An Objective Model of Localisation in Binaural Sound Reproduction Systems

Sontacchi A., Noisternig M., Majdak P., Höldrich R.

AES 21st International Conference

2002 June 1-3

St.Petersburg, Russia

Institute of Electronic Music and Acoustics

University of Music and Dramatic Arts
Graz, Austria

http://iem.kug.ac.at

in collaboration with

AKG-Acoustics Vienna, Austria

http://www.akg.com

Outline

- HRIR based Binaural Sound Reproduction Systems using Ambisonic
- Definitions
- Mathematical Model
- Results
- Outlook

Binaural Systems using Ambisonic

Binaural Reproduction Systems:

- filtering virtual source signals with HRIRs
- incorporate head tracking to improve source localisation

Problem:

- time-varying interpolation between HRIRs

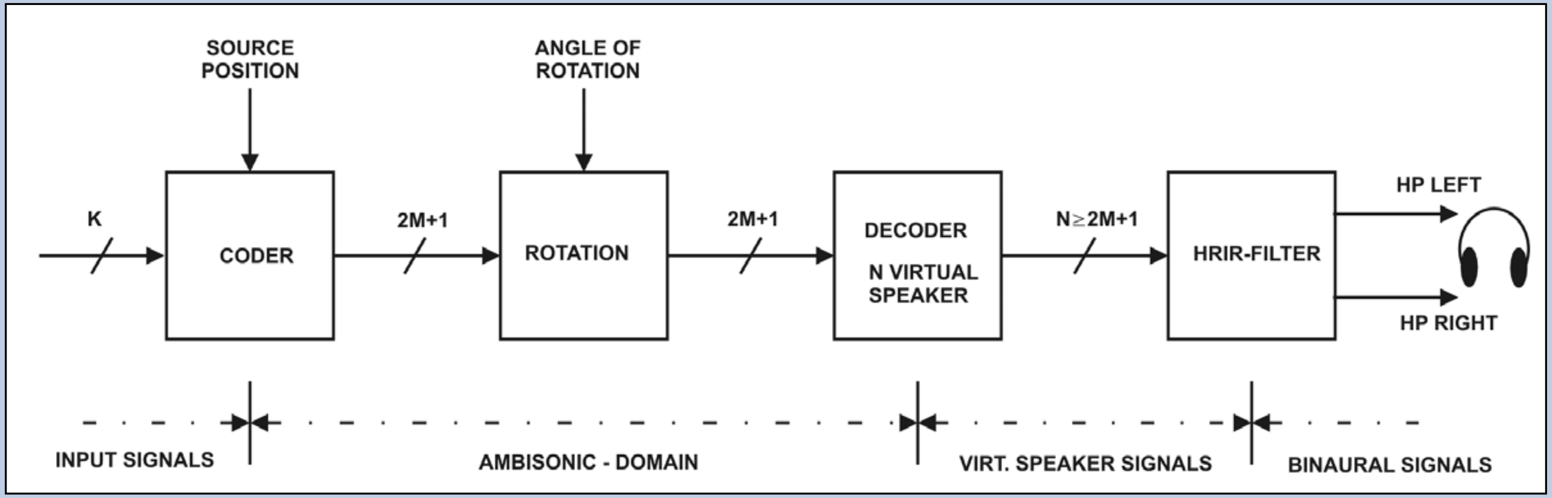
Binaural Systems using Ambisonic

Why Ambisonic?

- head movement:
  - time-invariant HRIR filter
  - cheap time-variant rotation matrix
- decoding is independent from coding
- number of transmit channels is independent from number of virtual sources

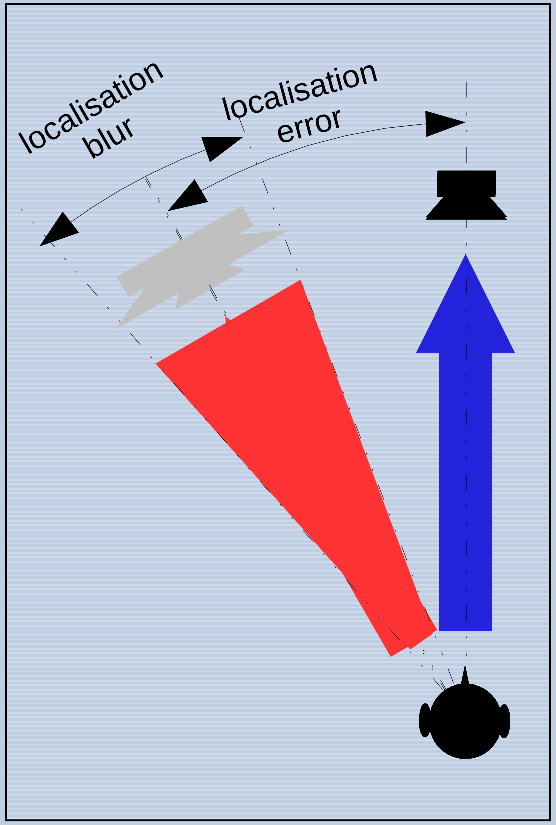
Binaural Systems using Ambisonic

- encode signals depending on source position
- rotate Ambisonic depending on head position
- decode Ambisonic to virtual speaker signals
- convolve speaker signals with HRIRs



Definitions

- localisation function
- localisation error
- localisation blur
- average localisation error
- average localisation blur



Mathematical Model

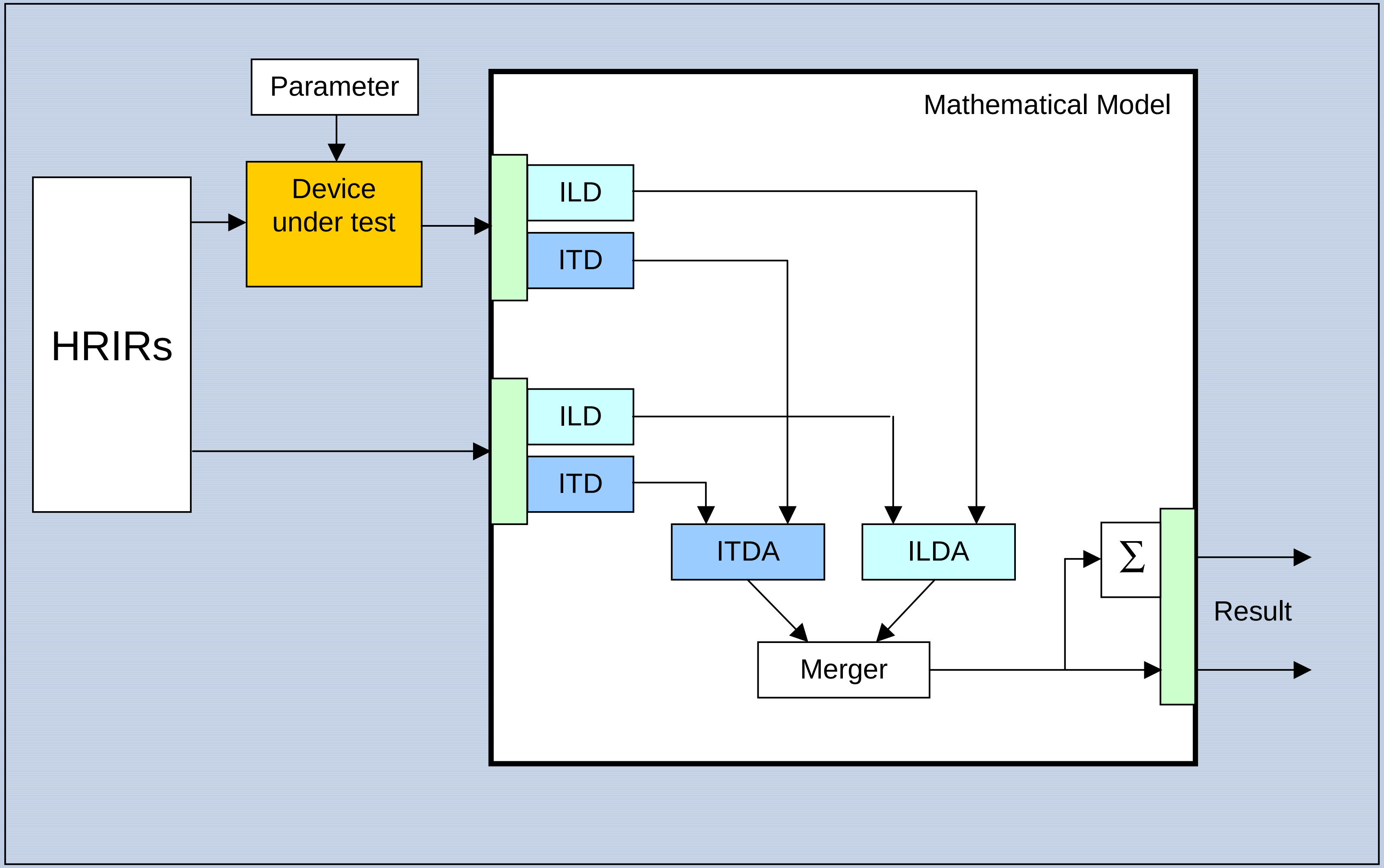
Intentions:

- classifying Binaural Reproduction Systems
- prediction of localisation performance

Main assumption:

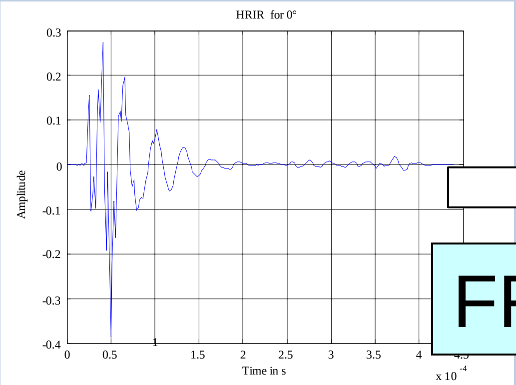
- the reference HRIRs result in optimal localisation
- 2D systems only

Mathematical Model

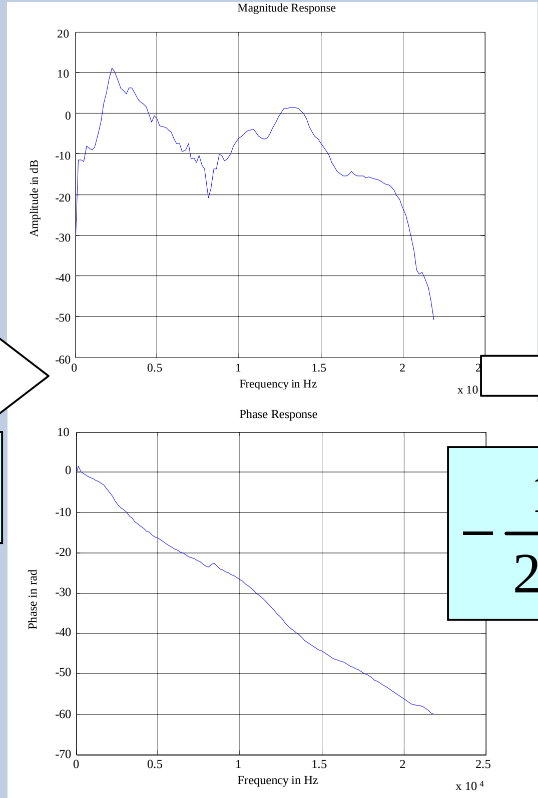


Interaural Time Difference (ITD)

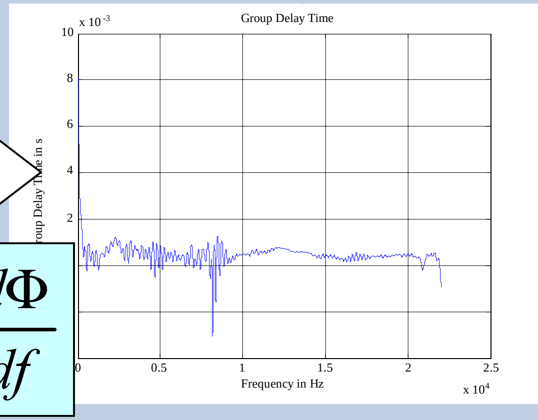
HRIR, φ=0°



FFT



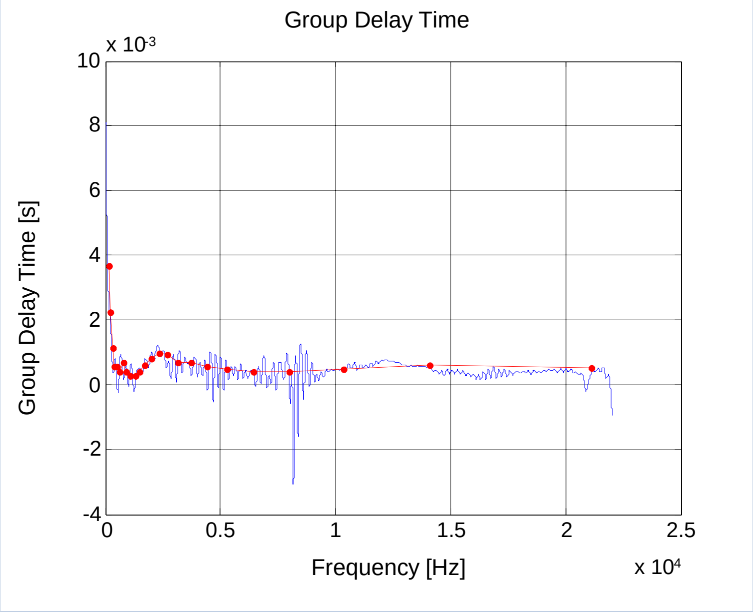
$$-\frac{1}{2\pi}\cdot\frac{d\Phi}{df}$$



Interaural Time Difference (ITD)

Calculation of the group delay time:

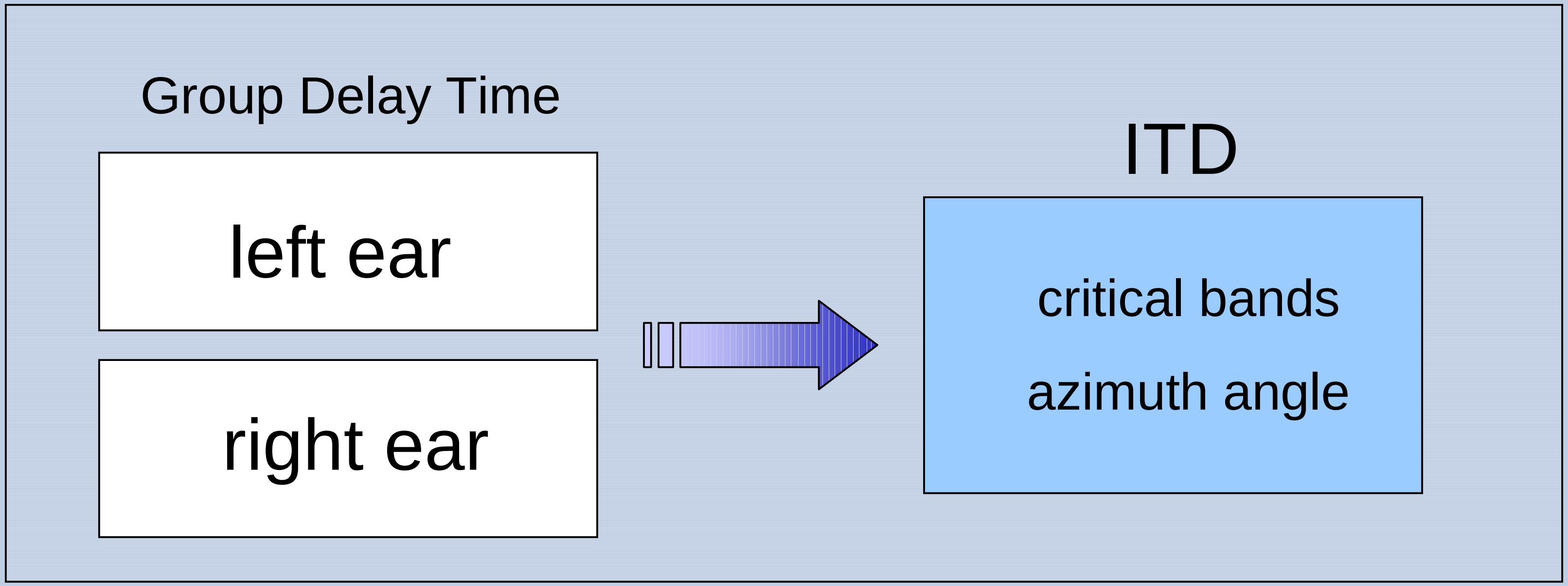
- impulse response (HRIR)
- zero phase filter over critical bands
- energy center of gravity per band
- group delay time per band



Interaural Time Difference (ITD)

ITD over critical bands:

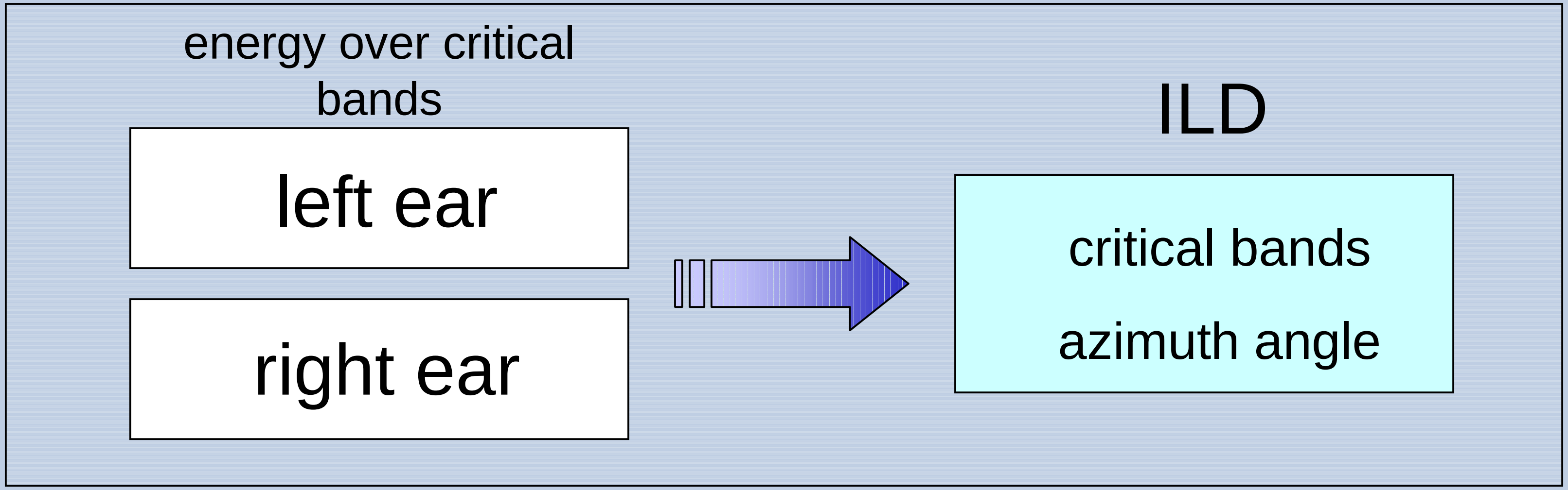
difference of the group delay time between ipsi- and contralateral eardrum



Interaural Level Difference (ILD)

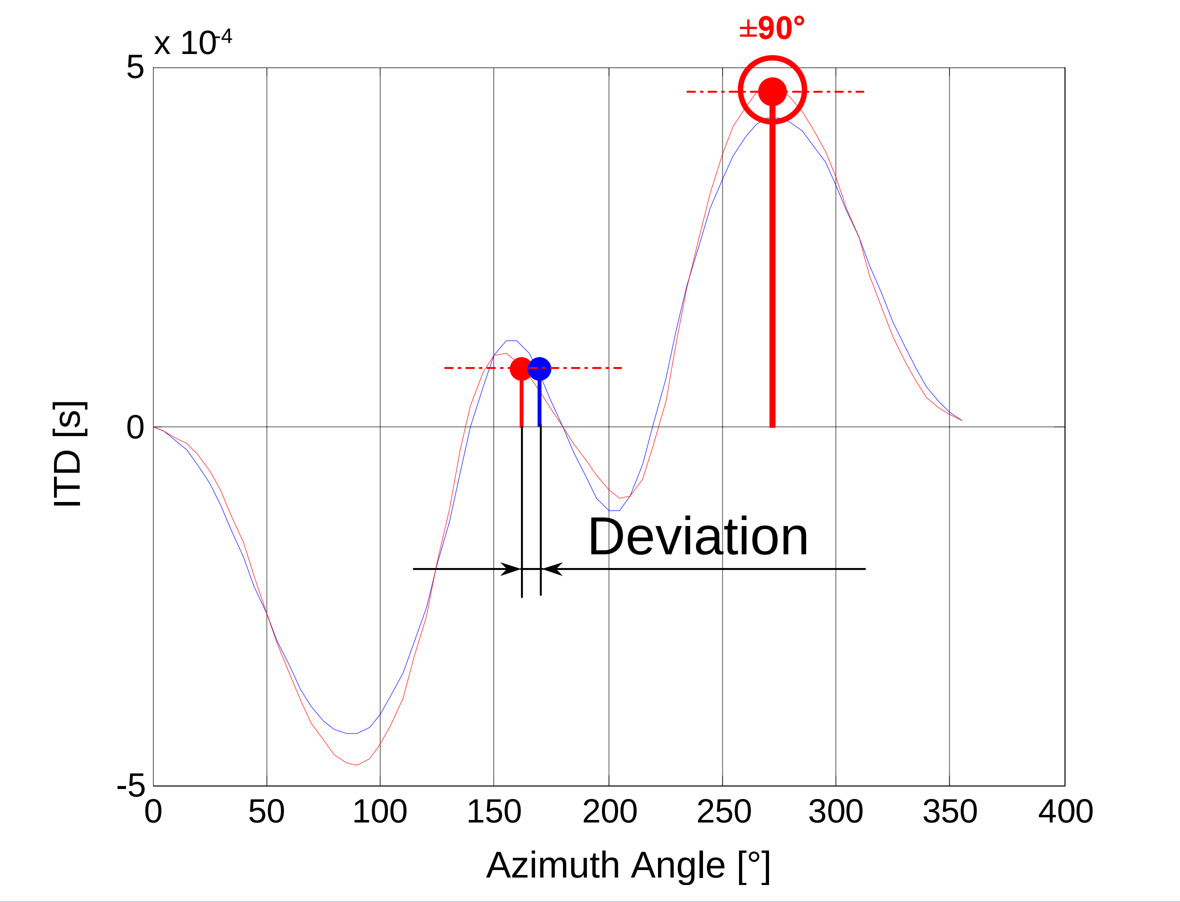
Calculation of the ILD:

- amplitude response (HRTF)
- energy-difference over critical bands between ipsi- and contralateral eardrums



Interaural Difference Angle (ITDA/ILDA)

Search Algorithm:



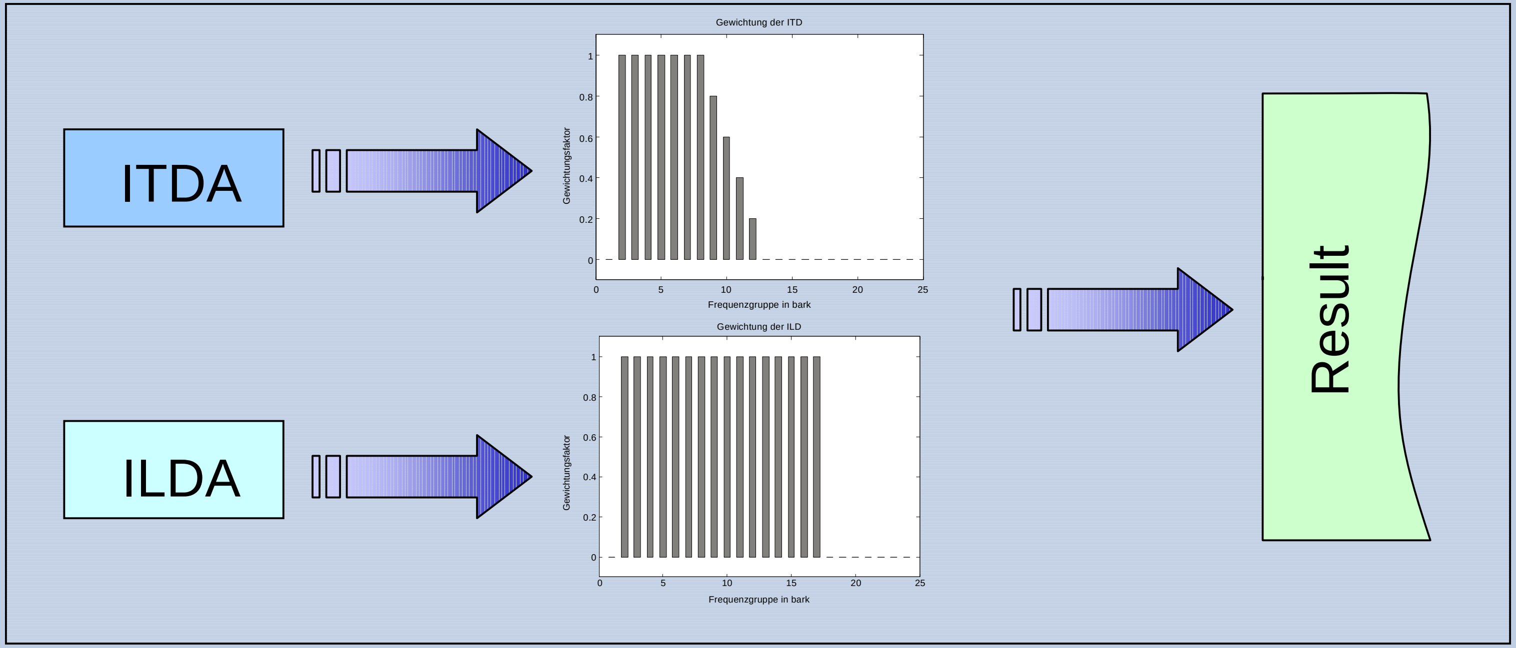
ITD for one critical band calculated from measured HRIR

distorted ITD

reference ITD

Merging of ITDA and ILDA

- ITDA: fade out above 800 Hz
- ILDA: fully weighted



Results (1)

- localisation function (averaged over critical bands)

$$L(\Theta) = \frac{1}{2} \cdot \left(\frac{1}{\sum w_{ITD}(z)} \cdot \sum_{z=1}^{24} w_{ITD}(z) \cdot \Theta_{ITD}(z,\Theta) \right) + \frac{1}{2} \cdot \left(\frac{1}{\sum w_{ILD}(z)} \cdot \sum_{z=1}^{24} w_{ILD}(z) \cdot \Theta_{ILD}(z,\Theta) \right)$$

- localisation blur (standard deviation over critical bands)

$$Bl(\Theta) = \sqrt{\frac{1}{2} \cdot \sum_{i=ILD}^{ITD} \frac{1}{\sum w_i(z)} \cdot \sum_{z=1}^{24} w_i(z) \cdot \left[\Theta_i(z,\Theta) - L(\Theta)\right]^2}$$

Results (2)

- average localisation blur
  (averaged over the azimuth angle)

$$\overline{L} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[L(\Theta_i) - \Theta_i\right]^2}$$

- average localisation error
  averaged over azimuth angle referring to the MAA

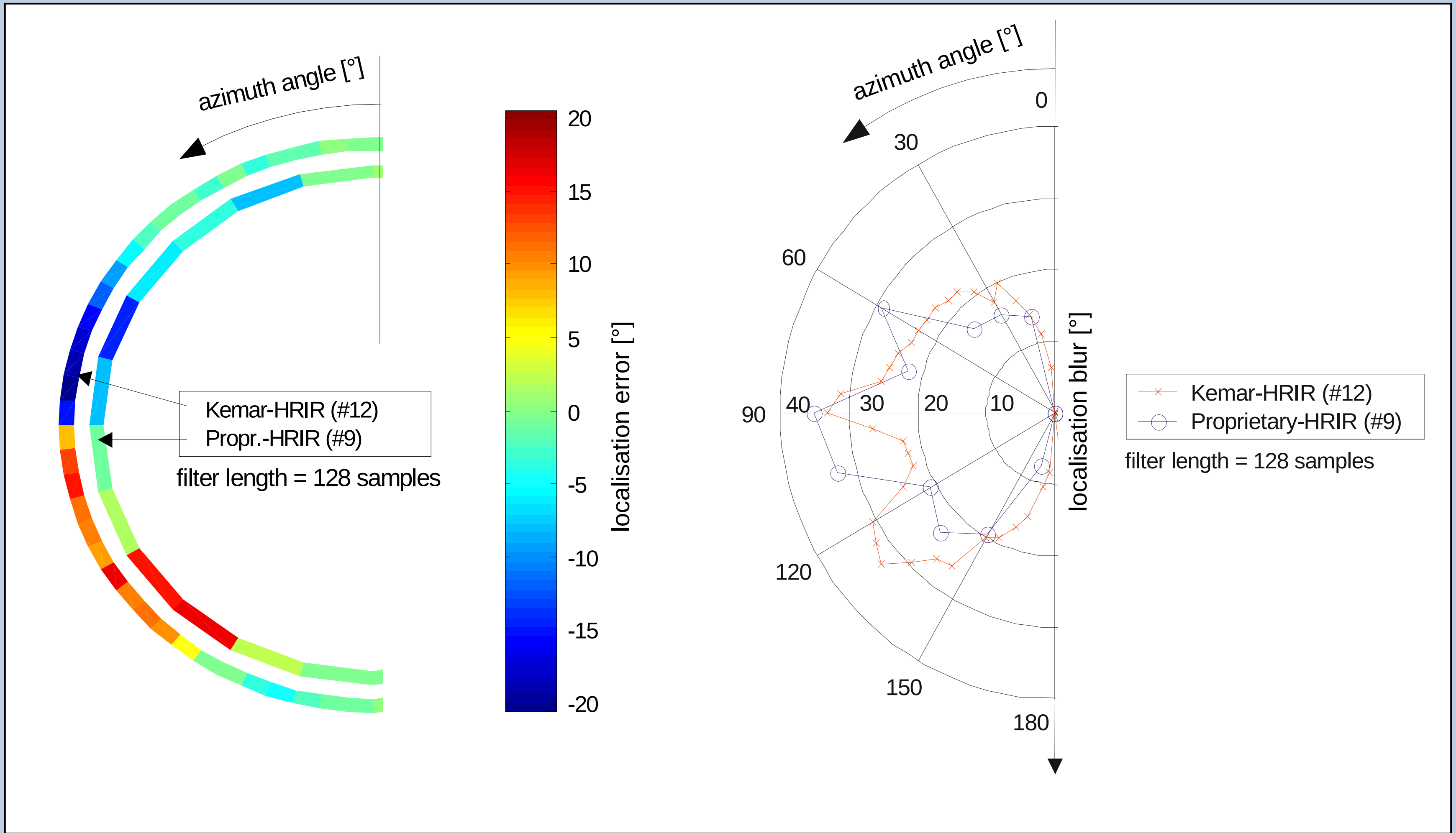
$$\overline{Bl} = \frac{1}{N} \cdot \sum_{\Theta=0}^{360°} \frac{Bl(\Theta)}{MAA(\Theta)}$$

Analysis

Variations of design parameters:

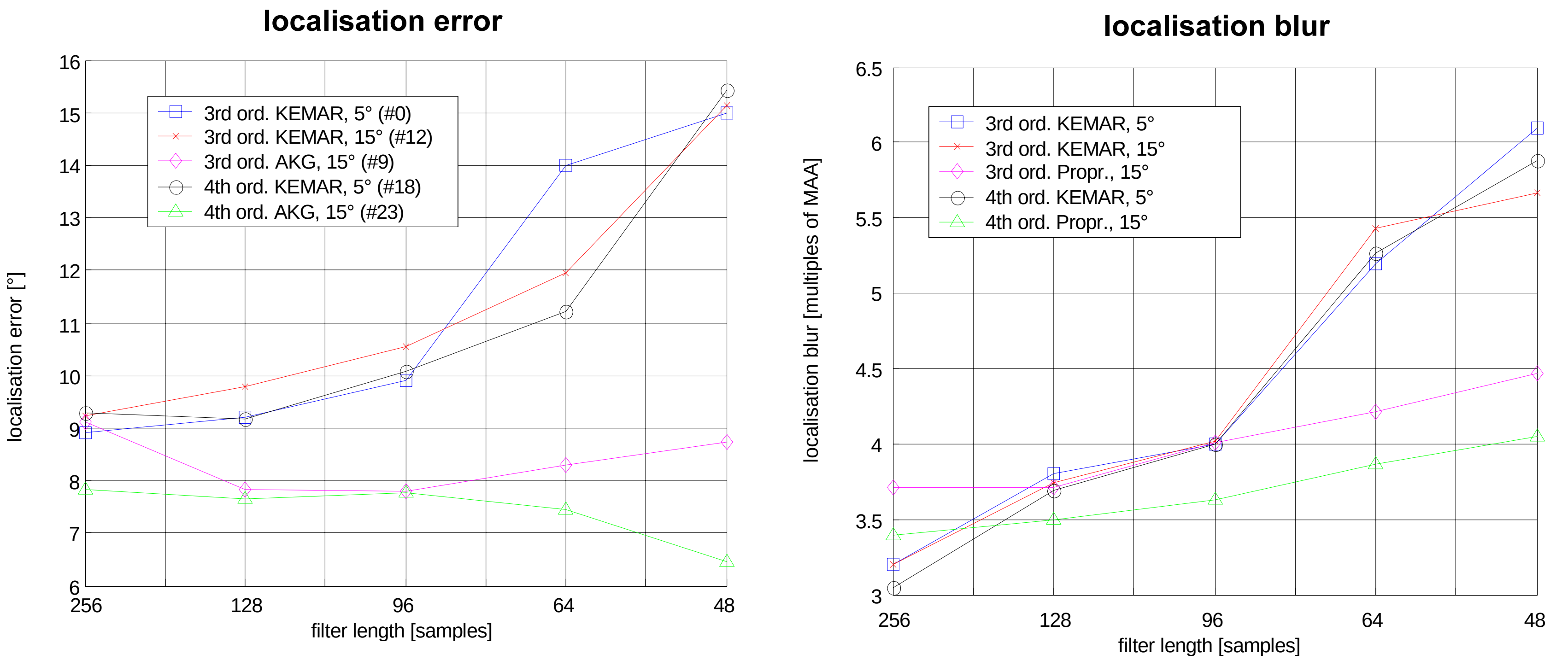
- reference HRIRs
- length of HRIR
- ambisonic order
- weighting of ambisonic channels
- number of virtual speaker

HRIR Type (1)

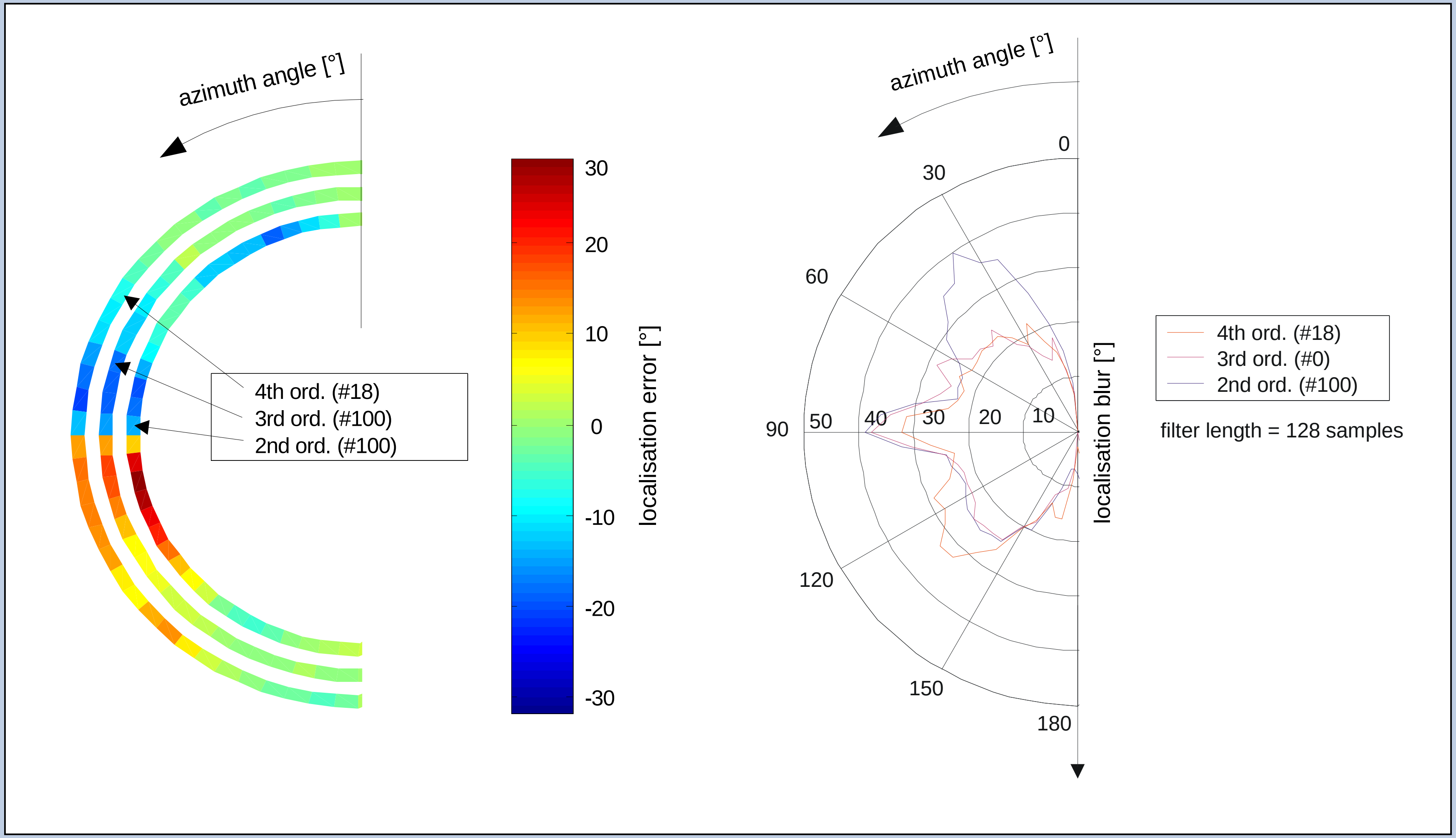


azimuth angle [°]
Kemar-HRIR (#12)
Propr.-HRIR (#9)
filter length = 128 samples
20
15
10
5
0
-5
-10
-15
-20
localisation error [°]
azimuth angle [°]
0
30
60
90
120
150
180
40
30
20
10
localisation blur [°]
Kemar-HRIR (#12)
Proprietary-HRIR (#9)
filter length = 128 samples

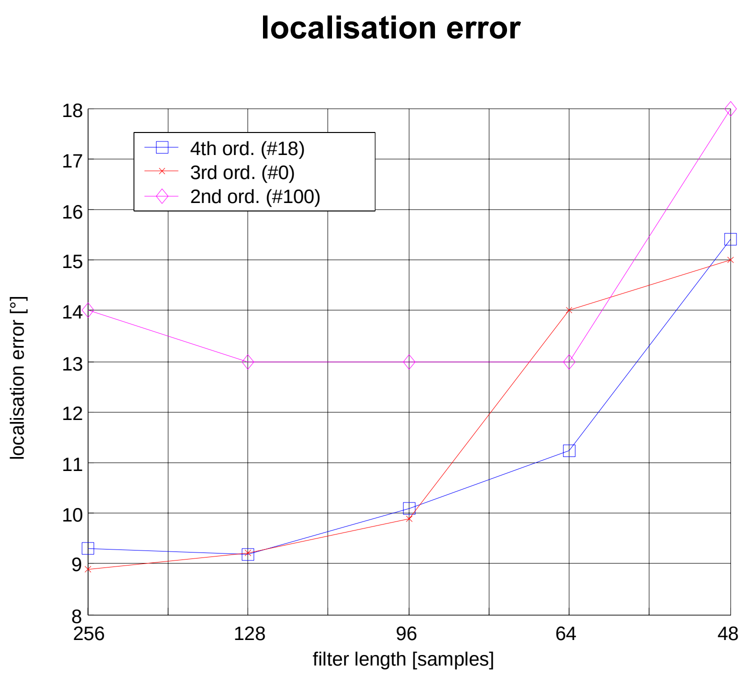
HRIR Type (2)

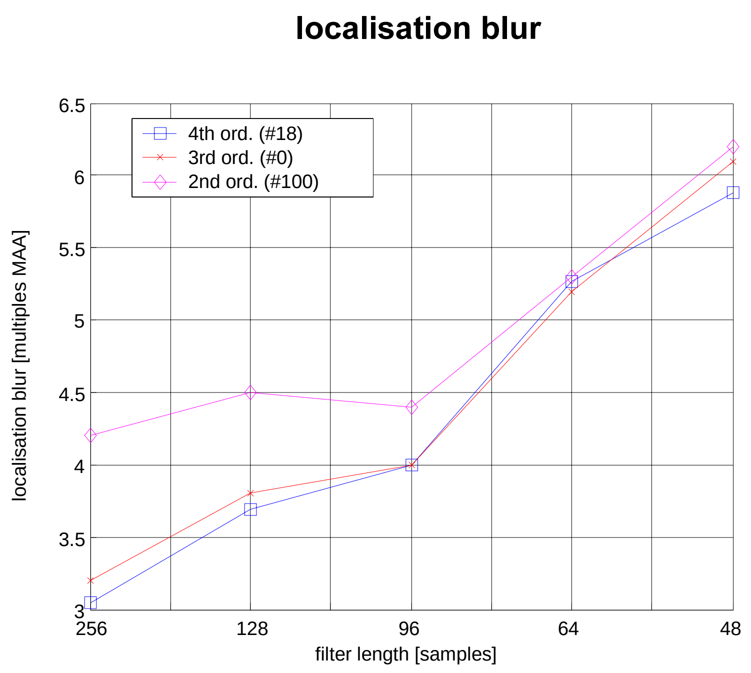

localisation error
3rd ord. KEMAR, 5° (#0)
3rd ord. KEMAR, 15° (#12)
3rd ord. AKG, 15° (#9)
4th ord. KEMAR, 5° (#18)
4th ord. AKG, 15° (#23)
localisation error [°]
filter length [samples]
256
128
96
64
48
6
7
8
9
10
11
12
13
14
15
16
localisation blur
3rd ord. KEMAR, 5°
3rd ord. KEMAR, 15°
3rd ord. Propr., 15°
4th ord. KEMAR, 5°
4th ord. Propr., 15°
localisation blur [multiples of MAA]
filter length [samples]
3
3.5
4
4.5
5
5.5
6
6.5

Ambisonic Order (1)

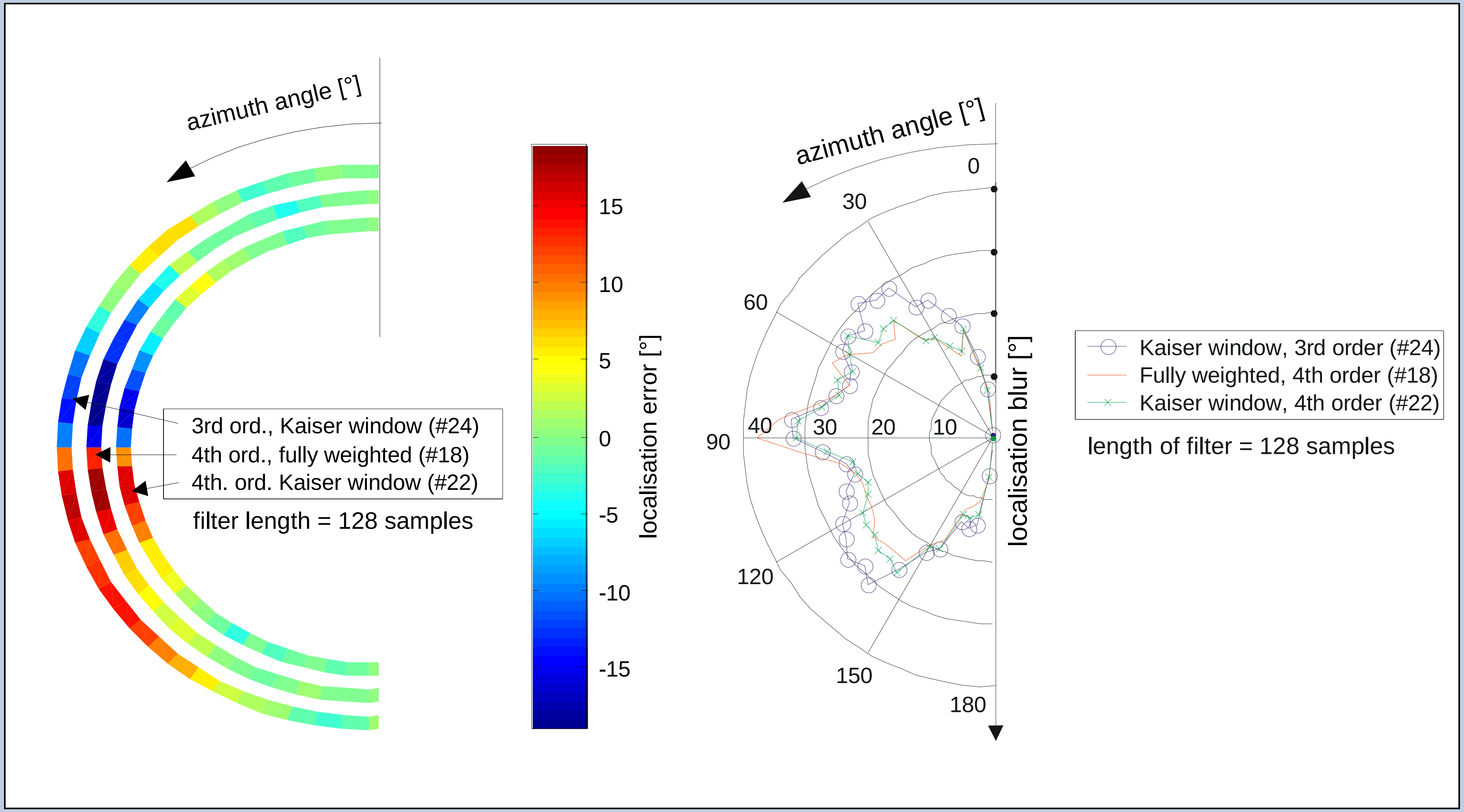


Ambisonic Order (2)

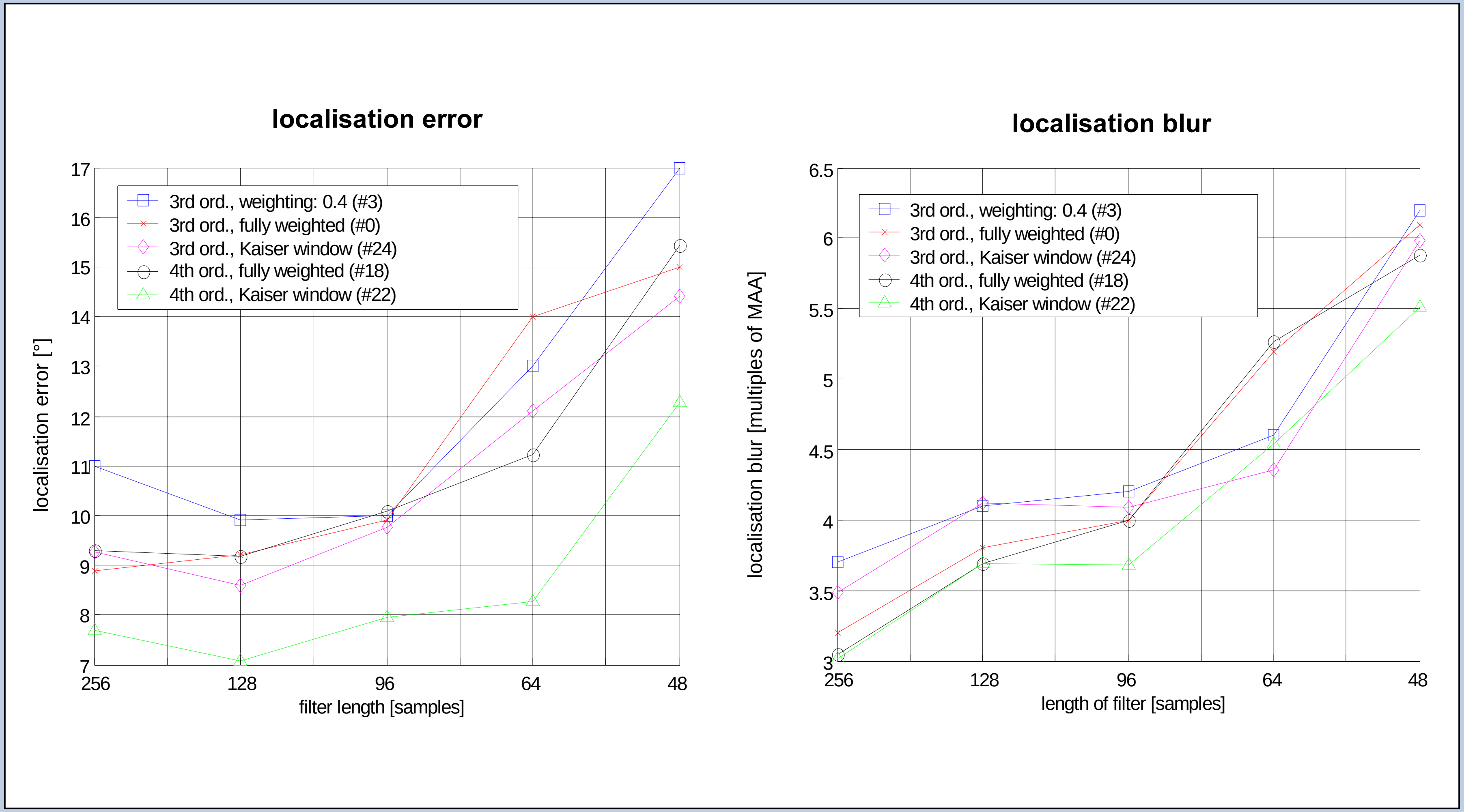




Ambisonic Weighting (1)



Ambisonic Weighting (2)



Conclusion

- prediction of localisation is possible
- results correspond with ambisonic theory
  - ambisonic not less than 3rd order for correct localisation
  - weighting of channels using Kaiser-window yields better localisation
  - filter length up to 128 samples
- model was evaluated using listening tests

Outlook

- incorporate IGD (*interaural group delay*) as a cue for high frequency localisation
- classification regarding coloration and externity
- extend the model for 3D applications
- warping tables to predict and minimise the localisation errors for further implemen-tations

Institute of Electronic Music and Acoustics

Thank you!

University of Music and Dramatic Arts

Graz, Austria

http://iem.kug.ac.at