PADERBORN UNIVERSITY
The University for the Information Society

Computer Science, Electrical Engineering and Mathematics

NT Communications Engineering
Prof. Dr.-Ing. Reinhold Häb-Umbach

On the Bias of Direction of Arrival Estimation Using Linear Microphone Arrays

Florian Jacob, Reinhold Haeb-Umbach

Paderborn University, Germany
{jacob, haeb}@nt.uni-paderborn.de
http://nt.uni-paderborn.de

Introduction

- Direction of Arrival (DoA) estimates are used to
  - steer a beamformer to a given direction
  - perform blind source separation (BSS)
  - localize speakers and/or events
- Practical applications often use linear arrays, while existing evaluations of DoA estimators often use circular arrays
- Goal: Develop a model for the DoA estimation error in a reverberant scenario
- Issue: A linear microphone arrangement leads to a limited field of view (FoV), which results in a bias, due to
  - the restriction of the measurements to one half-plane and
  - the reverberation profile being non uniform

Linear array DoA estimation

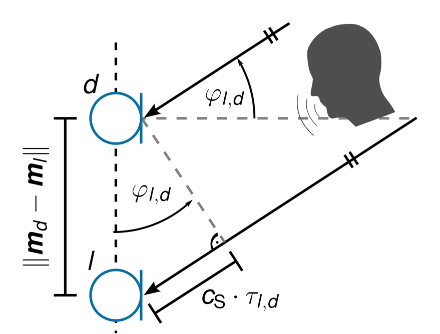
- Estimate Time Difference of Arrival (TDoA) of the microphone signals $X_d(m,k)$ and $X_l(m,k)$ using generalized cross correlation with phase transform (GCCPhat):

$$\hat{\tau}_{l,d} = \underset{\lambda}{\operatorname{argmax}}\{\mathrm{p}_{l,d}(\lambda)\} \cdot f_s^{-1} \text{ with } \mathrm{p}_{l,d}(m,\lambda) = \mathrm{IDFT}\left(\frac{X_l(m,k)X_d^*(m,k)}{|X_l(m,k)||X_d(m,k)|}\right) \quad (1)$$

- Microphone pair: Compute DoA estimate $\hat{\varphi}_{l,d}$ from TDoA $\hat{\tau}_{l,d}$:

$$\hat{\varphi}_{l,d} = \arcsin(\hat{\tau}_{l,d}/\tau_{\max}) \quad (2)$$

$$\text{with } \tau_{\max} = \|\boldsymbol{m}_l - \boldsymbol{m}_d\| \cdot c_S^{-1} \quad (3)$$



- Steered Response Power with Phase Transform (SRPPhat) evaluates score function for each candidate direction φ:

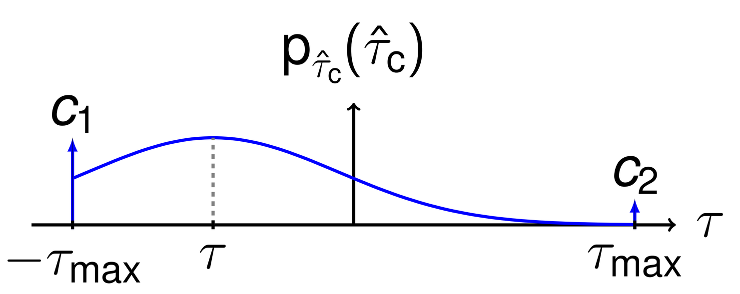
$$P_{\mathrm{SRP}}(\varphi) = \sum_{d=1}^{D}\sum_{l=d+1}^{D} \mathrm{p}_{l,d}(\tilde{\tau}_{l,d}(\varphi) \cdot f_s) \text{ with } \tilde{\tau}_{l,d}(\varphi) = \begin{bmatrix}\cos(\varphi)\\ \sin(\varphi)\end{bmatrix}^{\mathrm{T}} \frac{(\boldsymbol{m}_l - \boldsymbol{m}_d)}{c_S} \quad (4)$$

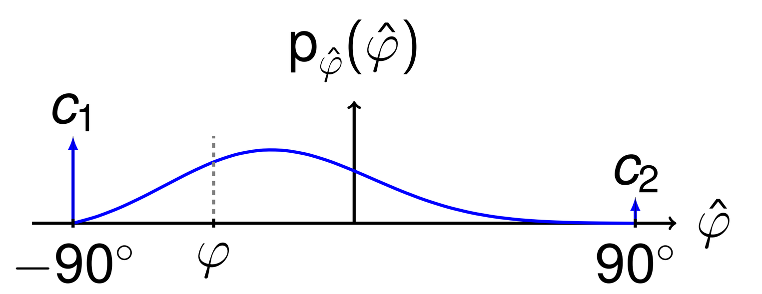
Model for TDoA clipping

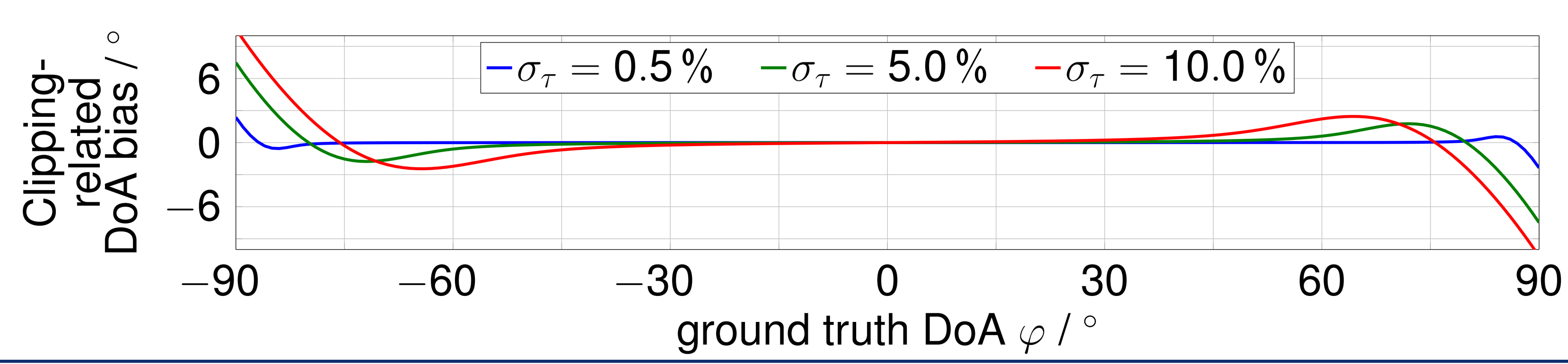
- TDoA cannot be larger than $\tau_{\max}$ (geometrical constraints)
- Due to reverberation and noise or other causes the computed TDoA estimate $\hat{\tau}$ can be larger than $\tau_{\max}$
- TDoA estimate $\hat{\tau}$ need to be clipped to apply (2):

$$\hat{\tau}_c = \max(-\tau_{\max}, \ \min(\hat{\tau}, \tau_{\max})) \quad (5)$$

- Assumption: TDoA estimation error follows zero-mean normal distribution: $\hat{\tau} \sim \mathcal{N}(\hat{\tau}; \tau, \sigma_\tau^2)$
- Clipped TDoA follows truncated normal distribution
- Corresponding distribution of the DoA from (2) via random variable transformation







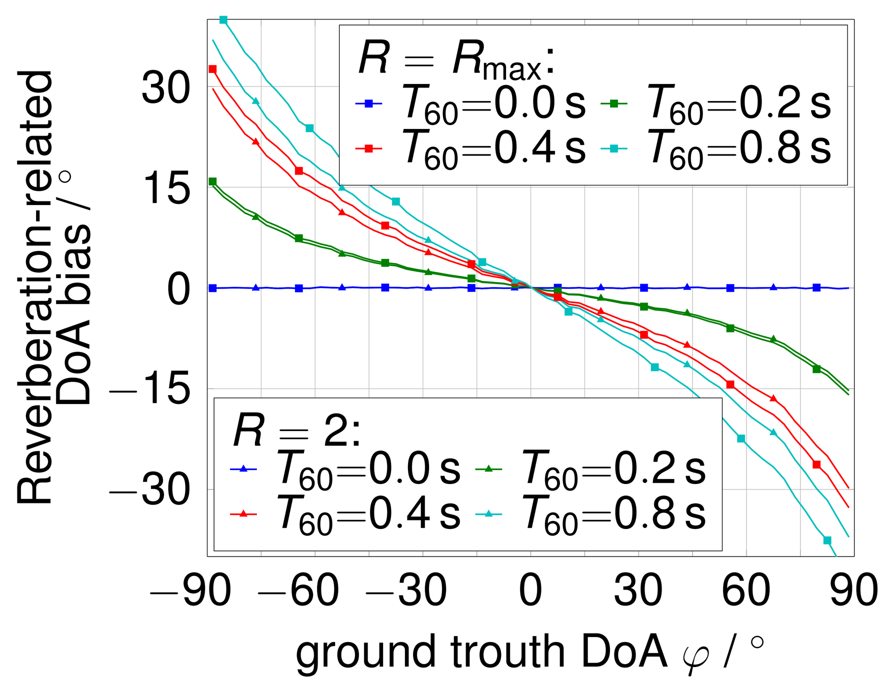
Impact of reverberation

- Uniform distribution of the phase used in statistical room acoustics takes late reverberation into account, but neglects impact of early reflections
- First- and second-order reflections exhibit higher energy
  $\Rightarrow$ Amplitude and phase need to be considered together
- Use room impulse response (RIR) to model impact of reverberation

$$h_{\mathrm{RIR}} = \sum_{r=0}^{R_{\max}} \beta_{d,r} \cdot \delta(t - \xi_{d,r}) \quad (6)$$

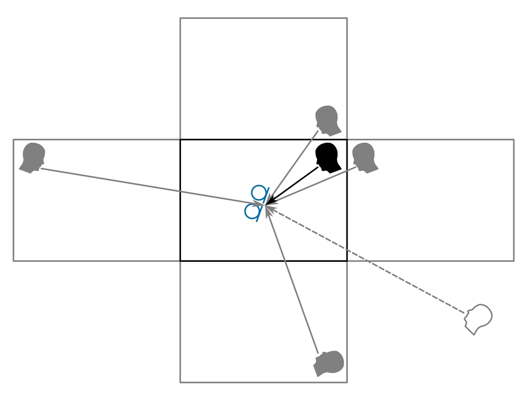
- Average TDoA of RIR

$$\hat{\tau}_{d,l}^{\mathrm{RIR}} = \frac{\sum_{r=1}^{R}\sum_{q=1}^{R} \beta_{d,r}\beta_{l,q}(\xi_{d,r} - \xi_{l,q})}{\sum_{r=1}^{R}\sum_{q=1}^{R} \beta_{d,r}\beta_{l,q}} \quad (7)$$



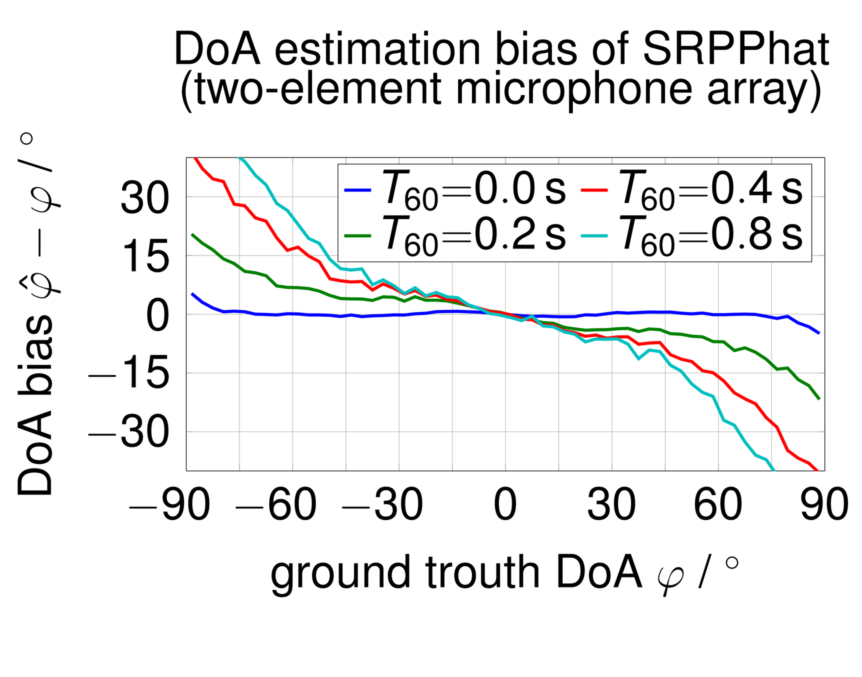
Microphones not arranged in a line

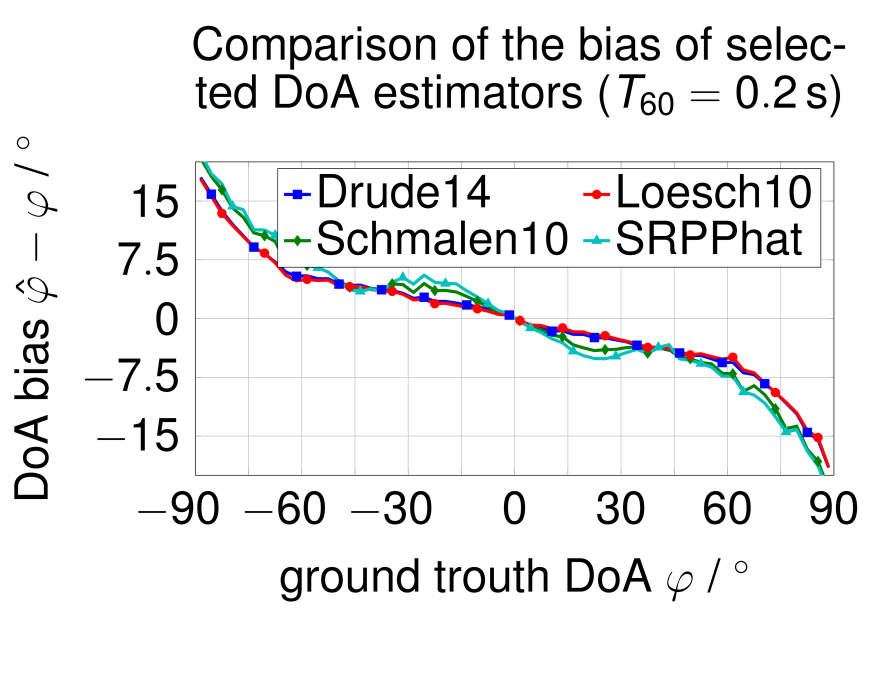
- Clipping and non uniform reverberation profile are both a consequence of the linear arrangement $\Rightarrow$ Bias vanishes if microphones are not arranged in a line
- Possible solution to avoid the bias: Use DoA estimator that models reflections



Simulations

- Speech samples from TIMIT database, convolved with RIR generated by the image method (IM)
- 10 000 random configurations
  - Room size between $4.0 \times 4.0\,\mathrm{m}^2$ and $8.0 \times 8.0\,\mathrm{m}^2$
  - Two-element array approximately in the room center





Conclusions

- Linear arrangement causes a limited FoV, which leads to
  - a clipping of the measurements
  - a mapping of the reflections to one half-plane
  - $\Rightarrow$ Significant bias, that depends on impinging direction
- Simulations verify that these items explain the bias
- Bias vanishes if microphones are not arranged on a line
- Bias compensation requires knowledge of RIR