Acoustic field estimation with differentiable physics

Samuel A. Verburg and **Peter Gerstoft** *Department of Electrical and Photonics Engineering, Technical University of Denmark (DTU)*
{saveri,pegers}@dtu.dk
Efren Fernandez-Grande *Universidad Politecnica de Madrid*
efren.fernandez@upm.es

Abstract

Differentiable physics is used to estimate acoustic fields from a limited number of spatially distributed observations. The initial conditions of the wave equation are approximated with a neural network, and the differential operator is computed with a differentiable numerical solver. We introduce an additional sparsity-promoting constraint to achieve meaningful solutions even under severe undersampling conditions. Numerical experiments demonstrate that the approach can reconstruct sound fields under extreme data scarcity.

Keywords: Differentiable physics, neural network.

Sound field estimation refers to the inverse problem of estimating a sound field over time and space from a limited number of spatially distributed observations. A main challenge is the large number of measurements required, which increases with the domain size and the frequency. Efforts are devoted to developing efficient models to reconstruct sound fields with minimal data [1, 2, 4, 5, 7, 7, 9, 13].

This study introduces a differentiable physics (DP) [11, 12] approach for sound field estimation, where the initial condition is modeled with a neural network, and a differentiable finite-difference solver is used to solve the wave equation. We demonstrate that even if the network is trained for a given discretization, the sound field can be reconstructed at higher resolutions, as the network can be queried at any point in the domain. Furthermore, we propose a sparsity-promoting constraint to the initial condition. In a series of experiments we show that the proposed DP approach is robust, presents good convergence, while achieving small errors.

1 Sound field estimation using differentiable physics

Let us consider the acoustic pressure field $p(\mathbf{r}, t)$ in the spatio-temporal domain $\Omega \times [0, T]$, where $\Omega \subset \mathbb{R}^2$, $\mathbf{r} \in \Omega$, and $t \in [0, T]$. The pressure field is the solution of the wave equation

$$\mathcal{D}[p] := \nabla^2 p(\mathbf{r}, t) - \frac{1}{c^2}\frac{\partial^2 p(r, t)}{\partial t^2} = 0, \tag{1}$$

with initial conditions, $p(\mathbf{r}, 0) = g(\mathbf{r})$ and $\frac{\partial p}{\partial t}(\mathbf{r}, 0) = 0$. In (1)$\mathcal{D}[\cdot]$ is a differential operator expressing the PDE, and $c \in \mathbb{R}$ is the medium wave speed, assumed to be a known constant. The domain is considered unbounded, with no reflected waves arriving from outside. To express this, a first-order absorptive boundary condition [3] is considered

$$\mathcal{B}[p] := \nabla p(\mathbf{r}, t) \cdot \mathbf{n} + \frac{1}{c}\frac{\partial p}{\partial t} = 0 \quad \text{at } \mathbf{r} \in \partial\Omega, \tag{2}$$

where $\mathbf{n}$ is the unit vector normal to the boundary $\partial\Omega$.

The goal of sound field reconstruction is to estimate the entire pressure field from noisy observations,

$$\hat{p}_{mn} = p(\mathbf{r}_m, t_n) + e_{mn} \quad \text{for } m = 0, \ldots, M_{\text{ob}} - 1 \text{ and } n = 0, \ldots, N-1, \tag{3}$$

where $\mathbf{r}_0, \ldots, \mathbf{r}_{M_{\text{ob}}-1}$ are the sensor locations, $t_0, \ldots, t_{N-1}$ are the time samples, and e_{mn} is additive noise. Measurements are typically performed using microphone arrays or distributed sensors. Therefore, the pressure is finely sampled over time, but only a few positions are sampled over space.

In the proposed DP approach a neural network $g_{\text{dp}}(\mathbf{r}; \boldsymbol{\theta})$ models the unknown initial pressure. The physical constraints are imposed by applying a numerical PDE solver to the network output. Training the DP neural network amounts to solving the optimization problem

$$\min_{\boldsymbol{\theta}} \left\{ \lambda_{\text{data}} \mathcal{L}_{\text{data}}(\mathbf{p}^0, \mathbf{p}^1 \ldots, \mathbf{p}^{N-1}) + \lambda_{\text{sp}} \mathcal{L}_{\text{sp}}(\mathbf{p}^0) \right\}, \tag{4}$$

where $\mathbf{p}^1, \ldots, \mathbf{p}^{N-1}$ is the numerical solution of the PDE computed in a M_{grid}-dimensional discretization grid, and $\mathbf{p}^0(\boldsymbol{\theta})$ is obtained by sampling the neural network $g_{\text{dp}}(\mathbf{r}; \boldsymbol{\theta})$ at the grid positions. Therefore, $\mathbf{p}^n$ is a vector in $\mathbb{R}^{M_{\text{grid}}}$ instead of being a continuous function. For simplicity and without loss of generality, it is assumed that for each observation position, $\mathbf{r}_0, \ldots, \mathbf{r}_{M_{\text{ob}}-1}$, there is a point in the discretization grid. A data fitting function, $\mathcal{L}_{\text{data}}$, that operates on the discrete pressure can be expressed as

$$\mathcal{L}_{\text{data}}\left(\mathbf{p}^0, \ldots, \mathbf{p}^{N-1}\right) = \frac{1}{M_{\text{ob}} N} \sum_{n=0}^{N-1} \left\| \mathbf{M}\mathbf{p}^n - \hat{\mathbf{p}}^n \right\|^2, \tag{5}$$

where $\mathbf{M}$ is a $M_{\text{ob}} \times M_{\text{grid}}$ binary matrix that extracts the pressure values at the observation positions, and $\hat{\mathbf{p}}^n \in \mathbb{R}^{M_{\text{ob}}}$ denotes the observations in Eq. (3) arranged as a vector.

The finite difference method [6] solves the PDE numerically. The solution is obtained by applying the explicit time integration scheme

$$\mathbf{p}^1 = \mathbf{p}^0 + 0.5\mathbf{L}\mathbf{p}^0, \text{ and } \mathbf{p}^{n+1} = 2\mathbf{p}^n - \mathbf{p}^{n-1} + \mathbf{L}\mathbf{p}^n \text{ for } n = 1, \ldots, N-1, \tag{6}$$

where $\mathbf{L} = (c\Delta t/\Delta r)^2 \mathbf{L}_\Delta$ and $\mathbf{L}_\Delta \in \mathbb{R}^{M_{\text{grid}} \times M_{\text{grid}}}$ is the central difference approximation of the Laplace operator $\nabla^2[\cdot]$. The scalars Δt and Δr are the sampling period and grid spacing, respectively. To handle the unbounded domain, the boundary condition of (2) is incorporated into the numerical solver. The finite difference approximation of the absorptive boundary computes the values of $\mathbf{p}^{n+1}$ at the boundary based on $\mathbf{p}^n$ at the boundary and adjacent points, as well as $\mathbf{p}^{n+1}$ at the adjacent points.

Central to the proposed DP approach is the neural network, $g_{\text{dp}}(\mathbf{r}; \boldsymbol{\theta})$, to model the initial condition as a continuous, smooth function that maps any input coordinate within the domain to a corresponding output value. Therefore, even if the PDE is solved on a fixed discrete grid during training, the resolution can be increased by sampling the neural network on a finer grid, and then solve the PDE with a higher resolution numerical solver.

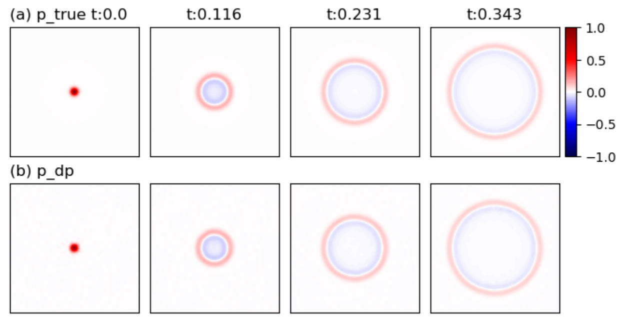

Figure 1: Sound field consisting on a single pulse at the domain center. Each column corresponds to a time frame. Row (a): reference solution. Row (b): DP model estimation.

2 Numerical experiments

A 2+1D domain is defined, where the spatial domain is a square of side length $L = 1$, the temporal domain has a duration of $T = 0.343$, and the speed of sound is $c = 1$. The temporal domain is divided into $n = 50$ samples, giving a sampling period $\Delta t = 7.0 \times 10^{-3}$. For the finite difference solver a regular discretization grid of $M_{\text{grid}} = 100^2$ is defined.

Single pulse: Synthetic data for multiple sound fields is generated. For the first one, the initial condition is a single Gaussian pulse of unit amplitude and scale $\sigma = 0.02$ placed at the center of the domain,

$$g(\mathbf{r}) = \exp\left(-0.5\|\mathbf{r} - \mathbf{r}_0\|^2/\sigma^2\right), \tag{7}$$

where $\mathbf{r}_0 = (L/2, L/2)$. The observations used for the reconstruction conform a pseudo-random array, shown in the first panel of Fig. 1(a), where the sensor locations are sampled from the discretization grid within $[0.1L, 0.9L] \times [0.1L, 0.9L]$ and with a minimum distance of 0.05 between sensors. The number of time samples is $N = 50$ and the number of sensors is $M_{\text{obs}} = 20$. Additive Gaussian noise is added to the data such that the SNR is 20 dB.

The observed data and reference values of the sound field are obtained from the analytical solution of the acoustic wave equation in free field with a Gaussian pulse as initial condition [10]. The reference sound field is computed on a grid of twice the spatial and temporal resolutions of the DP finite difference grid.

Figure 1 shows the reference sound field and estimation for the single Gaussian pulse. The DP results show an accurate reconstruction throughout the spatio-temporal domain, with only noticeable differences at $t = 0$. The normalized mean square error (NMSE) computed over all the spatio-temporal points on the evaluation grid is 5.3×10^{-3}.

The reconstruction performance is analyzed by training the model in different scenarios. As benchmark, a conventional phyics-informed neural network (PINN) [8] is trained to solve the same estimation problem. The proposed DP model largely outperforms the PINN for all tested SNRs, see Fig. 2(a), presenting

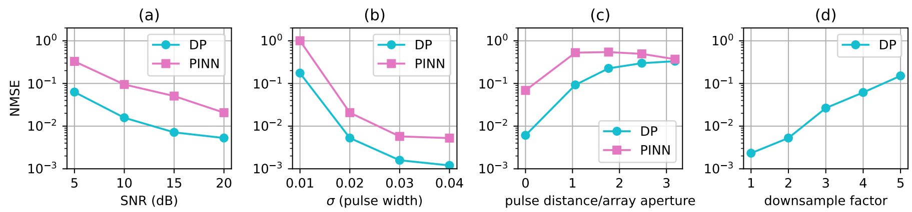

Figure 2: Normalized Mean squared error vs (a) the SNR, (b) the source width, (c) the pulse distance to the array center, and (d) the downsample factor between the evaluation and training grids.

errors almost one order of magnitude smaller. Figure 2(b) shows the NMSE as a function of the pulse scale σ, which is directly related to the frequency content of the acoustic field. The PINN fails to reconstruct the sound field of highest frequency (smallest σ), presenting a NMSE close to 1, while the DP model consistently achieves lower errors. Note that the DP network has one layer less and half the number of units per layer that the PINN. Figure 2(c) shows the NMSE vs. distance between the pulse and the array center normalized by the array aperture. The experiment serves to evaluate the extrapolation capabilities of the models to areas where there is no observed data and the estimation relies only on the physics of wave propagation. The PINN presents a large error as soon as the source is outside the array aperture as the physics are only included as a weak constraint. Conversely, the DP model output stratifies the underlying physics by design. Since the initial conditions are approximated with a continuous function we can upscale the estimation to any desired resolution. The NMSE increases for lower training resolutions, which is caused by the accumulation of numerical errors and the fact that coarser grids are not able to represent high spatial frequencies present in the initial condition.

3 Conclusion

We propose a differentiable physics approach for sound field reconstruction. Integration of a numerical solver in the training of a neural network enables the incorporation of hard physical constraints robustly. The optimization is more stable than in conventional PINNs, and convergence is achieved in a fraction of the optimizer steps. Formulating the solver in a differentiable way using AD makes the training process very simple since only the forward solver is required. The DP approach is generalizable beyond the training discretization, and the solutions obtained can be scaled to higher resolutions. The experiments show that the DP model achieves accurate reconstructions and low errors even in challenging, highly undersampled problems.

References

[1] H. Bi and T. D. Abhayapala. Point neuron learning: A new physics-informed neural network architecture. *EURASIP J. Audio Speech Music Process.*, 2024(1):56, 2024.

[2] D. Caviedes-Nozal, N. A. Riis, F. M. Heuchel, J. Brunskog, P. Gerstoft, and E. Fernandez-Grande. Gaussian processes for sound field reconstruction. *J. Acoust. Soc. Am.*, 149(2):1107–1119, 2021.

[3] R. Clayton and B. Engquist. Absorbing boundary conditions for acoustic and elastic wave equations. *B. seismological society of America*, 67(6): 1529–1540, 1977.

[4] E. Fernandez-Grande, X. Karakonstantis, D. Caviedes-Nozal, and P. Gerstoft. Generative models for sound field reconstruction. *J. Acoust. Soc. Am.*, 153(2):1179–1190, 2023.

[5] S. Koyama, J. G. Ribeiro, T. Nakamura, N. Ueno, and M. Pezzoli. Physics-informed machine learning for sound field estimation: Fundamentals, state of the art, and challenges. *IEEE Signal Process. Mag.*, 41(6):60–71, 2025.

[6] H. P. Langtangen and S. Linge. *Finite difference computing with PDEs: a modern software approach*. Springer Nature, 2017.

[7] R. Liu and P. Gerstoft. Deep-learning-based spatial acoustic properties recovery from incomplete signals. In *2024 IEEE 13rd Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5. IEEE, 2024.

[8] B. Moseley, A. Markham, and T. Nissen-Meyer. Solving the wave equation with physics-informed deep learning. *arXiv preprint arXiv:2006.11894*, 2020.

[9] J. G. Ribeiro, S. Koyama, R. Horiuchi, and H. Saruwatari. Sound field estimation based on physics-constrained kernel interpolation adapted to environment. *IEEE Trans. Audio Speech Lang. Process.*, 2024.

[10] C. K. Tam and J. C. Webb. Dispersion-relation-preserving finite difference schemes for computational acoustics. *J. Comput. Phys.*, 107(2):262–281, 1993.

[11] N. Thuerey, P. Holl, M. Mueller, P. Schnell, F. Trost, and K. Um. Physics-based deep learning. *arXiv preprint arXiv:2109.05237*, 2021.

[12] S. A. Verburg, E. Fernandez-Grande, and P. Gerstoft. Differentiable physics for sound field reconstruction. *arXiv preprint arXiv:2510.04459*, 2025.

[13] S. Yoon, Y. Park, P. Gerstoft, and W. Seong. Predicting ocean pressure field with a physics-informed neural network. *J. Acoust. Soc. Am.*, 155(3): 2037–2049, 2024.