RUHR-UNIVERSITÄT BOCHUM

RUB

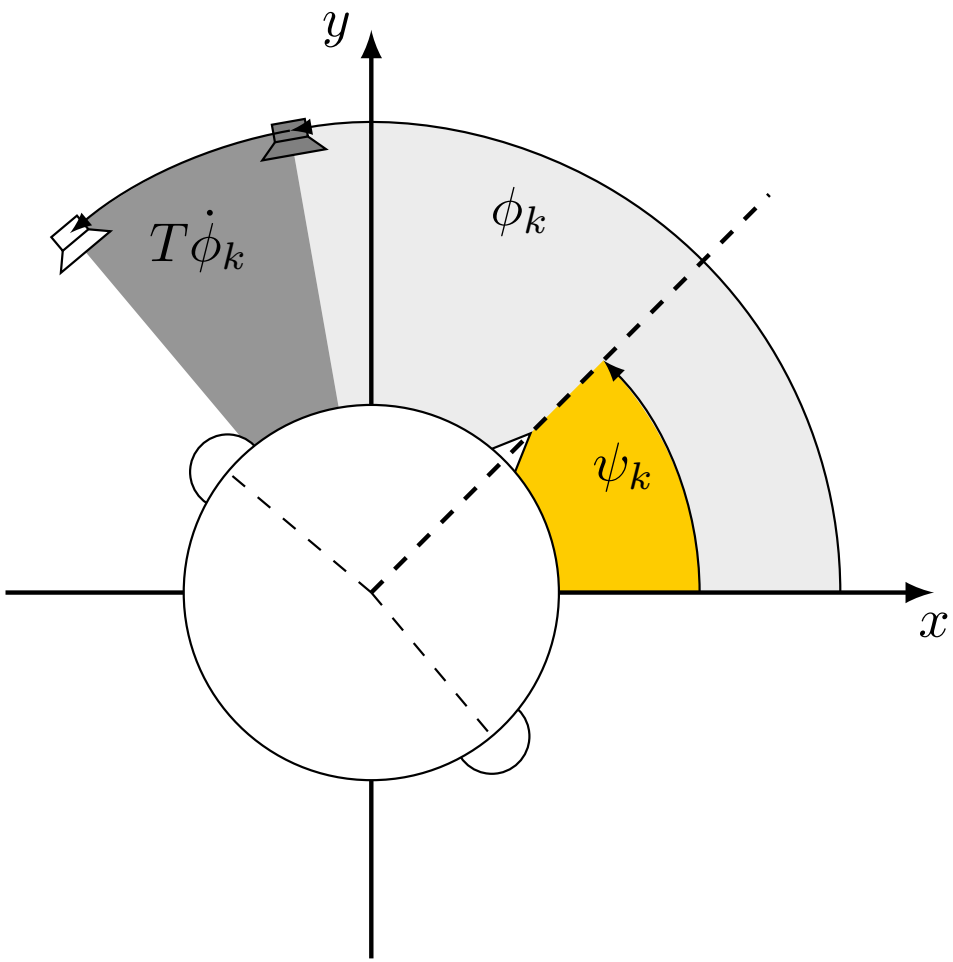
Binaural Sound Source Localisation and Tracking using a Dynamic Spherical Head Model

Christopher Schymura, **Fiete Winter, Dorothea Kolossa, Sascha Spors**

September 7, 2015

Introduction

Task: Tracking a moving sound source



Introduction

Some existing approaches for sound source tracking:

- [Portello et al. (2011), Traa & Smaragdis (2013)] using Kalman filters
- [Ward et al. (2003), Lu & Cooke (2011)] using particle filters

Introduction

Some existing approaches for sound source tracking:

- [Portello et al. (2011), Traa & Smaragdis (2013)] using Kalman filters
- [Ward et al. (2003), Lu & Cooke (2011)] using particle filters

Head rotations can improve localisation by resolving front-back ambiguities (see [Wallach (1940), Blauert(1997)]).

Introduction

Some existing approaches for sound source tracking:

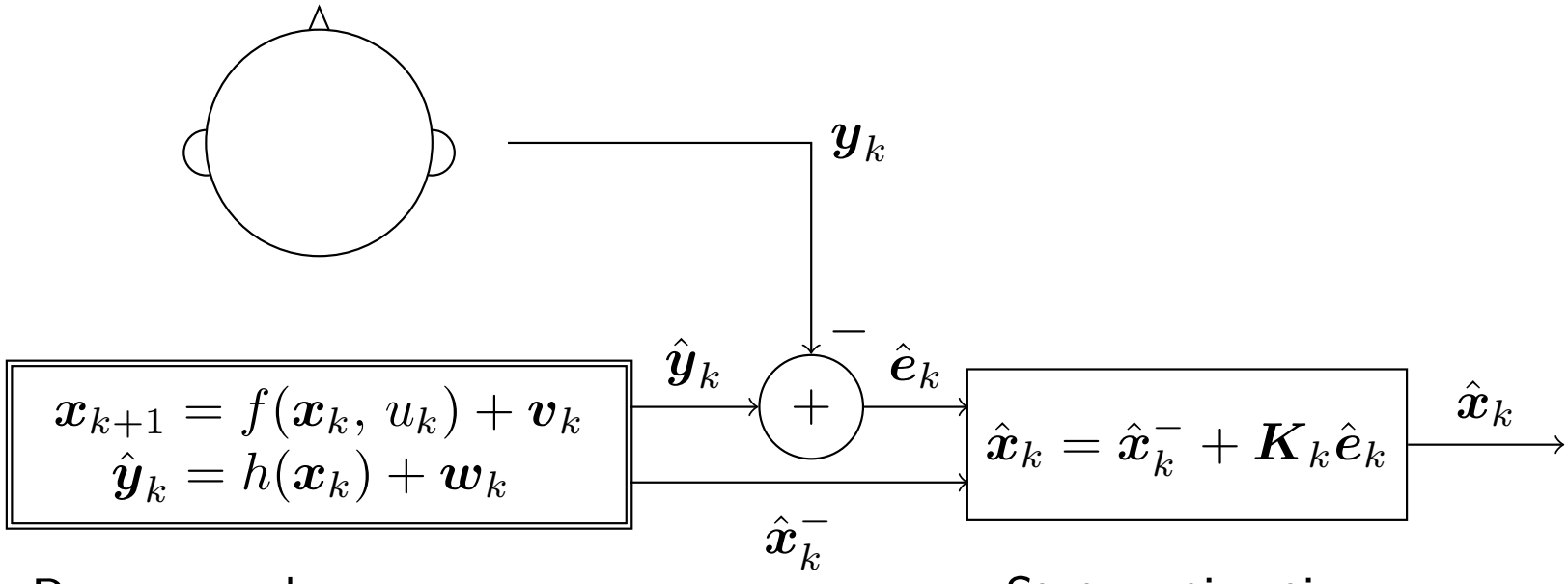
- [Portello et al. (2011), Traa & Smaragdis (2013)] using Kalman filters
- [Ward et al. (2003), Lu & Cooke (2011)] using particle filters

Head rotations can improve localisation by resolving front-back ambiguities (see [Wallach (1940), Blauert(1997)]).

Computational models investigating the effects of head movements:

- [Schymura et al. (2014), May et al. (2015), Ma et al. (2015)]

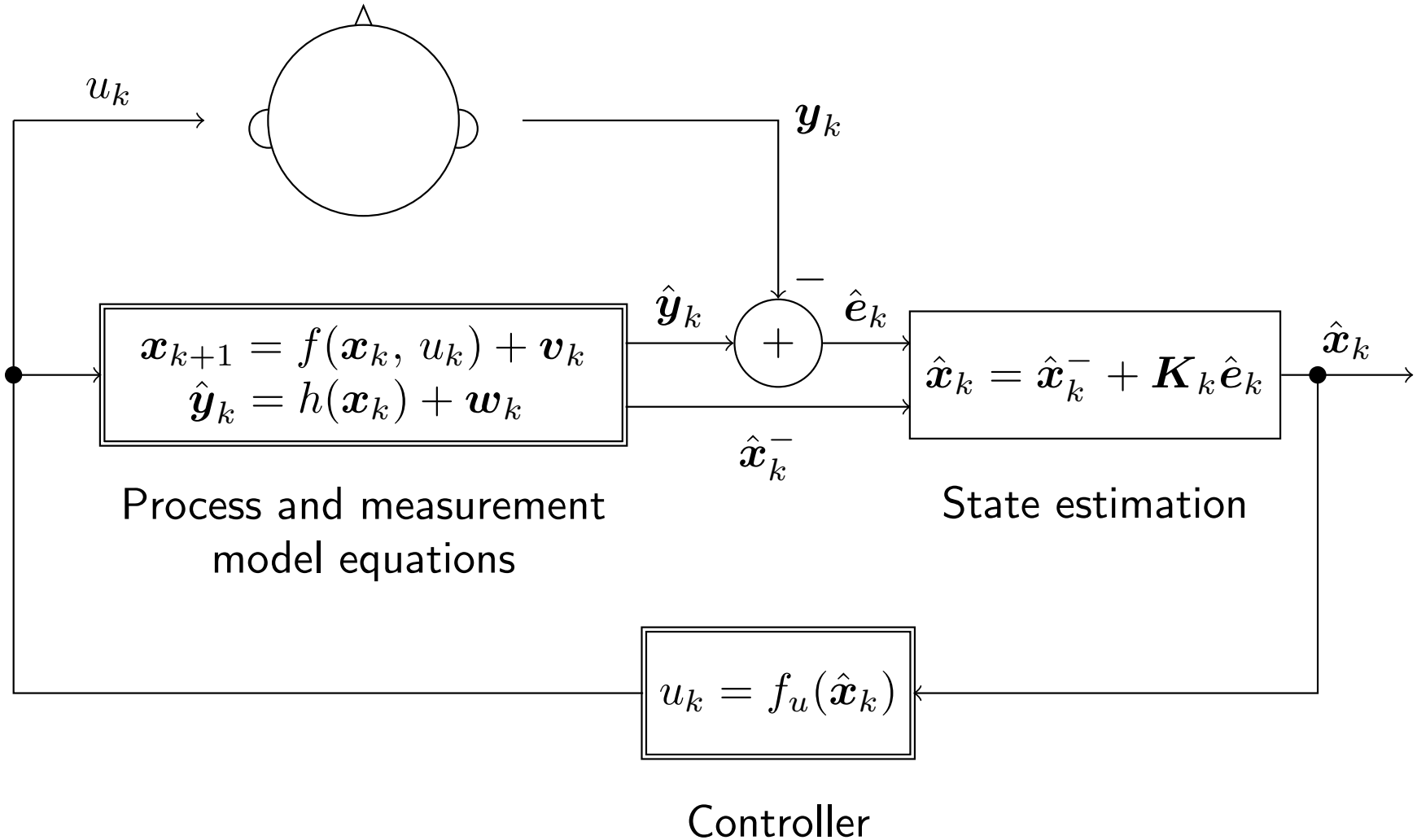
System overview



Process and measurement model equations

State estimation

System overview



Process model

State space:

$$\boldsymbol{x}_k = \begin{bmatrix} \phi_k & \dot{\phi}_k & \psi_k \end{bmatrix}^T$$

Process model

State space:

$$\boldsymbol{x}_k = \begin{bmatrix} \phi_k & \dot{\phi}_k & \psi_k \end{bmatrix}^T$$

Process model:

$$\boldsymbol{x}_{k+1} = \begin{bmatrix} \phi_{k+1} \\ \dot{\phi}_{k+1} \\ \psi_{k+1} \end{bmatrix} = \begin{bmatrix} \phi_k + T\dot{\phi}_k + v_{\phi,\,k} \\ \dot{\phi}_k + v_{\dot{\phi},\,k} \\ \mathrm{sat}(\psi_k + T\dot{\psi}_{\max} u_k,\, \psi_{\max}) + v_{\psi,\,k} \end{bmatrix}$$

Process model

State space:

$$\boldsymbol{x}_k = \begin{bmatrix} \phi_k & \dot{\phi}_k & \psi_k \end{bmatrix}^T$$

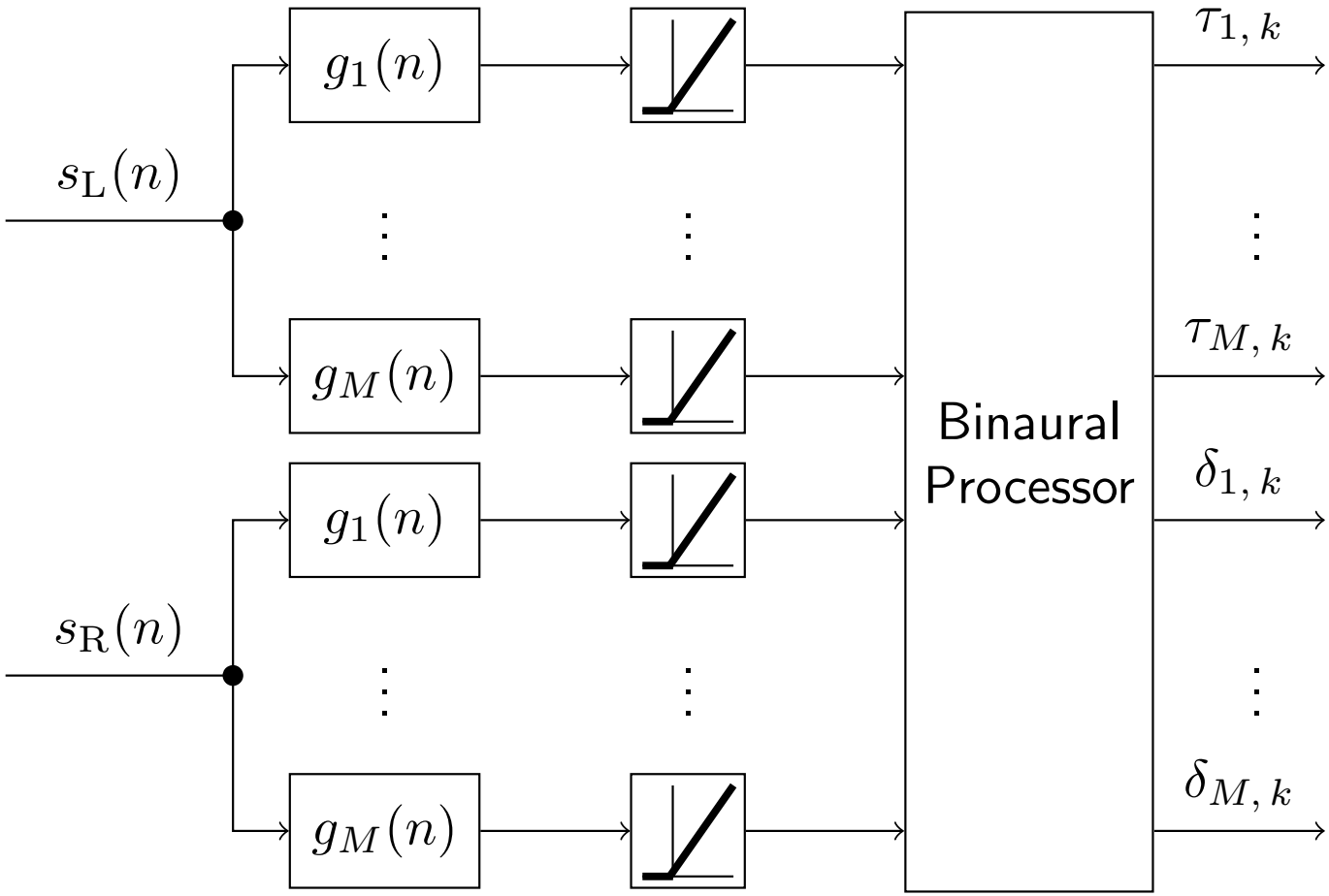
Process model:

$$\boldsymbol{x}_{k+1} = \begin{bmatrix} \phi_{k+1} \\ \dot{\phi}_{k+1} \\ \psi_{k+1} \end{bmatrix} = \begin{bmatrix} \phi_k + T\dot{\phi}_k + v_{\phi,\,k} \\ \dot{\phi}_k + v_{\dot{\phi},\,k} \\ \mathrm{sat}(\psi_k + T\dot{\psi}_{\max} u_k,\, \psi_{\max}) + v_{\psi,\,k} \end{bmatrix}$$

$$v_{\phi,\,k} \sim \mathcal{N}(0,\, \sigma_\phi^2),\;\; v_{\dot{\phi},\,k} \sim \mathcal{N}(0,\, \sigma_{\dot{\phi}}^2),\;\; v_{\psi,\,k} \sim \mathcal{N}(0,\, \sigma_\psi^2)$$

$$\mathrm{sat}(x,\, x_{\max}) = \min(|x|,\, x_{\max}) \cdot \mathrm{sgn}(x), \quad u_k \in [-1,\, 1]$$

Binaural front-end



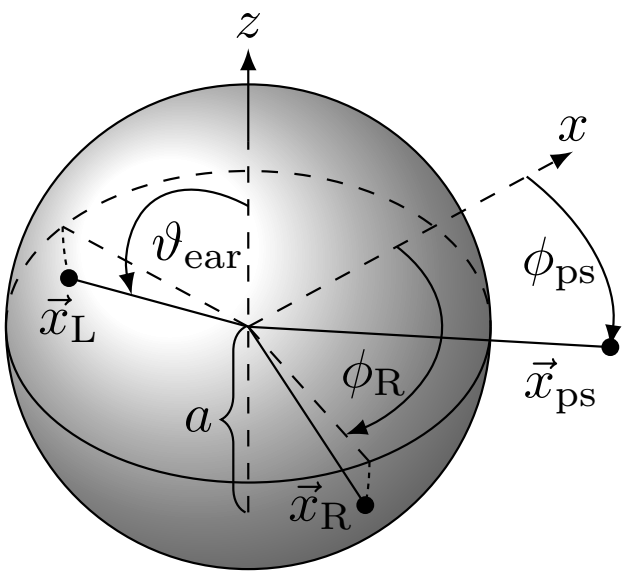
$$\boldsymbol{y}_k = \left[\tau_{1,\,k}, \cdots, \tau_{M,\,k},\, \delta_{1,\,k}, \cdots,\, \delta_{M,\,k}\right]^T$$

Measurement model

Spherical head model [Brungart (1999), Algazi et al. (2001)]:

$$R_i(\boldsymbol{x}_k, \omega) = \frac{c}{4\pi\omega a^2} \sum_{\nu=0}^{\infty} \frac{h_\nu\left(\frac{\omega}{c}d\right)}{h'_\nu\left(\frac{\omega}{c}a\right)} (2\nu+1)\, L_\nu\Big(\sin(\vartheta_{\text{ear}})\cos\left(\phi_k - \psi_k - \phi_i\right)\Big)$$

$$i \in \{\text{R, L}\}$$

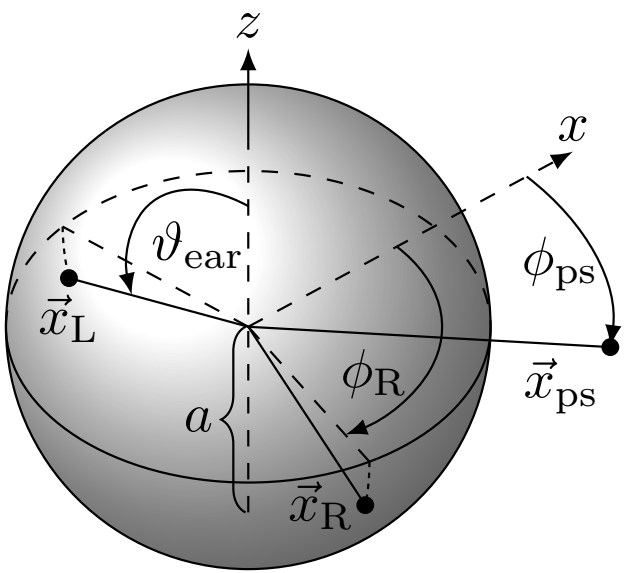


Measurement model

Spherical head model [Brungart (1999), Algazi et al. (2001)]:

$$R_i(\boldsymbol{x}_k, \omega) = \frac{c}{4\pi\omega a^2} \sum_{\nu=0}^{\infty} \frac{h_\nu\left(\frac{\omega}{c}d\right)}{h'_\nu\left(\frac{\omega}{c}a\right)} (2\nu+1)\, L_\nu\Big(\sin(\vartheta_{\text{ear}})\cos\left(\phi_k - \psi_k - \phi_i\right)\Big)$$

$$i \in \{\text{R, L}\}$$



Spherical head parameters, taken from [Algazi et al. (2001)]:

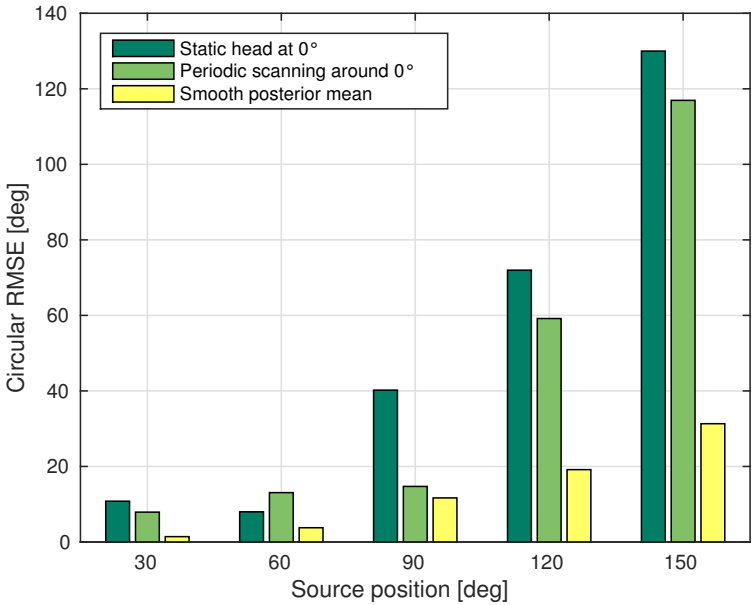
- Head radius a: $8.5\,\text{cm}$
- Ear's azimuth angle ϕ_i: $93.60°$
- Ear's polar angle ϑ_{ear}: $110.67°$

Head rotation strategies

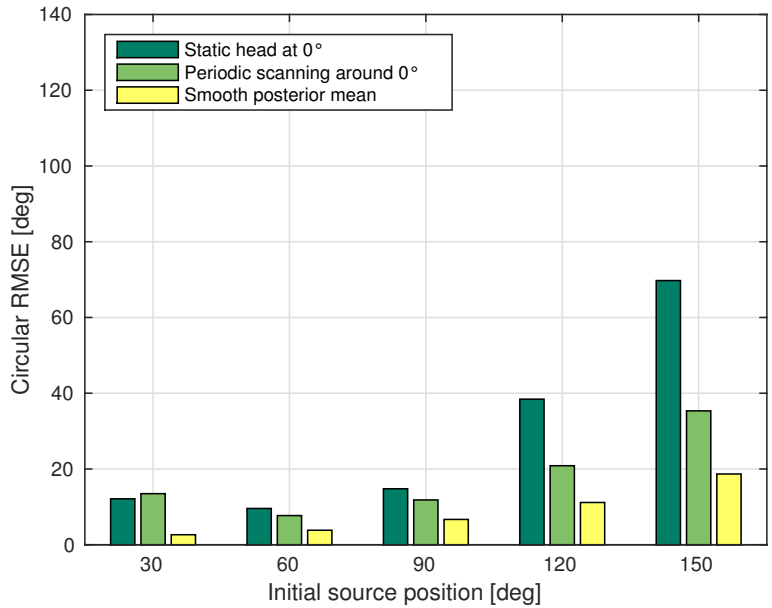
Evaluation of three different approaches:

	No head rotation	Periodic sweeping	Smooth posterior mean
f_u	0	$\sin\left(2\pi k \frac{T}{T_\mathrm{p}}\right)$	$\left(\frac{\lvert\phi_k - \psi_k\rvert}{1+\lvert\phi_k - \psi_k\rvert}\right) \mathrm{sgn}\left(\phi_k - \psi_k\right)$
Type	-	feed-forward	feedback

Evaluation results



Static scenario



Dynamic scenario

Evaluation metric:

$$\mathrm{cRMSE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\min_{l\in\mathbb{Z}}\left(\hat{\phi}_k - \phi_k + 2\pi l\right)^2}$$

Summary

- A binaural model for localisation and tracking of moving sound sources using continuous head rotations was proposed.

Summary

- A binaural model for localisation and tracking of moving sound sources using continuous head rotations was proposed.
- The model allows for treating the localisation task as a closed-loop control problem.

Summary

- A binaural model for localisation and tracking of moving sound sources using continuous head rotations was proposed.
- The model allows for treating the localisation task as a closed-loop control problem.
- Future extensions of the model may aim at investigating alternative control strategies, estimation of source distance and introducing additional degrees of freedom (e.g. translatory movements).

Summary

- A binaural model for localisation and tracking of moving sound sources using continuous head rotations was proposed.
- The model allows for treating the localisation task as a closed-loop control problem.
- Future extensions of the model may aim at investigating alternative control strategies, estimation of source distance and introducing additional degrees of freedom (e.g. translatory movements).

Thank you for your attention!