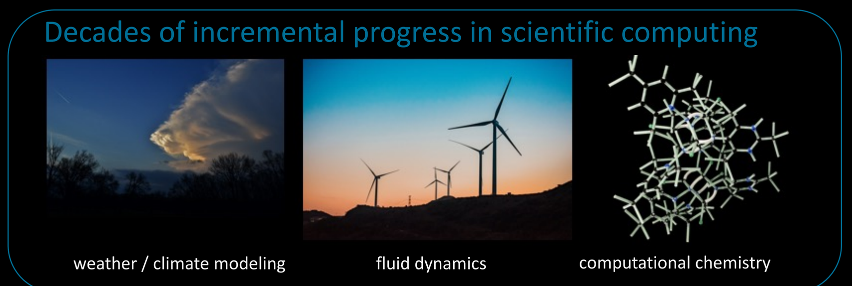
Using physics-informed regularization to improve extrapolation capabilities of neural networks

David Davini* — University of California, Los Angeles
Bhargav Samineni* — New Jersey Institute of Technology
Benjamin Thomas* — Louisiana State University
Amelia Huong Tran* — Mount Holyoke College
Cherlin Zhu* — Johns Hopkins University
Kyung Ha — University of California, Los Angeles
Ganesh Dasika — Advanced Micro Devices, Inc.
Laurent White — Advanced Micro Devices, Inc.

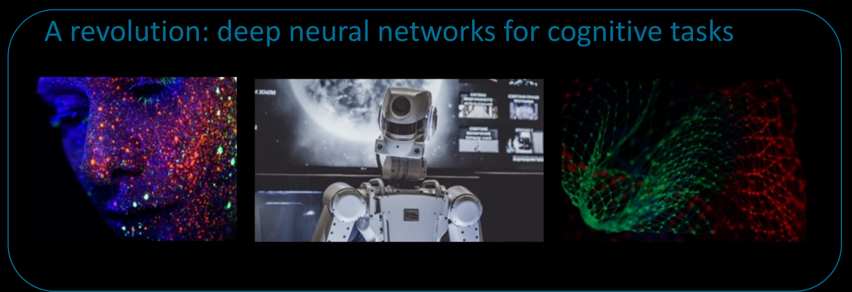
Scientific Computing Meets Machine Learning



Scientific computing community

- Sees opportunity for revolutionary improvement

Accelerating scientific applications with machine learning ("AI for Science")



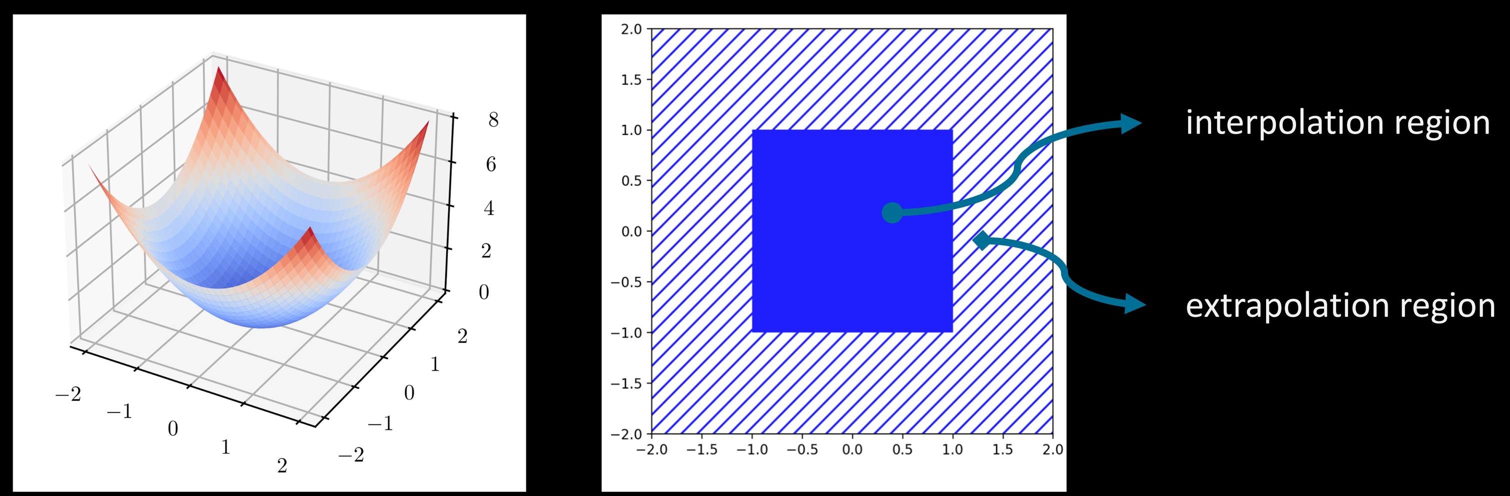
Machine learning community

- Building on prior success, wants to expand

Neural-network-based physics emulators suffer from lack of extrapolation capabilities. Exploring physics-based regularization to address this challenge.

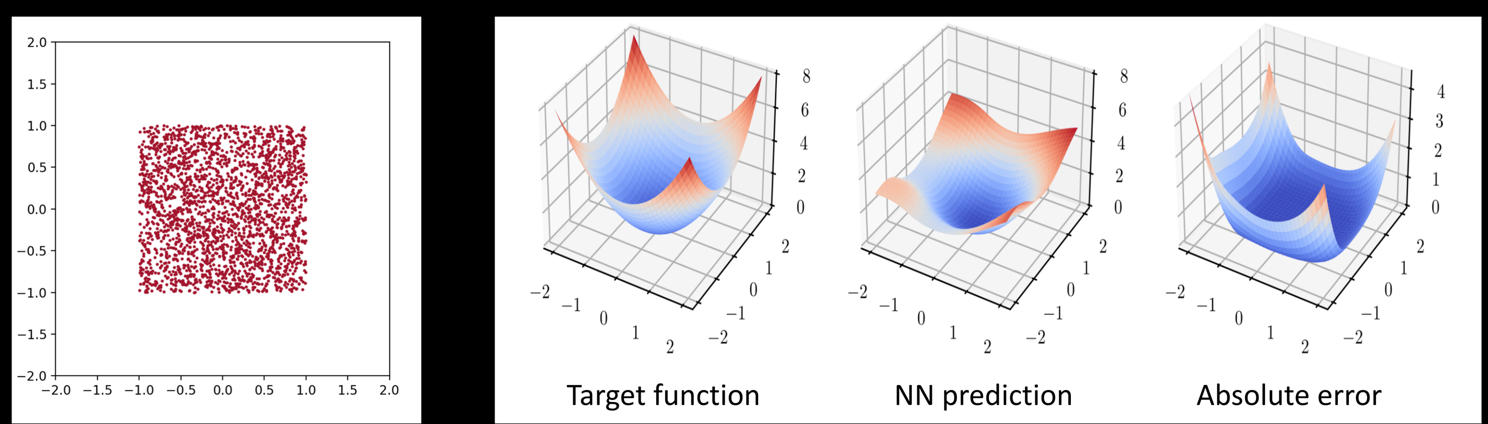
Improving Extrapolation Capabilities via Regularization: developing a strategy based on a two-dimensional function (paraboloid)

$f(x,y) = x^2 + y^2$ **Goal**: improve accuracy of neural network in extrapolation region while only using labeled data in interpolation region



Baseline: using labeled data in interpolation region $[-1,1]\times[-1,1]$ and no regularization

Loss function: $\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left\| f\left(\mathbf{x}^{\{i\}}\right) - \hat{f}\left(\mathbf{x}^{\{i\}}\right)\right\|_2^2$



Target function | NN prediction | Absolute error

Seeking to regularize the neural network based on information embedded in the function we're trying to predict, yet without using additional labeled data. Two examples:

"All second derivatives are constant"

$$E_{\text{second}}\left(\hat{f},\mathbf{x}\right) = \text{Var}\left(\hat{f}_{xx}(\mathbf{x})\right) + \text{Var}\left(\hat{f}_{yy}(\mathbf{x})\right) + 2\text{Var}\left(\hat{f}_{xy}(\mathbf{x})\right) \approx 0$$

"All third derivatives are zero"

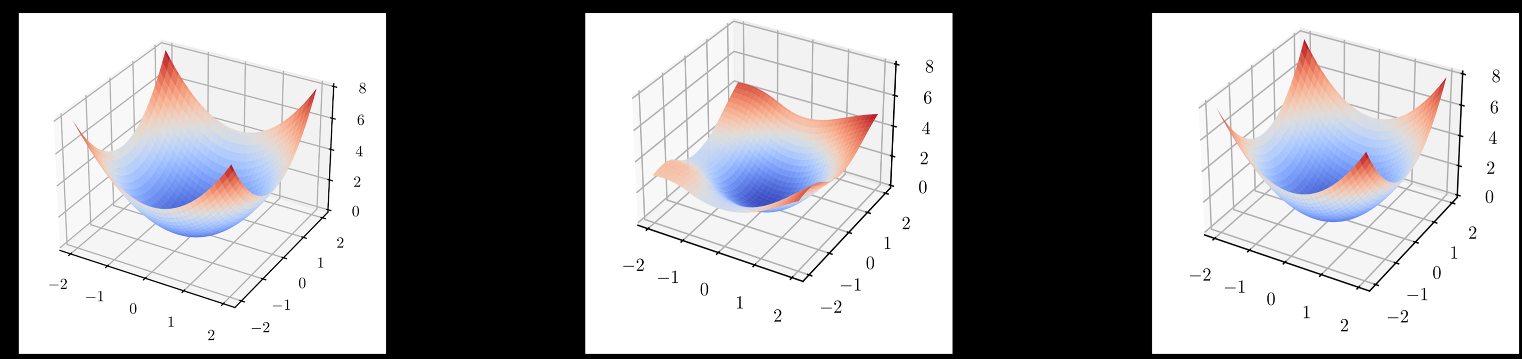
$$E_{\text{third}}\left(\hat{f},\mathbf{x}^{\{i\}}\right) = \left|\hat{f}_{xxx}\left(\mathbf{x}^{\{i\}}\right)\right| + \left|\hat{f}_{xxy}\left(\mathbf{x}^{\{i\}}\right)\right| + \ldots + \left|\hat{f}_{yyx}\left(\mathbf{x}^{\{i\}}\right)\right| + \left|\hat{f}_{yyy}\left(\mathbf{x}^{\{i\}}\right)\right| \approx 0$$

Adding third-order regularizer in $[-2,2]\times[-2,2]$

Loss function:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} L\left(\hat{f},\mathbf{x}^{\{i\}}\right) \text{ where}$$

$$L\left(\hat{f},\mathbf{x}^{\{i\}}\right) = \begin{cases} \left\| f\left(\mathbf{x}^{\{i\}}\right) - \hat{f}\left(\mathbf{x}^{\{i\}}\right)\right\|_2^2 + \lambda E_{\text{third}}\left(\hat{f},\mathbf{x}^{\{i\}}\right) & \text{Int.} \\ \lambda E_{\text{third}}\left(\hat{f},\mathbf{x}^{\{i\}}\right) & \text{Ext.} \end{cases}$$



Target function | Baseline (no regularization) Ext. error: 8.46×10^{-1} | Third-order regularizer Ext. error: 1.90×10^{-2}

Improving Extrapolation Capabilities via Regularization: two-dimensional acoustic wave equation

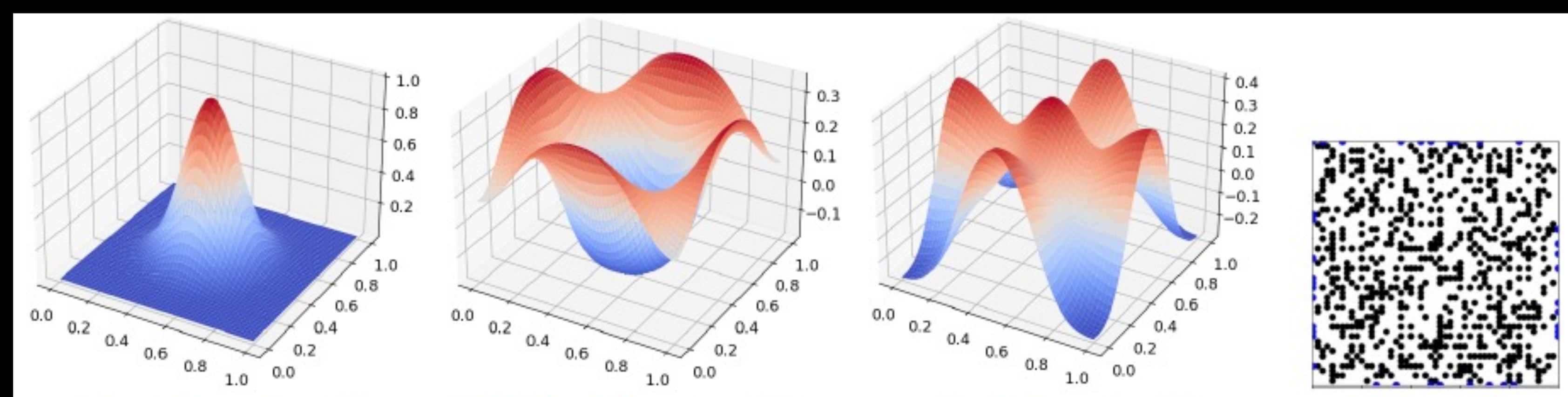
$$\frac{\partial p}{\partial t} + \kappa\nabla\cdot\mathbf{v} = 0$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla p = 0$$

Goal: improve accuracy of neural network in extrapolation region ($t > 1$) while only using labeled data in interpolation region ($t \leq 1$).

- Reflecting boundary conditions ($v_n = 0$)
- Labeled data: simulation data sampled in space every 10th step for $t \leq 1$ (time step is 0.001). Spatial samples based on random selection of 1% of discretization points (see Figure below).
- 10% of boundary points are also selected, where a boundary-condition regularizer can be applied.

The figure below shows simulation snapshots and an example of data collection (here for t = 0.56), where markers inside the domain denote locations for labeled data while markers on the boundary (blue markers) denote locations where a boundary-condition regularizer can be applied.



(a) Simulation (t = 0.0) (b) Simulation (t = 0.4) (c) Simulation (t = 1.0) (d) Sampling

Augmenting the loss function with physics-based regularization terms that do not depend on labeled data

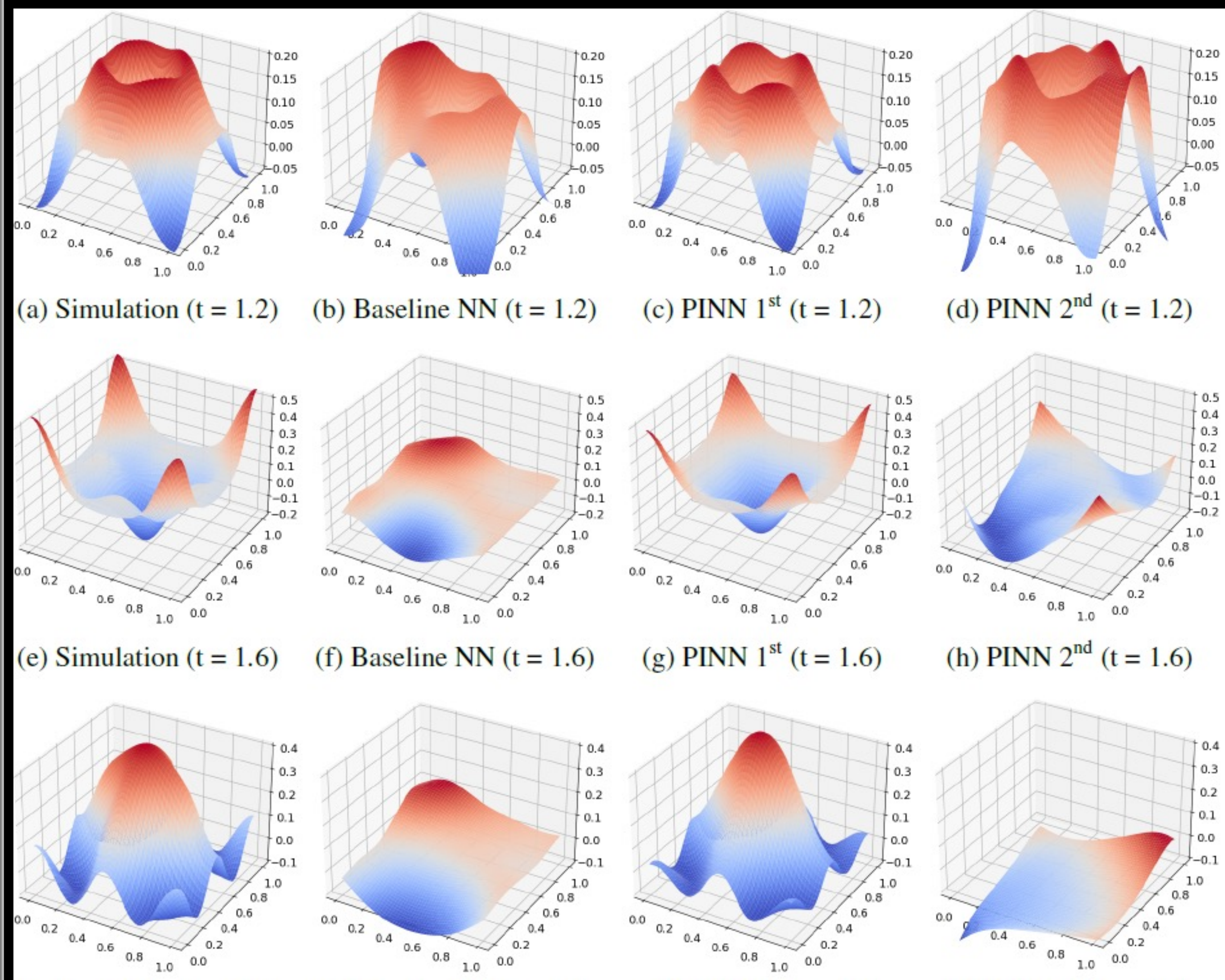
Loss function: $L = MSE + \lambda_r E_r + \lambda_b E_b$

- MSE: Mean squared error based on labeled data
- $\lambda_r E_r$: PDE-based regularizer (not based on labeled data)
- $\lambda_b E_b$: Boundary-condition regularizer (not based on labeled data)

The table below shows the performance of various physics-based regularization strategies (average of five runs, standard deviation shown in parentheses). PINN = Physics-informed neural network.

Name	E_r	E_b	Interp. error	Extrap. error
Baseline NN	None	None	1.8×10^{-3} (43%)	1.8×10^{-1} (16%)
PINN 1st	$\lvert\hat{p}_t + \nabla\cdot\hat{\mathbf{v}}\rvert + \lVert\hat{\mathbf{v}}_t + \nabla\hat{p}\rVert$	$\lVert\hat{\mathbf{v}}_n\rVert$	7.3×10^{-3} (9%)	3.1×10^{-2} (35%)
PINN 2nd	$\lvert\hat{p}_{tt} - \nabla^2\hat{p}\rvert$	None	1.8×10^{-2} (7%)	1.9×10^{-1} (25%)

The figure below compares the extrapolation performance between simulation (left column) and predictions from neural networks trained with different regularization strategies.

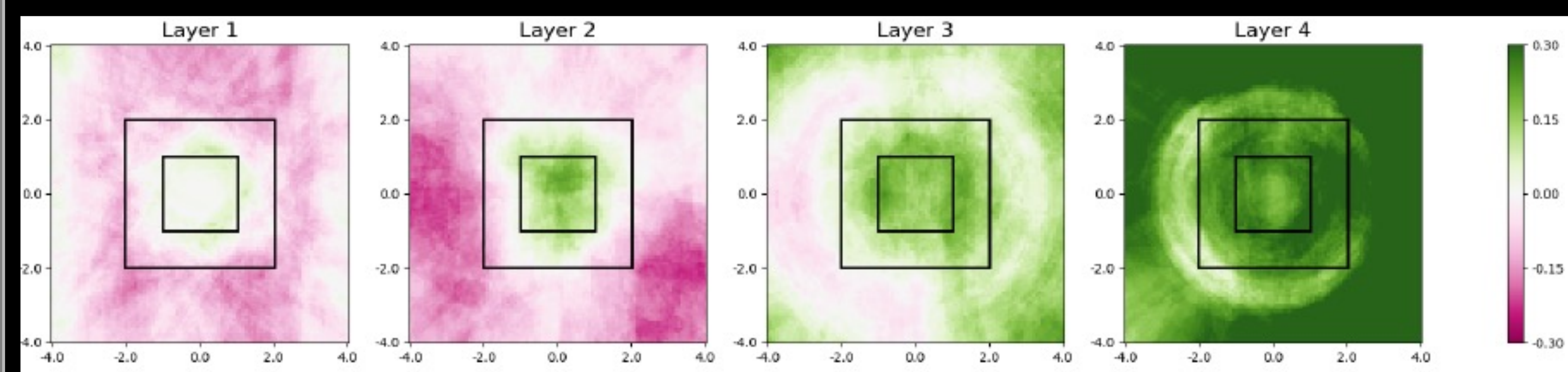


(a) Simulation (t = 1.2) (b) Baseline NN (t = 1.2) (c) PINN 1st (t = 1.2) (d) PINN 2nd (t = 1.2)
(e) Simulation (t = 1.6) (f) Baseline NN (t = 1.6) (g) PINN 1st (t = 1.6) (h) PINN 2nd (t = 1.6)
(i) Simulation (t = 2.0) (j) Baseline NN (t = 2.0) (k) PINN 1st (t = 2.0) (l) PINN 2nd (t = 2.0)

Future work

Develop a posteriori error estimates and more efficient sampling strategies based on saturation.

The figure below shows the difference in layer-wise saturation between the physics-informed NN and the baseline NN for the paraboloid target function. Negative values across the first two layers indicate that the PINN is less saturated in the extrapolation region (and therefore, more effective). This behavior could be leveraged to develop a posteriori error estimates or more efficient sampling strategies



(*) These five authors contributed equally. Work performed while undergraduate students Davini, Samineni, Thomas, Tran, and Zhu attended UCLA/IPAM's Research in Industrial Projects for Students summer program (Ha and White were academic and industrial mentors, respectively).