SMART Initiative

Science-informed Machine Learning to Accelerate Real Time (SMART) Decisions in Subsurface Applications

SAND2024-10174C

Transfer Learning for Multi-Physics Problems

Hongkyu Yoon,
Kadeethum Teeratorn, K-Won Chang
Sandia National Laboratories

2024 FECM / NETL Carbon Management Research Project Review Meeting

U.S. DEPARTMENT OF ENERGY

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories

NATIONAL ENERGY TECHNOLOGY LABORATORY

Types of Transfer Learning Approaches

Transfer learning with fine tuning/new layers

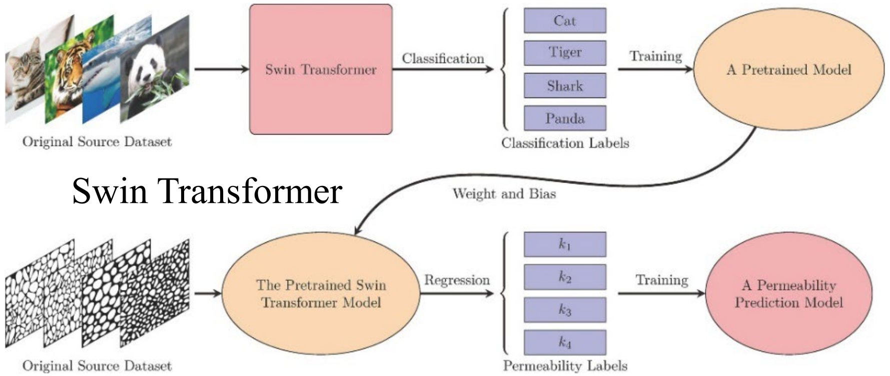
ImageNet
14M images, 200 classes

trained on

ResNet

add fresh layers, fine tuning, etc.

Predict 201st class

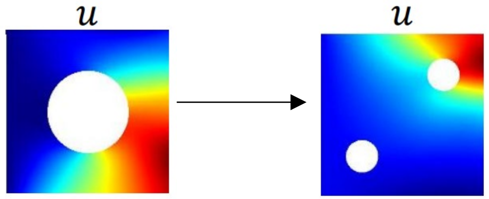


Swin Transformer

Physics based problems

Conditional embedding operator theory (CEOT)

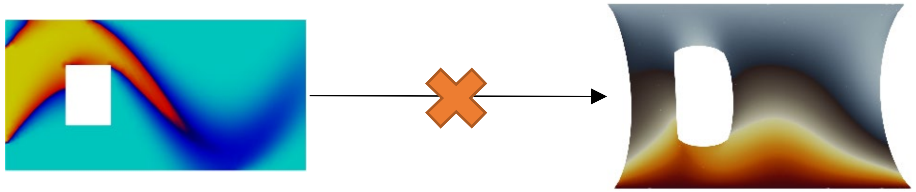
(Goswami et al., 2022)



Shortcomings

CEOT requires that underlying of source and target parameters share the same distribution

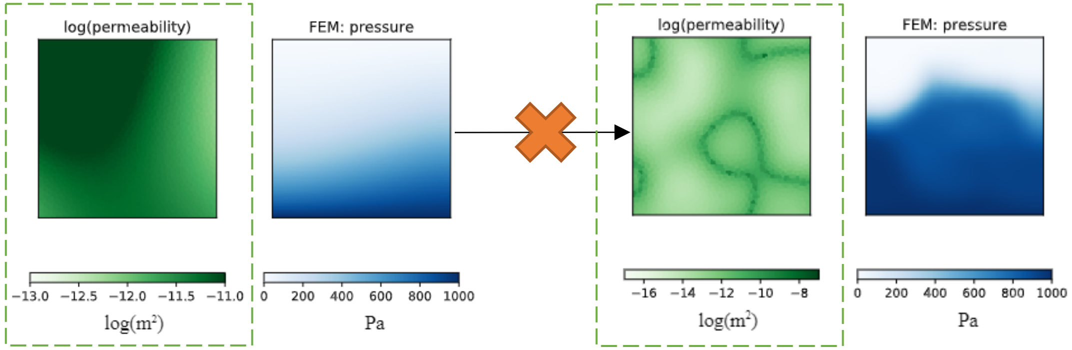
Different physics



Transport in porous media
Parameters: velocity on left BC

Hyperelasticity
Parameters: displacement sources

Different parameter distribution



Flow to coupled poroelasticity
Parameters: permeability fields

Transfer learning for multiple physics

Current Topic

Same geologic site

- Geologic features
- Operational conditions
- Number of wells

Multi-physics

- Single-phase flow
- Multi-phase flow
- Multi-phase flow and poroelasticity

Accuracy vs. training samples

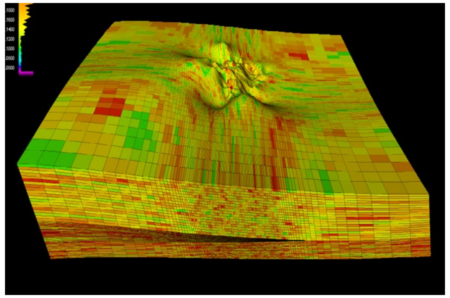
Can we achieve acceptable accuracy with less training data?

Our future work

Data: IBDP

- Illinois Basin – Decatur Project
- Heterogeneity in x-, y-, and z-directions
- Single injection well

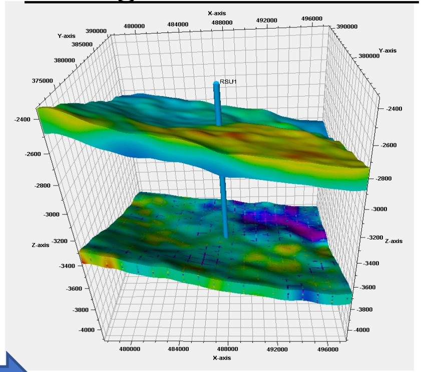
Surrogate: IBDP



Data: CarbonSAFE WY

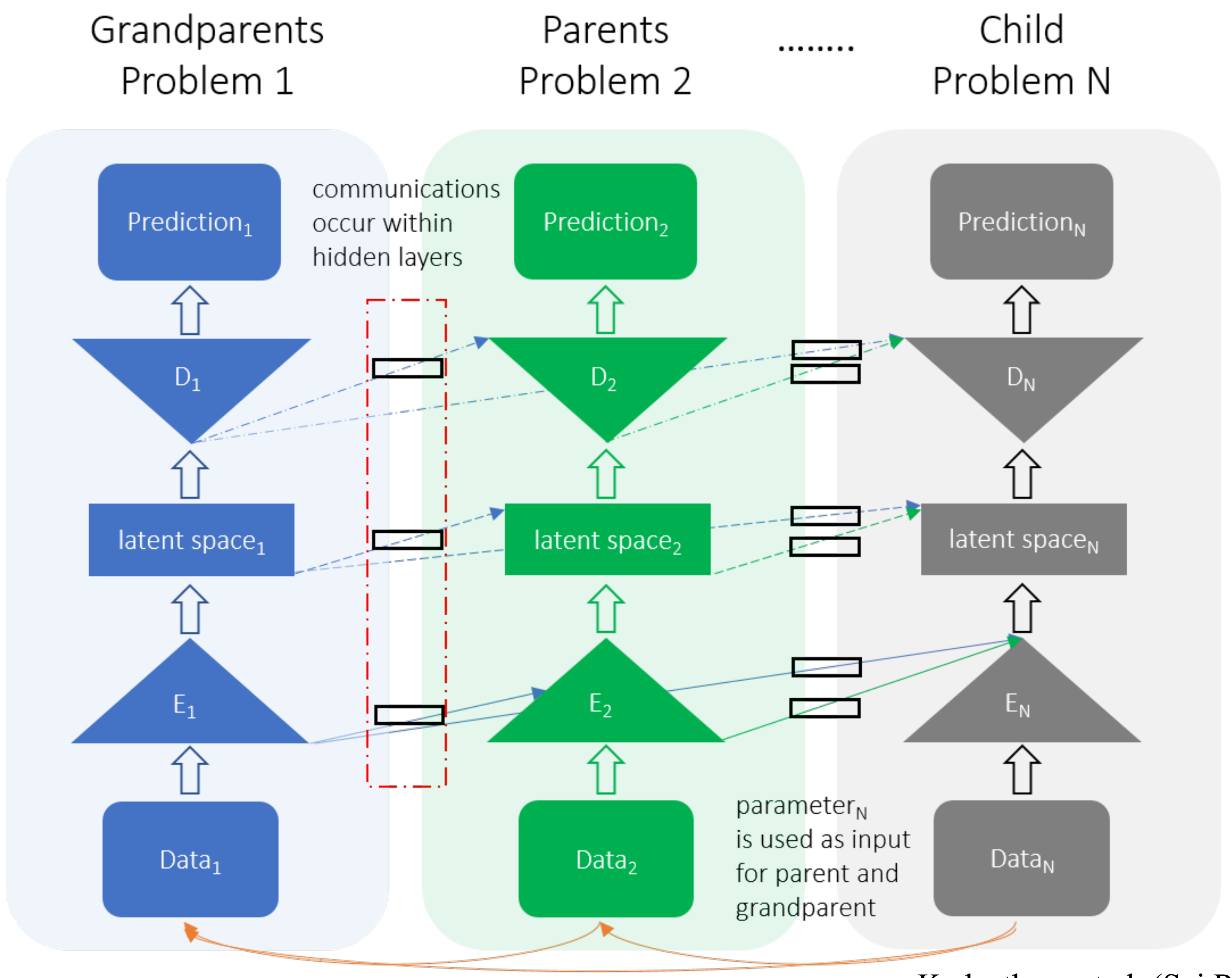
- Wyoming CarbonSAFE
- Different heterogeneity with geologic formations
- Multiple injection wells

Surrogate: CarbonSAFE



Utilize prior knowledge

Progressive learning



<u>Information transfer through gates in hidden layers</u>

Weights in hidden layers are used to control information flow from parents models to current generation, i.e., Weights are increased if output of each block is useful, and vice versa

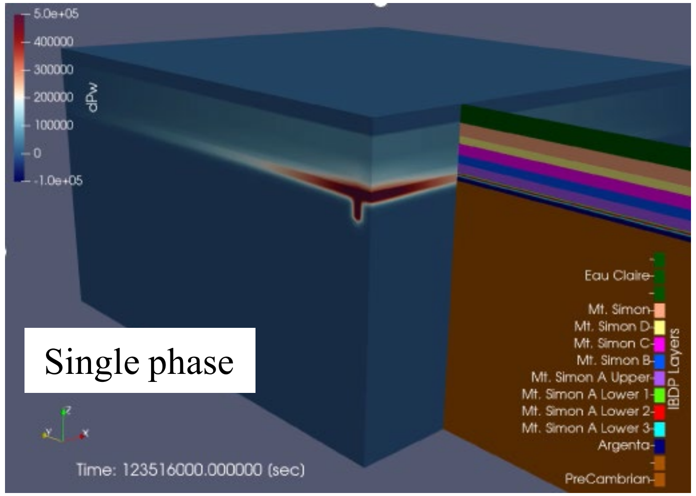
Different topology and sizes of input dimension can be handled with linear embedding layer through data input stream

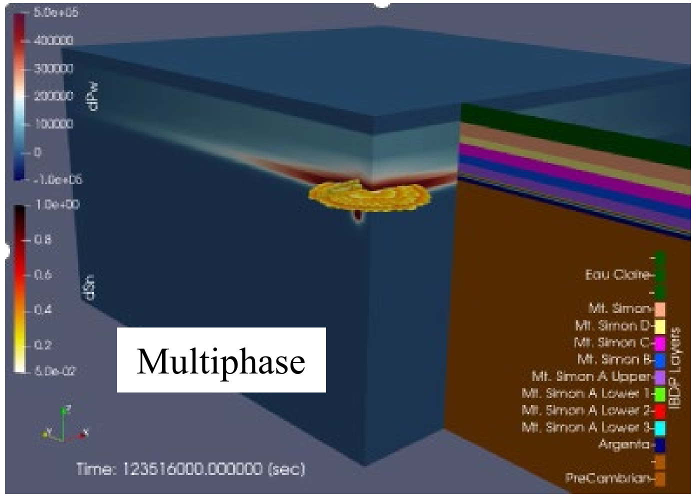
Kadeethum et al. (Sci Rep., 2024)

Rusu et al. (2016, Progressive Neural Networks)

Method: Sandia Sierra Mechanics for Multiphysics Simulations

- Sandia Sierra Mechanics Thermal/fluid module, ARIA (FE, **unstructured mesh ~ 1M DOF**)
- CO_2 Injection through CCS1 using a string function (better representation of well injection physics)
- Injection history is based on actual **daily** injection rates
- Simplified layered model domain with homogeneous/anisotropic domain and one single fault
- Single phase, multiphase, multiphase+poroelasticity (multiphase+thermoporoelasticity for EY24)





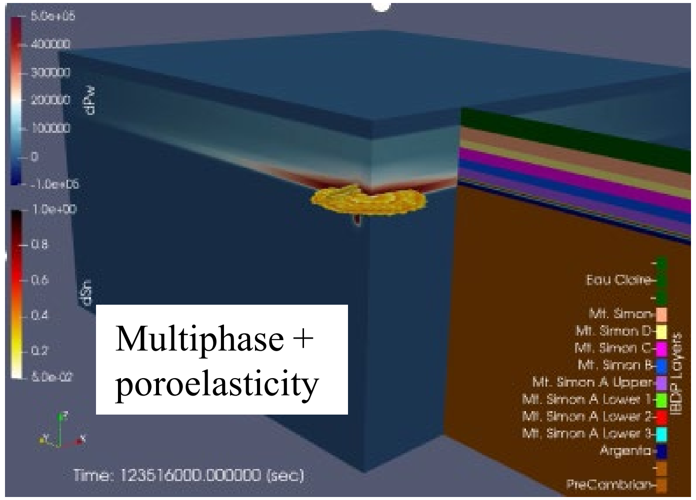


Table 1: Summary of varied parameters used in this data generation

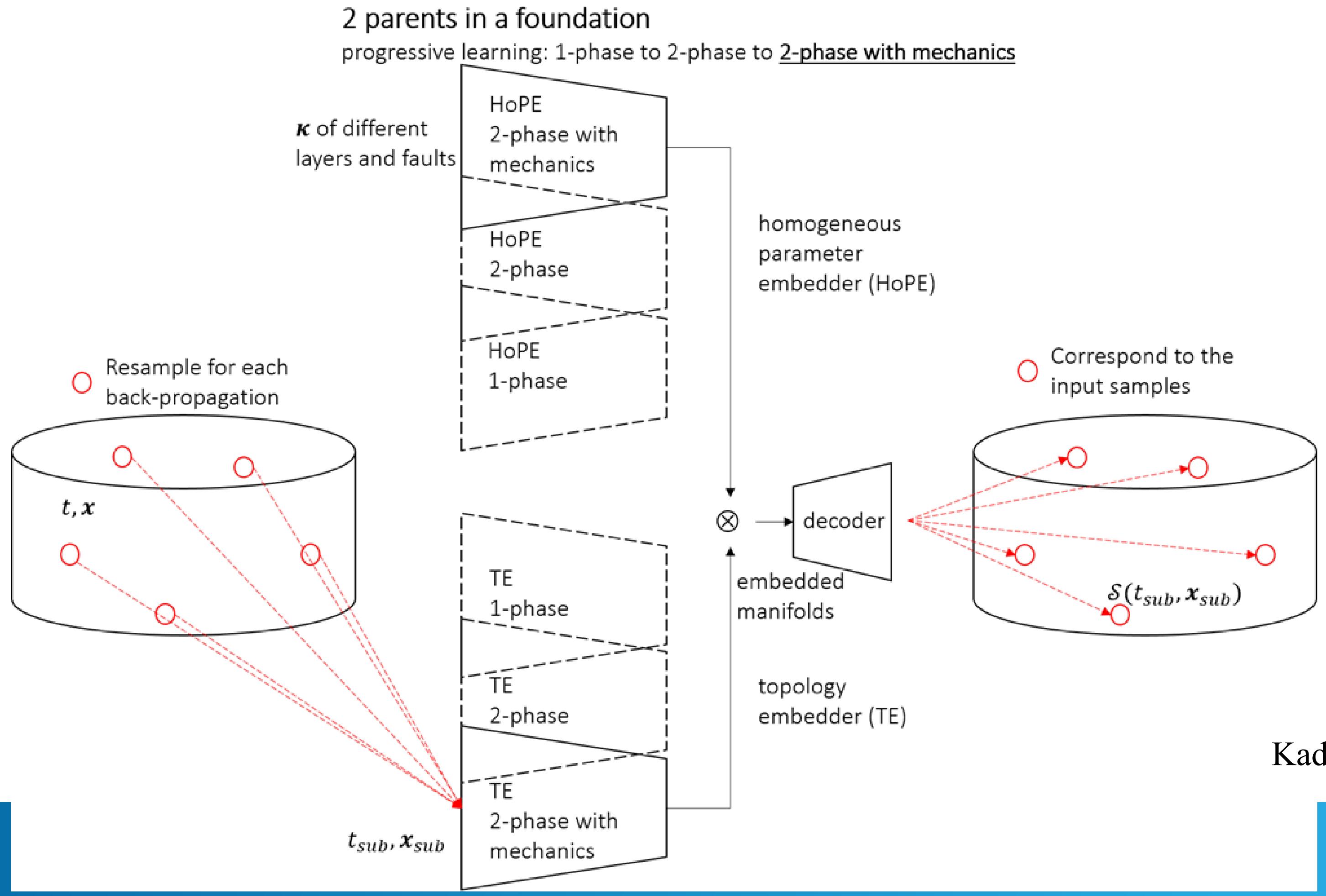
Parameter		Values	Count
κ_h (m^2)	reservoir (Mt. Simon A Upper)	2.47×10^{-14}, 1.35×10^{-14}, 4.93×10^{-14}	3
	interbedding (Mt. Simon A Lower 3)	2.96×10^{-13}, 1.97×10^{-13}, 1.48×10^{-13}, 1.13×10^{-13}, 4.93×10^{-14}	5
	basement (Argenta)	2.96×10^{-15}, 3.40×10^{-16}	2
	fault	9.87×10^{-15}, 4.93×10^{-15}, 9.87×10^{-16}	3
total			90

A total of 90 cases for training data generation

Kadeethum et al. (ARMA, 2024)

Progressive-Improved Neural Operator (p-INO)

- INO has been developed in Task 5 based on DeepONet architecture with subsampling based training to handle a large training data very efficiently (e.g., 3 hrs training time for IBDP case)
- INO can handle unstructured/structured data
- Trained INO model can predict at any positions and any time within interpolation regime

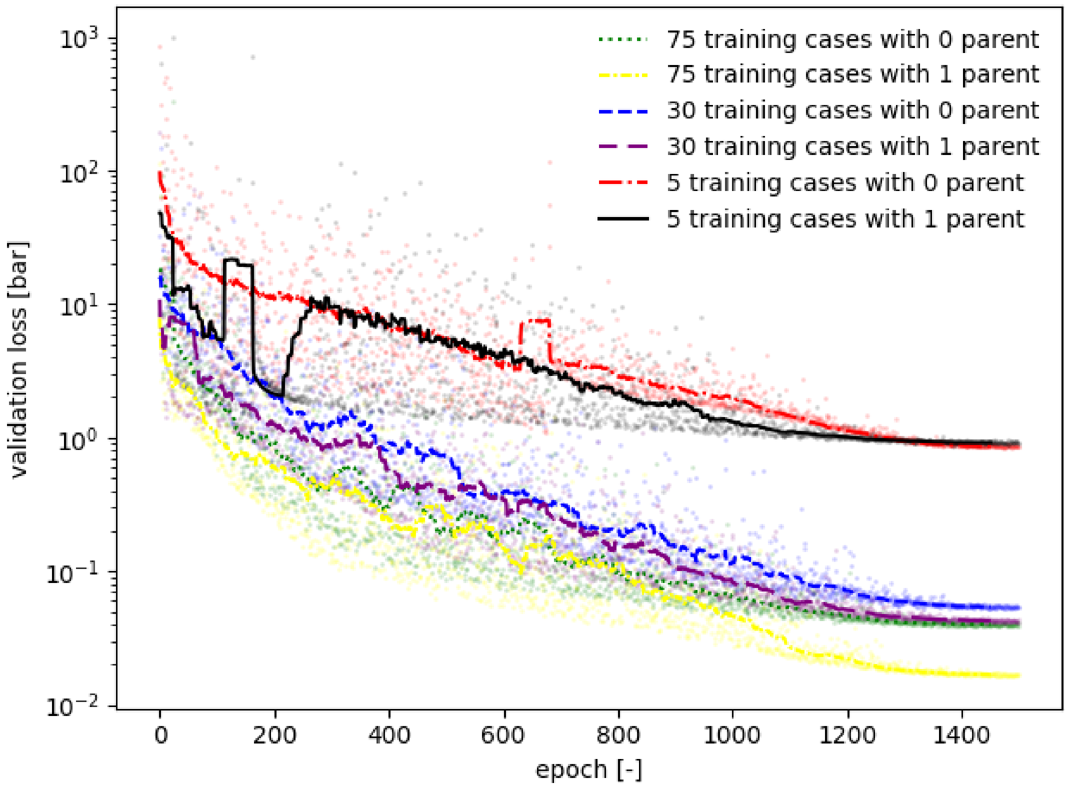


Kadeethum et al. (Geoenergy Sci. Eng., 2024)

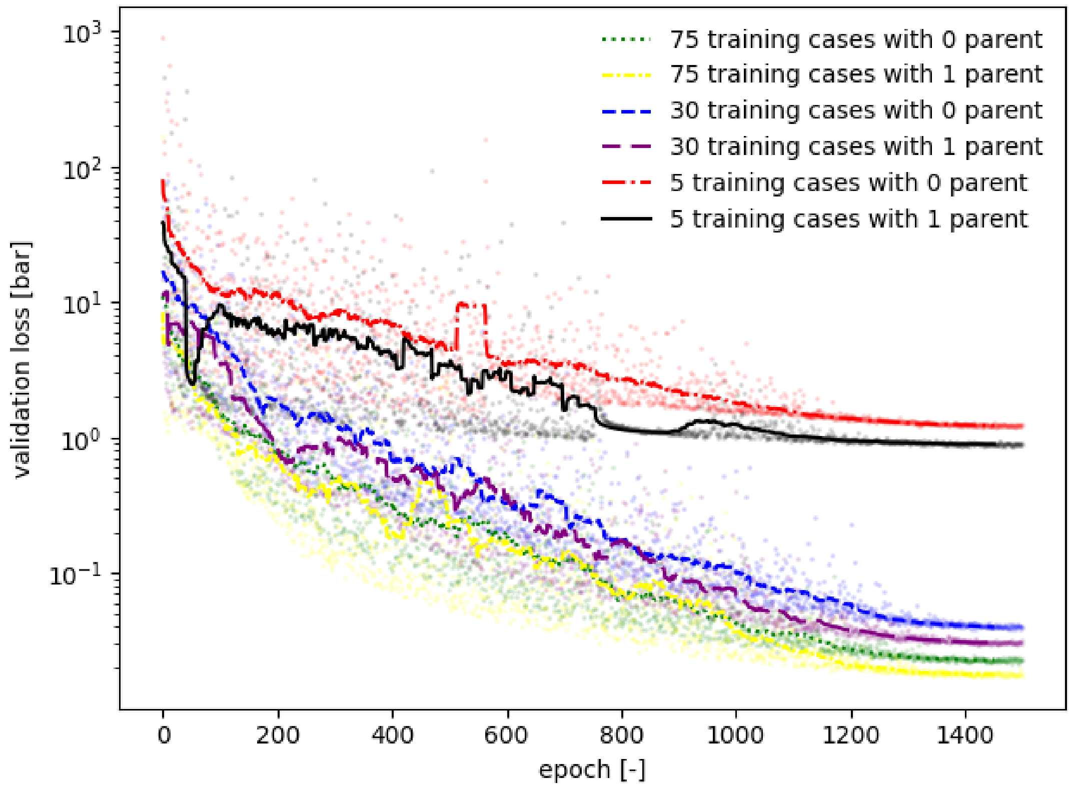
Results: Benefits of progressive learning

- Progressive learning can improve validation loss with 30/75 training cases for both cases
- Note that 0 parent means no progressive transfer learning

pressure: 1-phase to 2-phase



pressure: 1-phase to 2-phase with mechanics

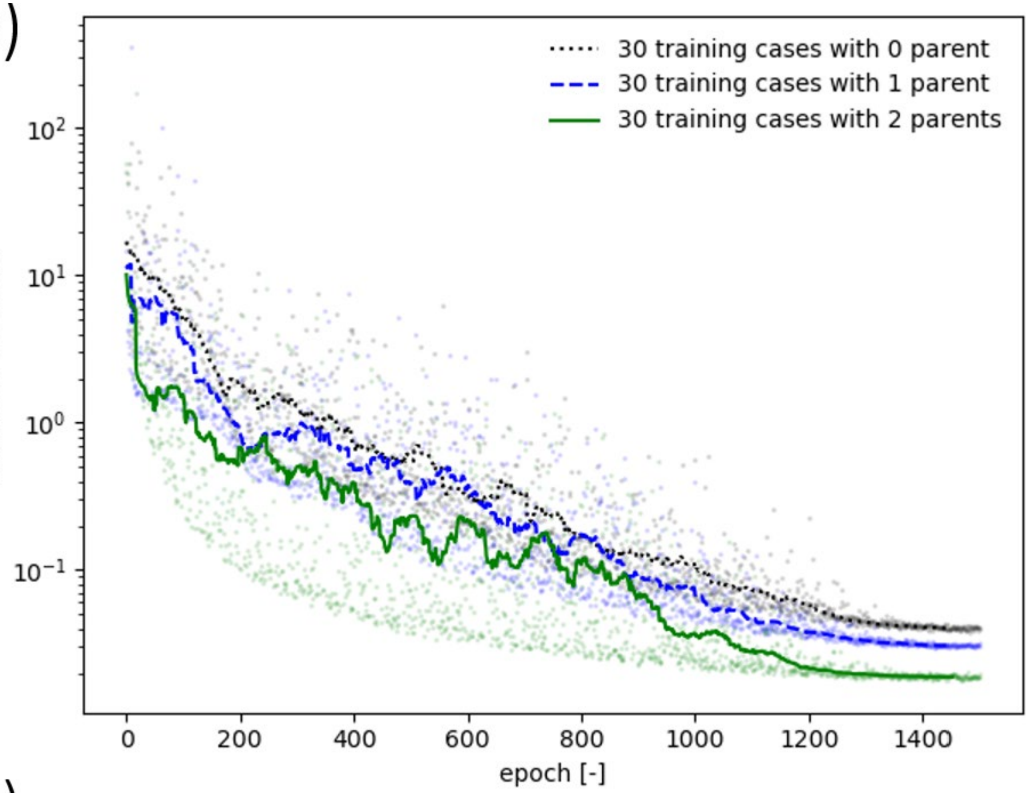


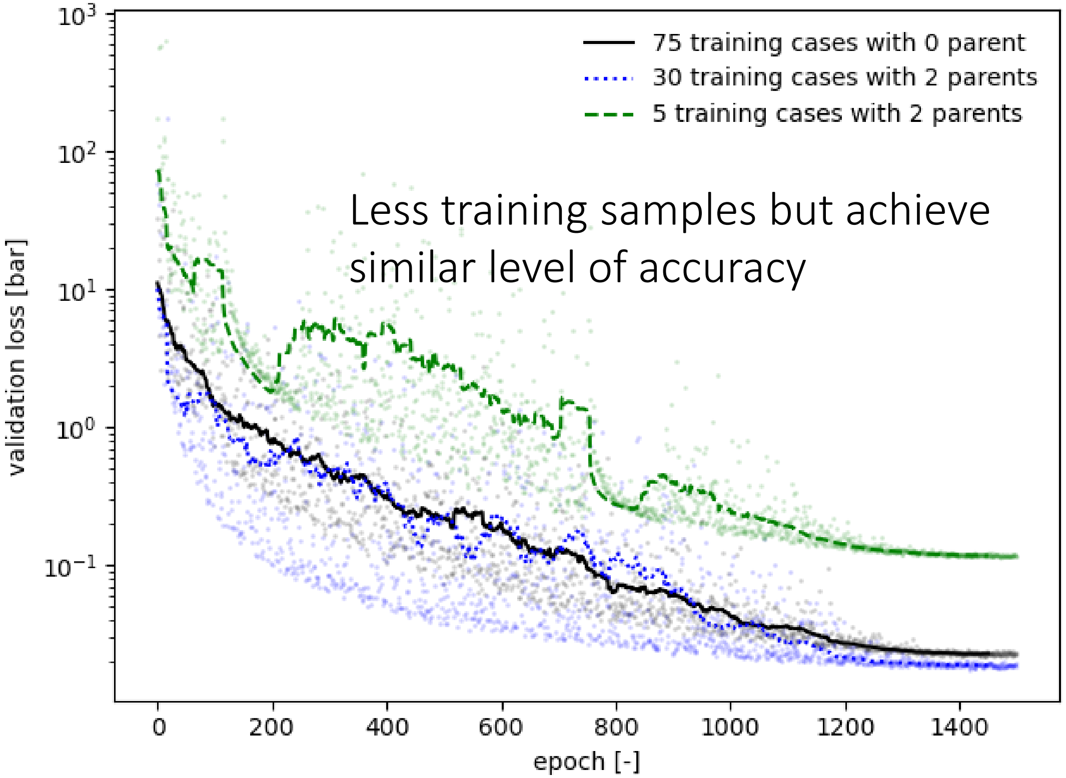
Results: Benefits of progressive learning

- More parents lead to higher accuracy with 30 training cases
- Two parents case with 30 training cases perform better than no parent case with 75 training cases, highlighting accuracy improvement as well as less data requirement

pressure: 1-phase to 2-phase with mechanics

b)

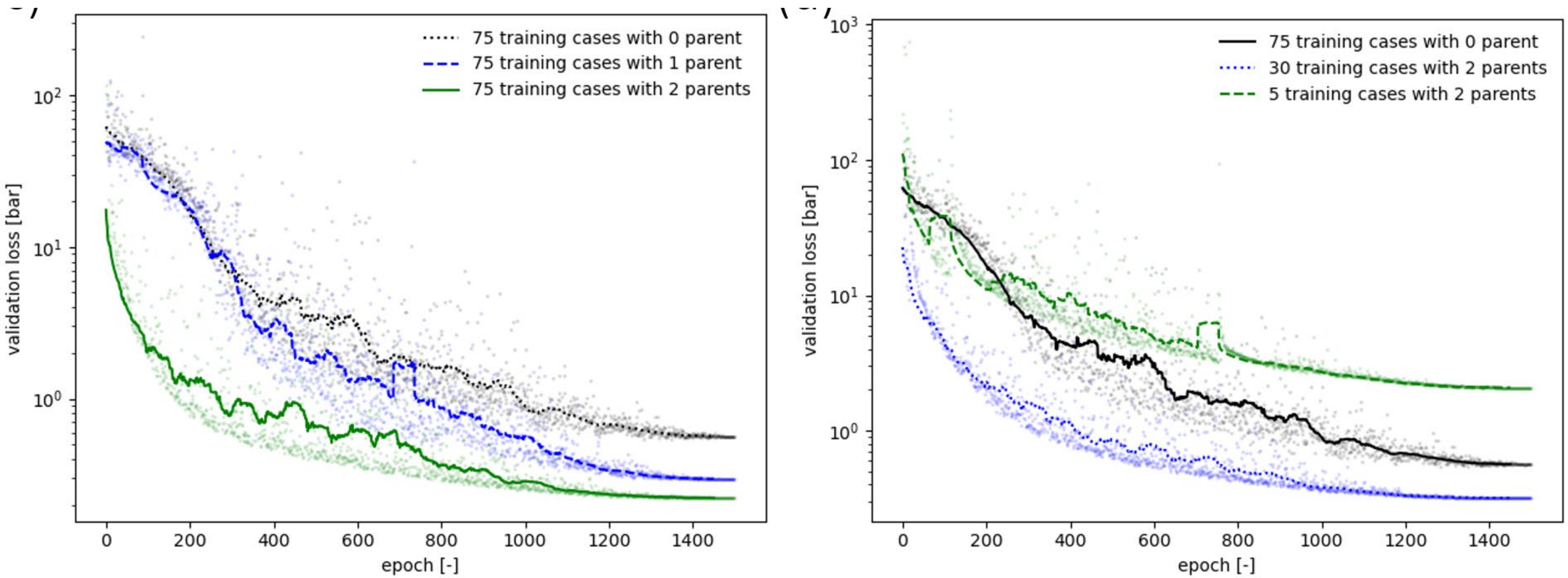




Results: around fault

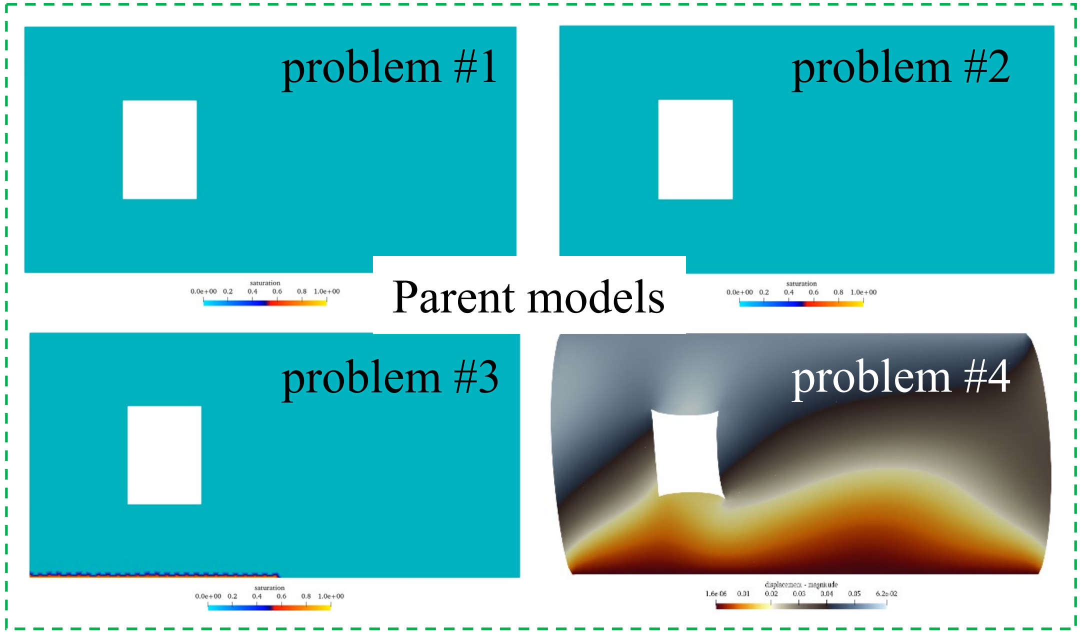
- More parents lead to higher accuracy with 75 training cases
- Two parents case with 30 training cases perform better than no parent case with 75 training cases, highlighting accuracy improvement as well as less data requirement for fault zone

pressure: 1-phase, 2-phase to 2-phase with mechanics



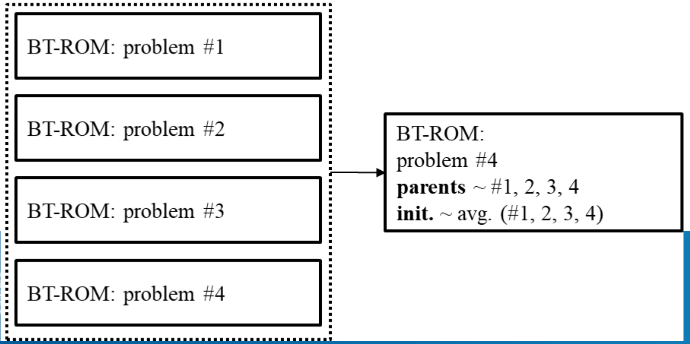
Results: Different geometries and physics

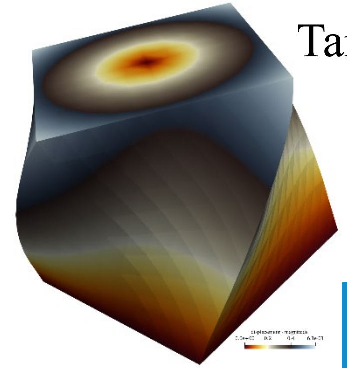
- Progress learning can be applicable for different geometries, boundary conditions, and different physics (e.g., from flow/transport to mechanics)



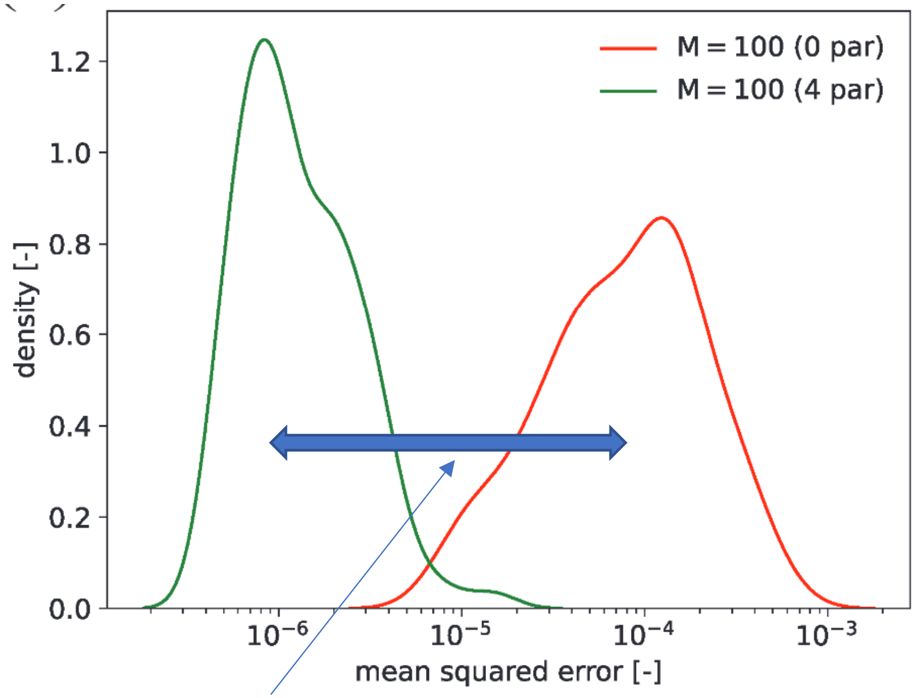
Parents:
Similar topology

Child:
Different topology





Target problem



accuracy gained by progressive learning

Kadeethum et al. (Sci Rep., 2024)

Summary

Primary benefits:

- Progressive learning can enhance accuracy with the same training data as in no progressive case
- Effectively reduce the training dataset generation requirements for more complex physics
- Current framework can be applicable for many different scenarios such as multiple well configurations and optimal injection cases

We aim to improve the following questions:

- Can we train base (parents) models using data generated by simple physics model rather than full physics model(s)? (e.g., Eikonal equation for pressure, percolation model for saturation)
- Current framework becomes expensive with many parent models. Can we prune unnecessary parts of parent models through more efficient neural networks such as attention mechanisms as in transformer architecture?