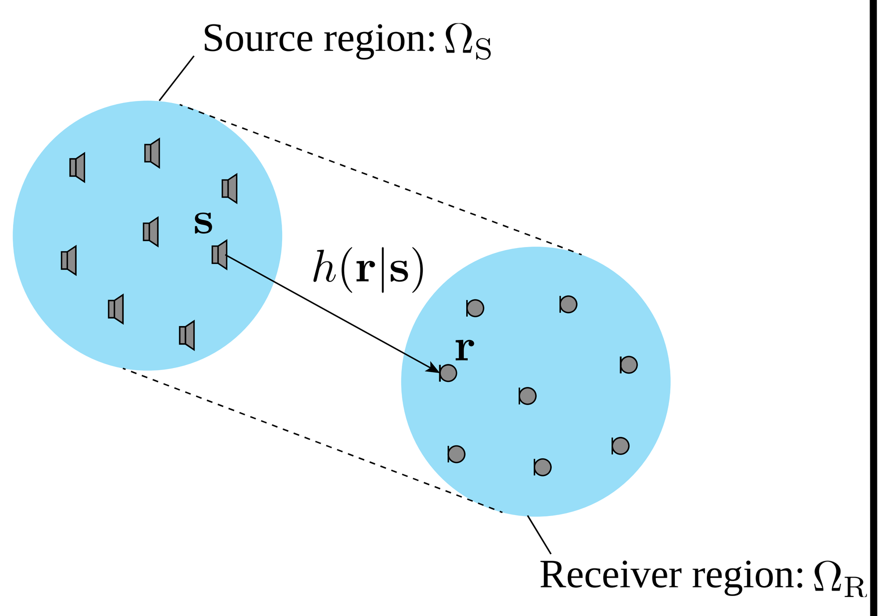
AUD-19.2

Region-to-region kernel interpolation of acoustic transfer function with directional weighting

Juliano G. C. Ribeiro, Shoichi Koyama, and Hiroshi Saruwatari
(The University of Tokyo, Tokyo, Japan)

Background

- The environment effects signal recordings.
  ⇒ Reflections, diffraction and other physical phenomena.
- We calculate the **acoustic transfer function** (ATF) to represent these effects.
- We wish to represent the ATF in a **region-to-region** basis, using ATF measurements alone.
- The ATF can be represented using sound field analysis techniques.



Problem statement

- The ATF can be divided into a direct component $h_D(\mathbf{r}|\mathbf{s},k)$ and a reverberant component $h_R(\mathbf{r}|\mathbf{s},k)$.
- $$h_D(\mathbf{r}|\mathbf{s},k) = \frac{e^{ik\|\mathbf{r}-\mathbf{s}\|}}{4\pi\|\mathbf{r}-\mathbf{s}\|}$$ Assumed to be point source in the free-field
- The reverberant component can be approximated using kernel ridge regression.
  ⇒ We have shown this method performs better than the wavefunction expansion in [Ribeiro+, 2020].
- The interpolation function will be:

$$\hat{h}_R(\mathbf{r}|\mathbf{s}) = \boldsymbol{\kappa}(\mathbf{r}|\mathbf{s})(\mathbf{K}+\lambda\mathbf{I})^{-1}\mathbf{y}$$ Measurements

where

$$\boldsymbol{\kappa}(\mathbf{r}|\mathbf{s}) = [\kappa(\mathbf{r}|\mathbf{s},\mathbf{q}_1),\ldots,\kappa(\mathbf{r}|\mathbf{s},\mathbf{q}_N)]$$

and

$$\mathbf{K} = \begin{bmatrix} \kappa(\mathbf{q}_1,\mathbf{q}_1) & \kappa(\mathbf{q}_1,\mathbf{q}_2) & \ldots & \kappa(\mathbf{q}_1,\mathbf{q}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa(\mathbf{q}_N,\mathbf{q}_1) & \kappa(\mathbf{q}_N,\mathbf{q}_2) & \ldots & \kappa(\mathbf{q}_N,\mathbf{q}_N) \end{bmatrix}.$$

But what kernel function should we use?

⇒ Embedding directionality into the kernel has been shown to improve sound field interpolation results in [Ito+, 2020].

Proposed method

- We employed the Herglotz wavefunction:

$$h_R(\mathbf{r}|\mathbf{s}) = \mathcal{I}\left(\tilde{h}_R;\mathbf{r}|\mathbf{s}\right),$$

$$\mathcal{I}(f;\mathbf{r}|\mathbf{s}) := \int_{\mathbb{S}^2\times\mathbb{S}^2} e^{ik(\hat{\mathbf{r}}\cdot\mathbf{r}+\hat{\mathbf{s}}\cdot\mathbf{s})} f(\hat{\mathbf{r}},\hat{\mathbf{s}})\mathrm{d}\hat{\mathbf{r}}\mathrm{d}\hat{\mathbf{s}}$$ Superposition of plane waves

- Which allowed us to define the fllowing function set and inner product, forming a Hilbert space:

$$\mathscr{H} = \left\{h_R = \mathcal{I}\left(\tilde{h}_R;\mathbf{r}|\mathbf{s}\right) : \tilde{h}_R \in L^2(W,\mathbb{S}^2\times\mathbb{S}^2), \tilde{h}_R(\hat{\mathbf{r}},\hat{\mathbf{s}}) = \tilde{h}_R(\hat{\mathbf{s}},\hat{\mathbf{r}})\ \forall \hat{\mathbf{r}},\hat{\mathbf{s}}\in\mathbb{S}^2\right\}$$ Acoustic reciprocity

$$\langle f,g\rangle_{\mathscr{H}} = \int_{\mathbb{S}^2\times\mathbb{S}^2} \frac{\overline{\tilde{f}(\hat{\mathbf{r}},\hat{\mathbf{s}})}\tilde{g}(\hat{\mathbf{r}},\hat{\mathbf{s}})}{W(\hat{\mathbf{r}},\hat{\mathbf{s}})}\mathrm{d}\hat{\mathbf{r}}\mathrm{d}\hat{\mathbf{s}},\ \forall f,g\in\mathscr{H}$$

informing this kernel function:

$$\kappa(\mathbf{r}|\mathbf{s},\mathbf{r}'|\mathbf{s}') = \mathcal{I}\left(W(\hat{\mathbf{r}},\hat{\mathbf{s}})\left(\frac{e^{-ik(\hat{\mathbf{r}}\cdot\mathbf{r}'+\hat{\mathbf{s}}\cdot\mathbf{s}')}+e^{-ik(\hat{\mathbf{r}}\cdot\mathbf{s}'+\hat{\mathbf{s}}\cdot\mathbf{r}')}}{2}\right);\mathbf{r}|\mathbf{s}\right)$$

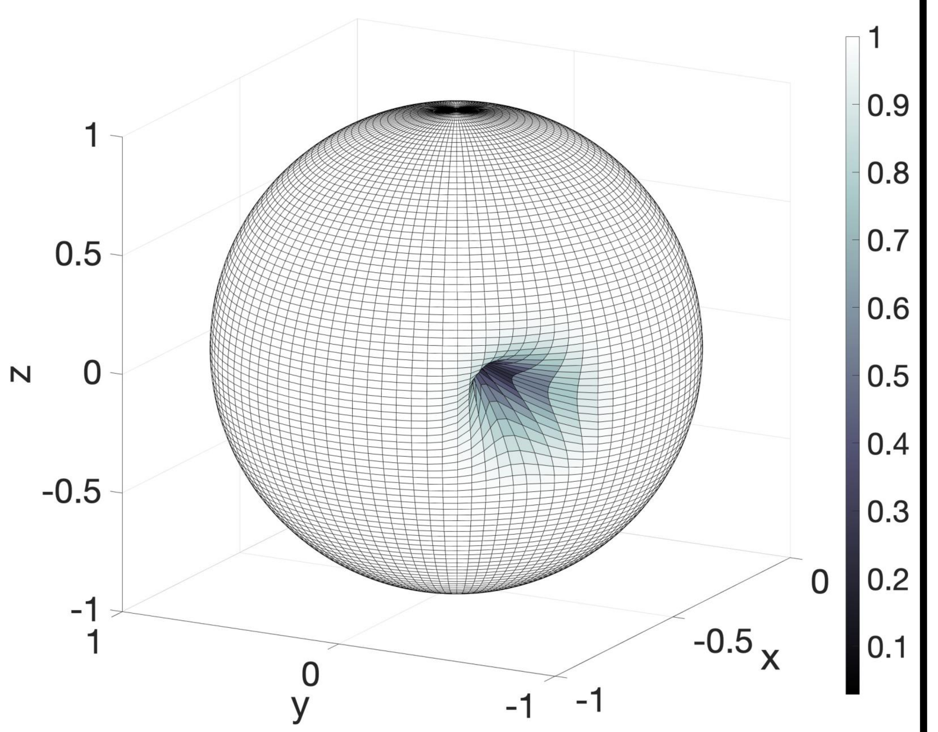
- Making the space a reproducing kernel Hilbert space(RKHS).
- We chose the following weight function:

$$W(\hat{\mathbf{r}},\hat{\mathbf{s}}) = w(\hat{\mathbf{r}})w(\hat{\mathbf{s}})$$

$$w(\hat{\mathbf{v}}) = \frac{1}{4\pi}\left(1+\gamma^2-\frac{\cosh(\beta\hat{\mathbf{v}}\cdot\hat{\mathbf{v}}_0)}{\cosh(\beta)}\right)$$ Favors early reflections

- Considering the direct component is removed, the weight was made minimal in the direction connecting the centers.

⇒ The hyperparameter β controls the width of the cavity, γ the depth.



- Parameters were obtained by optimizing leave-one-out (LOO) cross-validation of the square error(SQE) and Tukey biweight loss.

$$\mathrm{SQE}(z) = |z|^2$$

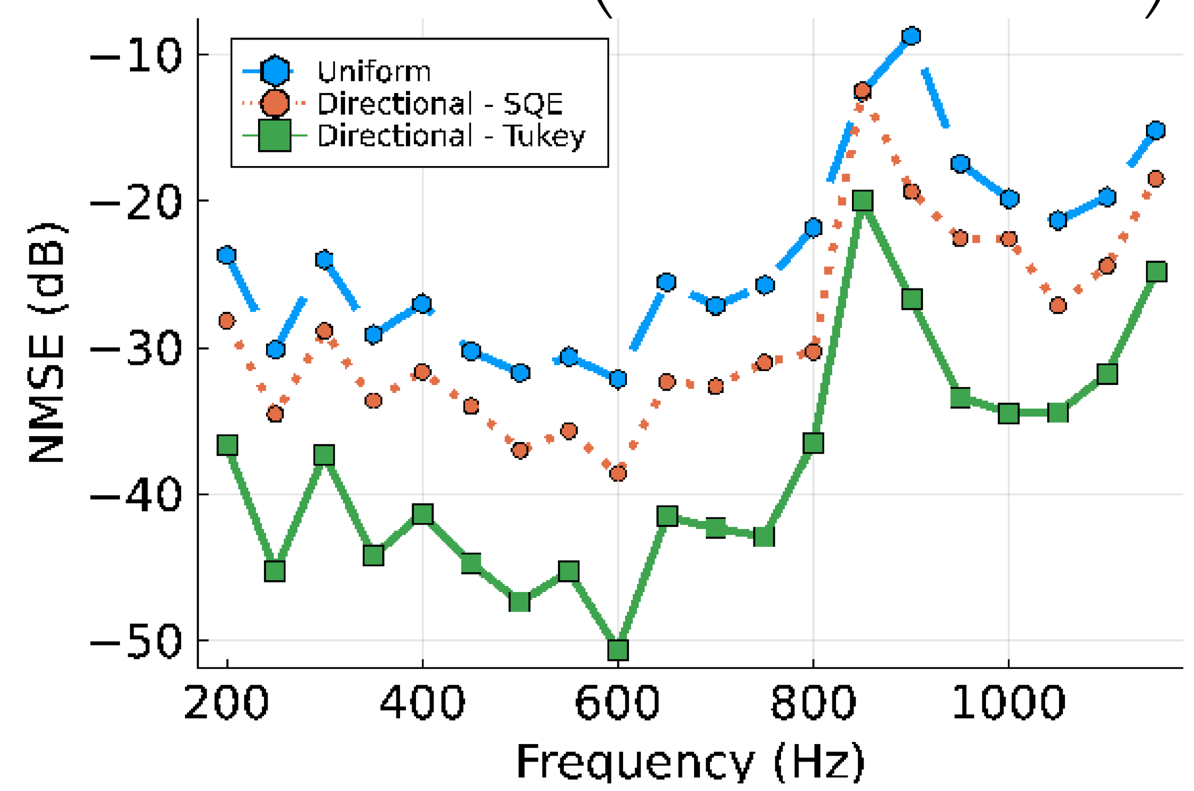
$$\mathrm{Tukey}(z) = \begin{cases} \frac{\sigma^2}{6}\left(1-\left(1-\frac{|z|^2}{\sigma^2}\right)^3\right), & |z|\le\sigma \\ \frac{\sigma^2}{6}, & |z|>\sigma \end{cases}$$ Robust loss function

Experimental results

- Simulations with the image source method.
  - -Room dimensions: $3.2\ \mathrm{m}\times 4.0\ \mathrm{m}\times 2.7\ \mathrm{m}$.
  - -Reverberation time: $T_{60} = 0.45\ \mathrm{s}$.
  - -Radius of both regions: $0.2\ \mathrm{m}$.
  - -Centers of $\Omega_{S,R}$: $\pm(0.35, 0.43, 0.29)\ \mathrm{m}$.
- We compared the proposed directionally-weighted kernels to the **uniform** kernel derived when $w \equiv (4\pi)^{-1}$.

Normalized mean square error (NMSE)

$$\mathrm{NMSE} = 10\log_{10}\left(\frac{\sum_n \left|\hat{h}(\mathbf{q}_n')-h(\mathbf{q}_n')\right|^2}{\sum_n |h(\mathbf{q}_n')|^2}\right)$$



Reconstruction of impulse response (950Hz)