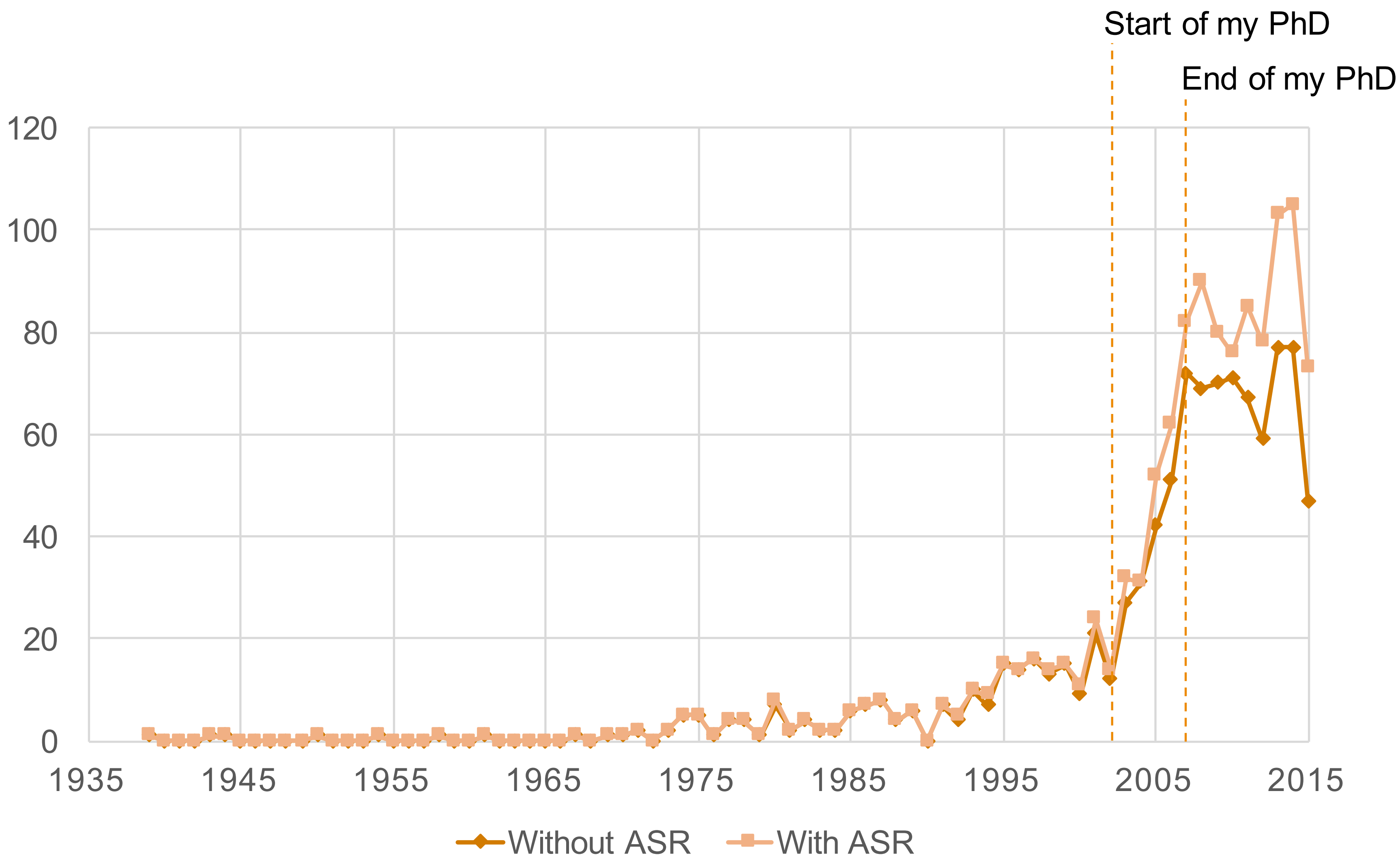
- Current Work
  - Developing even more robust detectors
  - Performing distributed signal processing

Outline

- Acoustic Signal Extraction
- Dereverberation*
  - Reverberation Cancellation
  - Reverberation Suppression
- Conclusions and Future Challenges

*In collaboration with Sharon Gannot, Boaz Schwartz, and Ofer Schwartz from Bar-Ilan University, Israel

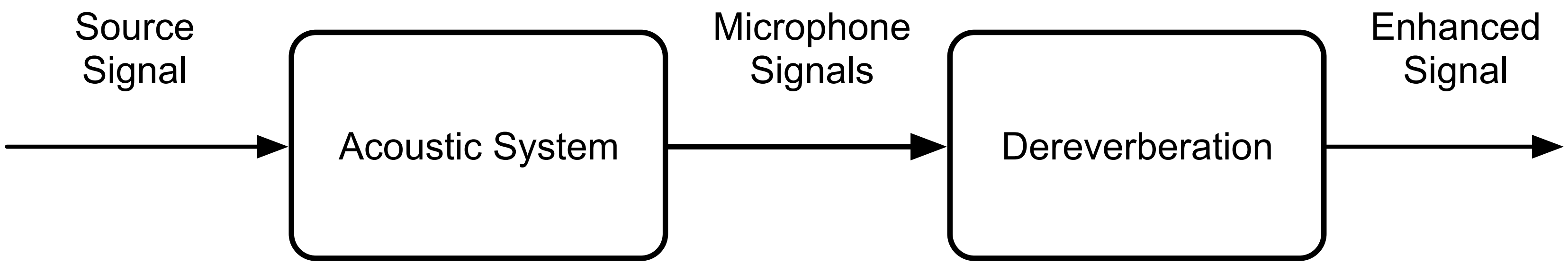
Dereverberation Publications



Source: Scopus

Dereverberation Approaches

- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown

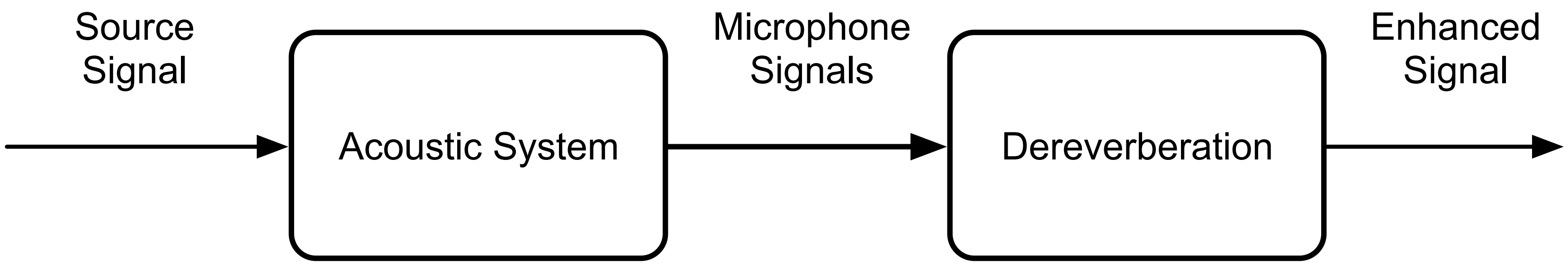


Dereverberation Approaches

- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal

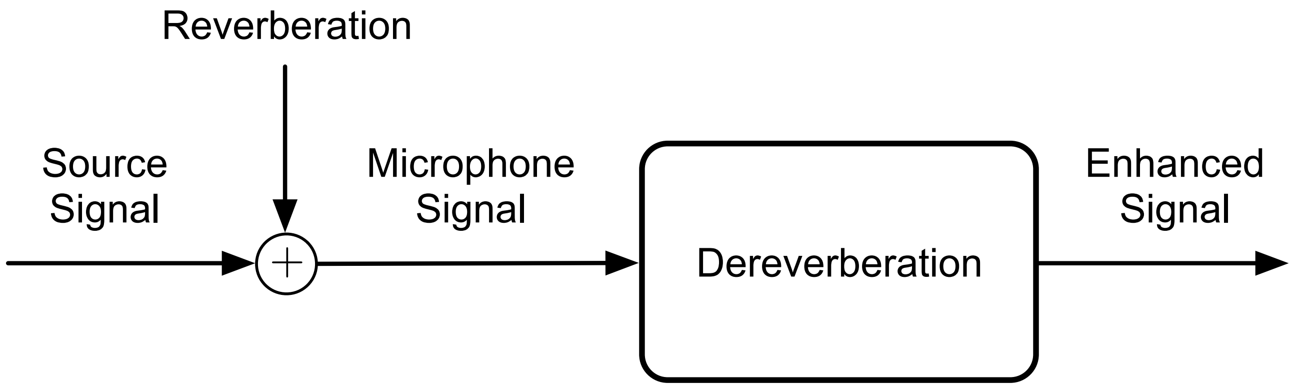
     Reverberation Cancellation

  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown



Dereverberation Approaches

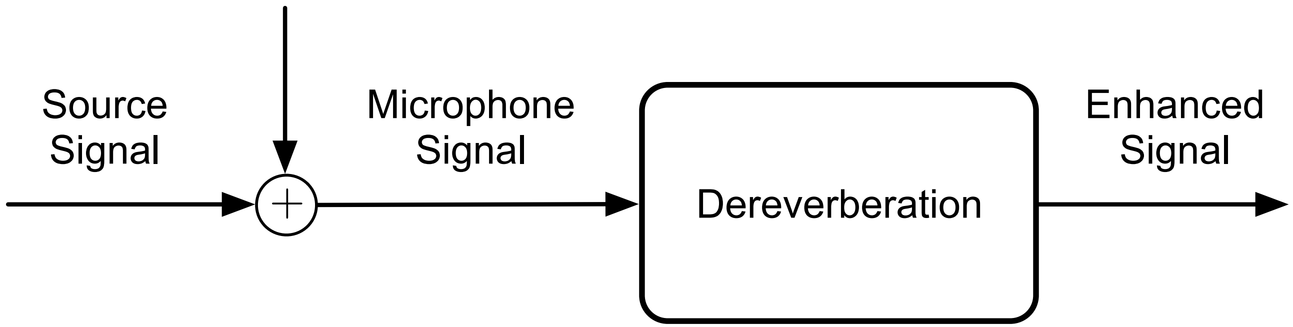
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown



Dereverberation Approaches

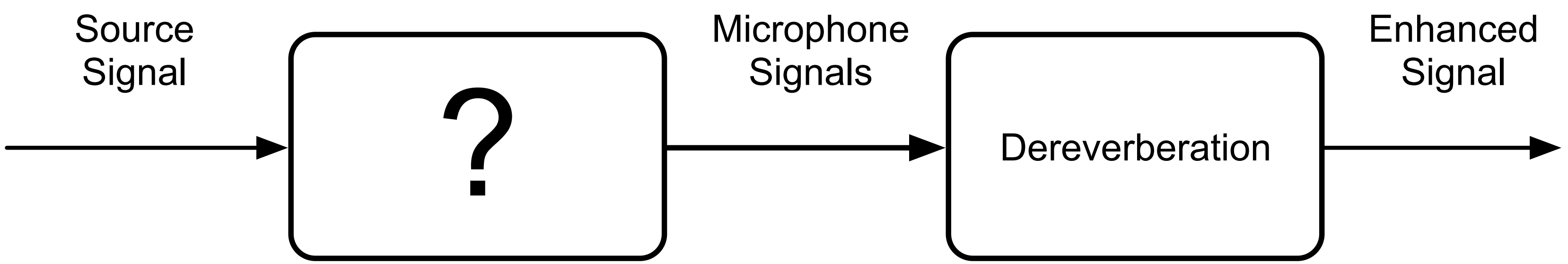
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown

Reverberation Suppression



Dereverberation Approaches

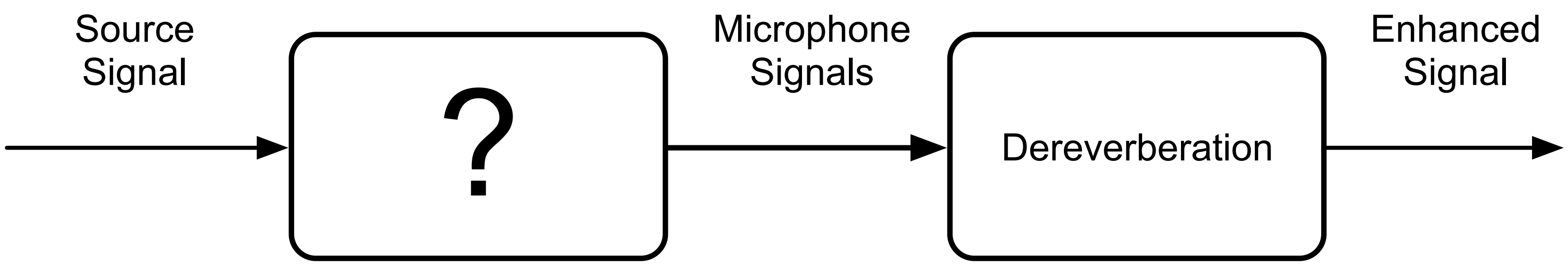
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown



Dereverberation Approaches

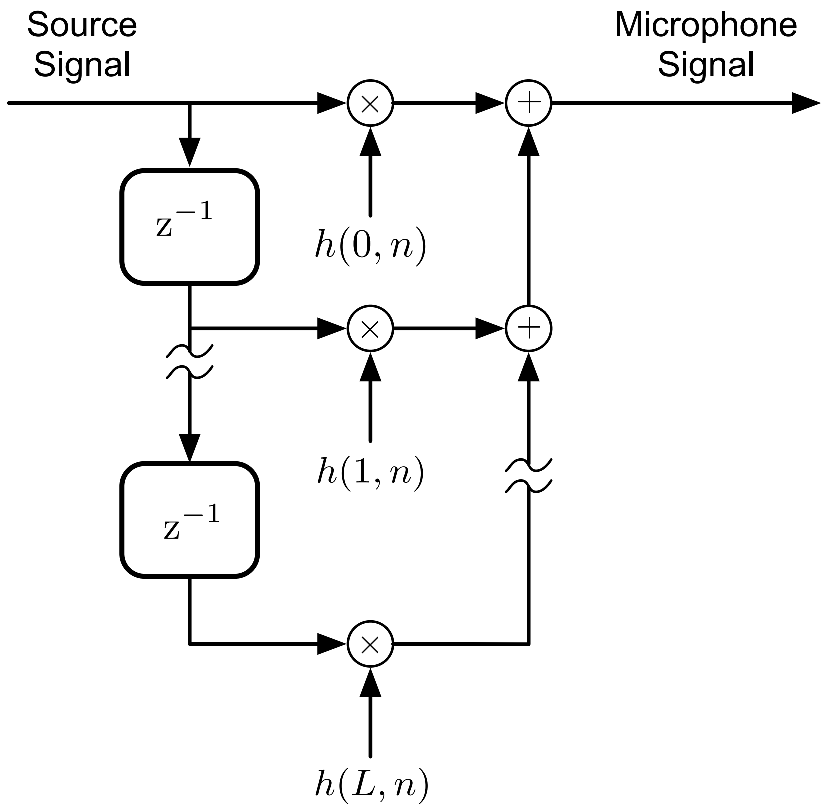
- Three fundamentally different approaches
  1. Model the acoustic system, estimate the model parameters by treading the source signal as a nuisance, and then estimate the source signal
  2. Model the reverberation as an additive process, and then estimate the source signal
  3. Directly estimate the source signal from the microphone signals by treading the acoustic system as unknown

Direct Estimation



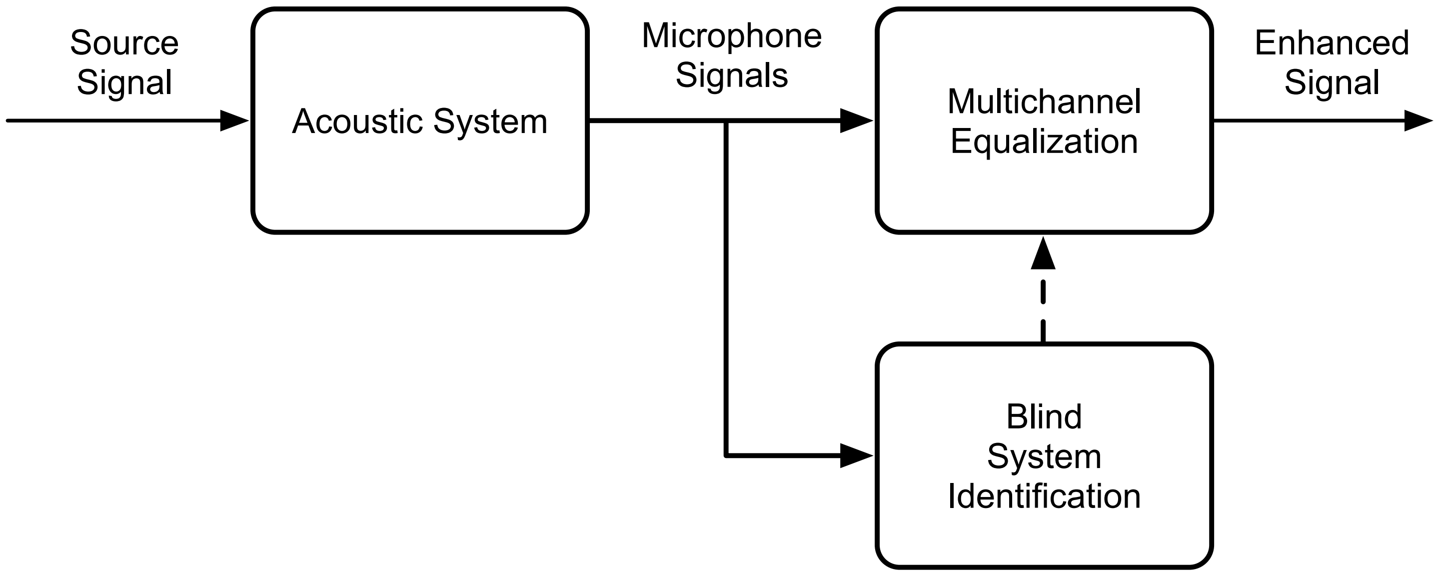
Reverberation Cancellation
Models

- Acoustic models
  - Finite Impulse Response
  - Infinite Impulse Response
- Signal models
  - Moving average process
  - Autoregressive process
- The models can be described in the time-domain or time-frequency domain



Reverberation Cancellation
Moving Average Process (Time-Domain)

- The desired signal is a delayed or filtered version of the source signal
- To obtain an estimate of the desired signal:
  1. Blindly identify the model parameters of the acoustic system
  2. Estimate the desired signal by applying a multichannel equalizer



Reverberation Cancellation
Moving Average Process (TF-Domain)

- In (B. Schwartz et al., 2015) the microphone signals were modeled in the TF domain as a moving average process

$$Y_m(n,k) = \sum_{\ell=0}^{L} H_m(\ell,k)\, S(n-\ell,k)$$

- In the context of binaural hearing aids (B. Schwartz et al., 2015):

$$Z_{\mathrm{L}}(n,k) = \sum_{\ell=0}^{L} W(\ell,k)\, H_{\mathrm{L}}(\ell,k)\, S(n-\ell,k) \qquad W(\ell,k) = \mathrm{e}^{-\alpha(k)\,\ell}$$

$$Z_{\mathrm{R}}(n,k) = \sum_{\ell=0}^{L} W(\ell,k)\, H_{\mathrm{R}}(\ell,k)\, S(n-\ell,k)$$

Reverberation Cancellation
Moving Average Process (TF-Domain)

- In (B. Schwartz et al., 2015) the microphone signals were modeled in the TF domain as a moving average process

$$Y_m(n,k) = \underbrace{H_m(0,k)\,S(n,k)}_{X_m(n,k)} + \sum_{\ell=1}^{L} H_m(\ell,k)\,S(n-\ell,k)$$

- In the context of binaural hearing aids (B. Schwartz et al., 2015):

$$Z_{\mathrm{L}}(n,k) = \sum_{\ell=0}^{L} W(\ell,k)\,H_{\mathrm{L}}(\ell,k)\,S(n-\ell,k) \qquad W(\ell,k) = \mathrm{e}^{-\alpha(k)\,\ell}$$

$$Z_{\mathrm{R}}(n,k) = \sum_{\ell=0}^{L} W(\ell,k)\,H_{\mathrm{R}}(\ell,k)\,S(n-\ell,k)$$

Reverberation Cancellation
Moving Average Process (TF-Domain)

- In (B. Schwartz et al., 2015) the microphone signals were modeled in the TF domain as a moving average process

$$Y_m(n,k) = \underbrace{H_m(0,k)\,S(n,k)}_{X_m(n,k)} + \sum_{\ell=1}^{L} H_m(\ell,k)\,S(n-\ell,k)$$

- In the context of binaural hearing aids (B. Schwartz et al., 2015):

$$Z_{\mathrm{L}}(n,k) = \sum_{\ell=0}^{L} W(\ell,k)\,\widetilde{H}_{\mathrm{L}}(\ell,k)\,\underbrace{\boxed{X_{\mathrm{L}}(n-\ell,k)}}_{\text{Early Speech at the Reference Microphone}} \qquad W(\ell,k) = \mathrm{e}^{-\alpha(k)\,\ell}$$

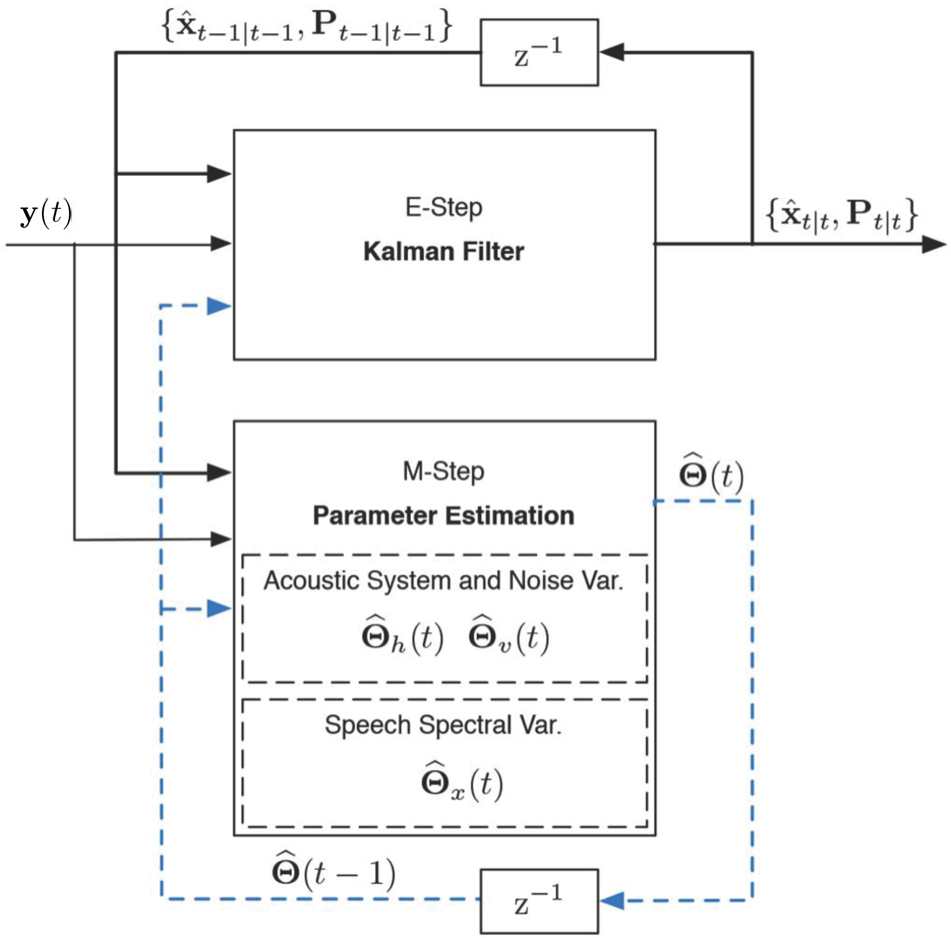
$$Z_{\mathrm{R}}(n,k) = \sum_{\ell=0}^{L} W(\ell,k)\,\widetilde{H}_{\mathrm{R}}(\ell,k)\,X_{\mathrm{L}}(n-\ell,k)$$

Relative CTFs

$$\boxed{\widetilde{H}_m(\ell,k)} = H_{\mathrm{L}}^{-1}(0,k)H_{\mathrm{m}}(\ell,k)$$

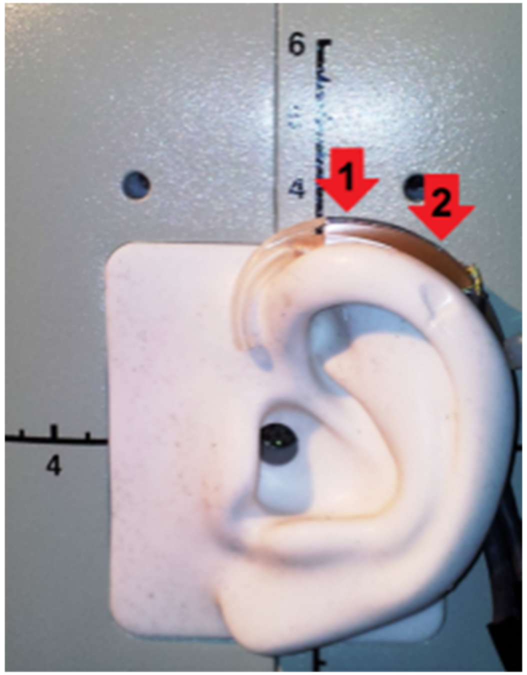
Reverberation Cancellation
Moving Average Process (TF-Domain)

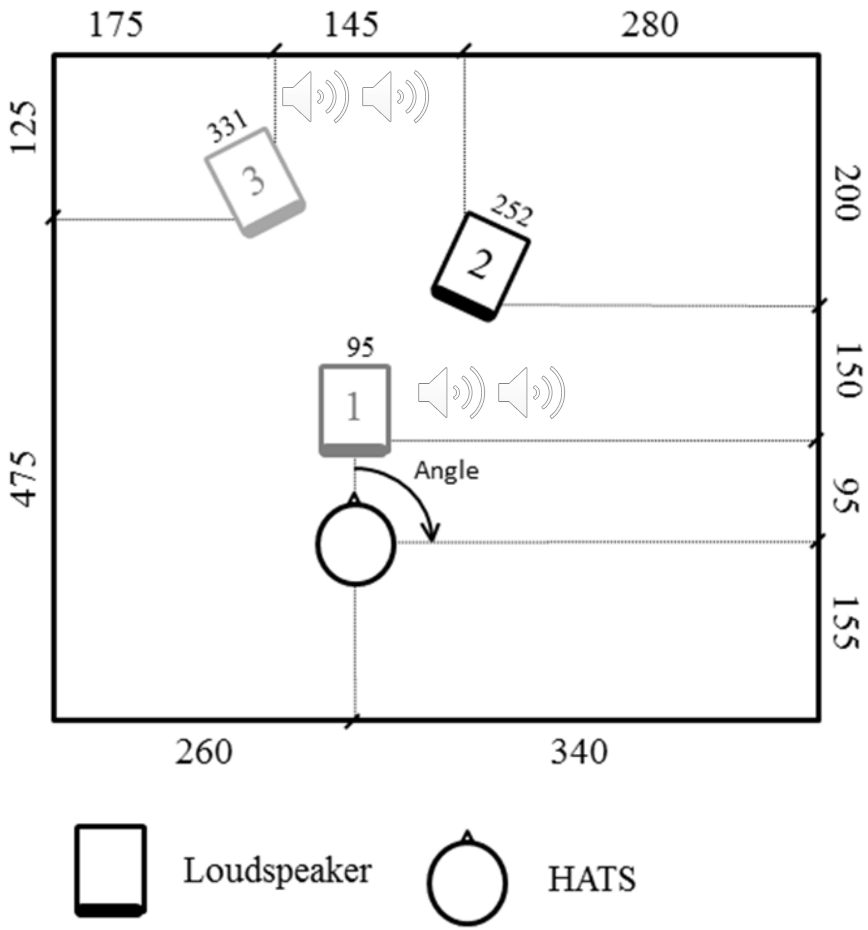
- A recursive expectation-maximization scheme is used to estimate online the acoustic system, speech, and noise parameters
- In the E-Step, a Kalman filter is used to estimate the desired speech signal (and the error covariance matrix)



Source: (B. Schwartz et al., 2015)

Reverberation Cancellation
Binaural Hearing Aids

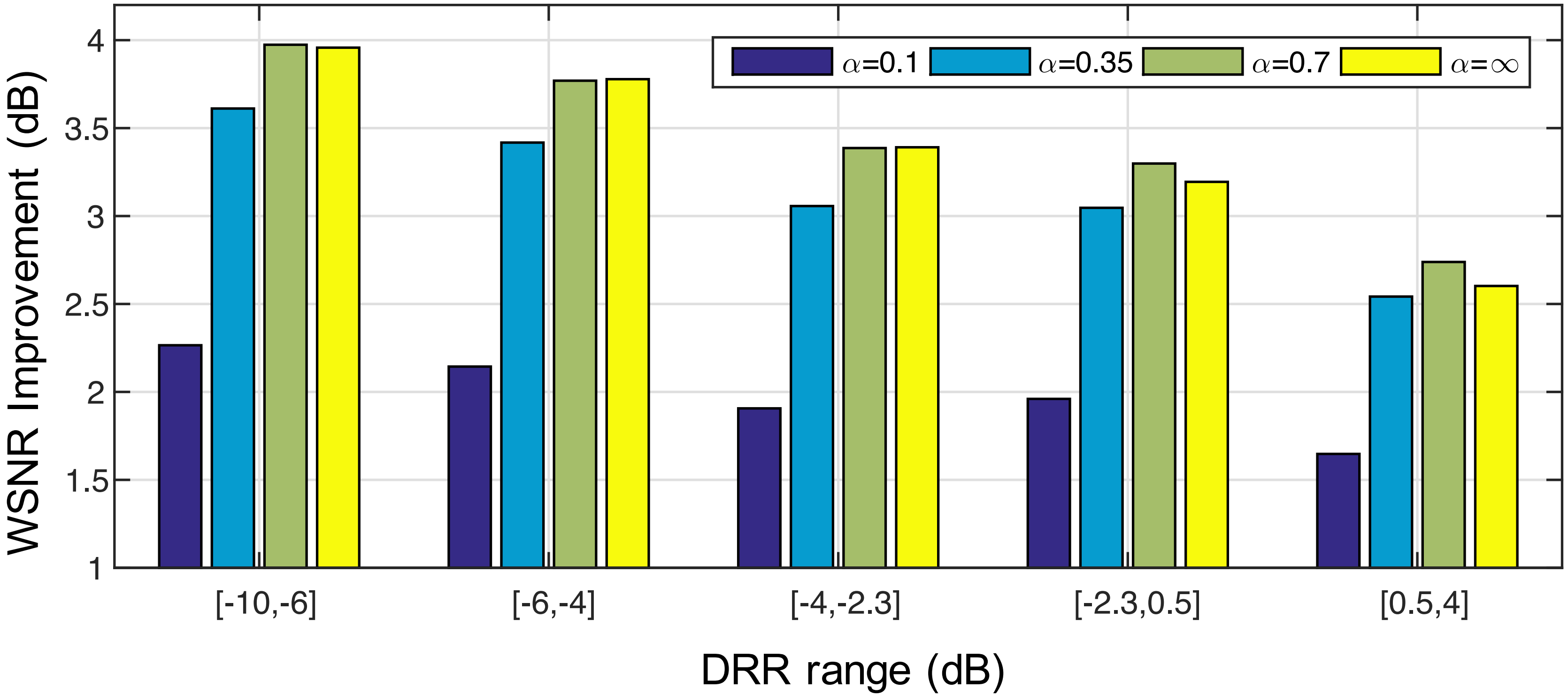




Source: (B. Schwartz et al., 2015)

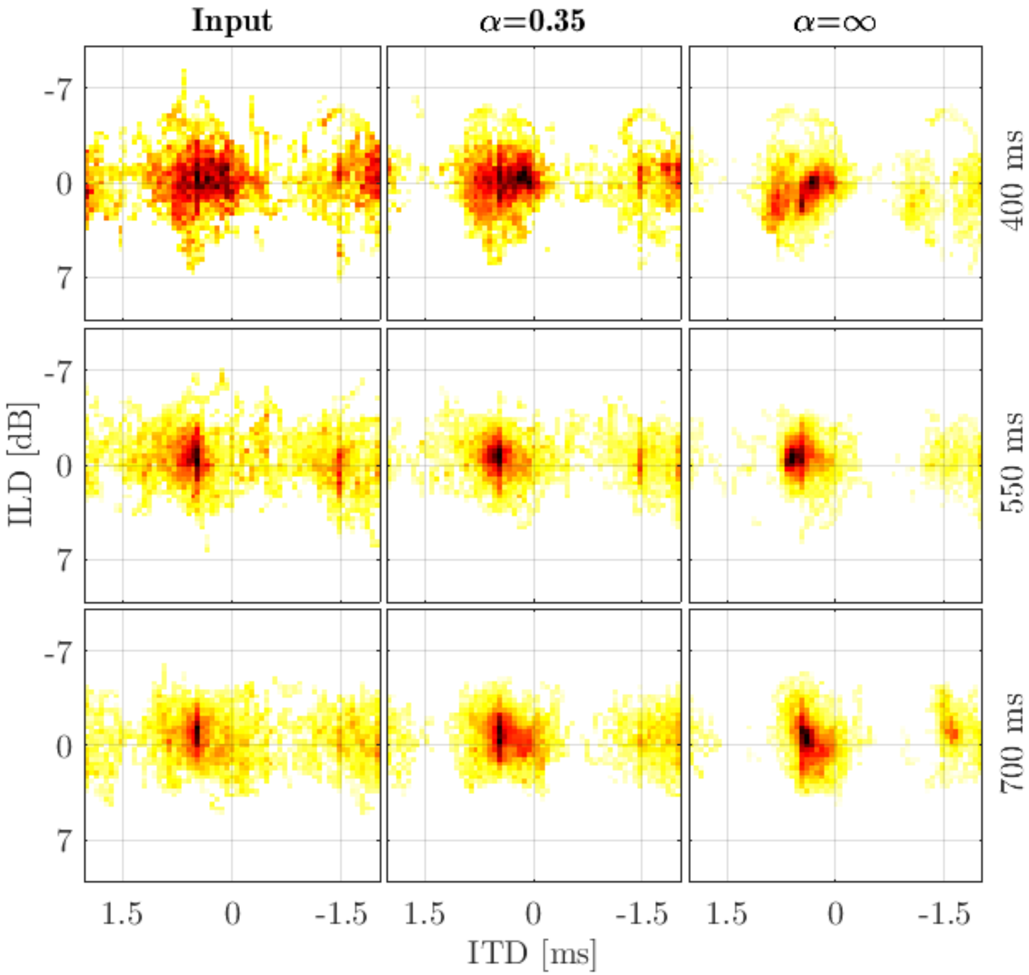
Reverberation Cancellation

Binaural Hearing Aids



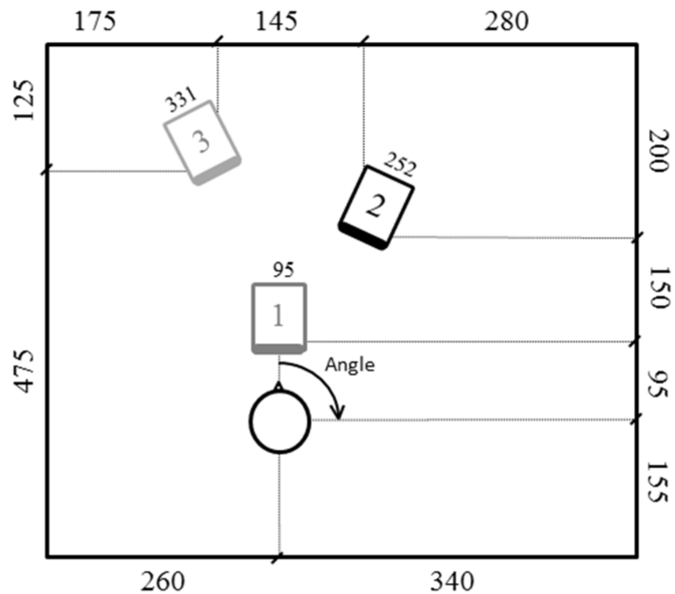
Source: (B. Schwartz et al., 2015)

Reverberation Cancellation
Binaural Hearing Aids



Source: (B. Schwartz et al., 2015)

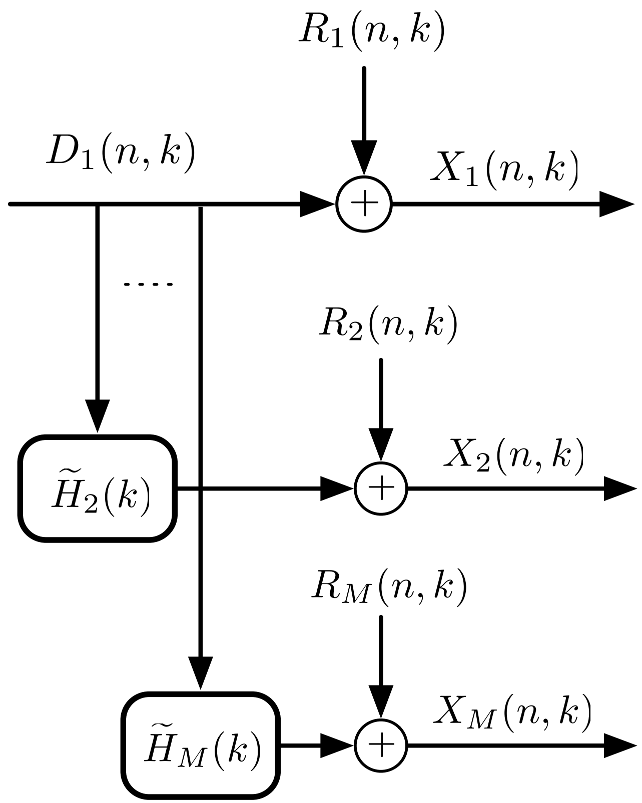
- ITD and ILD distribution for the different reverberation levels and window functions
- These plots relate to Position 3



Reverberation Suppression

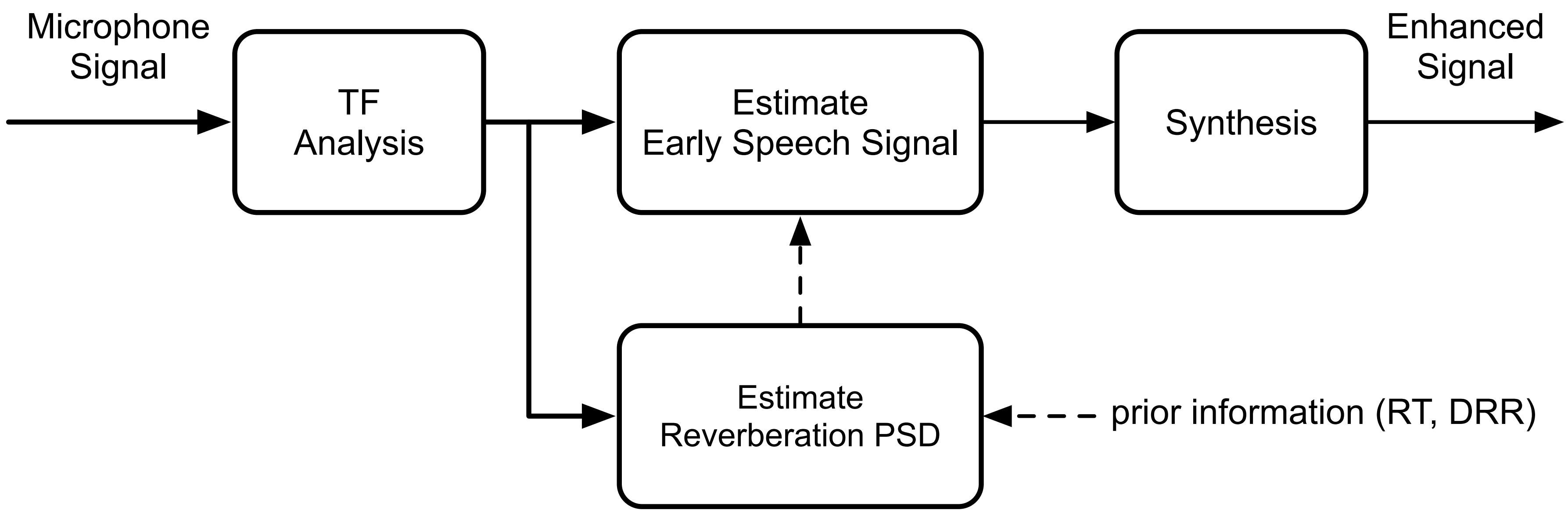
- It is assumed that
  1. reverberation is an additive process
  2. the desired signal and the reverberant signal are uncorrelated
  3. the reverberant signal can be modeled as

$$\mathbf{r}(n,k) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \phi_R(n,k)\boldsymbol{\Gamma}(k))$$



Reverberation Suppression
Single-Channel Spectral Enhancement

- Single-channel spectral enhancement techniques commonly require an estimate of the clean speech PSD and the interference PSD
- Statistical models for the acoustic channel can be used to derive estimators for the reverberation PSD



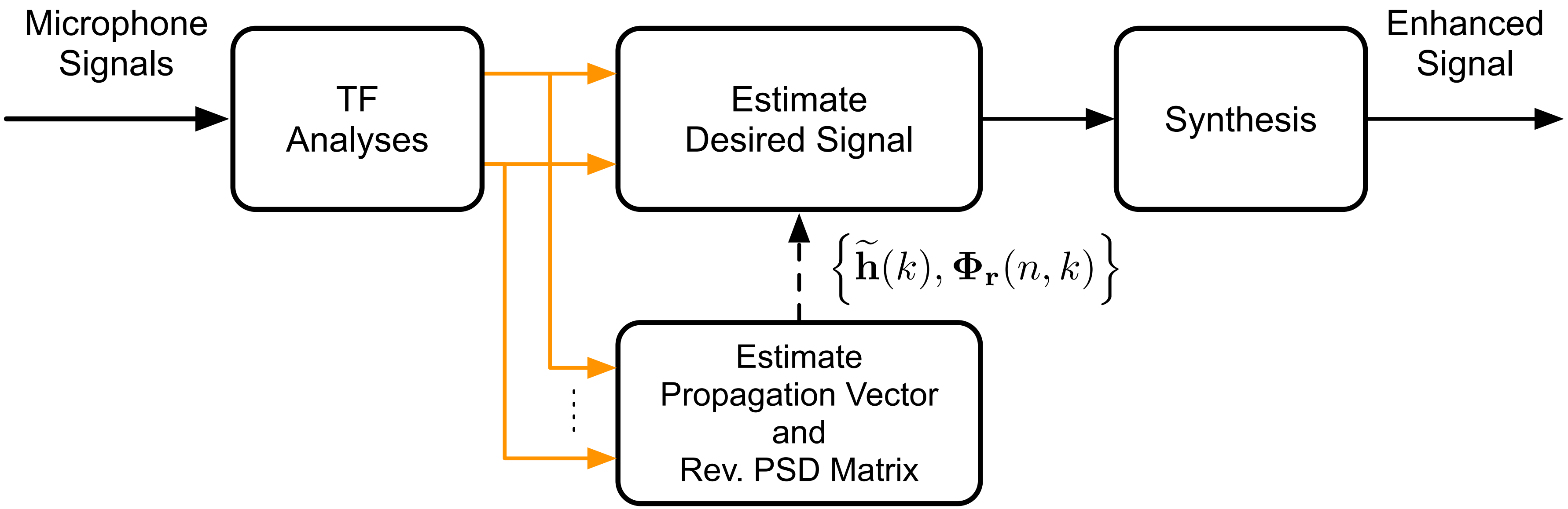
Reverberation Suppression
Single-Channel Spectral Enhancement

- Selected approaches to estimate the reverberation PSD:
  - Lebart et al. (2001) used Moorer's model and a frequency independent reverberation time (RT)
  - Habets and Sommen (2003) used Polack's model and a frequency dependent RT
  - Habets et al. (2007/2009) proposed a generalized statistical model that depends on the direct-to-reverberation ratio (DRR) and RT
  - Erkelens et al. (2010) proposed a correlation-based PSD estimator
- This let to new challenges such as blindly estimating the DRR and RT which were also part of the recent ACE 2015 Challenge

Reverberation Suppression
Data-Dependent Spatial Filtering

- Fully exploit the spatial diversity of the desired and undesired signals

- Estimate the propagation vector and reverberation PSD matrix



Reverberation Suppression
Data-Dependent Spatial Filtering

- In (O. Schwartz et al., 2016) we used a multi-channel MMSE filter
- The **relative early transfer functions**, as well as the **level and spatial coherence matrix of the reverberation** were iteratively estimated using an Expectation-Maximization scheme
- Example (Distance = 2 m, RT_{60}=0.61 s, Noiseless)

	Female	Male
Input		
Singe-channel Dereverberation (Habets, 2007)		
Four-channel Dereverberation (2016)		

Some audio demos can be found at
http://www.eng.biu.ac.il/gannot/speech-enhancement/wiener-em

Outline

- Acoustic Signal Extraction
- Dereverberation
  - Reverberation Cancellation
  - Reverberation Suppression
- Conclusions and Future Challenges

Conclusions and Future Challenges

- Significant advances have been made in the areas of acoustic signal extraction and dereverberation
- Using newly developed acoustic signal processing techniques we are starting to see a **true benefit of multi-microphone processing**

- Future Challenges
  - Lower signal-to-noise ratios and higher reverberation times
  - Incorporating perceptual models and knowledge
  - Automatic adaptation of the desired spatial response

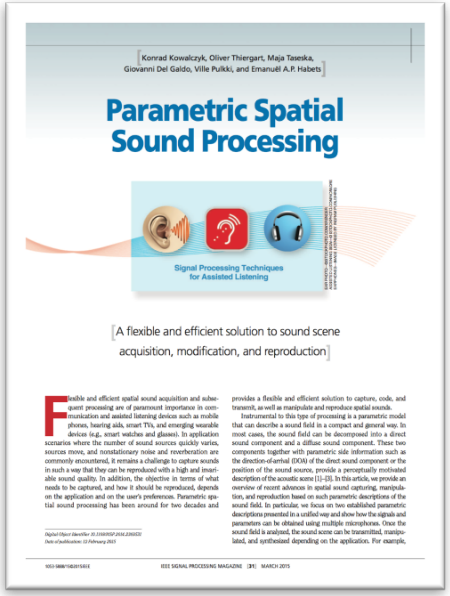
Special thanks to

- Oliver Thiergart (AudioLabs)
- Maja Taseska (AudioLabs)
- Sharon Gannot (Bar-Ilan University, Israel)
- Boaz Schwartz (Bar-Ilan University, Israel)
- Ofer Schwartz (Bar-Ilan University, Israel)

Thank you for your attention….

References

- O. Thiergart, M. Taseska and E.A.P. Habets, “An informed parametric spatial filter based on instantaneous direction-of-arrival estimates,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 22, Issue 12, pp. 2182-2196, Dec. 2014.
- K. Kowalczyk, O. Thiergart, M. Taseska, G. Del Galdo, V. Pulkki and E.A.P. Habets, "Parametric spatial sound processing: A flexible and efficient solution to sound scene acquisition, modification and reproduction," IEEE Signal Processing Magazine, Vol. 32, Issue 2, pp. 31-42, Mar. 2015.
- M. Taseska and E.A.P. Habet, "Minimum Bayes risk signal detection for speech enhancement based on a narrowband DOA model," Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brisbane, Australia, 2015.
- M. Taseska and E.A.P. Habets, "Spotforming using distributed microphone arrays," Best Student Paper Award, Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, USA, Oct. 20-23, 2013.
- B. Schwartz, S. Gannot and E.A.P. Habets, “An online dereverberation algorithm for hearing aids with binaural cues preservation,” Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015.
- O. Schwartz, S. Gannot and E.A.P. Habets, "An expectation-maximization algorithm for multi-microphone speech dereverberation and noise reduction with coherence matrix estimation," IEEE/ACM Transactions on Audio, Speech, and Language Processing," to appear.


Konrad Kowalczyk, Oliver Thiergart, Maja Taseska, Giovanni Del Galdo, Ville Pulkki, and Emanuël A.P. Habets

Parametric Spatial Sound Processing

Signal Processing Techniques for Assisted Listening

A flexible and efficient solution to sound scene acquisition, modification, and reproduction