Rate-Distributed Binaural LCMV Beamforming for Assistive Hearing in Wireless Acoustic Sensor Networks

Jie Zhang, Richard Heusdens, and Richard C. Hendriks

Circuits and Systems (CAS) Group, Delft University of Technology, the Netherlands

July 5, 2018
Introduction

Problem statement:

\[
\begin{align*}
\text{minimise} & \quad \text{total transmission costs} \\
\text{subject to} & \quad \text{noise reduction performance} \\
& \quad \text{spatial cue preservation}
\end{align*}
\]

General binaural HAs configuration: HAs are a part of a bigger wireless acoustic sensor network.
Motivation

• Energy efficiency is essential in the design of algorithms in WASNs since usually each sensor has a limited energy budget (e.g., battery).

• In general, there are two ways to reduce energy costs:
  - Microphone subset selection
  - Bit-rate allocation

• Rate allocation is more general than sensor selection, i.e., sensor selection (hard/binary decision) can be seen as a special case of rate allocation (soft/multiple decision).
Fundamentals

• Signal model in STFT domain:

$$\hat{y} = y + q \in \mathbb{C}^M,$$

where

$$y = x + z + v$$

$$= \sum_{i=1}^{I} a_i s_i + \sum_{j=1}^{J} h_j u_j + v$$

$$= A s + H u + v,$$

where $a_i = [a_{i1}, a_{i2}, \ldots, a_{iM}]^T$, $h_j = [h_{j1}, h_{j2}, \ldots, h_{jM}]^T$, $A = [a_1, \ldots, a_I]$, $s = [s_1, \ldots, s_I]^T$, $H = [h_1, \ldots, h_J] \in \mathbb{C}^{M \times J}$, $u = [u_1, \ldots, u_J]^T$. 
Fundamentals

- **Second-order statistics:**

\[
R_{yy} = \mathbb{E}\{yy^H\} = R_{xx} + \underbrace{R_{zz} + R_{vv}}_{R_{nn}} \in \mathbb{C}^{M \times M},
\]

where \(R_{xx} = \sum_{i=1}^{I} \mathbb{E}\{x_i x_i^H\}\) and \(R_{zz} = \sum_{j=1}^{J} \mathbb{E}\{z_i z_i^H\}\).

- **Assumption:** the target sources, interfering sources and quantisation noise are mutually uncorrelated, such that

\[
R_{n+q} = R_{nn} + R_{qq}.
\]

- **Uniform quantisation:**

\[
R_{qq} = \frac{1}{12} \text{diag} \left( \left[ \frac{A_1^2}{4b_1}, \frac{A_2^2}{4b_2}, \ldots, \frac{A_M^2}{4b_M} \right] \right),
\]

where \(A_k = \max\{|y_k|\}\) and \(b_k, \forall k\) denotes the rate for the \(k\)th sensor node.
Fundamentals

- **Transmission energy model:**
  - SNR over communication channels: \( \text{SNR}_k = d_k^{-2} \frac{E_k}{V_k} \).
  - Channel capacity:
    \[
    b_k = \frac{1}{2} \log_2 (1 + \text{SNR}_k).
    \]
  - Transmission energy:
    \[
    E_k = d_k^2 V_k (4^{b_k} - 1),
    \]
    which holds under two conditions: 1) for spectrum-limited applications; 2) quantize at the channel capacity (i.e., the upper bound).
Binaural LCMV beamforming

- **Binaural cue preservation:** $w_{BLCMV} = [w^T_L \ w^T_R]^T$
  
  - Preserving target sources:
    $$\begin{align*}
    w^H_L a_i &= a_i L \\
    w^H_R a_i &= a_i R
    \end{align*}\quad \implies \quad ITF_{in}^{x_i} = ITF_{out}^{x_i} \quad \implies \quad \Lambda_1^H w = f_1$$

  - Preserving interfering sources:
    $$\begin{align*}
    ITF_{in}^{n_j} &= ITF_{out}^{n_j} \quad \implies \quad \frac{h_j L}{h_j R} = \frac{w^H_L h_j}{w^H_R h_j}, \forall j, \\
    &\implies \quad w^H_L h_j h_j R - w^H_R h_j h_j L = 0 \quad \implies \quad \Lambda_2^H w = f_2
    \end{align*}$$

- Binaural cues: \(ILD = |ITF|^2, \quad IPD = \angle ITF.\)
Binaural LCMV beamforming

- **Binaural LCMV (BLCMV) beamforming** for joint noise reduction and spatial cue preservation can be formulated as

  \[
  \hat{\mathbf{w}}_{\text{BLCMV}} = \arg \min_{\mathbf{w}} \mathbf{w}^H \tilde{\mathbf{R}}_{n+q} \mathbf{w}, \quad \text{s.t.} \quad \Lambda^H \mathbf{w} = \tilde{\mathbf{f}},
  \]

  where

  \[
  \tilde{\mathbf{R}}_{n+q} = \begin{bmatrix} \mathbf{R}_{n+q} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{n+q} \end{bmatrix} \in \mathbb{C}^{2M \times 2M},
  \]

  \[
  \Lambda = \begin{bmatrix} \Lambda_1 & \Lambda_2 \end{bmatrix} \in \mathbb{C}^{2M \times (2I+J)},
  \]

  \[
  \tilde{\mathbf{f}} = \begin{bmatrix} \mathbf{f}_1^H \\ \mathbf{f}_2^H \end{bmatrix}^T \in \mathbb{R}^{2I+J}.
  \]

- **BLCMV beamformer**: \( \hat{\mathbf{w}}_{\text{BLCMV}} = \tilde{\mathbf{R}}_{n+q}^{-1} \Lambda (\Lambda^H \tilde{\mathbf{R}}_{n+q} \Lambda)^{-1} \tilde{\mathbf{f}}. \)
Problem formulation

- **General problem formulation:**

$$\min_{\mathbf{w}, \mathbf{b}} \sum_{k=1}^{M} d_k^2 V_k (4^{b_k} - 1)$$

subject to

$$\mathbf{w}^H \mathbf{R}_{n+q} \mathbf{w} \leq \frac{\beta}{\alpha}$$

$$\Lambda^H \mathbf{w} = \mathbf{f},$$

$$b_k \in \mathbb{Z}^+, \quad b_k \leq b_0, \forall k,$$

where

- $\beta$: minimum output noise power;
- $\alpha \in (0, 1]$: performance controller;
- $b_0$: maximum rate, e.g., 16 bits per sample.
• Substitution of the BLCMV beamformer, we can obtain

\[
\min_b \sum_{k=1}^{M} d_k^2 V_k (4^{b_k} - 1)
\]

\[
\text{s.t. } \tilde{f}^H (\Lambda^H \tilde{R}_{n+q}^{-1} \Lambda)^{-1} \tilde{f} \leq \frac{\beta}{\alpha}
\]

\[
b_k \in \mathbb{Z}_+, \quad b_k \leq b_0, \forall k.
\]

\[\text{(P2)}\]

• Convex optimisation:

\[
\Lambda^H \tilde{R}_{n+q}^{-1} \Lambda = Z, \quad \text{(1)}
\]

\[
\tilde{f}^H Z^{-1} \tilde{f} \leq \frac{\beta}{\alpha}, \quad \text{(2)}
\]

where \( Z \in S^{2I+J}_+ \) is Hermitian.
Rate-distributed BLCMV beamforming

- Define a constant vector \( \mathbf{e} = \left[ \frac{12}{A_1^2}, \cdots, \frac{12}{A_M^2} \right] \) and introduce a variable change \( t_k = 4^{b_k} \in \mathbb{Z}_+, \forall k \), such that \( \mathbf{R}_{qq}^{-1} = \text{diag} (\mathbf{e} \odot \mathbf{t}) \),

\[
\begin{align*}
\min_{\mathbf{t}, \mathbf{Z}} & \sum_{k=1}^{M} d_k^2 V_k(t_k - 1) \\
\text{s.t.} & \left[ \begin{array}{cc}
\mathbf{Z} & \mathbf{f} \\
\mathbf{f}^H & \frac{\beta}{\alpha}
\end{array} \right] \succeq \mathbf{O}_{2\mathcal{I}+\mathcal{J}+1} \\
& \left[ \tilde{\mathbf{R}}_{nn}^{-1} + \tilde{\mathbf{R}}_{qq}^{-1} \quad \tilde{\mathbf{R}}_{nn}^{-1} \Lambda \\
& \Lambda^H \tilde{\mathbf{R}}_{nn}^{-1} \quad \Lambda^H \tilde{\mathbf{R}}_{nn}^{-1} \Lambda - \mathbf{Z} \right] \succeq \mathbf{O}_{2M+2\mathcal{I}+\mathcal{J}} \\
& 1 \leq t_k \leq 4^{b_0}, \forall k.
\end{align*}
\]

- The integer rates can be resolved by \( b_k = \log_4 t_k, \forall k \) and randomised rounding technique.
Simulation results

- Setting: 6 mics in (4 × 3)m 2D room, $f_s=16$kHz, $T_{60}=200$ms, $\alpha = 0.8$
Simulation results

- Setting: 6 mics in $(4 \times 3)$m 2D room, $f_s=16$kHz, $T_{60}=200$ms, $\alpha = 0.8$
Simulation results

- Setting: 6 mics in \((4 \times 3)\)m 2D room, \(f_s=16\)kHz, \(T_{60}=200\)ms, \(\alpha = 0.8\)
Simulation results

- Noise reduction performance and energy usage ratio (EUR):

![Graph showing noise reduction performance and energy usage ratio (EUR)]
Simulation results

- Performance of spatial cue preservation:

\[
\Delta \text{ILD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{ILD}_j(\omega) - \tilde{\text{ILD}}_j(\omega) \right)^2 ; \quad \Delta \text{IPD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{IPD}_j(\omega) - \tilde{\text{IPD}}_j(\omega) \right)^2
\]
Simulation results

- Performance of spatial cue preservation:

\[
\Delta \text{ILD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{ILD}_j(\omega) - \tilde{\text{ILD}}_j(\omega) \right)^2;
\]

\[
\Delta \text{IPD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{IPD}_j(\omega) - \tilde{\text{IPD}}_j(\omega) \right)^2.
\]
Simulation results

- Performance of spatial cue preservation:

\[
\Delta \text{ILD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{ILD}_j(\omega) - \tilde{\text{ILD}}_j(\omega) \right)^2 ; \quad \Delta \text{IPD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{IPD}_j(\omega) - \tilde{\text{IPD}}_j(\omega) \right)^2
\]
Simulation results

- Performance of spatial cue preservation:

\[ \Delta \text{ILD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{ILD}_j(\omega) - \tilde{\text{ILD}}_j(\omega) \right)^2 ; \quad \Delta \text{IPD} = \sum_{j=1}^{J} \sum_{\omega} \left( \text{IPD}_j(\omega) - \tilde{\text{IPD}}_j(\omega) \right)^2 \]
Conclusion

- We studied rate-distributed BLCMV beamforming for wireless binaural hearing aids. The proposed method was formulated by minimizing the energy usage and constraining the noise reduction performance.

- Under the utilization of a BLCMV beamformer, the problem was solved by semi-definite programming with the capability of joint noise reduction and binaural cue preservation.

- The proposed method can achieve better energy efficiency by distributing bit rates, and preserve more interferers’ spatial cues by activating more sensors as compared to sensor selection approaches.
Conclusion

- We studied rate-distributed BLCMV beamforming for wireless binaural hearing aids. The proposed method was formulated by minimizing the energy usage and constraining the noise reduction performance.

- Under the utilization of a BLCMV beamformer, the problem was solved by semi-definite programming with the capability of joint noise reduction and binaural cue preservation.

- The proposed method can achieve better energy efficiency by distributing bit rates, and preserve more interferers’ spatial cues by activating more sensors as compared to sensor selection approaches.

Thank you!
Appendix

• Matrix inversion lemma:

\[ \tilde{R}_{n+q}^{-1} = (\tilde{R}_{nn} + \tilde{R}_{qq})^{-1} = \tilde{R}_{nn}^{-1} - \tilde{R}_{nn}^{-1}(\tilde{R}_{nn}^{-1} + \tilde{R}_{qq}^{-1})^{-1}\tilde{R}_{nn}^{-1}, \]

where

\[ \tilde{R}_{nn} = \begin{bmatrix} R_{nn} & 0 \\ 0 & R_{nn} \end{bmatrix}, \quad \tilde{R}_{qq} = \begin{bmatrix} R_{qq} & 0 \\ 0 & R_{qq} \end{bmatrix}. \]

• Convex relaxation:

\[ \Lambda^H R_{n+q}^{-1} \Lambda \succeq Z \]

\[ \implies \Lambda^H \tilde{R}_{nn}^{-1} \Lambda - Z \succeq \Lambda^H \tilde{R}_{nn}^{-1}(\tilde{R}_{nn}^{-1} + \tilde{R}_{qq}^{-1})^{-1}\tilde{R}_{nn}^{-1} \Lambda \]

\[ \implies \begin{bmatrix} \tilde{R}_{nn}^{-1} + \tilde{R}_{qq}^{-1} & \tilde{R}_{nn}^{-1} \Lambda \\ \Lambda^H \tilde{R}_{nn}^{-1} & \Lambda^H \tilde{R}_{nn}^{-1} \Lambda - Z \end{bmatrix} \succeq O_{2M+2I+J}. \]