



Dynamic Measurement of Binaural Room Impulse Responses Using an Optical Tracking System

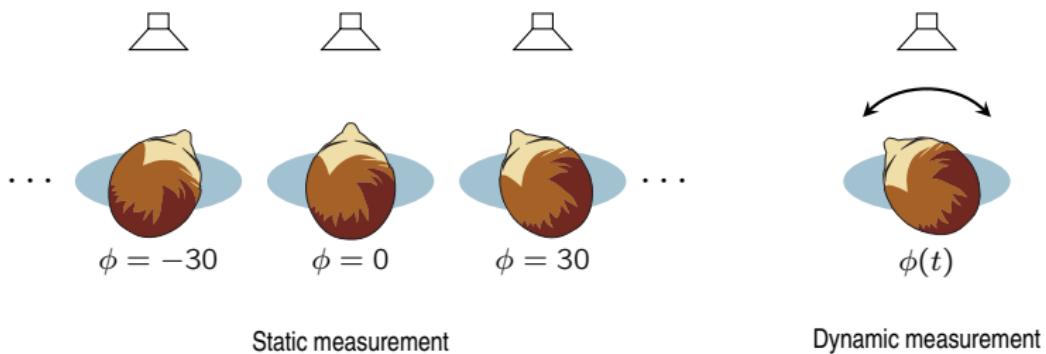
Nara Hahn, Wiebke Hahne, and Sascha Spors

Institute of Communications Engineering
University of Rostock, Germany

ICSA 4th International Conference on Spatial Audio

8. September 2017, Graz, Austria

Dynamic Measurement of HRIRs/BRIRs¹



- continuous head-movement during the measurement
- time-varying system identification problem

¹head-related impulse responses, binaural room impulse responses

Dynamic Measurement

Existing methods

- adaptive filtering [Antweiler and Enzner, 2009, He et al., 2016]
- analytic solution based on the projection-slice theorem for linear/circular trajectories [Ajdler et al., 2007]
- spatial interpolation [Hahn and Spors, 2014]

In the present work ...

- capture BRIRs for different head-orientations (azimuth-only)
- excitation by periodic perfect sequence
- time-varying system identification based on spatial interpolation
- optical tracking system
- comparison with a static measurement

Time-Invariant System Identification

- ear signal for fixed head orientation ϕ

$$\underbrace{y(\phi, n)}_{\text{ear signal}} = \sum_{k=0}^{N-1} \underbrace{x(n-k)}_{\text{excitation}} \underbrace{h(\phi, k)}_{\text{BRIR}}$$

- perfect sequence excitation (N -periodic, self-orthogonal)
e.g. maximum length sequence (MLS), perfect sweep

circular
auto-correlation

$$\sum_{n=0}^{N-1} x(n)x(n-m) = \delta(m \bmod N)$$

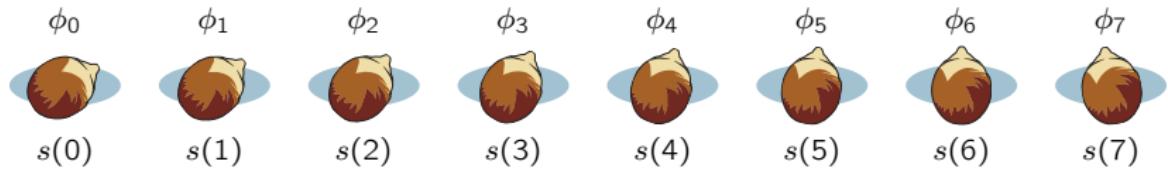
- impulse response

$$h(\phi, n) = \sum_{k=0}^{N-1} y(\phi, k)x(k-n) = y(\phi, n) \circledast_N x(-n)$$

\Rightarrow requires N samples of $y(\phi, n + \nu N)$, $n = 0, \dots, N - 1$, $\nu \in \mathbb{Z}$

Time-Varying System Identification

[Hahn and Spors, 2014]



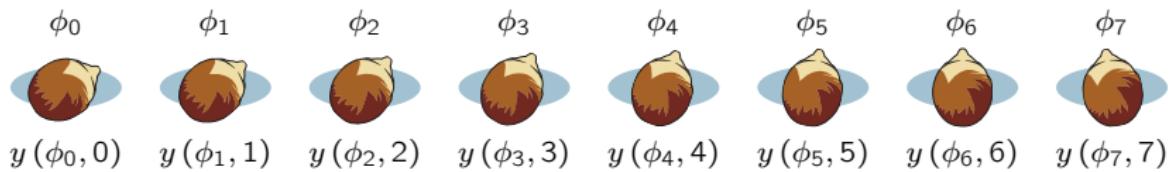
- ear signal captured by a rotating head

$$s(n) = y(\phi_n, n)$$

- constitutes a spatio-temporal sampling of $y(\phi, n)$
- only one sample available for each angle ϕ_n

Time-Varying System Identification

[Hahn and Spors, 2014]



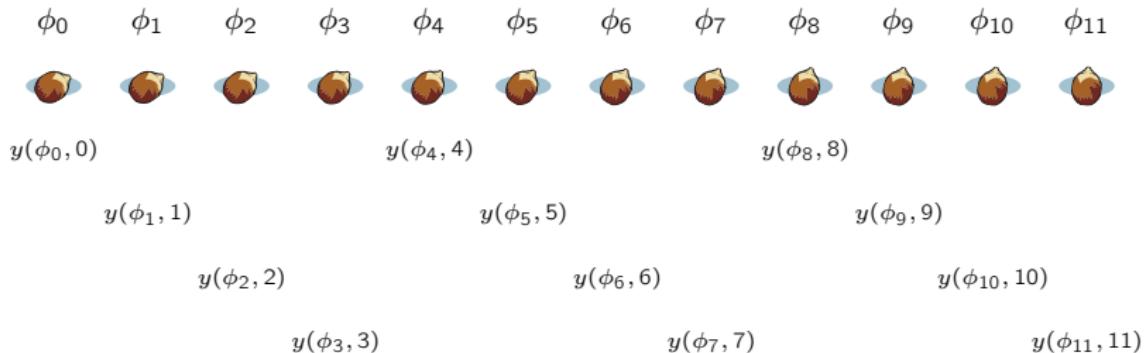
- ear signal captured by a rotating head

$$s(n) = y(\phi_n, n)$$

- constitutes a spatio-temporal sampling of $y(\phi, n)$
- only one sample available for each angle ϕ_n

Spatio-Temporal Sampling

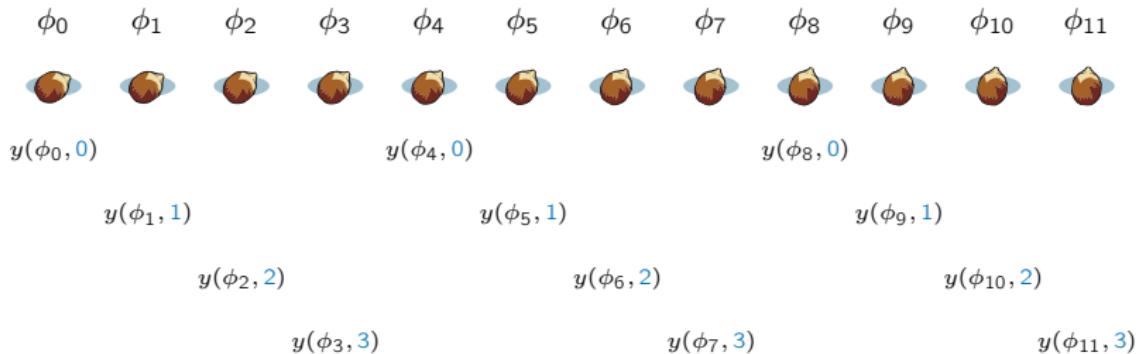
$$N = 4$$



- N -periodicity: $y(\phi_n, n) = y(\phi_n, n \bmod N)$
- $y(\phi, n)$ sampled at $\phi_{\{n+\mu N\}}$, $\mu \in \mathbb{Z}$
- spatial interpolation: $\hat{y}(\phi, n)$ estimated from $\{y(\phi_n, n + \mu N), \mu \in \mathbb{Z}\}$
- $\hat{h}(\phi, n) = \hat{y}(\phi, n) \circledast_N x(-n)$

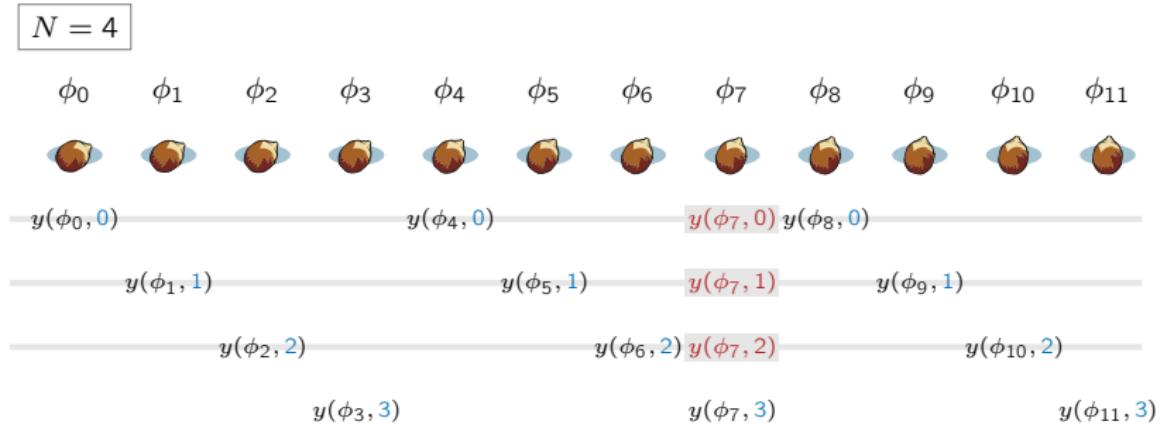
Spatio-Temporal Sampling

$$N = 4$$



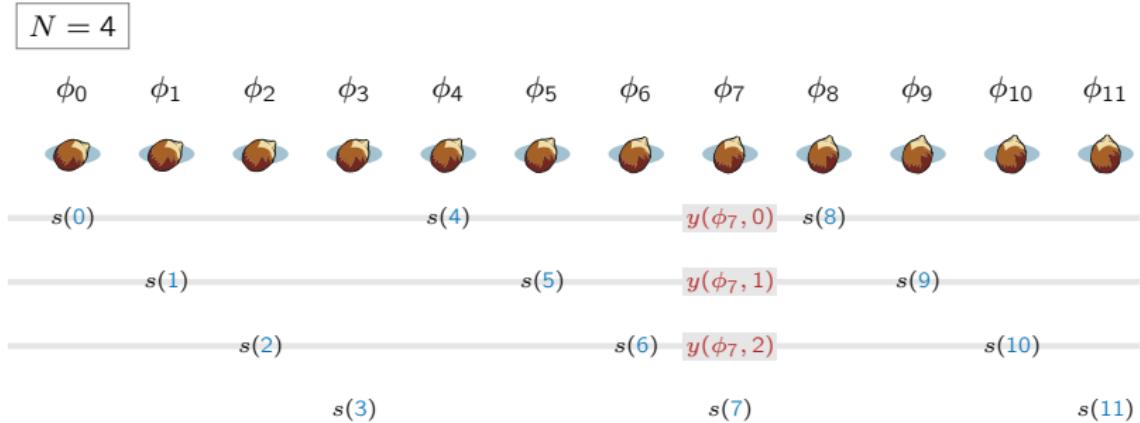
- N -periodicity: $y(\phi_n, n) = y(\phi_n, n \bmod N)$
- $y(\phi, n)$ sampled at $\phi_{\{n+\mu N\}}$, $\mu \in \mathbb{Z}$
- spatial interpolation: $\hat{y}(\phi, n)$ estimated from $\{y(\phi_n, n + \mu N), \mu \in \mathbb{Z}\}$
- $\hat{h}(\phi, n) = \hat{y}(\phi, n) \circledast_N x(-n)$

Spatio-Temporal Sampling



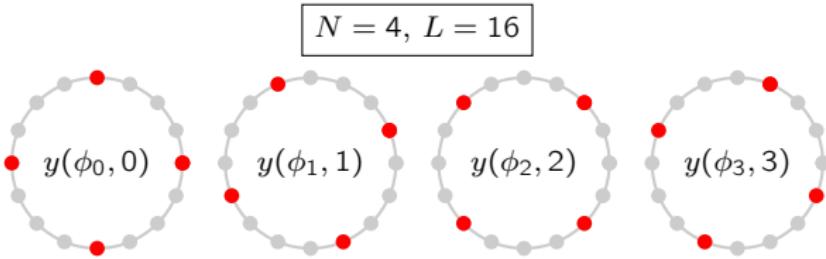
- N -periodicity: $y(\phi_n, n) = y(\phi_n, n \bmod N)$
- $y(\phi, n)$ sampled at $\phi_{\{n+\mu N\}}$, $\mu \in \mathbb{Z}$
- spatial interpolation: $\hat{y}(\phi, n)$ estimated from $\{y(\phi_n, n + \mu N), \mu \in \mathbb{Z}\}$
- $\hat{h}(\phi, n) = \hat{y}(\phi, n) \circledast_N x(-n)$

Spatio-Temporal Sampling



- N -periodicity: $y(\phi_n, n) = y(\phi_n, n \bmod N)$
- $y(\phi, n)$ sampled at $\phi_{\{n+\mu N\}}$, $\mu \in \mathbb{Z}$
- spatial interpolation: $\hat{y}(\phi, n)$ estimated from $\{s(n + \mu N), \mu \in \mathbb{Z}\}$
- $\hat{h}(\phi, n) = \hat{y}(\phi, n) \circledast_N x(-n)$

Anti-Aliasing Condition

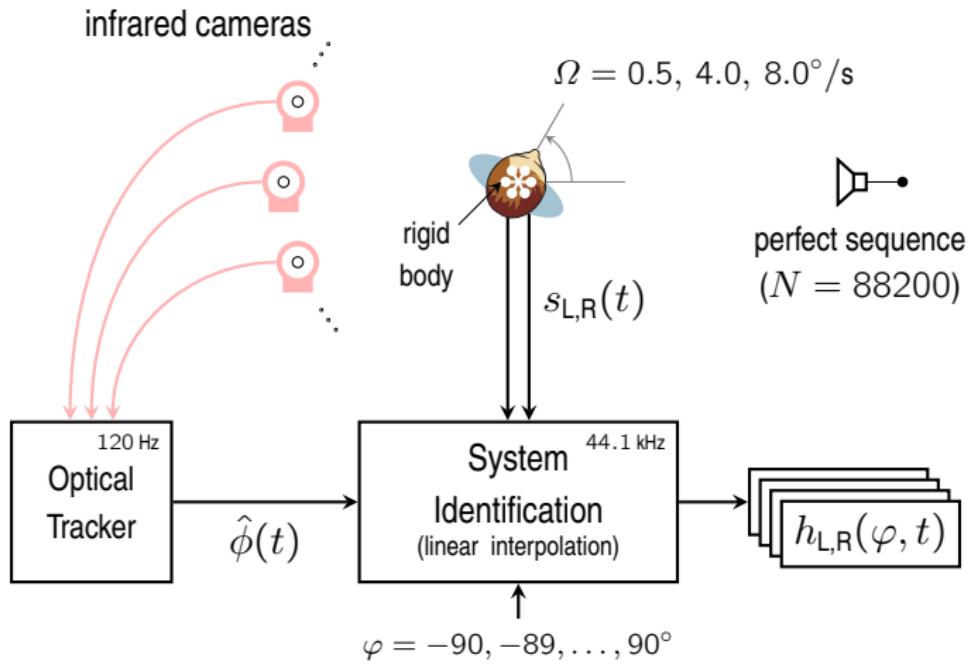


Uniform rotation on a circle $\phi(t) = \Omega t$

- number of captured samples: $L = \frac{2\pi}{\Omega} f_s$ (f_s : sampling rate)
- effective number of spatial sampling points: $M_{\text{eff}} = \frac{L}{N}$
- spatial bandwidth in the circular harmonics domain (approximated): $\frac{2\pi f}{c} R$ [Kuntz, 2008]
- anti-aliasing condition (maximum allowable angular speed) [Hahn and Spors, 2015]

$$\Omega < \frac{c}{RN}$$

Dynamic Measurement



- (approximate) anti-aliasing condition $\Omega < 2.2^\circ/\text{s}$

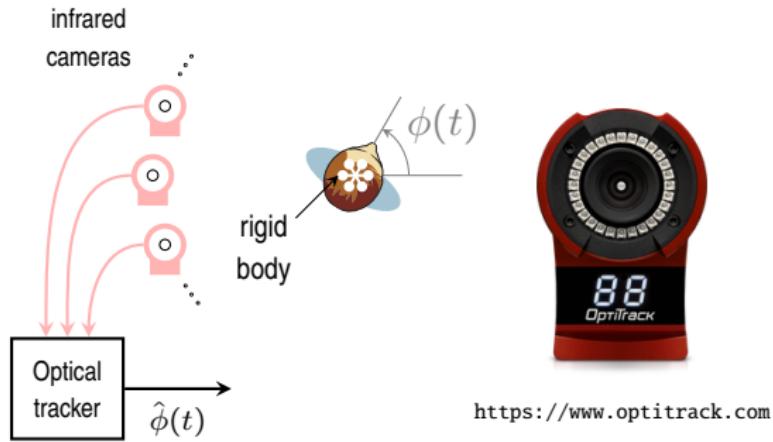
$\Omega (\text{ }^\circ/\text{s})$	M_{eff}	SNR (dB)
0.5	200	30.0
4.0	25	17.1
8.0	12	13.1

- system distance

$$\mathcal{D}(\varphi) = 10 \log_{10} \left[\frac{\sum_{n=0}^{N-1} |\hat{h}(\varphi, n) - h(\varphi, n)|^2}{\langle \sum_{n=0}^{N-1} |h(\varphi, n)|^2 \rangle_\varphi} \right] \text{dB}$$

$h(\varphi, n)$: static measurement, $\hat{h}(\varphi, n)$: dynamic measurement, $\langle \cdot \rangle_\varphi$: average over φ

Optical Tracking System

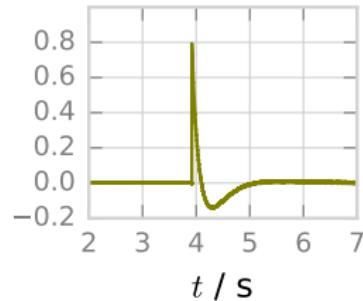
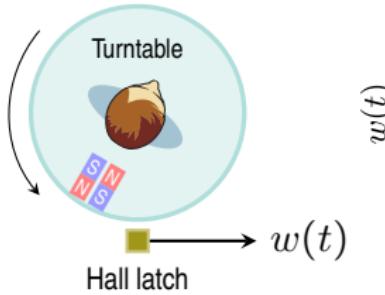
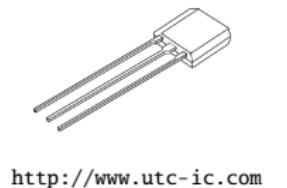


<https://www.optitrack.com>



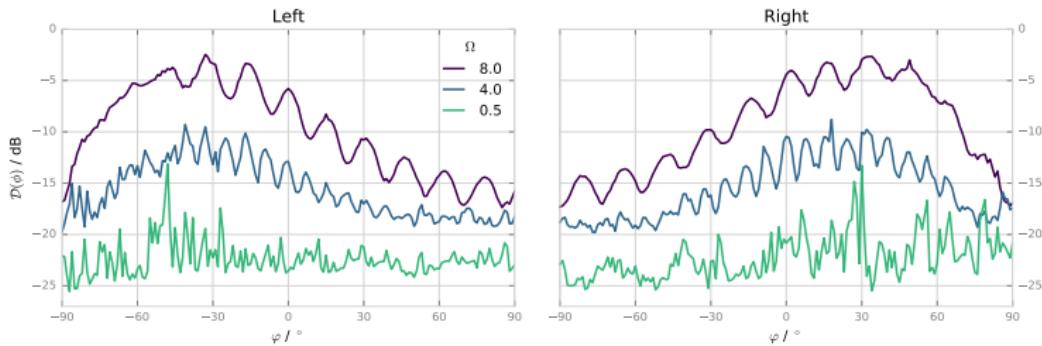
- 12 cameras (OptiTrack Flex 13)
 - rigid body consisting of 6 retro-reflective markers
 - Motive (motion capture software)
- ⇒ position & orientation streamed via TCP/IP

Bipolar Hall-effect sensor



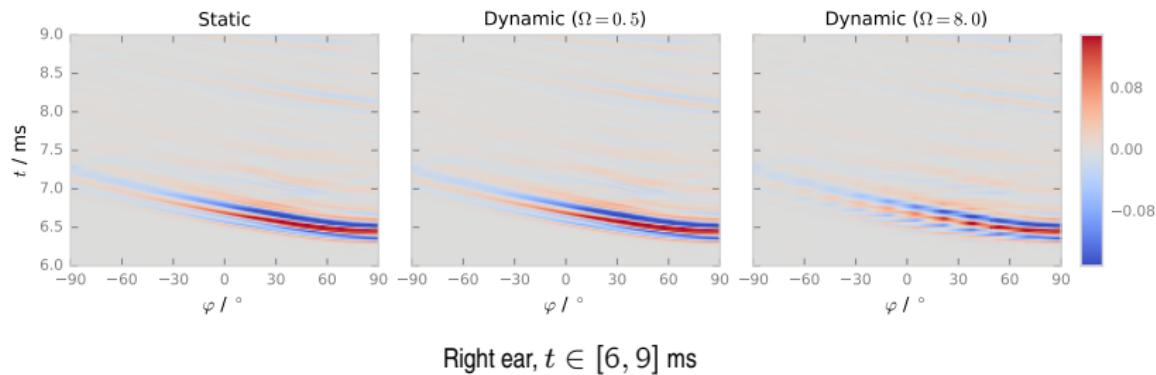
- Hall-effect sensor (Unisonic U18)
 - 2 permanent magnets (opposite polarities) at a reference angle ϕ_{ref}
 - sensor output $w(t)$ recorded synchronously with the audio signal $s(n)$
- ⇒ time-alignment of $\hat{\phi}(t)$ and $s(t)$

System Distance



- higher angular speed results in more deviations
- due to lower SNR and fewer spatial sampling points
- $\Omega = 4.0, 8.0$: depends on the system variability
- $\Omega = 0.2$: exhibits relatively uniform performance

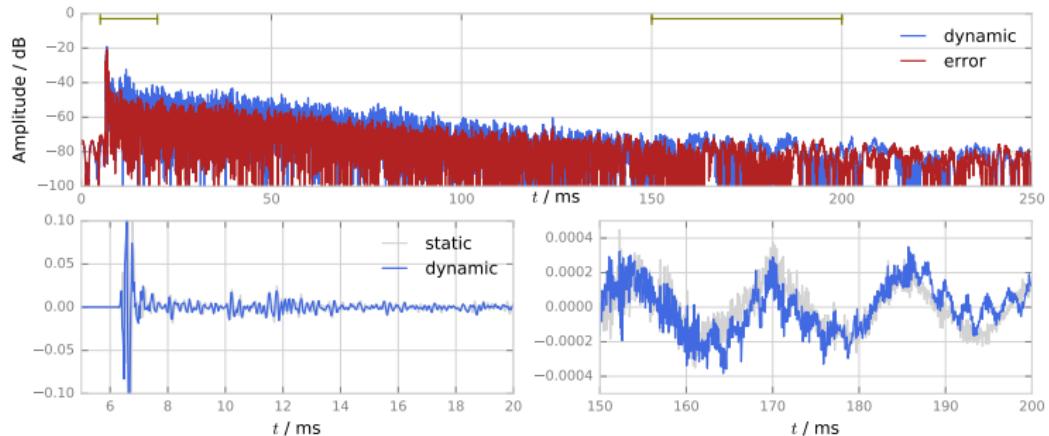
Spatial Aliasing Artifacts



- $\Omega = 0.5$: spatio-temporal structure successfully recovered
- $\Omega = 8.0$: identified as a piece-wise time-invariant system

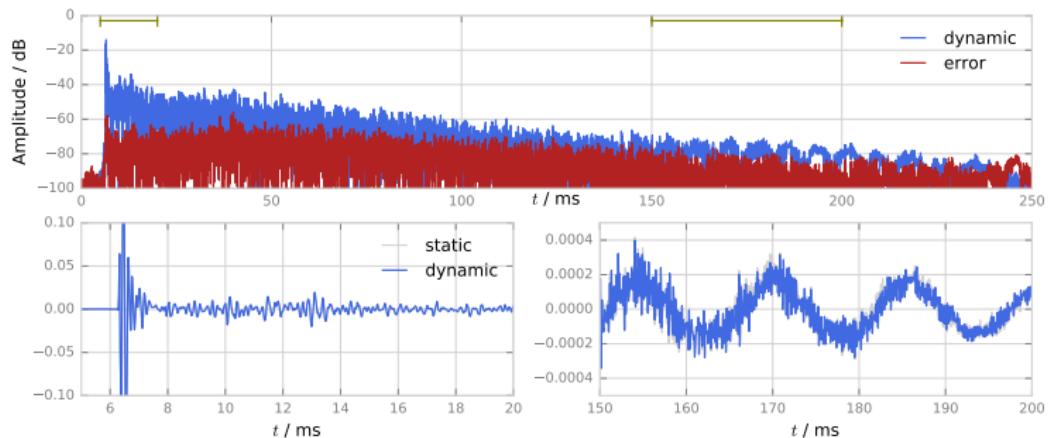
BIRRs - Worst Case

$$\Omega = 8.0^\circ/\text{s} \quad (\text{Left ear}, \varphi = -33, \mathcal{D}(\varphi) = -2.5 \text{ dB})$$



BIRRs - Best Case

$$\Omega = 0.5^\circ/\text{s} \quad (\text{Left ear}, \varphi = -87, \mathcal{D}(\varphi) = -25.5 \text{ dB})$$



Informal Listening

- speech [EBU, 2008], castanets [Frank, 2009]
- azimuthal localization not impaired in all cases
- $\Omega = 4.0, 8.0$: artifacts and coloration
- $\Omega = 0.5$: indistinguishable
- <http://spatialaudio.net/dynamic-brir-measurement/>

Conclusion

Summary

- dynamic measurement of BRIRs in a real room
- captured signal is a spatio-temporally sampled → interpolation
- anti-aliasing condition has to be met

Outlook

- individual HRIR/BRIR measurement
- extension to 2D and 3D trajectories
- coming soon!



<https://www.optitrack.com>

Hahn and Spors (2017): "*Continuous Measurement of Spatial Room Impulse Responses Using a Non-Uniformly Moving Microphone*", WASPAA.

THANK YOU!

References I

-  Ajdler, T., Sbaiz, L., and Vetterli, M. (2007).
Dynamic measurement of room impulse responses using a moving microphone.
J. Acoust. Soc. Ame. (JASA), 122(3):1636–1645.
-  Antweiler, C. and Enzner, G. (2009).
Perfect Sequence LMS for Rapid Acquisition of Continuous-azimuth Head Related Impulse Responses.
In *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, pages 281–284, New Paltz, NY, USA.
-  EBU (2008).
Sound quality assessment material recordings for subjective tests.
Compact Disc.
-  Frank, M. (2009).
Castanets dry.
Audio.
-  Hahn, N. and Spors, S. (2014).
Identification of dynamic acoustic systems by orthogonal expansion of time-variant impulse responses.
In *Proc. 6th Int. Symp. on Commun. Control Signal Process. (ISCCSP)*, Athens, Greek.
-  Hahn, N. and Spors, S. (2015).
Continuous measurement of impulse responses on a circle using a uniformly moving microphone.
In *Proc. the Eur. Signal Process. Conf. (EUSIPCO)*, Nice, France.
-  Hahn, N. and Spors, S. (2016).
Analysis of Time-Varying System Identification Using Normalized Least Mean Square (NLMS) in the Context of Data-Based Binaural Synthesis.
In *Proc. of the 42nd German Annual Conference on Acoustics (DAGA)*, Aachen, Germany.

References II



Hahn, N. and Spors, S. (2017).

Spatial aliasing in continuous measurement of spatial room impulse responses.

In *Proc. 43rd German Annu. Conf. on Acoust. (DAGA)*, Kiel, Germany.



He, J., Ranjan, R., and Gan, W.-S. (2016).

Fast continuous HRTF acquisition with unconstrained movements of human subjects.

In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pages 321–325. IEEE.



Kuntz, A. (2008).

Wave Field Analysis using Virtual Circular Microphone Arrays.

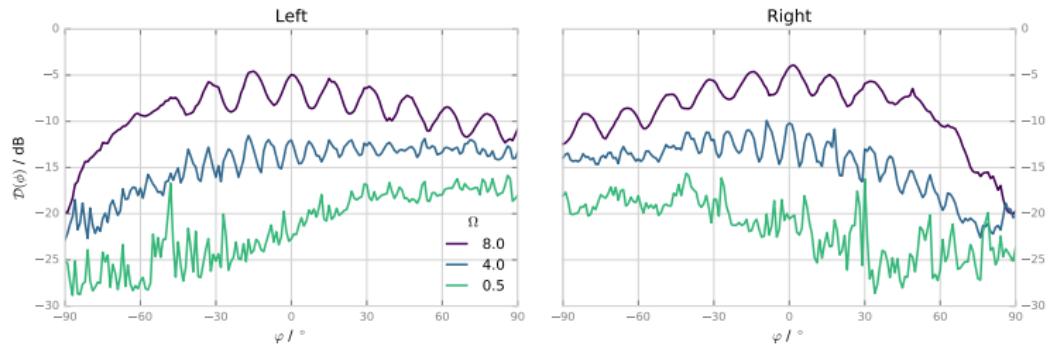
Verlag Dr. Hut.

Spatial Interpolation

[Hahn and Spors, 2015, Hahn and Spors, 2016]

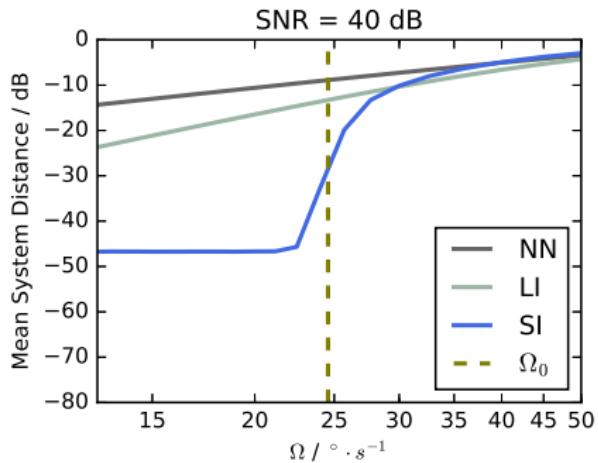
Approach	Equivalent interpolation filter
cross-correlation	rectangular function (Nearest neighbour)
NLMS ($\mu = 1$)	rectangular function (Nearest neighbour)
NLMS ($\mu \neq 1$)	rectangular pulses
projection-slice approach	sinc or periodic sinc

NMSE



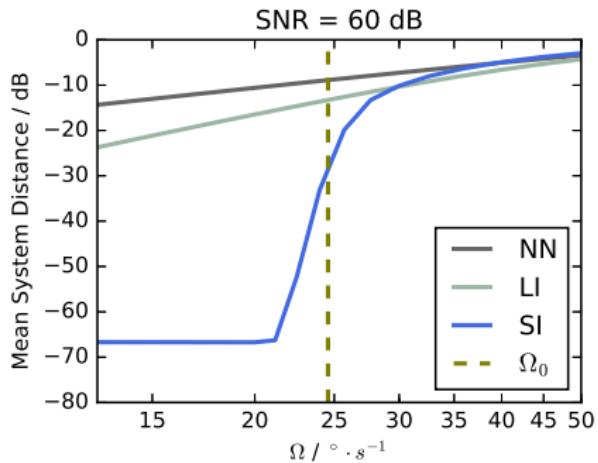
System Distance

[Hahn and Spors, 2017]



System Distance

[Hahn and Spors, 2017]



System Distance

[Hahn and Spors, 2017]

