Time-of-arrival estimation and phase unwrapping of head-related transfer functions with integer linear programming

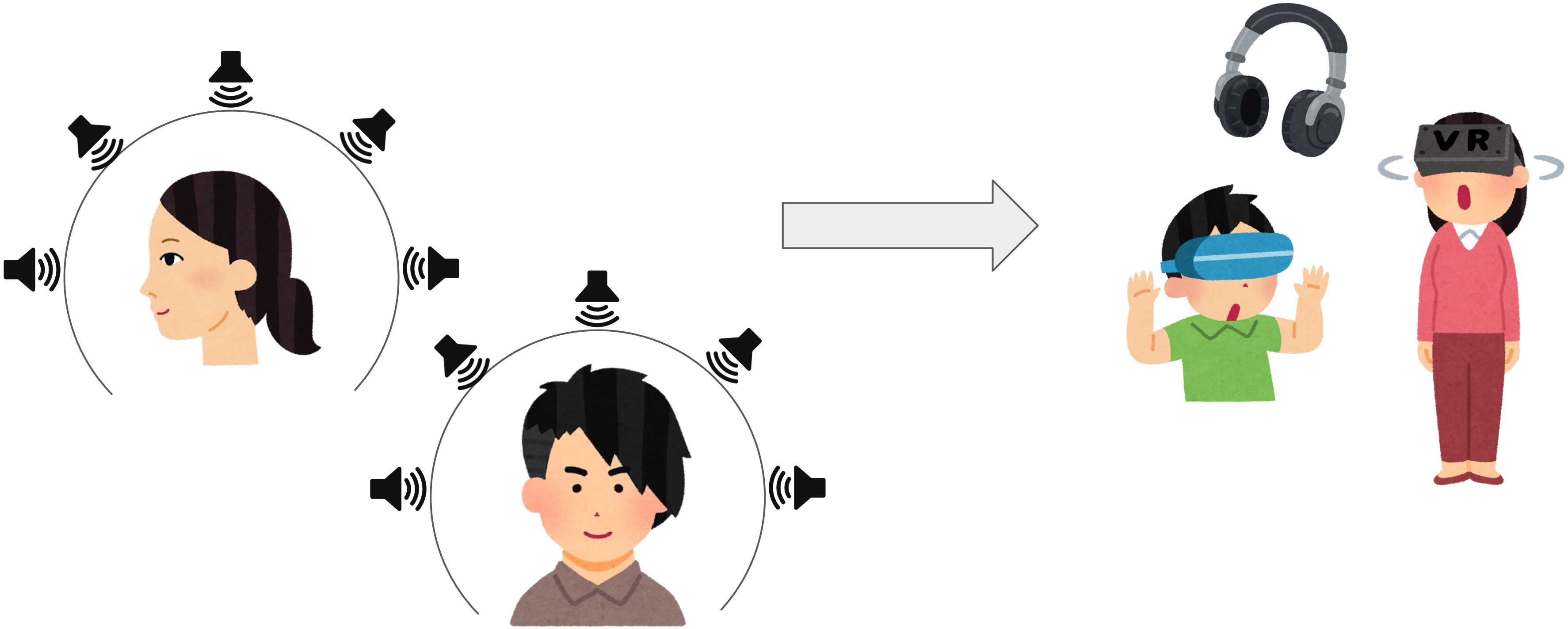
Chin-Yun Yu, Johan Pauwels, and György Fazekas
Centre for Digital Music, Queen Mary University of London

Chin-Yun Yu

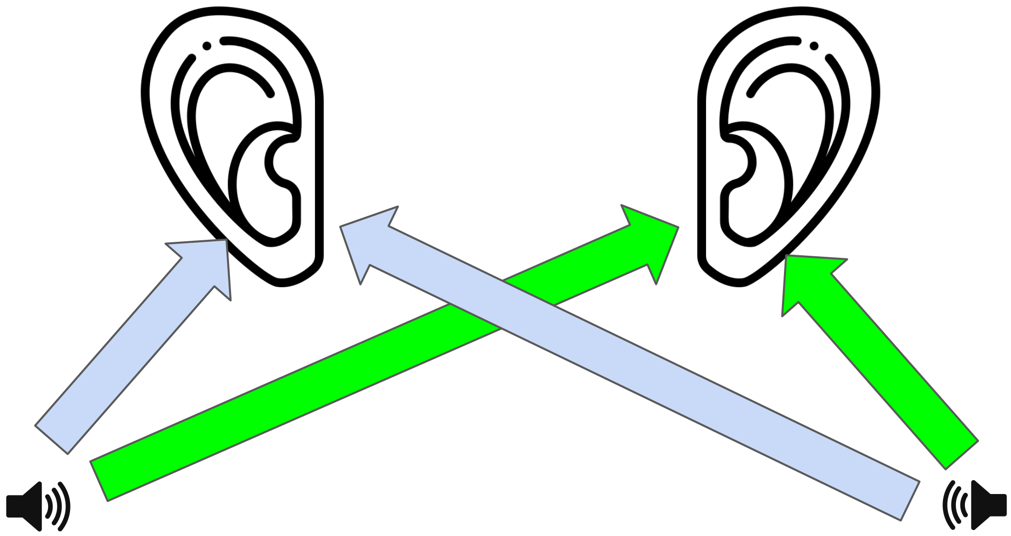
- Experiences
  - Bachelor of computer science, NCTU, Taiwan (2014~2018)
  - Engineer at HTC Vive (2019~2020)
  - Backend engineer at Rayark Inc. (2020~2022)
  - PhD student, C4DM, Queen Mary University of London (2022~)
- Voice synthesis, DSP, generative models, spatial audio
- GitHub/Twitter: @yoyololicon

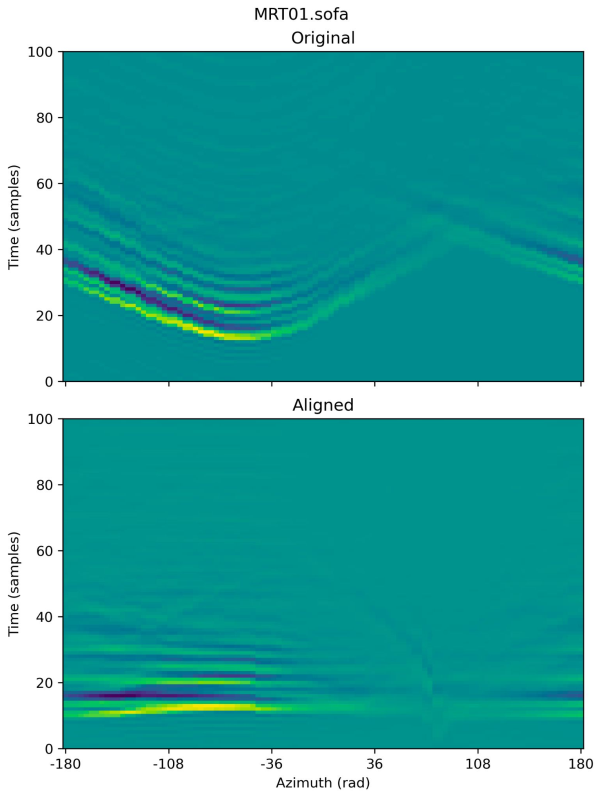
Introduction

Head-related Transfer Function (HRTF)



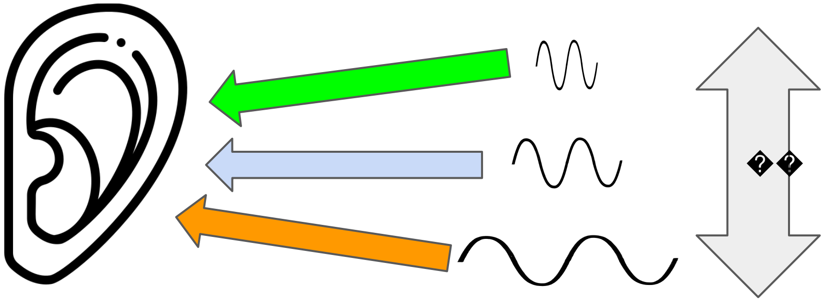
Time-of-arrival (TOA)





HRTF Phase Response

- Time-of-arrival of different **frequencies**
- Method: Phase unwrapping (PU)



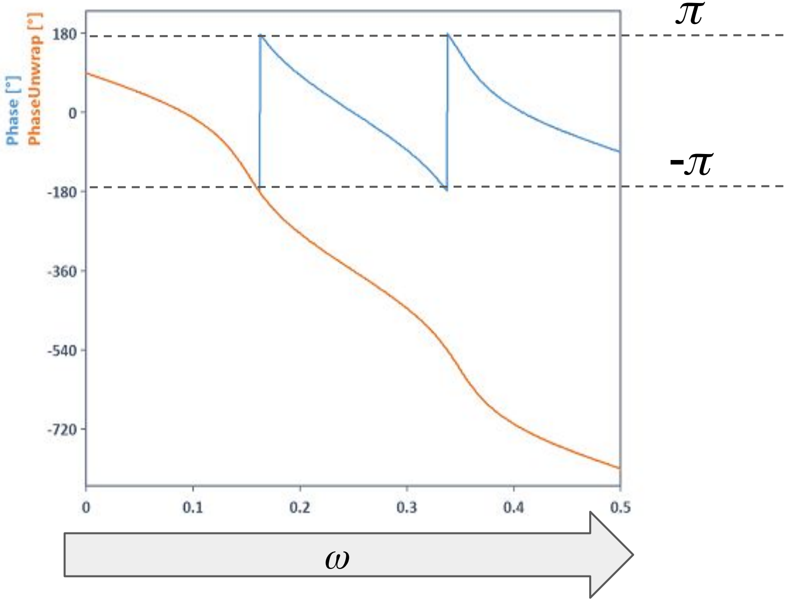
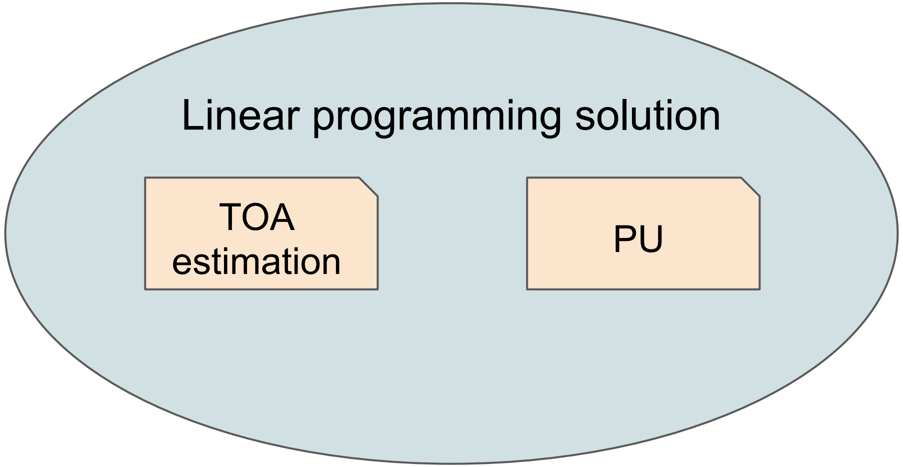


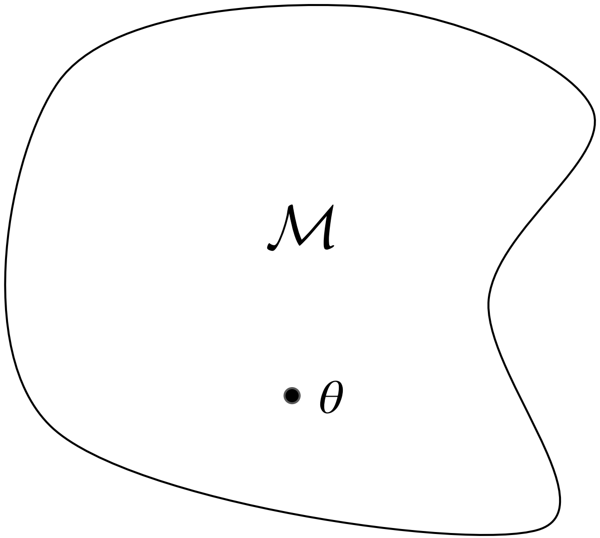
Image: https://www.weisang.com/en/documentation/phaseunwrap_en/

Motivation: A General Approach



Methodology

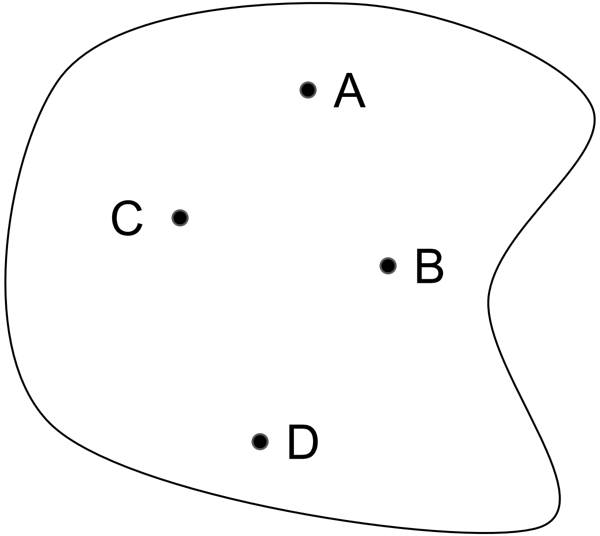
Sampling a Continuous Function f



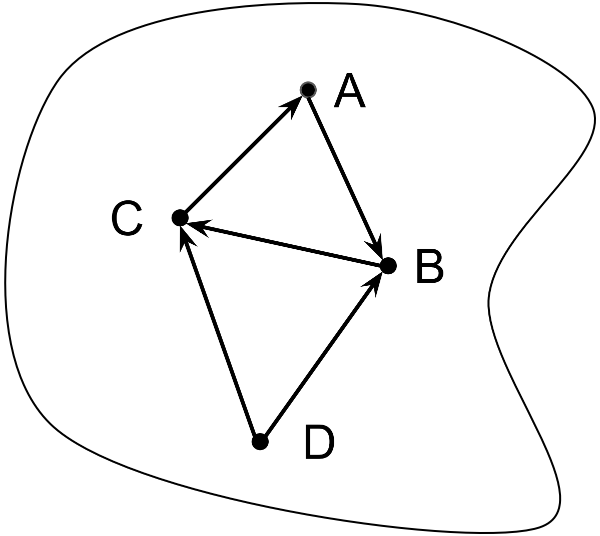
$$f(\theta), \theta \in \mathcal{M}$$

A four-point Toy Example

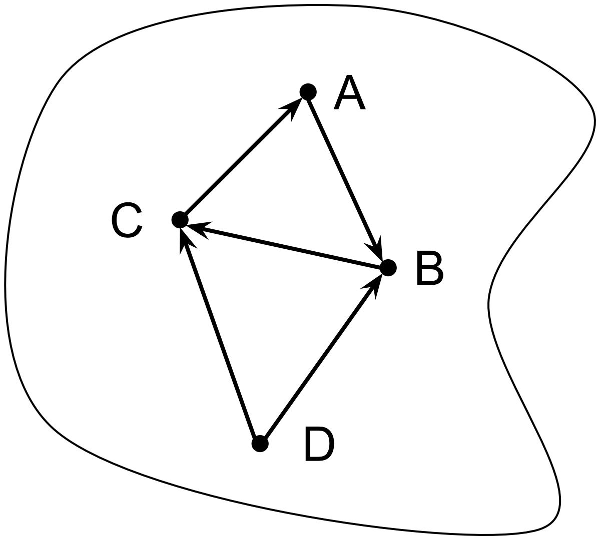
Direct estimation $f(\theta_i) \rightarrow$ 😞



Relative difference $f(\theta_j)$ - $f(\theta_i) \rightarrow$ 😊

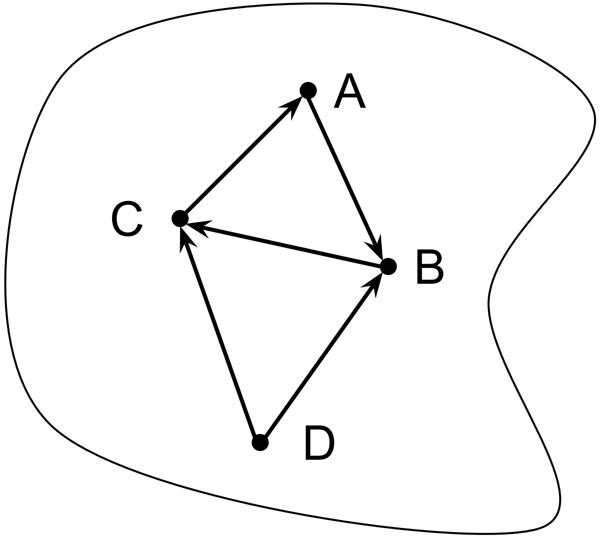


The Residuals **K**



$$\underbrace{\Gamma_{i,j}}_{\text{Measured}} = (f(\theta_j) - f(\theta_i)) + K_{i,j}$$

Expressing as Matrix Equation

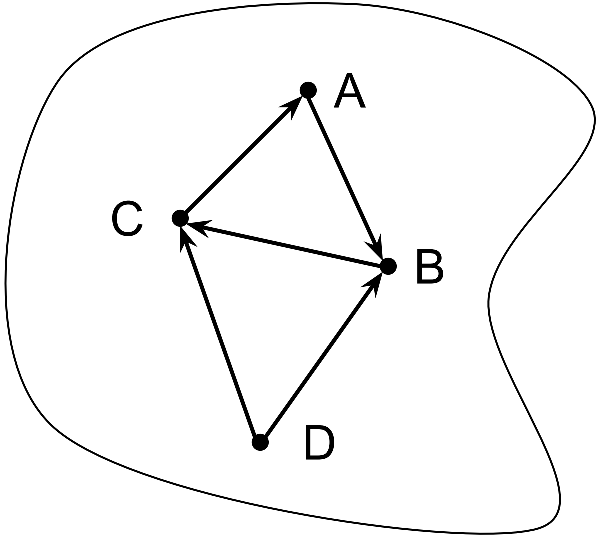


$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} f(A) \\ f(B) \\ f(C) \\ f(D) \end{bmatrix} + \begin{bmatrix} K_{C,A} \\ K_{A,B} \\ K_{B,C} \\ K_{D,B} \\ K_{D,C} \end{bmatrix} = \begin{bmatrix} \Gamma_{C,A} \\ \Gamma_{A,B} \\ \Gamma_{B,C} \\ \Gamma_{D,B} \\ \Gamma_{D,C} \end{bmatrix}$$

Variables Measurements

A. Piyush Shanker and Howard Zebker, "Edgelist phase unwrapping algorithm for time series InSAR analysis," J. Opt. Soc. Am. A 27, 605-612 (2010).

Solution with Linear Programming (LP)



$$\min_{\mathbf{k}} \underbrace{\mathbf{w}^T}_{\text{Weights}} |\mathbf{k}|, \ \mathbf{k} = \begin{bmatrix} K_{C,A} \\ K_{A,B} \\ K_{B,C} \\ K_{D,B} \\ K_{D,C} \end{bmatrix}$$

Comparing to Prior Work

Reijniers et al. $\min_{\mathbf{k}, f(\theta_*)} \underbrace{\mathbf{k}^T \operatorname{diag}(\mathbf{w}) \mathbf{k}}_{L_2} + \lambda \sum_{i=1}^{N} f(\theta_i)$

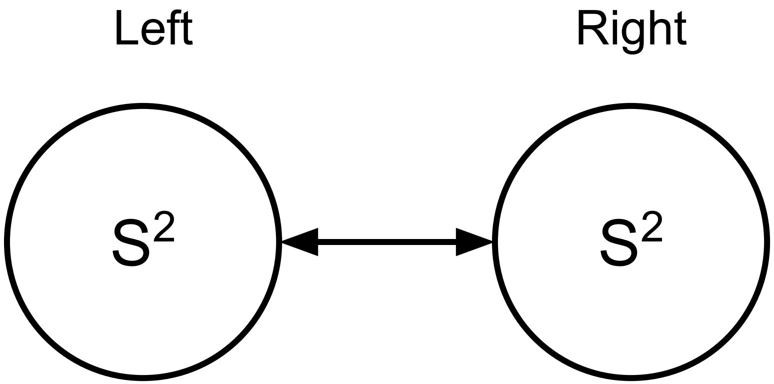
Ours $\min_{\mathbf{k}} \underbrace{\mathbf{w}^T |\mathbf{k}|}_{L_1}$

Reijniers, J., Partoens, B., and Peremans, H., "Noise-resistant correlation-based alignment of head-related transfer functions for high-fidelity spherical harmonics representation," in AES International Conference on Spatial and Immersive Audio, 2023.

Applications: Time-of-Arrival Estimation

$$\mathcal{M} \equiv \mathbb{S}^2 \times \{\text{Left}, \text{Right}\}$$
$$\Gamma_{i,j} \equiv \arg\max_t h_i[t] \star h_j[t]$$

Correlation indexes are **integers!**

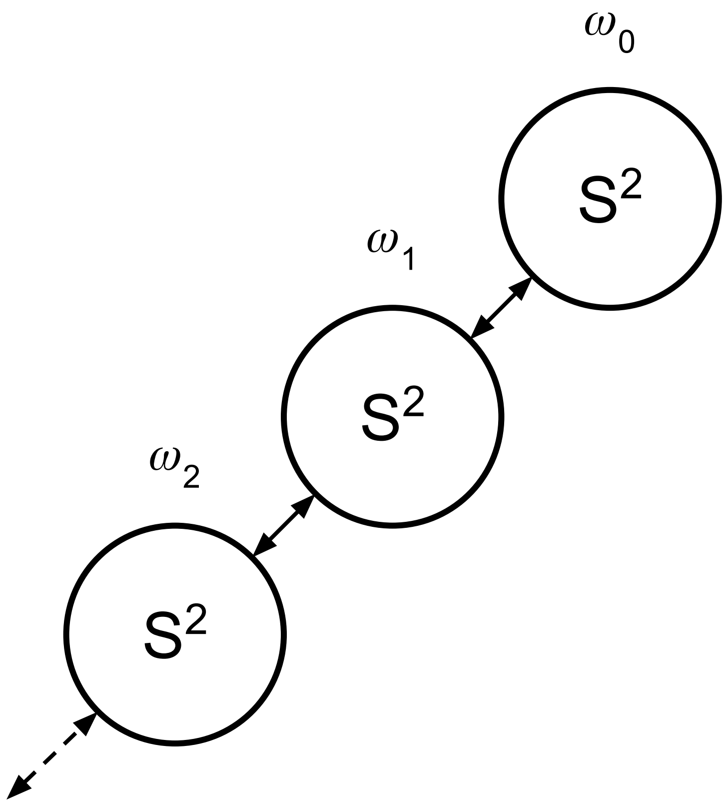


Applications: Phase Unwrapping

$$\mathcal{M} \equiv \mathbb{S}^2 \times [0, \pi]$$

$$\Gamma_{i,j} \equiv \frac{\mathrm{wrap}(\psi_j - \psi_i)}{2\pi}$$

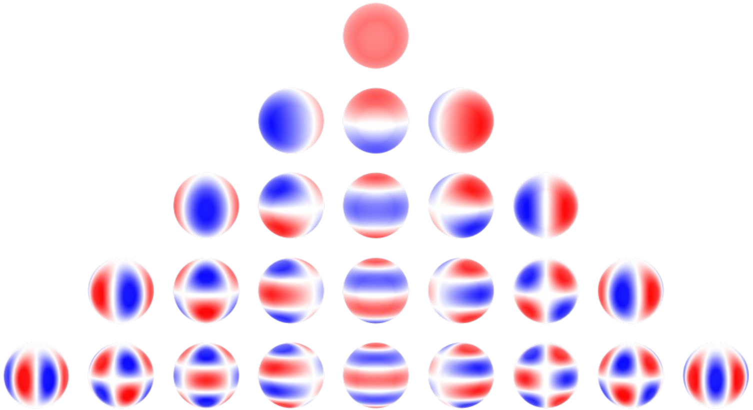
Phase jumps are **integers!**

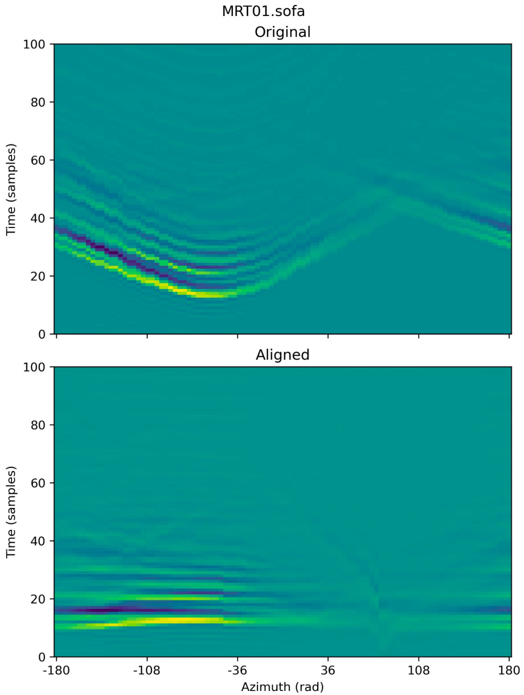


Evaluations

Feature Compression with Spherical Harmonics

- Interaural time delays (ITD)
- Aligned HRIRs





	AACHEN		ARI		CIPIC		RIEC	
Algorithm	**△ITD**	**LSD**	**△ITD**	**LSD**	**△ITD**	**LSD**	**△ITD**	**LSD**
Proposed EDGY	12.26	**3.80**	17.68	**4.08**	13.06	**3.34**	13.22	**3.35**
Proposed SIMP	12.30	**3.80**	17.70	**4.08**	13.09	**3.34**	13.26	**3.35**
LS	**11.32**	3.96	**17.22**	4.20	**12.42**	3.51	**12.65**	3.46

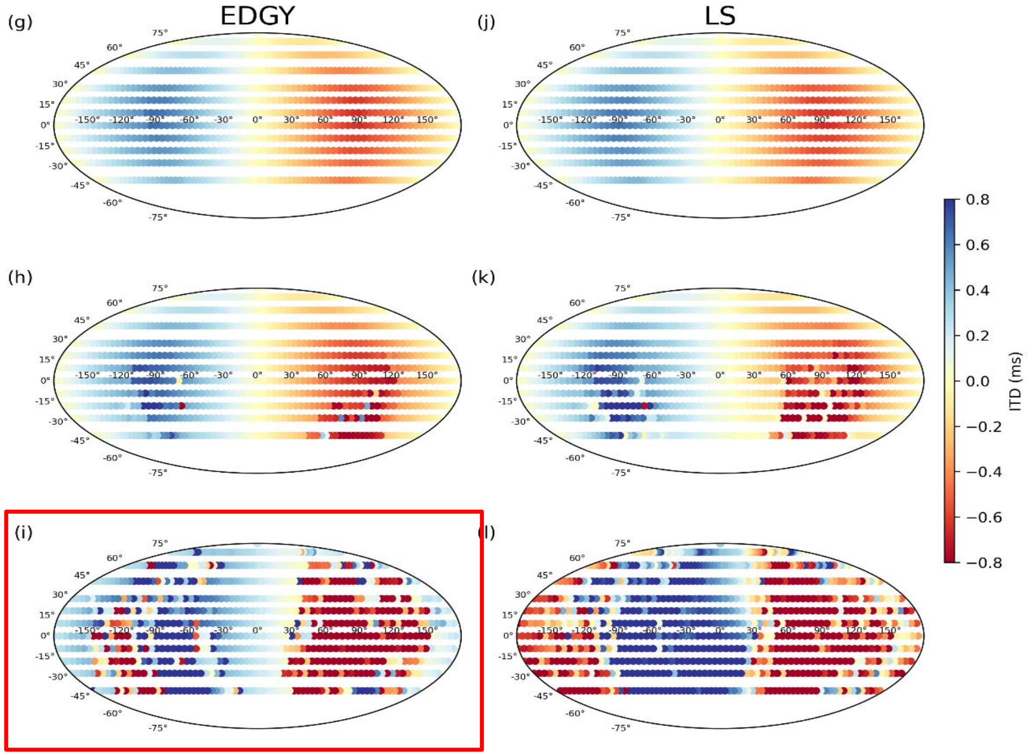
- ↓ITD distortion (μs) → smoother TOA
- ↓LSD (dB) → better alignment

ITD Noise Robustness

36 dB

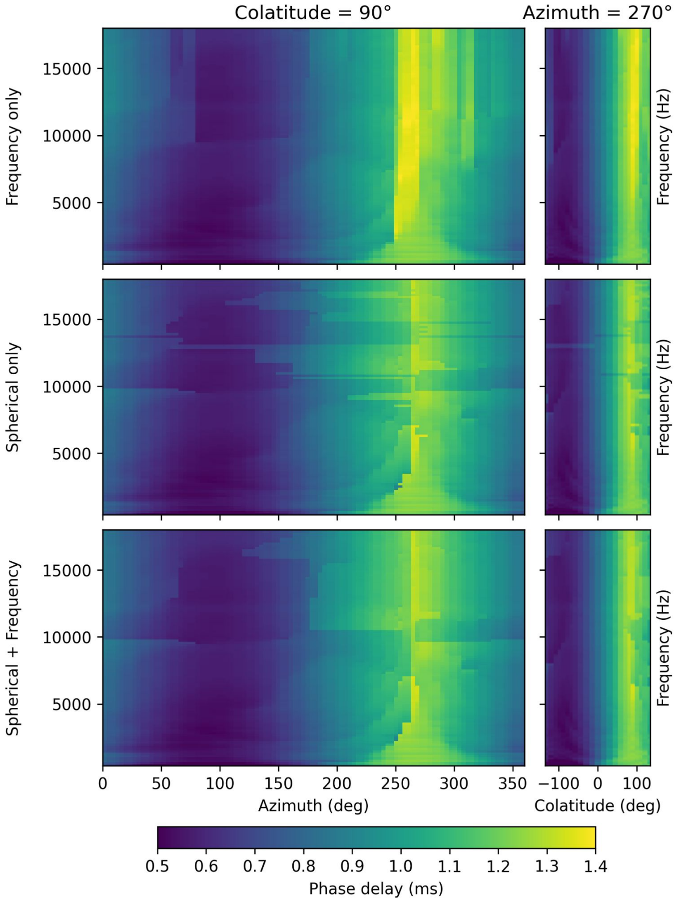
18 dB

10 dB



Proposed

Phase Unwrapping



$$\mathcal{M} \equiv [0, \pi]$$

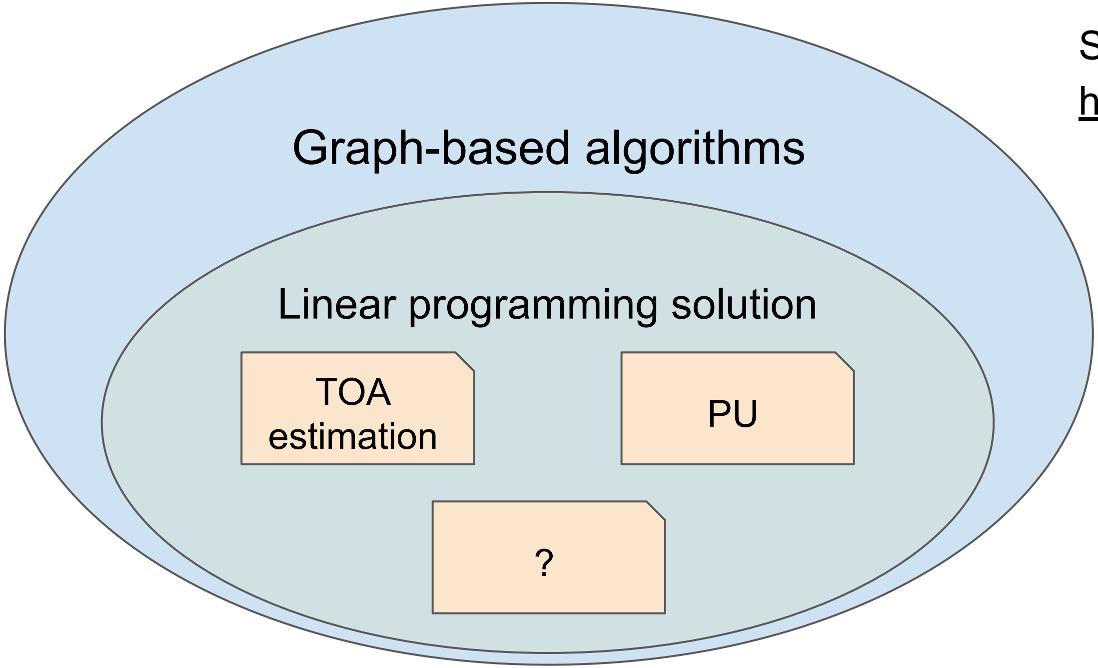
$$\mathcal{M} \equiv \mathbb{S}^2$$

$$\mathcal{M} \equiv \mathbb{S}^2 \times [0, \pi]$$

Summary

- One method solves two problems!
- More accurate and noise robust TOA
- Smoother HRTF phase response due to considering both dimensions

Future work



Source code:
https://github.com/yoyololicon/hrtf-ilp