



**HAL**  
open science

## Neural operator preconditioning for accelerating the solution of the parametric Helmholtz equations

Yanfei Xiang, Luc Giraud, Paul Mycek, Maksym Shpakovych, Carola Kruse

► **To cite this version:**

Yanfei Xiang, Luc Giraud, Paul Mycek, Maksym Shpakovych, Carola Kruse. Neural operator preconditioning for accelerating the solution of the parametric Helmholtz equations. 2024 SIAM Conference on Applied Linear Algebra (SIAM LA24), May 2024, Paris, France. pp.2. hal-04581208v2

**HAL Id: hal-04581208**

**<https://inria.hal.science/hal-04581208v2>**

Submitted on 21 May 2024 (v2), last revised 10 Jun 2024 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

**2024 SIAM Conference on Applied Linear Algebra (SIAM LA24). At MS85: Recent Advances in Iterative Solvers for Non SPD Problems: Neural operator preconditioning for accelerating the solution of the parametric Helmholtz equations**

Yanfei Xiang\*  
Inria, France  
yanfei.xiang@inria.fr

Luc Giraud  
Inria, Concace joint Inria team with Airbus CR&T and Cerfacs, France  
luc.giraud@inria.fr

Paul Mycek  
Cerfacs, France  
mycek@cerfacs.fr

Maksym Shpakovych  
Inria, France  
maksym.shpakovych@inria.fr

Carola Kruse  
Cerfacs, France  
carola.kruse@cerfacs.fr

\*Speaker

## **Abstract**

In recent decades, scientific machine learning (SciML), utilizing deep learning methodologies, has found widespread application in the fields of scientific computing and computational engineering. That includes learning the neural networks as a solver and learning functions by the neural operators. Neural network solver can be quite promising after a stroke of luck and proper training. However, they generally yield solution with limited accuracy and exhibit potential issues in network generalizability. Besides, unlike the classical numerical linear algebra solvers, purely data-driven network solvers lack theoretical convergence guarantee. In this work, we focus of training different neural operators to accelerate the solution of the parametric Helmholtz equations by the classical Krylov subspace methods, like the flexible GMRES method (FGMRES). To be specific, the learned neural operators will be used as a nonlinear preconditioner with the Krylov basis as its input to improve the effectiveness of the FGMRES method.

Given the goal is to learning a preconditioner rather than a solver, the required accuracy is not high and so does the training cost. Here the varying parameters in the parametric Helmholtz equations include the source term, the speed of sound and the discretization domain size. In order to learning an effective preconditioner for these parametric Helmholtz equations, we train various neural operators with different neural network architectures. Then we compare the preconditioning quality of these trained inferences algebraically to select the optimal neural operator to meet our goal. Given there is no information about the data structure of the Krylov basis, the neural operators are trained with randomly generated datasets. With a physics-based loss function that constructed by the residual of the linear systems, the training process is finished under the context of unsupervised way. This means no ground truth, which is not always available in practice, is required in advance. We also investigate the influence of the settings for the randomly generated datasets as well as the hyper-parameters of different neural operators. The selected models are trained on a smaller 2D domain  $64 \times 64$  with randomly generated datasets. While we successfully test their network generalizability from different perspectives (related to vary each of the involved parameters in the parametric Helmholtz equations). Furthermore, we also success in applying our trained inference to accelerating the solution of a challenging practical human head CT scan dataset. This is a challenging example since it involves all the unseen varying parameters on a 64 times larger domain  $512 \times 512$ . In short, this work demonstrates the efficiency of learning neural operators for preconditioning to accelerate the simulation of the parametric Helmholtz equations. More generally, it also illustrates that the performance of the SciML depends on the choosing of the neural network architecture, the setting of training datasets, and the tuning of hyper-parameters of the neural networks.