Paper ID: 1884

Kernel interpolation of acoustic transfer functions with adaptive kernel for directed and residual reverberations

Juliano G. C. Ribeiro[1], Shoichi Koyama[2], and Hiroshi Saruwatari[1]

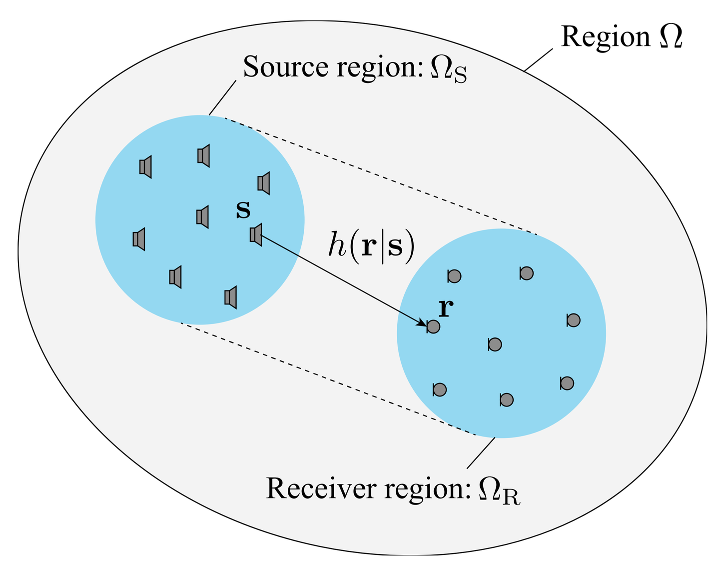
[1]The University of Tokyo, [2]National Institute of Informatics

Background

- Sound wave propagation is unpredictable.
  ⇒ Physical phenomena such as reflection and diffraction.
- **Acoustic transfer function** (ATF):
  ⇒ Describes space's influence on sound waves
- **Region-to-region**:
  ⇒ Interpolate ATF continuously for source/receiver.
- **Sound field analysis**:
  ⇒ Deeply related problem we can draw solutions from

Problem statement

- Objective: interpolate ATF values from measurements.



- ATF components $h(\mathbf{r}|\mathbf{s}) = h_R(\mathbf{r}|\mathbf{s}) + h_D(\mathbf{r}|\mathbf{s})$.
- Direct is known: $h_D(\mathbf{r}|\mathbf{s},k) = \dfrac{e^{ik\|\mathbf{r}-\mathbf{s}\|}}{4\pi\|\mathbf{r}-\mathbf{s}\|}$.
- Reverberant: $h_R(\mathbf{r}|\mathbf{s},k)$ is more involved.
  ⇒ [Ribeiro+, 2020]: physical constraint in **kernel ridge regression** with ATF kernel.
  ⇒ [Ribeiro+, 2022]: directionality improved performance.
- Generalized representation: Herglotz wave function.

$$h_R(\mathbf{r}|\mathbf{s}) = \mathcal{T}(\tilde{h}_R; \mathbf{r}|\mathbf{s})$$

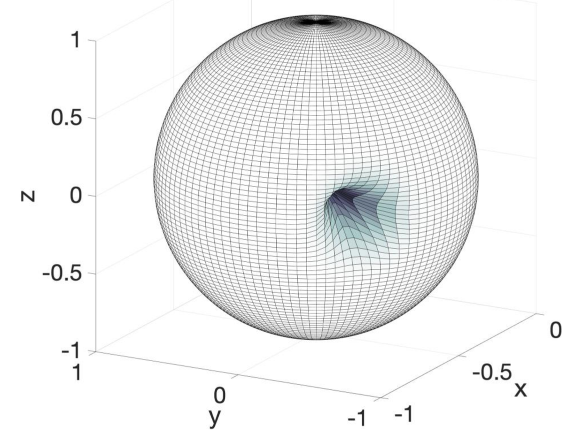
Plane wave superposition

$$:= \int_{\mathbb{S}^2\times\mathbb{S}^2} e^{ik(\hat{\mathbf{r}}\cdot\mathbf{r}+\hat{\mathbf{s}}\cdot\mathbf{s})}\tilde{h}_R(\hat{\mathbf{r}},\hat{\mathbf{s}})d\hat{\mathbf{r}}d\hat{\mathbf{s}}$$

- Representation that guarantees physics of the problem.
  ⇒Kernel function that learns weight function as data model.

$$\kappa(\mathbf{r}|\mathbf{s},\mathbf{r}'|\mathbf{s}') = \mathcal{T}\left(w(\hat{\mathbf{r}},\hat{\mathbf{s}})\frac{e^{-ik(\hat{\mathbf{r}}\cdot\mathbf{r}'+\hat{\mathbf{s}}\cdot\mathbf{s}')}+e^{-ik(\hat{\mathbf{r}}\cdot\mathbf{s}'+\hat{\mathbf{s}}\cdot\mathbf{r}')}}{2};\mathbf{r}|\mathbf{s}\right)$$

- Previous kernels can be expressed with Herglotz wave function.
- [Ribeiro+, 2020]: kernel function equivalent to a **uniform** weight.
  ⇒ Satisfies basic physical properties.
- [Ribeiro+, 2022]: **sunken sphere** weight function.
  ⇒ Introduces directionality to the estimation.
  ⇒ Physical model limited, assumes uniform gain on the sides.
- Data models for the weights are inflexible.
- We propose an **adaptive kernel** that learns the weight as a more general data model.



Proposed method

- Weighting represents directed and residual reverberations as the sum $w = w_{\mathrm{dir}} + w_{\mathrm{res}}$.

Directed weight w_{dir}

- High amplitudes on sparse set of directions
- Strong directionality.
- Combination of von Mises-Fisher distributions.

$$w_{\mathrm{dir}}(\hat{\mathbf{r}},\hat{\mathbf{s}}) = \varphi_{\mathrm{dir}}(\hat{\mathbf{r}})\varphi_{\mathrm{dir}}(\hat{\mathbf{s}}),$$

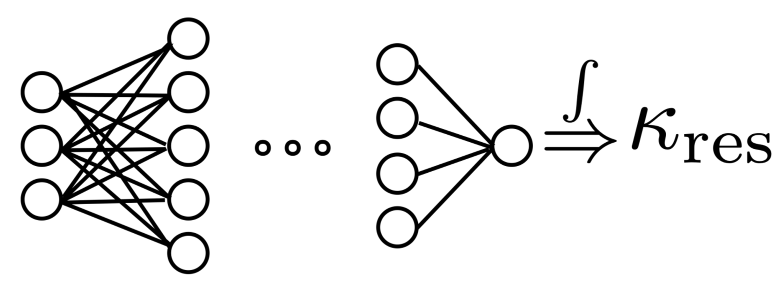
$$\varphi_{\mathrm{dir}}(\hat{\mathbf{v}}) = \sum_{d=1}^{D}\alpha_d\frac{e^{\beta_d\hat{\mathbf{v}}\cdot\hat{\mathbf{v}}_d}}{4\pi C(\beta_d)},$$

$$\|\boldsymbol{\alpha}\|_1 = 1,$$

$$C(\beta_d) = \begin{cases}\frac{\sinh(\beta_d)}{\beta_d}, & \beta_d \neq 0\\ 1, & \beta_d = 0\end{cases}.$$

Residual weight w_{res}

- Low amplitude on a dense set of directions
- Unpredictable behavior.
- Represented by neural network with parameters $\boldsymbol{\theta}$.
- Integral operator approximated numerically.



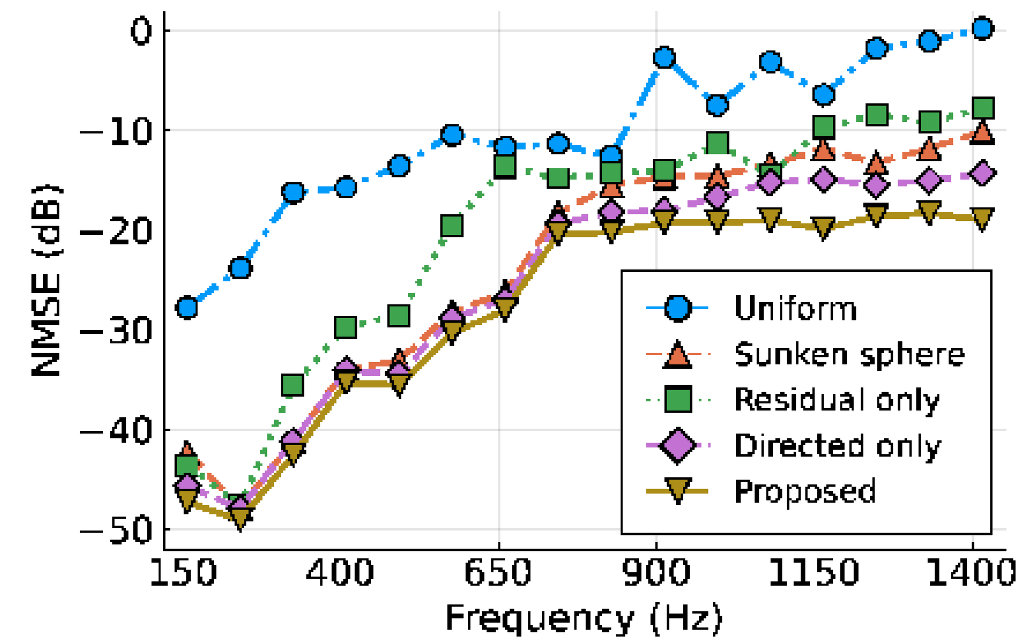
- The resulting adaptive kernel will be the superposition of the kernels.
- Kernel optimized as to minimize leave-one-out cross validation loss.
- Model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ optimized using gradient descent.
- Parameter $\boldsymbol{\alpha}$ optimized with reduced gradient descent to guarantee restrictions are upheld.
- The adaptive kernel learns particularities of the ATF in question without compromising model dynamics.

Experiments

- 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 $\Omega_{S,R}$: $\pm[0.65,\ 0.8,\ 0.48]^{\mathsf{T}}$.
- **Proposed** compared to **uniform** and **sunken sphere** kernels.

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 ATF (1150Hz)