中国科学技术大学 University of Science and Technology of China
上海人工智能实验室 Shanghai Artificial Intelligence Laboratory
清華大學 Tsinghua University

AOT-POT

SCIENTIFIC MACHINE LEARNING
FOUNDATION MODELS FOR PDEs

Adaptive Operator Transformation for Large-Scale PDE Pre-training

Qitan Lv[1 2*] Hong Wang[1*] Zhongkai Hao[3] Wen Wu[2] Xuenan Xu[2] Bowen Zhou[1 2 3] Feng Wu[1] Chao Zhang[2 3†]

[1] University of Science and Technology of China [2] Shanghai AI Laboratory [3] Department of Electronic Engineering, Tsinghua University

NeurIPS

40.9%
mean per-dataset reduction
AOT-POT-S vs. DPOT-S; 12 datasets

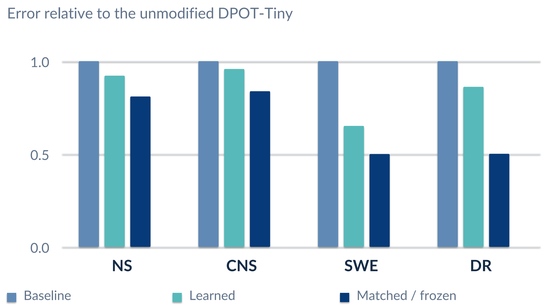
77.6%
largest error reduction
vs. DPOT-S on FNO ν = 10^{-3}

31M
vs. 122M parameters
AOT-POT-S wins 11/12 vs. DPOT-M

16
vs. 64 fine-tuning samples
Beats DPOT-S in NS + Wave-Layer

01 WHY TRANSFORM OPERATORS?

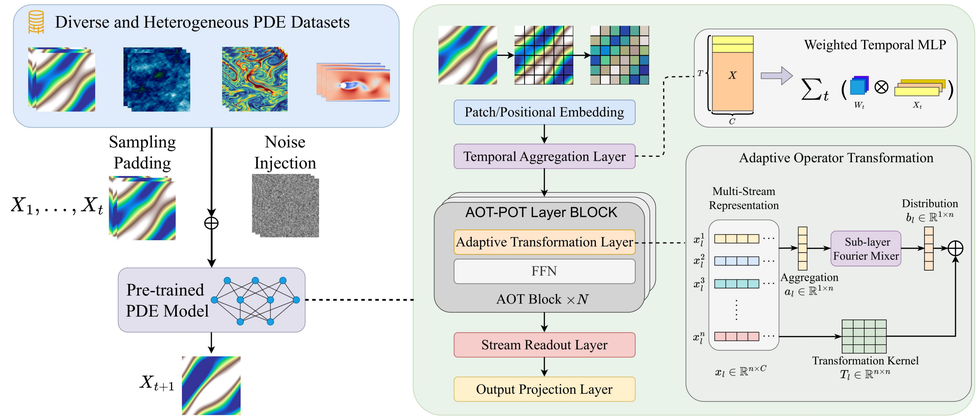
Different PDEs have different solution operators. Learn a transformation that makes each easier to approximate.



Only 40 added parameters. PDE-matched transforms consistently give the lowest error.

02 INPUT-ADAPTIVE OPERATOR TRANSFORMATION

Adapt the latent representation to the input PDE, then share the same Fourier operator backbone.



A
AGGREGATE
Select an input-dependent mixture of latent streams.

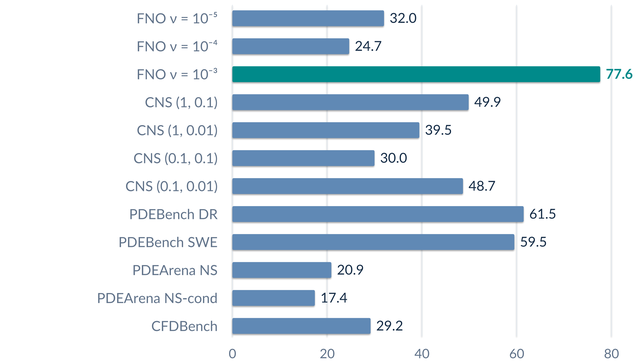
R
REDISTRIBUTE
Send the Fourier / MLP update back to all streams.

T
Transformed RESIDUALS
Mix streams with a Sinkhorn-normalized matrix.

T ≥ 0; rows and columns sum to 1. No PDE-type labels are needed to compute the transformations.

03 ACCURACY AT SMALL SCALE

AOT-POT-S vs. DPOT-S: relative L2 error reduction (%)



12 pre-training datasets; no per-dataset fine-tuning.

04 TRANSFER WITH FEW SAMPLES

Fewer fine-tuning examples on both seen and unseen PDE families.

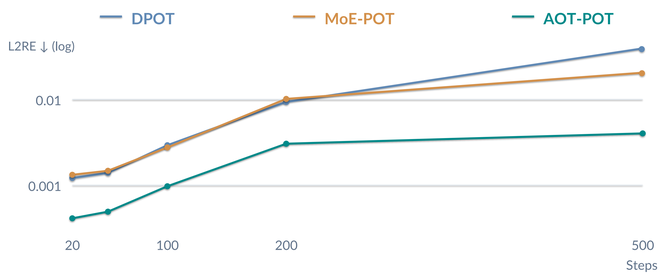
PDE task	AOT-POT · 16	DPOT · 64
NS (ν = 10^{-4})	**0.11**	0.13
Wave-Layer	**0.21**	0.26

Fine-tuned Small models · L2RE ↓ · 500 epochs

4× fewer fine-tuning examples

05 LONG-HORIZON PREDICTION

NS (ν = 10^{-3}) · fine-tuned Small models



86.2% lower error at 500 steps vs. DPOT.

THE TAKEAWAY

Learn better operator representations to improve accuracy, parameter efficiency and transfer

* Equal contribution. † Corresponding author.

PAPER
arxiv.org/abs/2605.15793
qitanlv@mail.ustc.edu.cn

CODE
OpenTSLab/
NeurIPS2026_AOT-POT