Bayesian Inference and Neural Estimation of Acoustic Wave Propagation

Yongchao Huang[1], Yuhang He[2], Hong Ge[3]

[1]University of Aberdeen [2]University of Oxford [3]University of Cambridge

RoboAcoustics poster @ ICRA 2025

Summary

We present a novel framework integrating physics and machine learning to estimate frequency-domain acoustic wave propagation coefficients [1]. Using acoustic waveforms captured at speaker-receiver pairs, we estimate attenuation and wavenumber via: (1) Bayesian inference for uncertainty-aware learning from small and noisy data, (2) a neural-physical model trained with forward-backward physical losses, and (3) non-linear least squares as baseline. With inferred propagation coefficients, room impulse responses (RIRs) are derived, enabling robot relocalisation with uncertainty quantification.

Acoustic Wave Propagation

The frequency-domain wave equation:

$$\frac{\partial^2}{\partial x^2}\tilde{P}(x,\omega) + \frac{\omega^2}{c^2}\tilde{P}(x,\omega) = 0$$

One-way solution: $\tilde{P}(x,\omega) = \tilde{P}(x_0,\omega)e^{-\gamma(x-x_0)}$, with $\gamma(\omega) = \alpha(\omega) + i\kappa(\omega)$. Goal: infer $\gamma(\omega)$ from signal pairs (x_1, x_2).

Bayesian Inference

Prior distributions:

$$\alpha \sim \mathcal{N}(\mathbf{1}, I), \kappa \sim \mathcal{N}(\mathbf{0}, 10^2 I), \sigma \sim HalfNormal(std = 1)$$

Likelihood:

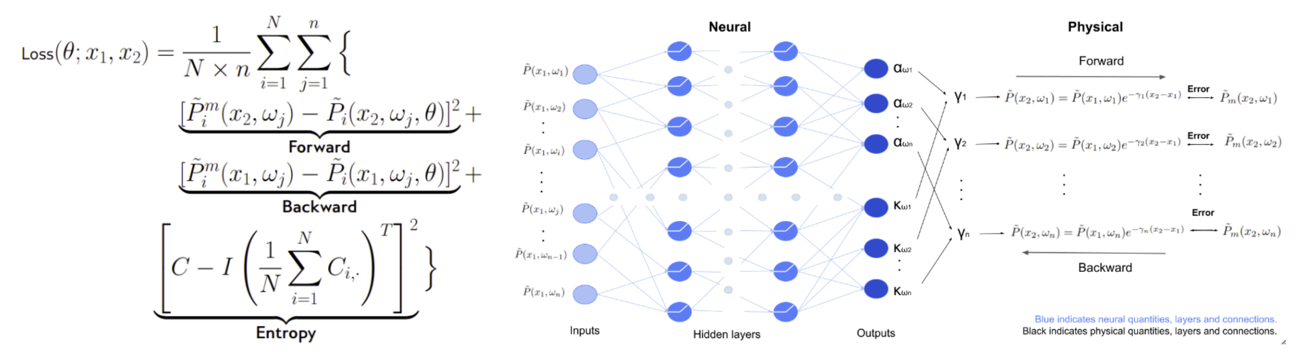
$$\begin{bmatrix} \tilde{P}^m(x_2,\omega) - \tilde{P}^m(x_1,\omega)e^{-\gamma(\omega)(x_2-x_1)} \\ \tilde{P}^m(x_1,\omega) - \tilde{P}^m(x_2,\omega)e^{\gamma(\omega)(x_2-x_1)} \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 \\ \sigma^2 \end{bmatrix})$$

where $\tilde{P}^m(x_1,\omega)$ and $\tilde{P}^m(x_2,\omega)$ are the measured waves at speaker x_1 and receiver x_2. σ^2 is noise variance.
Posterior distributions are approximated via Hamiltonian Monte Carlo sampling, enabling prediction and uncertainty quantification.

Neural Estimation

A neural network learns (α, κ) from measured waves $\tilde{P}^m(x_1,\omega)$, with a loss function comprised of:

- **Forward loss:** measured $\tilde{P}^m(x_2,\omega)$ *vs* predicted $\tilde{P}(x_2,\omega)$
- **Backward loss:** $\tilde{P}^m(x_1,\omega)$ *vs* reversed prediction $\tilde{P}(x_1,\omega)$
- **Entropy regularisation:** promotes coefficient consistency across instances



Non-linear Least Squares Estimation

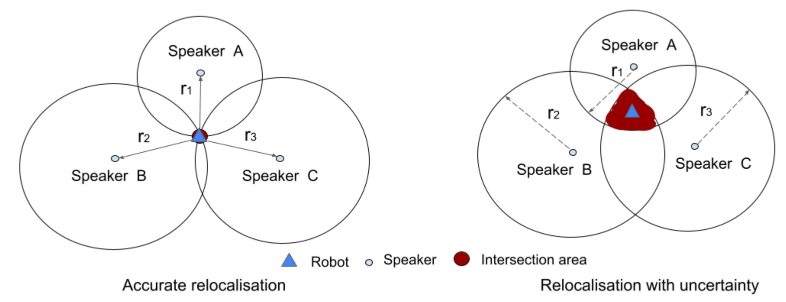
Log-transform the measured waves at the speaker x_1 and receiver x_2: $\log\tilde{P}^m(x_2,\omega) - \log\tilde{P}^m(x_1,\omega) \approx -\gamma(\omega)(x_2 - x_1)$.
Stacking N wave pairs and n frequencies yields the matrix form: $\Delta\log\tilde{P}^m_{N\times n} \approx -\Delta X_{N\times N}\Gamma_{N\times n}$, where $\Delta X = \mathrm{diag}(x_{i2} - x_{i1})$, and Γ contains γ duplicated row-wise.
Solving the least squares problem: $\hat{\gamma} = \arg\min_\gamma \|\Delta\log\tilde{P}_{N\times n} - (-\Delta X_{N\times N}\Gamma_{N\times n})\|_2^2$ yields the closed-form solution: $\hat{\Gamma} = -(\Delta X)^{-1}\Delta\log\tilde{P}^m$.
Rows of $\hat{\Gamma}$ should ideally be equal in noise-free settings. Variability across rows reflects noise-induced uncertainty.

Conclusion & Future Work

We propose a physics-informed framework for inferring acoustic wave propagation in frequency domain. All methods effectively infer $\gamma(\omega)$, offer RIR estimation and relocalisation using acoustic signals alone. Specifically: **Bayesian inference** learns from small and noisy data, provides data efficiency and principled uncertainty quantification; **Neural estimation** is fast, generalisable, and learns physics-consistent estimates; **Non-linear least squares** is efficient and interpretable. **Future work:** Extend to complex wave models, online inference, and real-world deployment.

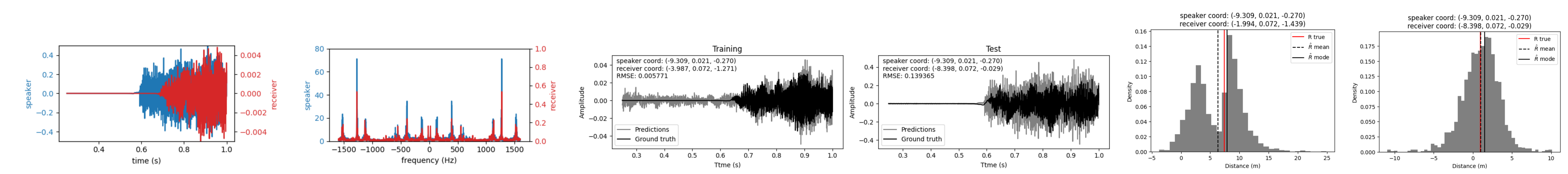
RIR Estimation and Relocalisation

The room impulse response $RIR(x,t)$ characterises how a wave propagates between two locations in time domain: $P^m(x_1,t) \circledast RIR(x_2 - x_1, t) = P^m(x_2,t)$.
Applying Fourier transform yields the frequency-domain $R\tilde{I}R$: $\tilde{P}^m(x_1,\omega) \odot R\tilde{I}R(x_2 - x_1,\omega) = \tilde{P}^m(x_2,\omega)$.
Solving for $R\tilde{I}R$: $R\tilde{I}R(x_2 - x_1,\omega) = \frac{\tilde{P}^m(x_2,\omega)}{\tilde{P}^m(x_1,\omega)} = e^{-\gamma(x_2-x_1)}$.
This allows estimation of $R\tilde{I}R$ from estimated γ. Specially, if $\tilde{P}^m(x_1,\omega)$ is a Dirac delta (i.e. flat in frequency), then $R\tilde{I}R(x_2 - x_1,\omega)$ equals $\tilde{P}^m(x_2,\omega)$. The time domain $RIR(x_2 - x_1, t)$ can be obtained by inverting the frequency domain $R\tilde{I}R$.
Relocalisation: Distance can be recovered via $\Delta x = -\frac{1}{\gamma(\omega)}\log\left(\frac{\tilde{P}^m(x_2,\omega)}{\tilde{P}^m(x_1,\omega)}\right)$.
Averaging over n frequencies gives a robust estimate: $\widehat{\Delta x} = \frac{1}{n}\sum_{i=1}^{n} -\frac{1}{\gamma(\omega_i)}\log\left(\frac{\tilde{P}^m(x_2,\omega_i)}{\tilde{P}^m(x_1,\omega_i)}\right)$.
In 2D/3D, intersecting circular shells from multiple fixed-position speakers localises the receiver position:

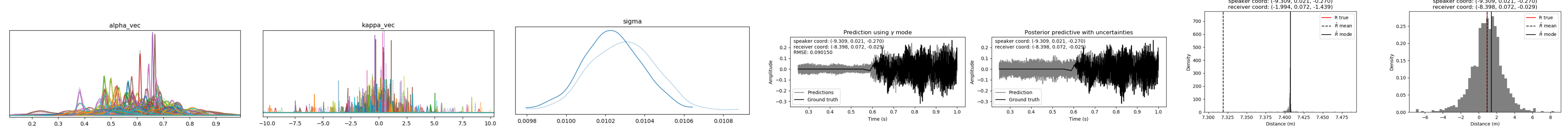


Experimental validations

Simulation setup: We use the *SoundSpaces 2.0* simulator with the *Matterport3D* dataset. A fixed-position speaker emits 16kHz signals captured by 9 spatially distributed receivers. Bayesian inference and non-linear least squares use 1 speaker-receiver pair; the neural model trains on 8 pairs and tests on 1.
Neural results: Example speaker and receiver signals are shown below (left 2 figures). The neural-physical model estimates the propagation coefficients and predicts waveforms (middle 2 figures) using MLP / auto-encoder architectures. The distance between speaker and receiver is estimated (right 2 figures).



Bayesian results: HMC (NUTS) is used to sample α, κ, and σ (Fig.1-3). Given the posterior samples of γ, we predict the waveform at the test receiver (Fig.4-5) and estimate the speaker-receiver distance (Fig.6-7). The predicted waveform shows credible intervals; the relocalisation results demonstrate consistency across receivers.



Performance comparison: Distance estimations across all methods. Bayesian MAP yields the most accurate and stable estimates:

	Test receiver 1			Test receiver 2		
Statistics	Mean	Mode	Std	Mean	Mode	Std
Bayesian MAP	**7.319**	**7.407**	**0.466**	**0.936**	1.425	1.715
Bayesian mean	2.749	2.896	1.311	0.416	0.487	**0.855**
Neural (MLP)	6.290	7.883	3.955	0.986	1.509	2.395
Neural (Autoencoder)	6.308	7.623	3.880	0.961	1.869	2.375
Least squares	5.458	4.998	2.968	0.854	**1.318**	2.016
Ground truth	7.407	-	-	0.943	-	-

References

[1] Yongchao Huang, Yuhang He, Hong Ge (2023). *Bayesian inference and neural estimation of acoustic wave propagation*, https://arxiv.org/abs/2305.17749.