Quality Analysis and Enhancement of Reverberated Speech using Microphone Arrays

Reuven Berkun

Department of Electrical Engineering
Technion – Israel Institute of Technology

Supervised by Prof. Israel Cohen and Prof. Jacob Benesty

November 18, 2015
Motivation

- Reverberation
  - Late & uncorrelated reflections
  - Quality & intelligibility degradation, temporal smearing
Motivation

- Reverberation
  - Late & uncorrelated reflections
  - Quality & intelligibility degradation, temporal smearing

(a)  
(b)
General Overview

- Unsolved problem:
  - Reverberation in speech signals
- The solution we propose:
  - Use microphone arrays
- The topic we will discuss:
  - Enhancement & reverberation suppression
  - Quality analysis of speech signals
General Outline

I Robust Regularized Superdirective Beamforming

II The Tunable Superdirective Beamformer

III Microphone Array Reverberation Assessment

IV Conclusions
General Outline

I Robust Regularized Superdirective Beamforming

II The Tunable Superdirective Beamformer

III Microphone Array Reverberation Assessment

IV Conclusions
General Outline

I Robust Regularized Superdirective Beamforming

II The Tunable Superdirective Beamformer

III Microphone Array Reverberation Assessment

IV Conclusions
Part I

Robust Regularized Superdirective Beamforming

Based on:


Background

- Multi-channel vs. single microphone
  - Preserve spatial information
  - Higher processing flexibility

- Microphone array benefits:
  - Source of interest extraction
  - Handling multiple sources
  - Superior noise suppression

- Impact factors:
  - Number of sensors
  - Array geometry
  - Processing algorithm
Background

- **Multi-channel vs. single microphone**
  - Preserve spatial information
  - Higher processing flexibility

- **Microphone array benefits:**
  - Source of interest extraction
  - Handling multiple sources
  - Superior noise suppression

- **Impact factors:**
  - Number of sensors
  - Array geometry
  - Processing algorithm
Background

- **Multi-channel vs. single microphone**
  - Preserve spatial information
  - Higher processing flexibility

- **Microphone array benefits**:
  - Source of interest extraction
  - Handling multiple sources
  - Superior noise suppression

- **Impact factors**:
  - Number of sensors
  - Array geometry
  - Processing algorithm
Microphone array beamforming:
- Class of Multichannel signal processing algorithms
- Extraction & enhancement of desired signals
- Noise & reverberation suppression

Delay & Sum (DS) beamformer
- Low gain for diffuse noise
- Inferior performance for reverberation
Background – Contd.

- Microphone array beamforming:
  - Class of Multichannel signal processing algorithms
  - Extraction & enhancement of desired signals
  - Noise & reverberation suppression

- Delay & Sum (DS) beamformer
  - Low gain for *diffuse* noise
  - Inferior performance for reverberation
Motivation

- **Reverberation**
  - Late & uncorrelated reflections
- **Superdirective beamforming:**
  - Supergain for diffuse noise
  - White noise amplification
- **The optimal beamformer:**
  - High gain for diffuse noise
  - Control of the white noise amplification
Motivation

- Reverberation
  - Late & uncorrelated reflections

- Superdirective beamforming:
  - Supergain for diffuse noise
  - White noise amplification

- The optimal beamformer:
  - High gain for diffuse noise
  - Control of the white noise amplification
Motivation

- Reverberation
  - Late & uncorrelated reflections

- Superdirective beamforming:
  - Supergain for diffuse noise
  - White noise amplification

- The optimal beamformer:
  - High gain for diffuse noise
  - Control of the white noise amplification
Signal Model

- Uniform Linear Array (ULA) with $M$ omnidirectional microphones
- Fixed beamformers with small $\delta$
  - Azimuth angle $\theta$
  - Mainlobe at endfire direction ($\theta = 0^\circ$)
Signal Model

- Uniform Linear Array (ULA) with $M$ omnidirectional microphones
- Fixed beamformers with small $\delta$
- Azimuth angle $\theta$
- Mainlobe at endfire direction ($\theta = 0^\circ$)

**Equations and Diagram**: 

\[ Y_1(\omega) = V_1(\omega) \]
\[ Y_M(\omega) = V_M(\omega) \]
\[ Y_2(\omega) = (M - 1)d \cos \theta \]
\[ X(\omega) = Y_1(\omega) - Y_M(\omega) \]

**Diagram Note**: 
- Plane wavefront
- Microphone positions
- Distance $\delta$
- Azimuth angle $\theta$
The steering vector (of length $M$):
\[ d(\omega, \theta) = \begin{bmatrix} 1 & e^{-j\omega \tau_0 \cos \theta} & \ldots & e^{-j(M-1)\omega \tau_0 \cos \theta} \end{bmatrix}^T \]

The observation signal vector:
\[ y(\omega) = \begin{bmatrix} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{bmatrix}^T = x(\omega) + v(\omega) = d(\omega)X(\omega) + v(\omega) \]
The steering vector (of length $M$):

\[
d(\omega, \theta) = \begin{bmatrix} 1 & e^{-j\omega \tau_0 \cos \theta} & \cdots & e^{-j(M-1)\omega \tau_0 \cos \theta} \end{bmatrix}^T
\]

The observation signal vector:

\[
y(\omega) = \begin{bmatrix} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{bmatrix}^T
\]

\[
y(\omega) = x(\omega) + v(\omega)
\]

\[
y(\omega) = d(\omega)X(\omega) + v(\omega)
\]
The objective:

- Apply $h(\omega)$ to the observation vector $y(\omega)$
- The beamformer output:

$$Z(\omega) = h^H(\omega)y(\omega)$$
$$= h(\omega)d(\omega)X(\omega) + h^H(\omega)v(\omega)$$

The distortionless constraint

$$h^H(\omega)d(\omega) = 1$$
SNR Gain:

\[
G[h(\omega)] = \frac{oSNR[h(\omega)]}{iSNR(\omega)} = \frac{|h^H(\omega)d(\omega)|^2}{h^H(\omega)\Gamma_v(\omega)h(\omega)}
\]

\[
\Gamma_v(\omega) = \frac{\Phi_v(\omega)}{\phi_{V_1}(\omega)}
\]
White Noise Gain (WNG):

\[
\mathcal{W}[h(\omega)] = G\left[h(\omega) \mid \Gamma_v(\omega) = I_M\right] \\
\leq M, \quad \forall h(\omega)
\]

For WNG: \( \Gamma_v(\omega) = I_M \)

WNG - robustness measure against:
- Array weight imperfections
- Sensor position mismatches
- Uncorrelated noise
- Microphone thermal noise

\[\mathcal{W}[h_S(\omega)] \approx 0\]
White Noise Gain (WNG):

\[
\mathcal{W}[h(\omega)] = G[h(\omega) \mid \Gamma_v(\omega) = I_M] \leq M, \forall h(\omega)
\]

For WNG: \( \Gamma_v(\omega) = I_M \)

- **WNG** - robustness measure against:
  - Array weight imperfections
  - Sensor position mismatches
  - Uncorrelated noise
  - Microphone thermal noise

\[
\mathcal{W}[h_S(\omega)] \approx 0
\]
White Noise Gain

- **White Noise Gain (WNG):**

\[
\mathcal{W}[h(\omega)] = G \left[ h(\omega) \mid \Gamma_v(\omega) = I_M \right] \\
\leq M, \ \forall h(\omega)
\]

- For **WNG**: \( \Gamma_v(\omega) = I_M \)
- **WNG** - robustness measure against:
  - Array weight imperfections
  - Sensor position mismatches
  - Uncorrelated noise
  - Microphone thermal noise

- \( \mathcal{W}[h_S(\omega)] \approx 0 \)
White noise gain of the DS (solid line) and superdirective (dashed line) beamformers, with $M = 10$ microphones, and $\delta = 1$ cm.
Directivity Factor

- Gain for **diffuse** noise
  - $\Gamma_v(\omega) = \Gamma_d(\omega)$
- Directivity Factor (DF):
  
  \[
  D[h(\omega)] = g\left[h(\omega) \left| \Gamma_v(\omega) = \Gamma_d \right.\right] \\
  \leq M^2, \ \forall h(\omega)
  \]
- $D[h_s(\omega)] = D_{\text{max}}(\omega) \xrightarrow{\delta \to 0} M^2$
Gain for *diffuse* noise

- $\Gamma_v(\omega) = \Gamma_d(\omega)$

**Directivity Factor (DF):**

\[
\mathcal{D}[h(\omega)] = \mathcal{G}\left[h(\omega) \mid \Gamma_v(\omega) = \Gamma_d\right] \\
\leq M^2, \ \forall h(\omega)
\]

\[
\mathcal{D}[h_S(\omega)] = \mathcal{D}_{\text{max}}(\omega) \xrightarrow[\delta \to 0]{} M^2
\]
Directivity Factor

- **Gain for diffuse noise**
  - $\Gamma_v(\omega) = \Gamma_d(\omega)$
- **Directivity Factor (DF):**
  \[
  D[h(\omega)] = G \left[ h(\omega) \bigg| \Gamma_v(\omega) = \Gamma_d \right] 
  \leq M^2, \quad \forall h(\omega)
  \]
- $D[h_S(\omega)] = D_{\text{max}}(\omega) \underset{\delta \to 0}{\longrightarrow} M^2$
Directivity factor of the DS (solid line) and superdirective (dashed line) beamformers, with $M = 10$ microphones, and $\delta = 1$ cm.
Beampattern

- Angular directivity-factor response
  - Beamformer response as function of $d(\omega, \theta)$
- Beampattern:

$$B[h(\omega), \theta] = d^H(\omega, \theta)h(\omega)$$

$$= \sum_{m=1}^{M} H_m(\omega) e^{j(m-1)\omega_0 \cos \theta}$$
Beampattern – Contd.

(a) $f = 1\, \text{KHz}$  
(b) $f = 3\, \text{KHz}$  
(c) $f = 1\, \text{KHz}$  
(d) $f = 3\, \text{KHz}$

Beampattern of the DS (a,b) and superdirective (c,d) beamformers, with $M = 8$ microphones, and $\delta = 1\, \text{cm}$. 

R. Berkun  
Microphone Array Reverberation Analysis and Enhancement 19/74
The Delay & Sum (DS) Beamformer

- Maximizes the WNG
  - subject to distortionless constraint

\[
h_{DS}(\omega) = \frac{d(\omega)}{d^H(\omega)d(\omega)} = \frac{d(\omega)}{M}
\]

- \(\mathcal{W}[h_{DS}(\omega)] = M = \mathcal{W}_{\text{max}}\)
- \(\mathcal{D}[h_{DS}(\omega)] \geq 1\)
The Delay & Sum (DS) Beamformer

- Maximizes the WNG
  - subject to distortionless constraint
  \[ h_{DS}(\omega) = \frac{d(\omega)}{d^H(\omega)d(\omega)} = \frac{d(\omega)}{M} \]
  - \( \mathcal{W}[h_{DS}(\omega)] = \mathcal{M} = \mathcal{W}_{\text{max}} \)
  - \( D[h_{DS}(\omega)] \geq 1 \)
The Superdirective Beamformer

- Maximizes the DF
  - subject to distortionless constraint

\[ h_S(\omega) = \frac{\Gamma_d^{-1}(\omega)d(\omega)}{d^H(\omega) \Gamma_d^{-1}(\omega)d(\omega)} \]

- \[ \mathcal{D}[h_S(\omega)] = d^H(\omega) \Gamma_d^{-1}(\omega)d(\omega) = \mathcal{D}_{\text{max}}(\omega) \]
The Superdirective Beamformer

- Maximizes the DF
  - subject to distortionless constraint

$$h_S(\omega) = \frac{\Gamma_d^{-1}(\omega)d(\omega)}{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)}$$

- $$\mathcal{D}[h_S(\omega)] = d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega) = \mathcal{D}_{\text{max}}(\omega)$$
Prior Work

- **Superdirective beamformer with white noise gain constraint:**
  - Introduced by [Cox et al., 1987]
  - Optimal constrained solution
  - Multistep reconstruction algorithm

- Optimization approaches:
  - Linear programming [Vorobyov et al., 2003]
  - Different mismatch formulations [Doclo and Moonen, 2007]
  - Minimax & non-linear optimization [Doclo and Moonen, 2003]

- The proposed approach:
  - Closed-form solution
  - Effective tradeoff
    - High gain for diffuse noise (directivity factor)
    - Fine control of the white noise gain
Prior Work

- **Superdirective beamformer with white noise gain constraint:**
  - Introduced by [Cox et al., 1987]
  - Optimal constrained solution
  - Multistep reconstruction algorithm

- **Optimization approaches:**
  - Linear programming [Vorobyov et al., 2003]
  - Different mismatch formulations [Doclo and Moonen, 2007]
  - Minimax & non-linear optimization [Doclo and Moonen, 2003]

- **The proposed approach:**
  - Closed-form solution
  - Effective tradeoff
  - High gain for diffuse noise (directivity factor)
  - Fine control of the white noise gain
Prior Work

- **Superdirective beamformer with white noise gain constraint:**
  - Introduced by [Cox et al., 1987]
  - Optimal constrained solution
  - Multistep reconstruction algorithm

- **Optimization approaches:**
  - Linear programming [Vorobyov et al., 2003]
  - Different mismatch formulations [Doclo and Moonen, 2007]
  - Minimax & non-linear optimization [Doclo and Moonen, 2003]

- **The proposed approach:**
  - Closed-form solution
  - Effective tradeoff
    - High gain for diffuse noise (directivity factor)
    - Fine control of the white noise gain
The Regularized Superdirective Beamformer

- Maximizes the DF [Cox et al., 1987, Benesty et al., 2008]
  - Constraint on the WNG
  - subject to distortionless constraint
    \[ h_{S,\epsilon}(\omega) = \frac{[\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)}{d^H(\omega) [\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)} \]
  - Define: \( \Gamma_\epsilon(\omega) = \Gamma_d(\omega) + \epsilon I_M \)
The Regularized Superdirective Beamformer

- Maximizes the DF [Cox et al., 1987, Benesty et al., 2008]
  - Constraint on the WNG
  - Subject to distortionless constraint

\[
\mathbf{h}_{S,\epsilon}(\omega) = \frac{[\mathbf{\Gamma}_d(\omega) + \epsilon \mathbf{I}_M]^{-1} \mathbf{d}(\omega)}{\mathbf{d}^H(\omega) [\mathbf{\Gamma}_d(\omega) + \epsilon \mathbf{I}_M]^{-1} \mathbf{d}(\omega)}
\]

- Define: \( \mathbf{\Gamma}_\epsilon(\omega) = \mathbf{\Gamma}_d(\omega) + \epsilon \mathbf{I}_M \)

\[
\mathbf{h}_{S,\epsilon}(\omega) \bigg|_{\mathbf{\Gamma}_d(\omega) \leftarrow \mathbf{\Gamma}_\epsilon(\omega)}
\]
The Proposed Beamformer

- Linear combination of the DS (maximum WNG) and the regularized superdirective (high DF) beamformer:

\[ h_{\alpha, \epsilon}(\omega) = \frac{[\Gamma^{-1}_\epsilon(\omega) + \alpha(\omega)I_M]d(\omega)}{d^H(\omega)[\Gamma^{-1}_\epsilon(\omega) + \alpha(\omega)I_M]d(\omega)} \]

\[ = \frac{h_{S, \epsilon}(\omega)}{1 + \alpha_\epsilon(\omega)} + \frac{h_{DS}(\omega)}{1 + \alpha_\epsilon^{-1}(\omega)} \]

- \( \alpha_\epsilon(\omega) = \alpha(\omega) \frac{W_{\text{max}}}{D_{\text{max}, \epsilon}} \)

\( D_{\text{max}, \epsilon} = d^H(\omega)\Gamma^{-1}_\epsilon(\omega)d(\omega) \)

- Tradeoff: Control of both DF & WNG
- \( \epsilon \) controls the regularization
The Proposed Beamformer

- Linear combination of the DS (maximum WNG) and the regularized superdirective (high DF) beamformer:

\[
\mathbf{h}_{\alpha, \epsilon}(\omega) = \frac{[\mathbf{\Gamma}_{\epsilon}^{-1}(\omega) + \alpha(\omega)\mathbf{I}_M] \mathbf{d}(\omega)}{\mathbf{d}^H(\omega)[\mathbf{\Gamma}_{\epsilon}^{-1}(\omega) + \alpha(\omega)\mathbf{I}_M] \mathbf{d}(\omega)}
\]

\[
= \frac{\mathbf{h}_{S, \epsilon}(\omega)}{1 + \alpha_{\epsilon}(\omega)} + \frac{\mathbf{h}_{DS}(\omega)}{1 + \alpha_{\epsilon}^{-1}(\omega)}
\]

- \(\alpha_{\epsilon}(\omega) = \alpha(\omega) \frac{\mathcal{W}_{\text{max}}}{\mathcal{D}_{\text{max,} \epsilon}}\)

\[
\mathcal{D}_{\text{max,} \epsilon} = \mathbf{d}^H(\omega) \mathbf{\Gamma}_{\epsilon}^{-1}(\omega) \mathbf{d}(\omega)
\]

- Tradeoff: Control of both DF & WNG

- \(\epsilon\) controls the regularization
Array gains of $h_{\alpha, \epsilon}(\omega)$ beamformer (solid line), with $\alpha = 1$, $\epsilon = 1 \cdot 10^{-5}$, $M = 8$, and $\delta = 1$ cm.
Fixed-WNG Beamformer

\[ W[h_{\alpha,\epsilon}(\omega)] = f\{W[h_{DS}(\omega)], W[h_{S,\epsilon}(\omega)]\} \] (1)

- Design \( h_{\alpha,\epsilon}(\omega) \) such that
  - \( W_0 \) is a constant user-determined WNG response
  - \( W[h_{S,\epsilon}(\omega)] < W_0 < M, \forall \omega \)
- Substitute \( W_0 \) in (1)

\[
\Rightarrow \alpha_\epsilon(\omega) = \cdots \\
\Rightarrow h_{\alpha+,\epsilon}(\omega) = \frac{h_{S,\epsilon}(\omega)}{1 + \alpha_{\epsilon+}(\omega)} + \frac{h_{DS}(\omega)}{1 + \alpha_{\epsilon+}^{-1}(\omega)}
\]
Fixed-WNG Beamformer

\[ \mathcal{W}[\mathbf{h}_{\alpha,\epsilon}(\omega)] = f\{\mathcal{W}[\mathbf{h}_{DS}(\omega)], \mathcal{W}[\mathbf{h}_S,\epsilon(\omega)]\} \quad (1) \]

- Design \( \mathbf{h}_{\alpha,\epsilon}(\omega) \) such that
  - \( \mathcal{W}_0 \) is a constant user-determined WNG response
  - \( \mathcal{W}[\mathbf{h}_S,\epsilon(\omega)] < \mathcal{W}_0 < M, \forall \omega \)
- Substitute \( \mathcal{W}_0 \) in (1)

\[ \Rightarrow \alpha_{\epsilon}(\omega) = \cdots \]
\[ \Rightarrow \mathbf{h}_{\alpha+,\epsilon}(\omega) = \frac{\mathbf{h}_{S,\epsilon}(\omega)}{1 + \alpha_{\epsilon+}(\omega)} + \frac{\mathbf{h}_{DS}(\omega)}{1 + \alpha_{\epsilon+}^{-1}(\omega)} \]
Fixed-WNG Beamformer

\[ \mathcal{W}[\mathbf{h}_{\alpha,\epsilon}(\omega)] = f\{\mathcal{W}[\mathbf{h}_{DS}(\omega)], \mathcal{W}[\mathbf{h}_{S,\epsilon}(\omega)]\} \]

\[ \mathcal{W}_0 \]

- Design \( \mathbf{h}_{\alpha,\epsilon}(\omega) \) such that
  - \( \mathcal{W}_0 \) is a constant user-determined WNG response
  - \( \mathcal{W}[\mathbf{h}_{S,\epsilon}(\omega)] < \mathcal{W}_0 < M, \forall \omega \)
- Substitute \( \mathcal{W}_0 \) in (1)

\[ \Rightarrow \alpha_{\epsilon}(\omega) = \cdots \]

\[ \Rightarrow \mathbf{h}_{\alpha+,\epsilon}(\omega) = \frac{\mathbf{h}_{S,\epsilon}(\omega)}{1 + \alpha_{\epsilon+}(\omega)} + \frac{\mathbf{h}_{DS}(\omega)}{1 + \alpha_{\epsilon+}^{-1}(\omega)} \]
Fixed-DF Beamformer

\[ \mathcal{D}[h_{\alpha,\epsilon}(\omega)] = f\{\mathcal{D}[h_{DS}(\omega)], \mathcal{D}[h_{S,\epsilon}(\omega)]\} \quad (2) \]

- Design \( h_{\alpha,\epsilon}(\omega) \) such that
  - \( \mathcal{D}_0 \) is a fixed user-determined DF response
  - \( \mathcal{D}[h_{DS}(\omega)] < \mathcal{D}_0 < \mathcal{D}[h_{S,\epsilon}(\omega)], \forall \omega \)
- Substitute \( \mathcal{D}_0 \) in (2)

\[ \Rightarrow \alpha_{\epsilon}(\omega) = \cdots \]

\[ \Rightarrow h_{\tilde{\alpha}_+,\epsilon}(\omega) = \frac{h_{S,\epsilon}(\omega)}{1 + \tilde{\alpha}_{\epsilon+}(\omega)} + \frac{h_{DS}(\omega)}{1 + \tilde{\alpha}_{\epsilon+}^{-1}(\omega)} \]
Fixed-DF Beamformer

\[ D[h_{\alpha, \epsilon}(\omega)] = \mathcal{f}\{D[h_{DS}(\omega)], D[h_{S, \epsilon}(\omega)]\} \]

= \mathcal{D}_0

- **Design** \( h_{\alpha, \epsilon}(\omega) \) such that
  - \( \mathcal{D}_0 \) is a fixed user-determined DF response
  - \( \mathcal{D}[h_{DS}(\omega)] < \mathcal{D}_0 < \mathcal{D}[h_{S, \epsilon}(\omega)], \forall \omega \)

- **Substitute** \( \mathcal{D}_0 \) in (2)

\[ \Rightarrow \alpha_{\epsilon}(\omega) = \cdots \]

\[ \Rightarrow h_{\tilde{\alpha}+, \epsilon}(\omega) = \frac{h_{S, \epsilon}(\omega)}{1 + \tilde{\alpha}_{\epsilon+}(\omega)} + \frac{h_{DS}(\omega)}{1 + \tilde{\alpha}_{\epsilon+}^{-1}(\omega)} \]
Fixed-DF Beamformer

\[
\mathcal{D}[\mathbf{h}_{\alpha,\epsilon}(\omega)] = f\{\mathcal{D}[\mathbf{h}_{DS}(\omega)], \mathcal{D}[\mathbf{h}_{S,\epsilon}(\omega)]\}
\]

\[
\Rightarrow \mathcal{D}_0
\]

- **Design** \(\mathbf{h}_{\alpha,\epsilon}(\omega)\) such that
  - \(\mathcal{D}_0\) is a fixed user-determined DF response
  - \(\mathcal{D}[\mathbf{h}_{DS}(\omega)] < \mathcal{D}_0 < \mathcal{D}[\mathbf{h}_{S,\epsilon}(\omega)], \forall \omega\)

- **Substitute** \(\mathcal{D}_0\) in (2)

\[
\Rightarrow \alpha_{\epsilon}(\omega) = \cdots
\]

\[
\Rightarrow \mathbf{h}_{\tilde{\alpha}_+,\epsilon}(\omega) = \frac{\mathbf{h}_{S,\epsilon}(\omega)}{1 + \tilde{\alpha}_{\epsilon+}(\omega)} + \frac{\mathbf{h}_{DS}(\omega)}{1 + \tilde{\alpha}_{\epsilon+}^{-1}(\omega)}
\]
Fixed-DF Beamformer - Contd.

- Another interpretation of the DF:

\[
\mathcal{D}[\mathbf{h}(\omega)] = \frac{|\mathcal{B}[\mathbf{h}(\omega), 0^\circ]|^2}{\frac{1}{2} \int_0^\pi |\mathcal{B}[\mathbf{h}(\omega), \theta]|^2 \sin \theta \, d\theta}
\]

- For the fixed-DF beamformer:
  - \( \mathcal{D}_0^{-1} = \frac{1}{2} \int_0^\pi |\mathcal{B}[\mathbf{h}_\alpha, \epsilon(\omega), \theta]|^2 \sin \theta \, d\theta \)

  \Rightarrow \text{Frequency-invariant beampattern beamformer (FIBP)}

  [Traverso et al., 2014]
Another interpretation of the DF:

$$\mathcal{D}[h(\omega)] = \frac{|B[h(\omega), 0^\circ]|^2}{\frac{1}{2} \int_0^\pi |B[h(\omega), \theta]|^2 \sin \theta d\theta}$$

For the fixed-DF beamformer:

$$\mathcal{D}_0^{-1} = \frac{1}{2} \int_0^\pi |B[h_{\alpha+}, \epsilon(\omega), \theta]|^2 \sin \theta d\theta$$

⇒ Frequency-invariant beampattern beamformer (FIBP)  
[Traverso et al., 2014]
Another interpretation of the DF:

\[
D[h(\omega)] = \frac{|B[h(\omega), 0^\circ]|^2}{\frac{1}{2} \int_0^\pi |B[h(\omega), \theta]|^2 \sin \theta \, d\theta}
\]

For the fixed-DF beamformer:

\[
D_0^{-1} = \frac{1}{2} \int_0^\pi |B[h_{\alpha+}, \epsilon(\omega), \theta]|^2 \sin \theta \, d\theta
\]

⇒ Frequency-invariant beampattern beamformer (FIBP) [Traverso et al., 2014]

\[
\mathbb{E}_\theta \{|B[h(\omega), \theta]|^2\} = \text{const}, \ \forall \omega
\]
Fixed-WNG Simulation

\[ \mathcal{W}[h_{\alpha+\epsilon}(\omega)] = \mathcal{W}_0(\omega) = -10 \text{ dB} \]

Array gains of \( h_{\alpha+\epsilon}(\omega) \) (solid line). \( \epsilon = 1 \cdot 10^{-4}, M = 8 \) and \( \delta = 1 \text{ cm.} \)
Fixed-DF Simulation

- $D[h_{\alpha+}, \epsilon(\omega)] = D_0(\omega) = 9$ dB

Array gains of $h_{\alpha+}, \epsilon(\omega)$ (solid line). $\epsilon = 1 \cdot 10^{-4}$, $M = 8$ and $\delta = 1$ cm.
User-Determined DF Simulation

- Multiband $\mathcal{D}[\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega)] = \mathcal{D}_0(\omega)$

Array gains of $\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega)$ (solid line) with multiband $\mathcal{D}_0(\omega)$.

$\epsilon = 1 \cdot 10^{-4}$, $M = 8$, and $\delta = 1$ cm.
Beampattern Simulation

- Fixed-WNG & fixed-DF beampattern $B[h(\omega), \theta]$

The squared beampattern [dB] with $\epsilon = 1 \cdot 10^{-4}$, $M = 8$. (a) $|B[h_{\alpha+}, \epsilon(\omega), \theta]|^2$, with $\delta = 1$ cm. (b) $|B[h_{\tilde{\alpha}+}, \epsilon(\omega), \theta]|^2$ ($\delta = 1$ cm). (c) $|B[h_{\tilde{\alpha}+}, \epsilon(\omega), \theta]|^2$ of the multiband user determined DF beamformer ($\delta = 2$ cm).
WNG–DF Tradeoff

- The DF curve vs. WNG of $h_{\alpha, \epsilon}(\omega)$ for increasing $\alpha(\omega) = \alpha$

![Diagram showing the tradeoff between DF and WNG]

The DF curve versus WNG, of $h_{\alpha, \epsilon}(\omega)$. $M = 8$, $\delta = 1$ cm, $f = 2$ KHz, and $\alpha$ is monotonically increased from 0 to $\infty$. 
Part II
The Tunable Superdirective Beamformer

Based on:

Outline

1. Introduction
2. The Tunable Beamformer
3. Simulations
Motivation

- Constrained optimization
  - No closed-form expressions
- Setting the regularization factor:
  - Iterative optimization process [Cox et al., 1987]
  - Ad hoc methods [Van Trees, 2002]
  - Optimize the Signal-to-Interference-plus-Noise-Ratio (SINR) [Mestre and Lagunas, 2006]
- The proposed approach:
  - Maximize DF and WNG
Motivation

- Constrained optimization
  - No closed-form expressions
- Setting the regularization factor:
  - Iterative optimization process [Cox et al., 1987]
  - Ad hoc methods [Van Trees, 2002]
  - Optimize the Signal-to-Interference-plus-Noise-Ratio (SINR) [Mestre and Lagunas, 2006]
- The proposed approach:
  - Maximize DF and WNG
Motivation

- Constrained optimization
  - No closed-form expressions
- Setting the regularization factor:
  - Iterative optimization process [Cox et al., 1987]
  - Ad hoc methods [Van Trees, 2002]
  - Optimize the Signal-to-Interference-plus-Noise-Ratio (SINR) [Mestre and Lagunas, 2006]
- The proposed approach:
  - Maximize DF and WNG
The **DF**: 

\[
\mathcal{D}[h(\omega)] = \frac{|B[h(\omega), 0^\circ]|^2}{\frac{1}{2} \int_0^{\pi} |B[h(\omega), \theta]|^2 \sin \theta d\theta}
\]

\[
= \frac{|h^H(\omega)d(\omega)|^2}{h^H(\omega)\Gamma_{0,\pi}(\omega)h(\omega)}
\]

\[\Gamma_d = \Gamma_{0,\pi}(\omega) = \frac{1}{2} \int_0^{\pi} d(\omega, \theta)d^H(\omega, \theta) \sin \theta d\theta\]
The Optimization problem

- Minimize some **white noise** + some **diffuse noise** energy:

\[
\min_h h^H(\omega) \left[ \epsilon_\psi I_M + \Gamma_{\psi,\pi}(\omega) \right] h(\omega)
\]

subject to \( h^H(\omega)d(\omega) = 1 \)

- \( \epsilon_\psi = \frac{1-\cos \psi}{2}, \quad 0 \leq \psi \leq \pi \)

- \( \Gamma_{\psi,\pi}(\omega) = \frac{1}{2} \int_{\psi}^{\pi} d(\omega, \theta)d^H(\omega, \theta) \sin \theta d\theta \)

- \( \psi \) controls the compromise between DF & WNG
The Optimization problem

- Minimize some **white noise** + some **diffuse noise** energy:

\[
\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega) [\epsilon_\psi \mathbf{I}_M + \Gamma_{\psi, \pi}(\omega)] \mathbf{h}(\omega)
\]

subject to  \( \mathbf{h}^H(\omega) \mathbf{d}(\omega) = 1 \)

- \( \epsilon_\psi = \frac{1 - \cos \psi}{2}, \quad 0 \leq \psi \leq \pi \)

- \( \Gamma_{\psi, \pi}(\omega) = \frac{1}{2} \int_{\psi}^{\pi} \mathbf{d}(\omega, \theta)\mathbf{d}^H(\omega, \theta) \sin \theta d\theta \)

- \( \psi \) controls the compromise between **DF & WNG**
The Tunable Beamformer

The Tunable beamformer:

\[ h_{T, \psi}(\omega) = \frac{[\epsilon_\psi I_M + \Gamma_{\psi, \pi}(\omega)]^{-1}d(\omega)}{d^H(\omega)[\epsilon_\psi I_M + \Gamma_{\psi, \pi}(\omega)]^{-1}d(\omega)} \]

- \( \psi \) determines the regularization factor
- \( 0 \leq \epsilon_\psi \leq 1 \)
Generalization

- Additional linear constraints
  - A null in the direction $\pi$: $h^H(\omega)d(\omega, \pi) = 0$

- The optimization problem:

$$\min_{h(\omega)} h^H(\omega) [\epsilon_\psi I_M + \Gamma_\psi,\pi(\omega)]h(\omega)$$

subject to

$$h^H(\omega)C(\omega) = i^T$$

$$C(\omega) = \begin{bmatrix} d(\omega) & d(\omega, \pi) \end{bmatrix},$$

$$i = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$$
Generalization

- Additional linear constraints
  - A null in the direction $\pi$: $h_h(\omega)^Hd(\omega, \pi) = 0$
  - The optimization problem:

$$\begin{align*}
\min_{h(\omega)} & \quad h_h(\omega)^H[\epsilon_\psi I_M + \Gamma_{\psi,\pi}(\omega)]h(\omega) \\
\text{subject to} & \quad h_h(\omega)^HC(\omega) = i^T \\
C(\omega) & = \begin{bmatrix} d(\omega) & d(\omega, \pi) \end{bmatrix}, \\
i & = \begin{bmatrix} 1 & 0 \end{bmatrix}^T
\end{align*}$$
The generalized beamformer:

$$h_{C, \psi}(\omega) = \Upsilon_{\psi}^{-1}(\omega)C(\omega)[C^H(\omega)\Upsilon_{\psi}^{-1}(\omega)C(\omega)]^{-1}i$$

$$\Upsilon_{\psi}(\omega) = \epsilon_{\psi}I_M + \Gamma_{\psi, \pi}(\omega)$$
The Tunable beamformer WNG & DF simulation

Array gains of $h_{T,\psi}(\omega)$ (solid line) with $\psi = 10^\circ$. $M = 8$ and $\delta = 1$ cm.
The generalized beamformer WNG & DF simulation

Array gains of $h_{C,\psi}(\omega)$ (solid line) with $\psi = 0.5^\circ$. $M = 8$ and $\delta = 1$ cm.
Beampattern Simulation

- The Tunable beamformer and the generalized beamformer beampattern $B[h(\omega), \theta]$
The Tunable Beamformer Simulation

Beampattern Simulation

Array Gains vs. $\psi$ Simulation

Array Gains vs. $\psi$ Simulation

The Tunable beamformer WNG and DF vs. $\psi$

(a) WNG

(b) DF

(a)-(b) Gain curves of $h_{T,\psi}(\omega)$ versus $\psi$, with $M = 8$ microphones and $\delta = 1$ cm, at selected frequencies.
Based on:

Motivation

- Reverberation
  - Late & uncorrelated reflections
- Speech communication systems
  - Distributed microphones environment
  - Quality monitoring
    - Select the best-quality channel
    - Pick-up least reverberant signal
- Quality measurement of reverberated speech
Motivation

- Reverberation
  - Late & uncorrelated reflections
- Speech communication systems
  - Distributed microphones environment
  - Quality monitoring
    - Select the best-quality channel
    - Pick-up least reverberant signal
- Quality measurement of reverberated speech
Motivation

- Reverberation
  - Late & uncorrelated reflections
- Speech communication systems
  - Distributed microphones environment
  - Quality monitoring
    - Select the best-quality channel
    - Pick-up least reverberant signal
- Quality measurement of reverberated speech
Related Work

- Reverberation measurement methods:
  - Quantifying system characteristics
    - Direct-to-reverberation energy ratio [Kuttruff, 2009]
      \[
      \text{DRR} = \frac{E_d}{E_r} = \frac{\int_{0}^{T_d} h^2(\tau) d\tau}{\int_{T_d}^{\infty} h^2(\tau) d\tau}
      \]
  - Comparison to clean reference [Naylor and Habets, 2010]
  - Direct signal-based evaluation
    - Signal energy
    - Signal-to-noise ratio (SNR) [Wolf and Nadeu, 2014]
Reverberation measurement methods:

- Quantifying system characteristics
  - Direct-to-reverberation energy ratio [Kuttruff, 2009]
    \[ \text{DRR} = \frac{E_d}{E_r} = \frac{\int_0^{T_d} h^2(\tau) d\tau}{\int_{T_d}^{\infty} h^2(\tau) d\tau} \]

- Comparison to clean reference [Naylor and Habets, 2010]
- Direct signal-based evaluation
  - Signal energy
  - Signal-to-noise ratio (SNR) [Wolf and Nadeu, 2014]
Reverberation measurement methods:
  - Quantifying system characteristics
    - Direct-to-reverberation energy ratio [Kuttruff, 2009]
      \[
      DRR = \frac{E_d}{E_r} = \frac{\int_0^{T_d} h^2(\tau) d\tau}{\int_{T_d}^{\infty} h^2(\tau) d\tau}
      \]
  - Comparison to clean reference [Naylor and Habets, 2010]
  - Direct signal-based evaluation
    - Signal energy
    - Signal-to-noise ratio (SNR) [Wolf and Nadeu, 2014]
Related Work – Contd.

- **Signal-based quality measures:**
  - Signal-to-diffuse ratio estimation
    - Spatial complex coherence between microphones [Jeub et al., 2011]
    - Direct & diffuse part segregation using beamforming [Thiergart et al., 2014, Hioka et al., 2012]
  - Modulation spectral analysis:
    - Speech to reverberation modulation energy ratio (SRMR) [Falk et al., 2010]
  - Low correlation of signal-based measures with subjective listening tests [Goetze et al., 2010]
  - High correlation of the *Clarity* measure (C50) [Kuttruff, 2009]
Related Work – Contd.

- **Signal-based quality measures:**
  - Signal-to-diffuse ratio estimation
    - Spatial complex coherence between microphones [Jeub et al., 2011]
    - Direct & diffuse part segregation using beamforming [Thiergart et al., 2014, Hioka et al., 2012]
  - Modulation spectral analysis:
    - Speech to reverberation modulation energy ratio (SRMR) [Falk et al., 2010]
  - Low correlation of signal-based measures with subjective listening tests [Goetze et al., 2010]
  - High correlation of the Clarity measure (C50) [Kuttruff, 2009]
Related Work – Contd.

- **Signal-based quality measures:**
  - Signal-to-diffuse ratio estimation
    - Spatial complex coherence between microphones [Jeub et al., 2011]
    - Direct & diffuse part segregation using **beamforming** [Thiergart et al., 2014, Hioka et al., 2012]
  - Modulation spectral analysis:
    - Speech to reverberation modulation energy ratio (SRMR) [Falk et al., 2010]
  - Low correlation of **signal-based** measures with subjective listening tests [Goetze et al., 2010]
    - High correlation of the **Clarity** measure (C50) [Kuttruff, 2009]
Related Work – Contd.

- The optimal solution:
  - Simple & reliable measure
  - Estimate quality & *reverberation* level
  - Signal-based
  - Correlative to objective parameters
Problem Formulation

- The measured signal:
  \[ z(t) = \int_{-\infty}^{t} s(\tau) h(t - \tau) d\tau + v(t), \]
  - \( s(t) \) anechoic speech signal
  - \( h(t) \) causal room impulse response (RIR)
  - \( v(t) = 0 \) ambient noise

- Reverberated RIR model:
  \[ h(t) = \begin{cases} 
  h_d(t), & \text{for } 0 \leq t < T_r \\
  h_r(t), & \text{for } t \geq T_r \\
  0, & \text{otherwise} 
\end{cases} \]

- The objective:
  - Estimate the reverberation (quality & DRR)
Problem Formulation

- The measured signal:

\[ z(t) = \int_{-\infty}^{t} s(\tau)h(t - \tau)d\tau + v(t), \]

- \( s(t) \) anechoic speech signal
- \( h(t) \) causal room impulse response (RIR)
- \( v(t) = 0 \) ambient noise

- Reverberated RIR model:

\[
h(t) = \begin{cases} 
  h_d(t), & \text{for } 0 \leq t < T_r \\
  h_r(t), & \text{for } t \geq T_r \\
  0, & \text{otherwise},
\end{cases}
\]

- The objective:
  - Estimate the reverberation (quality & DRR)
Problem Formulation

- The measured signal:

\[ z(t) = \int_{-\infty}^{t} s(\tau) h(t - \tau) d\tau + v(t), \]

- \( s(t) \) anechoic speech signal
- \( h(t) \) causal room impulse response (RIR)
- \( v(t) = 0 \) ambient noise

- Reverberated RIR model:

\[ h(t) = \begin{cases} 
  h_d(t), & \text{for } 0 \leq t < T_r \\
  h_r(t), & \text{for } t \geq T_r \\
  0, & \text{otherwise}, 
\end{cases} \]

- The objective:
  - Estimate the reverberation (quality & DRR)
Signal Model

- Statistical Room Acoustics (SRA) model [Polack, 1988]:
  \[ h_d(t) = \begin{cases} 
  b_d(t)e^{-\delta t}, & \text{for } 0 \leq t < T_r \\
  0, & \text{otherwise},
  \end{cases} \]
  \[ b_d(t) \sim N(0, \sigma_d^2) \]
  \[ \delta = \frac{3 \ln 10}{T_{60}} \]
  \[ h_r(t) = \begin{cases} 
  b_r(t)e^{-\delta t}, & \text{for } t \geq T_r \\
  0, & \text{otherwise},
  \end{cases} \]
  \[ b_r(t) \sim N(0, \sigma_r^2) \]

⇒ The measured signal energy:
  \[ \mathbb{E}_z\{z^2(t)\} = \mathbb{E}_z\{z_d^2(t)\} + \mathbb{E}_z\{z_r^2(t)\} \]

- \[ \lambda_s(t) = \mathbb{E}_s\{s^2(t)\} \]
- \[ \mathbb{E}_z\{z_d^2(t)\} = f(\lambda_s(t), \sigma_d^2, T_r) \]
- \[ \mathbb{E}_z\{z_r^2(t)\} = f(\lambda_s(t), \sigma_r^2, T_r) \]
Signal Model

- Statistical Room Acoustics (SRA) model [Polack, 1988]:
  \[
  h_d(t) = \begin{cases} 
  b_d(t)e^{-\delta t}, & \text{for } 0 \leq t < T_r \\
  0 & \text{otherwise,}
  \end{cases}
  \]
  \[
  h_r(t) = \begin{cases} 
  b_r(t)e^{-\delta t}, & \text{for } t \geq T_r \\
  0 & \text{otherwise,}
  \end{cases}
  \]
  \[
  b_d(t) \sim \mathcal{N}(0, \sigma_d^2)
  \]
  \[
  \delta = \frac{3 \ln 10}{T_{60}}
  \]
  \[
  \lambda_s(t) = \mathbb{E}_s\{s^2(t)\}
  \]
  \[
  \mathbb{E}_z\{z^2_d(t)\} = f(\lambda_s(t), \sigma_d^2, T_r)
  \]
  \[
  \mathbb{E}_z\{z^2_r(t)\} = f(\lambda_s(t), \sigma_r^2, T_r)
  \]

⇒ The measured signal energy:

\[
\mathbb{E}_z\{z^2(t)\} = \mathbb{E}_z\{z^2_d(t)\} + \mathbb{E}_z\{z^2_r(t)\}
\]
Signal Model

- Statistical Room Acoustics (SRA) model [Polack, 1988]:
  
  $$h_d(t) = \begin{cases} b_d(t)e^{-\delta t}, & \text{for } 0 \leq t < T_r \\ 0 & \text{otherwise,} \end{cases}$$

  $$\delta = \frac{3 \ln 10}{T_{60}}$$

  $$b_d(t) \sim \mathcal{N}(0, \sigma_d^2)$$

  $$h_r(t) = \begin{cases} b_r(t)e^{-\delta t}, & \text{for } t \geq T_r \\ 0 & \text{otherwise,} \end{cases}$$

  $$b_r(t) \sim \mathcal{N}(0, \sigma_r^2)$$

- The measured signal energy:
  
  $$\mathbb{E}_z\{z^2(t)\} = \mathbb{E}_z\{z^2_d(t)\} + \mathbb{E}_z\{z^2_r(t)\}$$

- $$\lambda_s(t) = \mathbb{E}_s\{s^2(t)\}$$

- $$\mathbb{E}_z\{z^2_d(t)\} = f(\lambda_s(t), \sigma_d^2, T_r)$$

- $$\mathbb{E}_z\{z^2_r(t)\} = f(\lambda_s(t), \sigma_r^2, T_r)$$
Directional Array Response

- Expansion of the SRA model

- Unidirectional microphone array
  - Directional elements
  - Beamforming

- $g_{\text{dir/opp}}(\theta)$ - The microphone directional gain at angle $\theta$
Directional Array Response

- Expansion of the SRA model

- **Unidirectional** microphone array
  - Directional elements
  - Beamforming

- $g_{\text{dir/opp}}(\theta)$ - The microphone directional gain at angle $\theta$
Expansion of the SRA model

Unidirectional microphone array
  - Directional elements
  - Beamforming

$g_{\text{dir/opp}}(\theta)$ - The microphone directional gain at angle $\theta$
⇒ The direct microphone signal energy:

\[
\mathbb{E}_z \{[z_{\text{dir}}(t)]^2\} = [g_{\text{dir}}(\theta)]^2 \cdot \mathbb{E}_z \{z_d^2(t)\} \\
+ \frac{1}{\Omega} \int_\Omega [g_{\text{dir}}(\theta')]^2 d\theta' \cdot \mathbb{E}_z \{z_r^2(t)\}
\]

⇒ The opposite microphone signal energy:

\[
\mathbb{E}_z \{[z_{\text{opp}}(t)]^2\} = \frac{1}{\Omega} \int_\Omega [g_{\text{opp}}(\theta')]^2 d\theta' \cdot \mathbb{E}_z \{z_r^2(t)\}
\]
⇒ The direct microphone signal energy:

\[
\mathbb{E}_z \{ [Z^{\text{dir}}(t)]^2 \} = [g^{\text{dir}}(\theta)]^2 \cdot \mathbb{E}_z \{ z_d^2(t) \} \\
+ \frac{1}{\Omega} \int_{\Omega} [g^{\text{dir}}(\theta')]^2 d\theta' \cdot \mathbb{E}_z \{ z_d^2(t) \}
\]

⇒ The opposite microphone signal energy:

\[
\mathbb{E}_z \{ [Z^{\text{opp}}(t)]^2 \} = \frac{1}{\Omega} \int_{\Omega} [g^{\text{opp}}(\theta')]^2 d\theta' \cdot \mathbb{E}_z \{ z_d^2(t) \}
\]
The Proposed Measure

Assuming the microphones are calibrated:

\[ \bar{g}^2 = \frac{1}{\Omega} \int_{\Omega} [g_{dir}(\theta')]^2 d\theta' = \frac{1}{\Omega} \int_{\Omega} [g_{opp}(\theta')]^2 d\theta' \]

The Power Ratio between the direct & opposite microphones:

\[
\frac{\mathbb{E}_Z \{ [z_{dir}(t)]^2 \}}{\mathbb{E}_Z \{ [z_{opp}(t)]^2 \}} = \frac{[g_{dir}(\theta)]^2}{\bar{g}^2} \cdot \frac{\sigma_d^2}{\sigma_r^2} (e^{2\delta T_r} - 1) + 1
\]
The Proposed Measure

- Assuming the microphones are calibrated:
  - $\bar{g}^2 = \frac{1}{\Omega} \int_{\Omega} [g_{\text{dir}}(\theta')]^2 d\theta' = \frac{1}{\Omega} \int_{\Omega} [g_{\text{opp}}(\theta')]^2 d\theta'$
- The Power Ratio between the direct & opposite microphones:

$$\frac{\mathbb{E}_Z\{[z_{\text{dir}}(t)]^2\}}{\mathbb{E}_Z\{[z_{\text{opp}}(t)]^2\}} = \frac{[g_{\text{dir}}(\theta)]^2}{\bar{g}^2} \cdot \left[\frac{\sigma_d^2}{\sigma_r^2} (e^{2\delta T_r} - 1)\right] + 1$$
Replace $\mathbb{E}_z \{ \cdot \} \leftrightarrow$ temporal smoothing

$\Rightarrow$ The Directional Power Ratio quality measure:

$$PR(t) = \frac{P_{\text{dir}}(t)}{P_{\text{opp}}(t)} = \frac{\int_{t-T}^{t} [Z_{\text{dir}}(\tau)]^2 d\tau}{\int_{t-T}^{t} [Z_{\text{opp}}(\tau)]^2 d\tau} = \frac{[g_{\text{dir}}(\theta)]^2}{\bar{g}^2} \cdot \text{DRR}(t) + 1$$
Measuring the power ratio of a speech signal. source-microphone distance = 2 m, $T_{60} = 1$ sec.

\[
\begin{array}{cccc}
\text{opp} & \text{dir} \\
0 & 0 & 0.05 & 0.05 \\
0.5 & 1 & 1.5 & 2 \text{ sec} \\
2.5 & 3 & 3.5 & 4
\end{array}
\]

(a) The measured signal $z(t)$
The Proposed Measure - Contd.

Measuring the power ratio of a speech signal. source-microphone distance = 2 m, $T_{60} = 1$ sec.

\[ P(t) \]

(b) The directional measured power $P(t)$. 

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure}
\caption{The directional measured power $P(t)$.}
\end{figure}
The Proposed Measure - Contd.

Measuring the power ratio of a speech signal. Source-microphone distance = 2 m, $T_{60} = 1$ sec.

\[ \text{PR} [\text{dB}] \]

![Graph](image)

(c) The measured power ratio $\text{PR}(t)$ (solid line) and $\frac{g_{\text{dir}}(\theta)^2}{g^2} \cdot \text{DRR} + 1$ (dashed-dotted line).
The Proposed Measure - Contd.

- Blind DRR estimator - $\hat{\text{DRR}}(t)$:

\[
\text{PR-DRR}(t) = \frac{\bar{g}^2}{[g_{\text{dir}}(\theta)]^2} \cdot \left( \frac{P_{\text{dir}}(t)}{P_{\text{opp}}(t)} - 1 \right)
\]

\[
= \frac{\bar{g}^2}{[g_{\text{dir}}(\theta)]^2} \cdot \frac{P_{\text{dir}}(t) - P_{\text{opp}}(t)}{P_{\text{opp}}(t)}
\]
Simulations

Experiment Conditions

- Performed experiments:
  - Increasing source-microphone distance with fixed $T_{60}$
  - Increasing $T_{60}$ with fixed source-microphone distance

- Simulation environment:
Simulations

Experiment Conditions

- Performed experiments:
  - Increasing source-microphone distance with fixed $T_{60}$
  - Increasing $T_{60}$ with fixed source-microphone distance

- Simulation environment:
Quality Measure Performance

- Reference quality measures
  - SRMR [Falk et al., 2010]
  - Envelope Variance (EV) [Wolf and Nadeu, 2014]

- Correlation coefficients with:
  - Clarity (C50) [Kuttruff, 2009]
  - ITU-T P.563 [Recommendation, 2004]
  - ITU-T P.862 (PESQ) [Recommendation, 2001]

<table>
<thead>
<tr>
<th>Input type</th>
<th>White noise</th>
<th>Speech signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test type</td>
<td>Algorithm</td>
<td>Correlation ref.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C50</td>
</tr>
<tr>
<td>$T_{60} = 0.3$ sec,</td>
<td>PR</td>
<td>0.999</td>
</tr>
<tr>
<td>Vs. increasing distance</td>
<td>SRMR</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>EV</td>
<td>-0.66</td>
</tr>
<tr>
<td>distance = 0.5 m,</td>
<td>PR</td>
<td>0.944</td>
</tr>
<tr>
<td>Vs. increasing $T_{60}$</td>
<td>SRMR</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td>EV</td>
<td>0.235</td>
</tr>
</tbody>
</table>
Quality Measure Performance

- Reference quality measures
  - SRMR [Falk et al., 2010]
  - Envelope Variance (EV) [Wolf and Nadeu, 2014]
- Correlation coefficients with:
  - Clarity (C50) [Kuttruff, 2009]
  - ITU-T P.563 [Recommendation, 2004]
  - ITU-T P.862 (PESQ) [Recommendation, 2001]

<table>
<thead>
<tr>
<th>Input type</th>
<th>White noise</th>
<th>Speech signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test type</td>
<td>Algorithm</td>
<td>Correlation ref.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C50</td>
</tr>
<tr>
<td>$T_{60} = 0.3$ sec, Vs. increasing distance</td>
<td>PR</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>SRMR</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>EV</td>
<td>-0.66</td>
</tr>
<tr>
<td>distance = 0.5 m, Vs. increasing $T_{60}$</td>
<td>PR</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>SRMR</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td>EV</td>
<td>0.235</td>
</tr>
</tbody>
</table>
Quality Measure Performance

- Reference quality measures
  - SRMR [Falk et al., 2010]
  - Envelope Variance (EV) [Wolf and Nadeu, 2014]
- Correlation coefficients with:
  - Clarity (C50) [Kuttruff, 2009]
  - ITU-T P.563 [Recommendation, 2004]
  - ITU-T P.862 (PESQ) [Recommendation, 2001]

<table>
<thead>
<tr>
<th>Input type</th>
<th>White noise</th>
<th>Speech signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test type</td>
<td>Correlation ref.</td>
<td>Correlation ref.</td>
</tr>
<tr>
<td></td>
<td>C50</td>
<td>C50</td>
</tr>
<tr>
<td>$T_{60} = 0.3$ sec, Vs. increasing distance</td>
<td>PR</td>
<td>0.999</td>
</tr>
<tr>
<td>distance = 0.5 m, Vs. increasing $T_{60}$</td>
<td>PR</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>SRMR</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>EV</td>
<td>-0.66</td>
</tr>
<tr>
<td></td>
<td>SRMR</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td>EV</td>
<td>0.235</td>
</tr>
</tbody>
</table>
Blind DRR Estimator Performance

- Reference DRR measure
  - CDR-based DRR [Jeub et al., 2011]
- Correlation coefficient with:
  - DRR

<table>
<thead>
<tr>
<th>Test type</th>
<th>Algorithm</th>
<th>White noise</th>
<th>Speech signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{60} = 1$ sec, vs. increasing distance</td>
<td>PR-DRR</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>distance $= 2$ m, vs. increasing $T_{60}$</td>
<td>CDR</td>
<td>0.964</td>
<td>0.972</td>
</tr>
<tr>
<td></td>
<td>PR-DRR</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>CDR</td>
<td>0.852</td>
<td>0.913</td>
</tr>
</tbody>
</table>
Blind DRR Estimator Performance

- Reference DRR measure
  - CDR-based DRR [Jeub et al., 2011]
- Correlation coefficient with:
  - DRR

<table>
<thead>
<tr>
<th>Input type</th>
<th>White noise</th>
<th>Speech signals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test type</strong></td>
<td><strong>Algorithm</strong></td>
<td><strong>Correlation ref. DRR</strong></td>
</tr>
<tr>
<td>$T_{60} = 1$ sec, Vs. increasing distance</td>
<td>PR-DRR</td>
<td>0.999</td>
</tr>
<tr>
<td>distance $= 2$ m, Vs. increasing $T_{60}$</td>
<td>CDR</td>
<td>0.964</td>
</tr>
<tr>
<td></td>
<td>PR-DRR</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>CDR</td>
<td>0.852</td>
</tr>
</tbody>
</table>
Blind DRR Estimator Performance

- Reference DRR measure
  - CDR-based DRR [Jeub et al., 2011]
- Correlation coefficient with:
  - DRR

<table>
<thead>
<tr>
<th>Input type</th>
<th>White noise</th>
<th>Speech signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test type</td>
<td>Algorithm</td>
<td>Correlation ref. DRR</td>
</tr>
<tr>
<td>$T_{60} = 1$ sec, Vs. increasing distance</td>
<td>PR-DRR</td>
<td>0.999</td>
</tr>
<tr>
<td>distance = 2 m, Vs. increasing $T_{60}$</td>
<td>CDR</td>
<td>0.964</td>
</tr>
<tr>
<td></td>
<td>PR-DRR</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>CDR</td>
<td>0.852</td>
</tr>
</tbody>
</table>
The proposed DRR estimate vs. distance. PR-DRR [dB] (solid-circled line), and the true DRR [dB] (dashed-line), vs. source-microphone distance, with fixed $T_{60} = 0.3$ sec.
Performance of the DRR estimate vs. SNR. Absolute difference of the proposed DRR estimate PR-DRR [dB] (solid-circled line), and of Jeub et al. CDR-based DRR estimate [dB] (dashed-asterisk line), vs. SNR [dB]. $T_{60} = 0.3$ sec and source-microphone distance $= 0.5$ m.
The proposed DRR estimate vs. $T_{60}$ - Off main-lobe. PR-DRR [dB] (solid-circled line), and the true DRR [dB] (dashed line), vs. $T_{60}$. (source–receiver angle $\in [-30^\circ .. +30^\circ]$, source-microphone distance $= 2$ m)
Recorded speech PR measure vs. source location. The measured PR of all microphone arrays (1 – 6) vs. the source position (hall of size $15 \times 10 \times 6$ m, with 3 m spacing between adjacent arrays)

![Graph showing PR measure vs. source location.](image-url)
Conclusions
New approach to robust regularized beamforming

- Combination of conventional and regularized beamformers
- Achieve robust distortionless high-gain solution
- Filter parameters determine the frequency response
- Control of the WNG–DF tradeoff

Derived closed-form beamformers with user-determined WNG or DF

- A unique form of the FIBP
Conclusion

Robust Regularized Superdirective Beamforming

- New approach to robust regularized beamforming
  - Combination of conventional and regularized beamformers
  - Achieve robust distortionless high-gain solution
  - Filter parameters determine the frequency response
  - Control of the WNG–DF tradeoff

- Derived closed-form beamformers with user-determined WNG or DF
  - A unique form of the FIBP
Conclusion
The Tunable Superdirective Beamformer

- Novel method to control the white noise & diffuse noise energy minimization
- New perspective for beamforming and angular noise field analysis
- Tunable regularization parameter
  - Closed-form expression for the regularization parameter $0 \leq \epsilon_\psi \leq 1$
- Generalized multiple linear constraints beamformer
Conclusion

Microphone Array Reverberation Assessment

- New approach to **reverberation ratio estimation**
  - Expansion of a well-founded statistical model, using directional microphone arrays
  - **Signal-based** acoustic quality measure
  - Blind DRR estimator

- Robust performance
  - Correlative to state-of-the-art quality & DRR measures
  - Robust against additive noise
Future research

- **Regularized Superdirective Beamforming:**
  - Different noise fields
  - Tailored frequency-dependent regularization
  - Broadside beamforming, differential microphones, etc.

- **The Tunable Beamformer:**
  - Linking the tunable regularization to the WNG & DF
  - Noise field angular analysis

- **Reverberation assessment:**
  - Microphone beampattern analysis
  - Adapting temporal smoothing (e.g. to the voice activity level)
  - Integration in real-time quality monitoring systems
Future research

- Regularized Superdirective Beamforming:
  - Different noise fields
  - Tailored frequency-dependent regularization
  - Broadside beamforming, differential microphones, etc.

- The Tunable Beamformer:
  - Linking the tunable regularization to the WNG & DF
  - Noise field angular analysis

- Reverberation assessment:
  - Microphone beampattern analysis
  - Adapting temporal smoothing (e.g. to the voice activity level)
  - Integration in real-time quality monitoring systems
Future research

- Regularized Superdirective Beamforming:
  - Different noise fields
  - Tailored frequency-dependent regularization
  - Broadside beamforming, differential microphones, etc.

- The Tunable Beamformer:
  - Linking the tunable regularization to the WNG & DF
  - Noise field angular analysis

- Reverberation assessment:
  - Microphone beampattern analysis
  - Adapting temporal smoothing (e.g. to the voice activity level)
  - Integration in real-time quality monitoring systems
Thank You!

Questions?

- Contact: berkun@tx.technion.ac.il
- Prof. Israel Cohen: http://webee.technion.ac.il/people/IsraelCohen/
- Prof. Jacob Benesty: http://externe.emt.inrs.ca/users/benesty/
*Microphone array signal processing.*  
Springer-Verlag, Berlin, Germany.

Robust adaptive beamforming.  

Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics.  

Superdirective beamforming robust against microphone mismatch.  

A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech.  

Quality assessment for listening-room compensation algorithms.  
In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on,* pages 2450–2453. IEEE.

Estimation of direct-to-reverberation energy ratio based on isotropic and homogeneous propagation model.  
For Further Reading II

Blind estimation of the coherent-to-diffuse energy ratio from noisy speech signals.
In 19th European Signal Processing Conference (EUSIPCO 2011), pages 1347–1351.

Room acoustics.
Taylor & Francis Press, New York, USA.

Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays.

Models, measurement and evaluation.

La transmission de l'énergie sonore dans les salles.

P. 862, perceptual evaluation of speech quality (pesq): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs.
International Telecommunication Union, 23.
*International Telecommunication Union, Geneva.*

Power-based signal-to-diffuse ratio estimation using noisy directional microphones.
*In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, pages 7440–7444. IEEE.*

Design of frequency-invariant robust beam patterns by the oversteering of end-fire arrays.

Optimum array processing (detection, estimation, and modulation theory, part iv).
*John Wiley and Sons Inc.*

Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem.

Channel selection measures for multi-microphone speech recognition.
*Speech Communication, 57:170–180.*
The Delay & Sum (DS) Beamformer

- Maximizes the WNG
  - subject to distortionless constraint

\[ h_{DS}(\omega) = \frac{d(\omega)}{d^H(\omega)d(\omega)} = \frac{d(\omega)}{M} \]

\[ \mathcal{W}[h_{DS}(\omega)] = M = \mathcal{W}_{\text{max}} \]

\[ \mathcal{D}[h_{DS}(\omega)] = \frac{M^2}{d^H(\omega)\Gamma_d(\omega)d(\omega)} \geq 1 \]
The Delay & Sum (DS) Beamformer

- Maximizes the WNG
  - subject to distortionless constraint

\[
h_{DS}(\omega) = \frac{d(\omega)}{d^H(\omega)d(\omega)} = \frac{d(\omega)}{M}
\]

- \( \mathcal{W}[h_{DS}(\omega)] = M = \mathcal{W}_{\text{max}} \)
- \( \mathcal{D}[h_{DS}(\omega)] = \frac{M^2}{d^H(\omega)\Gamma_d(\omega)d(\omega)} \geq 1 \)
The Superdirective Beamformer

- Maximizes the DF
  - subject to distortionless constraint
  \[ h_S(\omega) = \frac{\Gamma_d^{-1}(\omega)d(\omega)}{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)} \]

- \( \mathcal{W}[h_S(\omega)] = \mathcal{W}_{\text{max}} \cos^2 \varphi(\omega) \)

- \( \cos \varphi(\omega) = \cos [d(\omega), \Gamma_d^{-1}(\omega)d(\omega)] = \frac{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)}{\sqrt{d^H(\omega)d(\omega)}\sqrt{d^H(\omega)\Gamma_d^{-2}(\omega)d(\omega)}} \)

- \( \mathcal{D}[h_S(\omega)] = d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega) = D_{\text{max}}(\omega) \)
The Superdirective Beamformer

- Maximizes the DF
  - subject to distortionless constraint
  \[
  h_S(\omega) = \frac{\Gamma_d^{-1}(\omega)d(\omega)}{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)}
  \]
- \( \mathcal{W}[h_S(\omega)] = \mathcal{W}_{\text{max}} \cos^2 \varphi(\omega) \)
- \( \cos \varphi(\omega) = \cos \left[ d(\omega), \Gamma_d^{-1}(\omega)d(\omega) \right] = \frac{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)}{\sqrt{d^H(\omega)d(\omega)}\sqrt{d^H(\omega)\Gamma_d^{-2}(\omega)d(\omega)}} \)
- \( \mathcal{D}[h_S(\omega)] = d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega) = \mathcal{D}_{\text{max}}(\omega) \)
The Superdirective Beam Former

- Maximizes the DF
  - subject to distortionless constraint
    \[ h_S(\omega) = \frac{\Gamma_d^{-1}(\omega)d(\omega)}{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)} \]

- \[ \mathcal{W}[h_S(\omega)] = \mathcal{W}_{\text{max}} \cos^2 \varphi(\omega) \]
  - \[ \cos \varphi(\omega) = \cos [d(\omega), \Gamma_d^{-1}(\omega)d(\omega)] = \frac{d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega)}{\sqrt{d^H(\omega)d(\omega)}\sqrt{d^H(\omega)\Gamma_d^{-2}(\omega)d(\omega)}} \]
  - \[ \mathcal{D}[h_S(\omega)] = d^H(\omega)\Gamma_d^{-1}(\omega)d(\omega) = \mathcal{D}_{\text{max}}(\omega) \]
The Regularized Superdirective Beamformer

- Maximizes the DF [Cox et al., 1987, Benesty et al., 2008]
  - Constraint on the WNG
  - subject to distortionless constraint

\[ h_{S,\epsilon}(\omega) = \frac{[\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)}{d^H(\omega) [\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)} \]

- Define: \( \Gamma_\epsilon(\omega) = \Gamma_d(\omega) + \epsilon I_M \)
- \( \mathcal{W}[h_{S,\epsilon}(\omega)] = \mathcal{W}_{\text{max}} \cos^2 \varphi_\epsilon(\omega) \)
- \( \cos \varphi_\epsilon(\omega) = \cos [d(\omega), \Gamma_\epsilon^{-1}(\omega)d(\omega)] = \frac{d^H(\omega)\Gamma_\epsilon^{-1}(\omega)d(\omega)}{\sqrt{d^H(\omega)d(\omega)} \sqrt{d^H(\omega)\Gamma_\epsilon^{-2}(\omega)d(\omega)}} \)
- \( \mathcal{D}[h_{S,\epsilon}(\omega)] = \frac{[d^H(\omega)\Gamma_\epsilon^{-1}(\omega)d(\omega)]^2}{d^H(\omega)\Gamma_\epsilon^{-1}(\omega)d(\omega)\Gamma_d(\omega)\Gamma_\epsilon^{-1}(\omega)d(\omega)} \)
The Regularized Superdirective Beamformer

- Maximizes the DF [Cox et al., 1987, Benesty et al., 2008]
  - Constraint on the WNG
  - subject to distortionless constraint

\[
h_{S,\epsilon}(\omega) = \frac{[\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)}{d^H(\omega) [\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)}
\]

- Define: \( \Gamma_{\epsilon}(\omega) = \Gamma_d(\omega) + \epsilon I_M \)

- \( \mathcal{W}[h_{S,\epsilon}(\omega)] = \mathcal{W}_{\text{max}} \cos^2 \varphi_{\epsilon}(\omega) \)
- \( \cos \varphi_{\epsilon}(\omega) = \cos \left[ d(\omega), \Gamma_{\epsilon}^{-1}(\omega)d(\omega) \right] = \frac{d^H(\omega)\Gamma_{\epsilon}^{-1}(\omega)d(\omega)}{\sqrt{d^H(\omega)d(\omega)}\sqrt{d^H(\omega)\Gamma_{\epsilon}^{-2}(\omega)d(\omega)}} \)

- \( \mathcal{D}[h_{S,\epsilon}(\omega)] = \frac{[d^H(\omega)\Gamma_{\epsilon}^{-1}(\omega)d(\omega)]^2}{d^H(\omega)\Gamma_{\epsilon}^{-1}(\omega)d(\omega)\Gamma_d(\omega)\Gamma_{\epsilon}^{-1}(\omega)d(\omega)} \)
The Regularized Superdirective Beamformer

- Maximizes the DF [Cox et al., 1987, Benesty et al., 2008]
  - Constraint on the WNG
  - subject to distortionless constraint

\[
h_{S,\epsilon}(\omega) = \frac{[\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)}{d^H(\omega) [\Gamma_d(\omega) + \epsilon I_M]^{-1} d(\omega)}
\]

- Define: \( \Gamma_\epsilon(\omega) = \Gamma_d(\omega) + \epsilon I_M \)
- \( \mathcal{W}[h_{S,\epsilon}(\omega)] = \mathcal{W}_{\text{max}} \cos^2 \varphi_\epsilon(\omega) \)
  - \( \cos \varphi_\epsilon(\omega) = \cos \left[ d(\omega), \Gamma_\epsilon^{-1}(\omega) d(\omega) \right] = \frac{d^H(\omega) \Gamma_\epsilon^{-1}(\omega) d(\omega)}{\sqrt{d^H(\omega) d(\omega)} \sqrt{d^H(\omega) \Gamma_\epsilon^{-2}(\omega) d(\omega)}} \)
- \( \mathcal{D}[h_{S,\epsilon}(\omega)] = \frac{[d^H(\omega) \Gamma_\epsilon^{-1}(\omega) d(\omega)]^2}{d^H(\omega) \Gamma_\epsilon^{-1}(\omega) \Gamma_d(\omega) \Gamma_\epsilon^{-1}(\omega) d(\omega)} \)