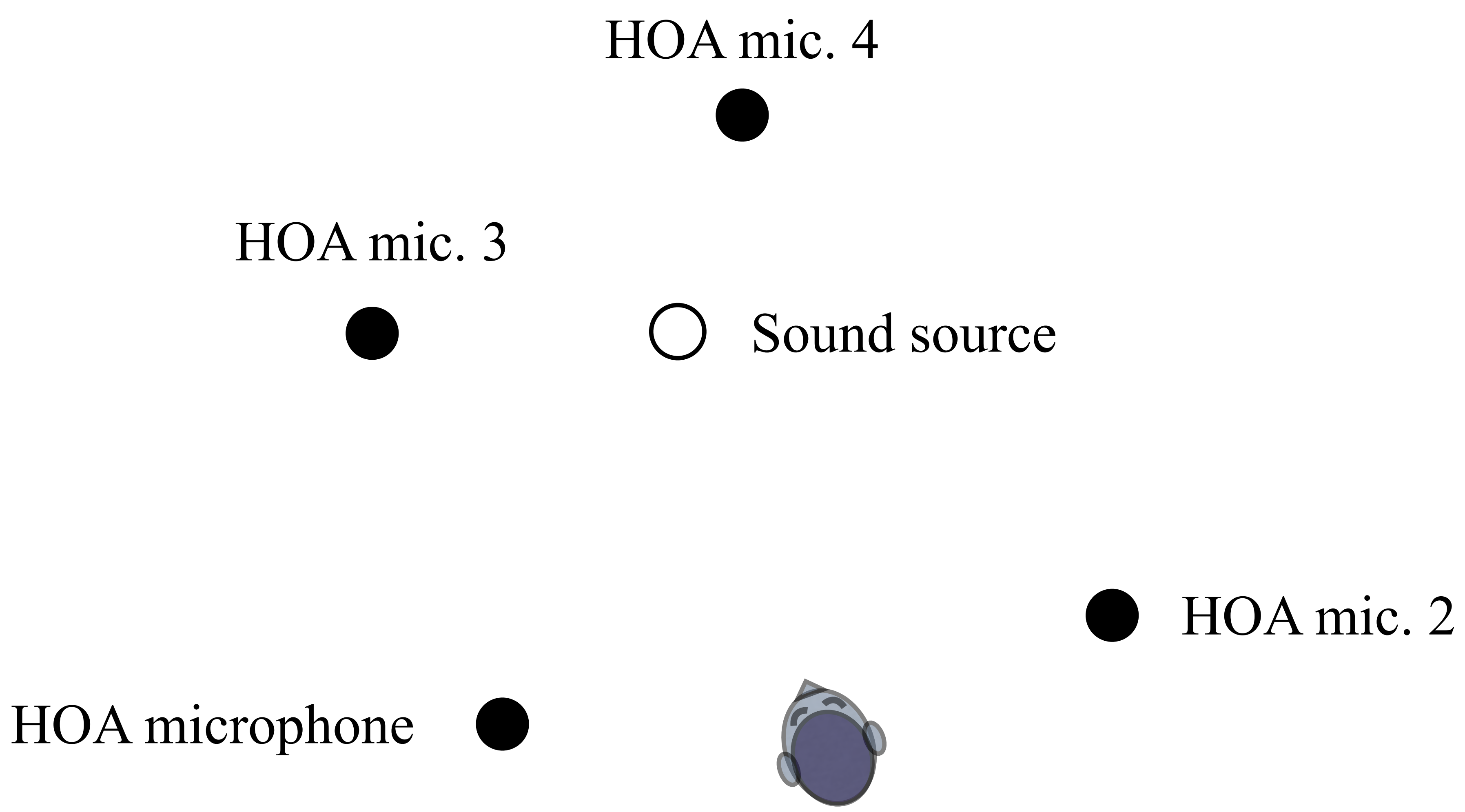
Evaluation of techniques for navigation of higher-order ambisonics

Acoustics ’17 Boston
Presentation 1pPPb4
June 25th, 2017

Joseph G. Tylka (presenter)
and Edgar Y. Choueiri

3D Audio and Applied Acoustics (3D3A) Laboratory
Princeton University
www.princeton.edu/3D3A

Sound Field Navigation



Sound Field Navigation

- Lots of different ways to navigate:
  - Plane-wave translation (Schultz & Spors, 2013)
  - Spherical-harmonic re-expansion (Gumerov & Duraiswami, 2005)
  - Linear interpolation/“crossfading” (Southern et al., 2009)
  - Collaborative blind source separation (Zheng, 2013)
  - Regularized least-squares interpolation (Tylka & Choueiri, 2016)

} HOA in ↓ HOA out

- Need a way to evaluate and compare them
  - Isolate navigational technique from binaural/ambisonic rendering
  - Subjective testing can be lengthy/costly ⟹ **Objective Metrics**

Overview

- For each quality (localization and coloration):
  - Existing metrics
  - Proposed metric
  - Listening test
  - Results
- Summary and outlook

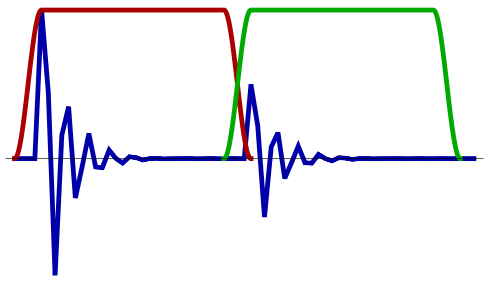
Source Localization

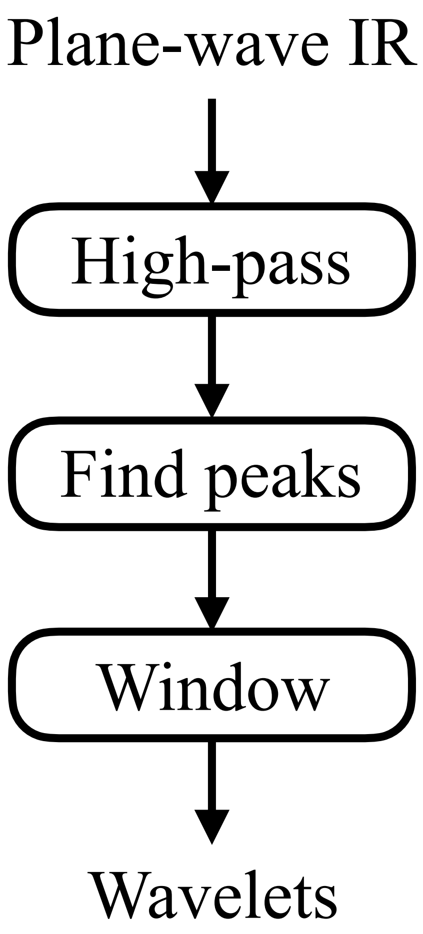
Existing Metrics

- Binaural models:
  - Lindemann (1986); Dietz et al. (2011); etc.
    - Predict perceived source azimuth given binaural impulse responses (IRs)
- Localization vectors:
  - Gerzon (1992) — for analyzing ambisonics
    - Low frequency (velocity) and high frequency (energy) vectors
    - Predict perceived source direction given speaker positions & gains
  - Stitt et al. (2016)
    - Incorporates precedence effect to Gerzon's energy vector
    - Model requires: direction-of-arrival, time-of-arrival, and amplitude for each source
    - Tylka & Choueiri (2016) generalized algorithm for ambisonics IRs

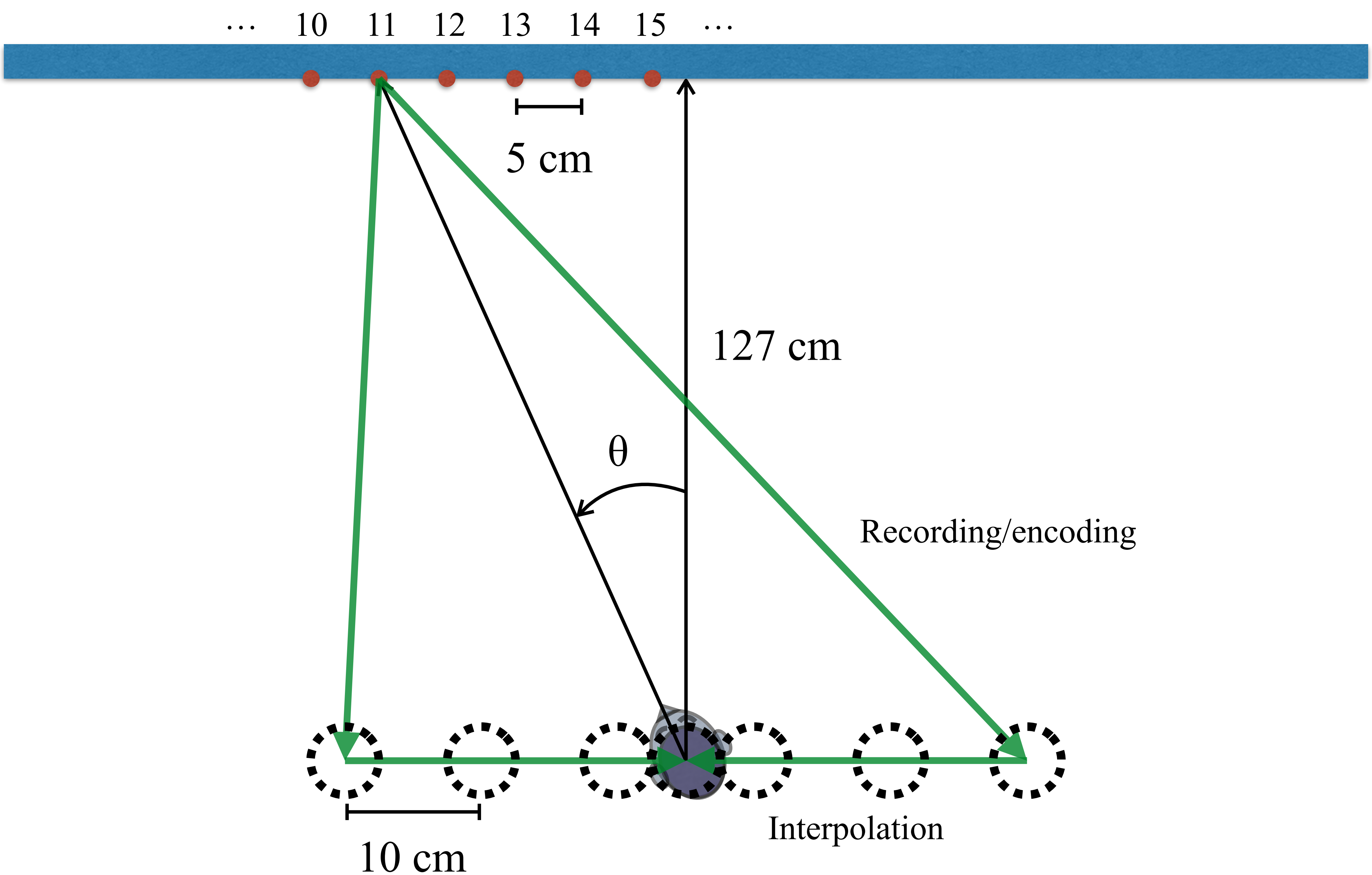
Proposed Metric

1. Transform to plane-wave impulse responses (IRs)
2. Split each IR into wavelets
3. Threshold to find onset times
4. Compute average amplitude in each critical band
5. Compute Stitt's energy vector in each band for $f \geq 700$ Hz
6. Similarly, compute velocity vector in each band for $f \leq 700$ Hz
7. Compute average vector weighted by stimulus energies in each band



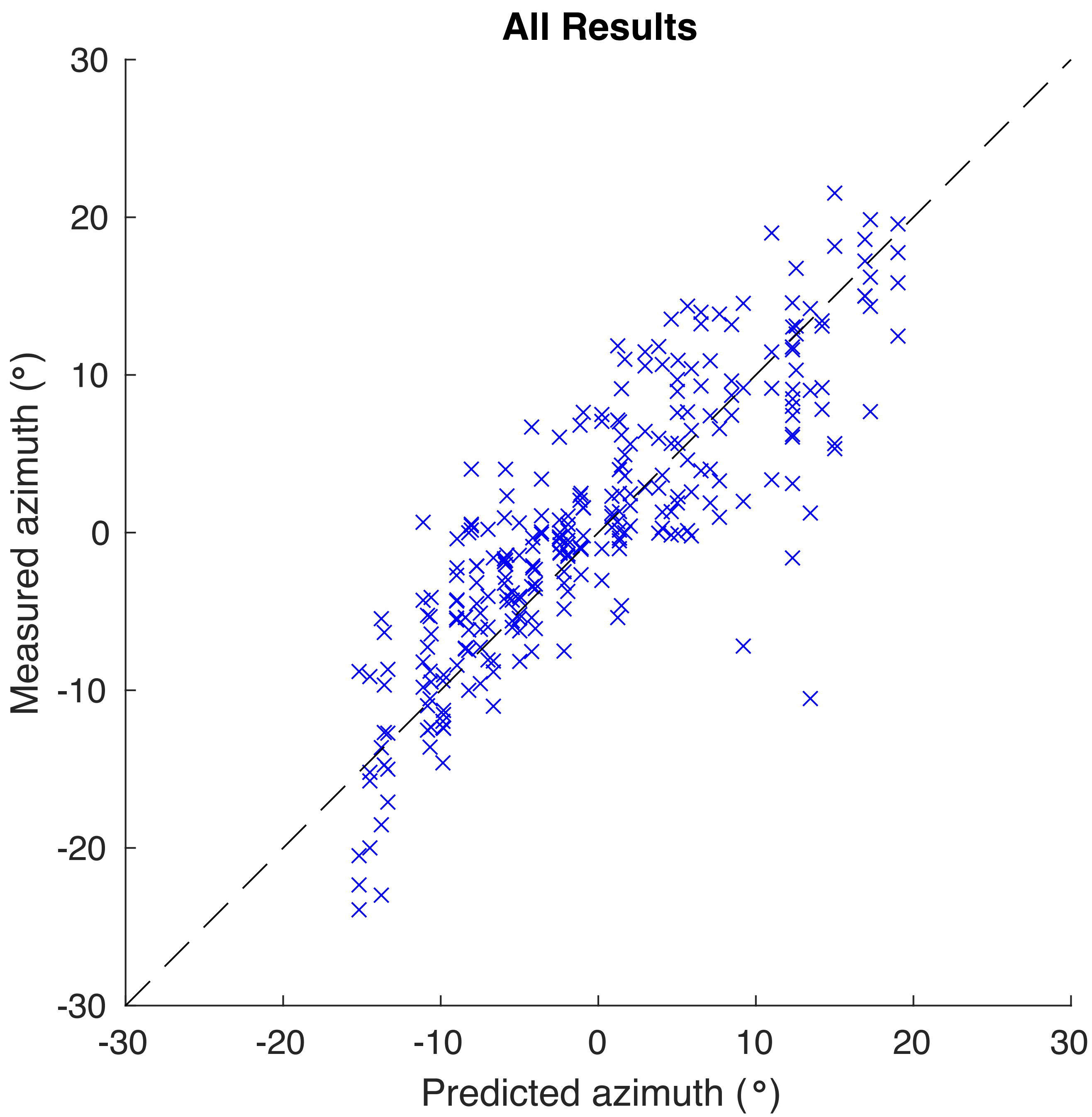


Localization Test



Localization Test Results

Test details:
- 70 test samples
- 4 trained listeners
- Speech signal



Pearson correlation coefficient: $r = 0.77$

Mean absolute error: $\varepsilon = 3.67°$

Spectral Coloration

Existing Metrics

- Auditory band error (Schärer & Lindau, 2009); peak and notch errors (Boren et al., 2015)
- Central spectrum (Kates, 1984; 1985)

} Free-field transfer functions

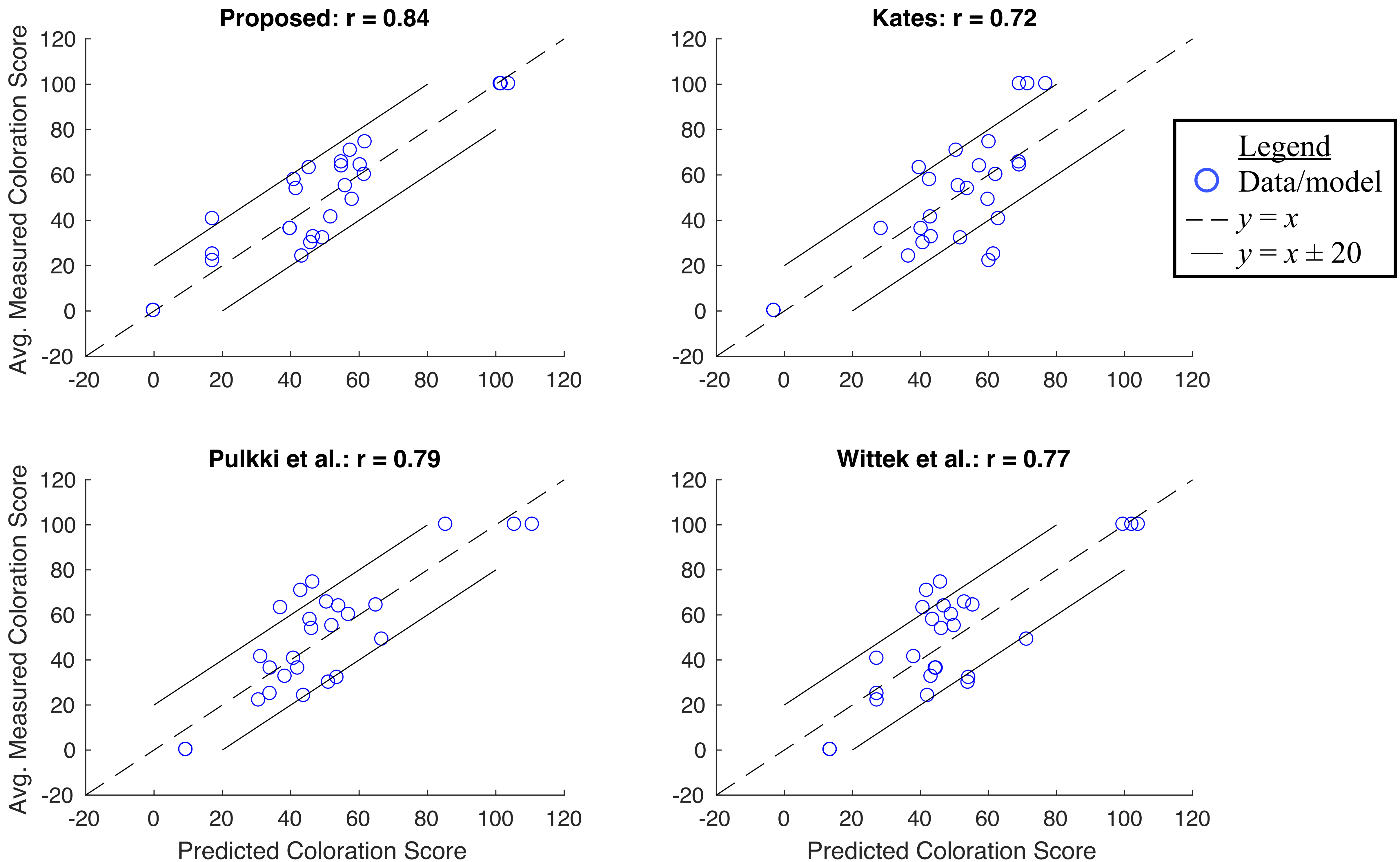
- Composite loudness level (Pulkki et al., 1999; Huopaniemi et al., 1999)
- Internal spectrum and A_0 measure (Salomons, 1995; Wittek et al., 2007)

} Binaural transfer functions

Methodology

- Perform multiple linear regression between ratings and various metrics
  - For spectral metrics: compute max−min & standard deviation
- **MU**ltiple **S**timuli with **H**idden **R**eference and **A**nchor (ITU-R BS.1534-3)
  - **Reference**: no navigation, pink noise
  - **Anchor 1**: 3.5 kHz low-passed version of **Ref**.
  - **Anchor 2**: +6 dB high-shelf above 7 kHz applied to **Ref**.
  - **Test samples**: vary interpolation technique and distance
- User rates each sample from 0–100: 100 = **Ref**.; 0 = **Anchor 1**
  - Coloration score = 100 − MUSHRA rating: 0 = **Ref**.; 100 = **Anchor 1**
- **Proposed model**: auditory band and notch errors only (Boren et al., 2015)

Regression Results



Summary and Outlook

- Presented objective metrics that predict localization and coloration
- Validated through comparisons with subjective test results

Next Steps:

1. Compare localization metric with binaural models
2. Validate metrics for other stimuli, directions, conditions
3. Verify generalization to other binaural rendering techniques

References

- Boren et al. (**2015**). "Coloration metrics for headphone equalization."
- Dietz et al. (**2011**). "Auditory model based direction estimation of concurrent speakers from binaural signals."
- Gerzon (**1992**). "General Metatheory of Auditory Localisation."
- Gumerov and Duraiswami (**2005**). *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*.
- Huopaniemi et al. (**1999**). "Objective and Subjective Evaluation of Head-Related Transfer Function Filter Design."
- ITU-R BS.1534-3 (**2015**). "Method for the subjective assessment of intermediate quality level of audio systems."
- Kates (**1984**). "A Perceptual Criterion for Loudspeaker Evaluation."
- Kates (**1985**). "A central spectrum model for the perception of coloration in filtered Gaussian noise."
- Lindemann (**1986**). "Extension of a binaural cross-correlation model by contralateral inhibition."
- Pulkki et al. (**1999**). "Analyzing Virtual Sound Source Attributes Using a Binaural Auditory Model."
- Salomons (**1995**). *Coloration and Binaural Decoloration of Sound due to Reflections*.
- Schärer and Lindau (**2009**). "Evaluation of Equalization Methods for Binaural Signals."
- Schultz and Spors (**2013**). "Data-Based Binaural Synthesis Including Rotational and Translatory Head-Movements."
- Southern, Wells, and Murphy (**2009**). "Rendering walk-through auralisations using wave-based acoustical models."
- Stitt, Bertet, and van Walstijn (**2016**). "Extended Energy Vector Prediction of Ambisonically Reproduced Image Direction at Off-Center Listening Positions."
- Tylka and Choueiri (**2016**). "Soundfield Navigation using an Array of Higher-Order Ambisonics Microphones."
- Wittek et al. (**2007**). "On the sound colour properties of wavefield synthesis and stereo."
- Zheng (**2013**). *Soundfield navigation: Separation, compression and transmission*.

Acknowledgments

- Binaural rendering was performed using M. Kronlachner's ambiX plug-ins: http://www.matthiaskronlachner.com/?p=2015
- The em32 Eigenmike by mh acoustics was used to measure the HOA RIRs: https://mhacoustics.com/products#eigenmike1
- Auditory filters were generated using the LTFAT MATLAB Toolbox: http://ltfat.sourceforge.net/
- P. Stitt's energy vector code can be found here: https://circlesounds.wordpress.com/matlab-code/