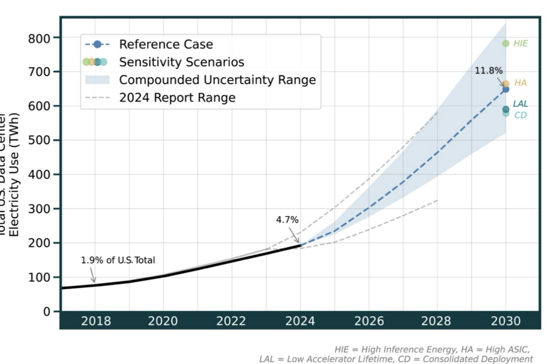
Computing with Chaos: a Compact Wave-Chaotic Neural Network

Ethan Valentino, John Guillamon, Tsampikos Kottos
Wave Transport in Complex Systems Lab, Department of Physics, Wesleyan University

Abstract

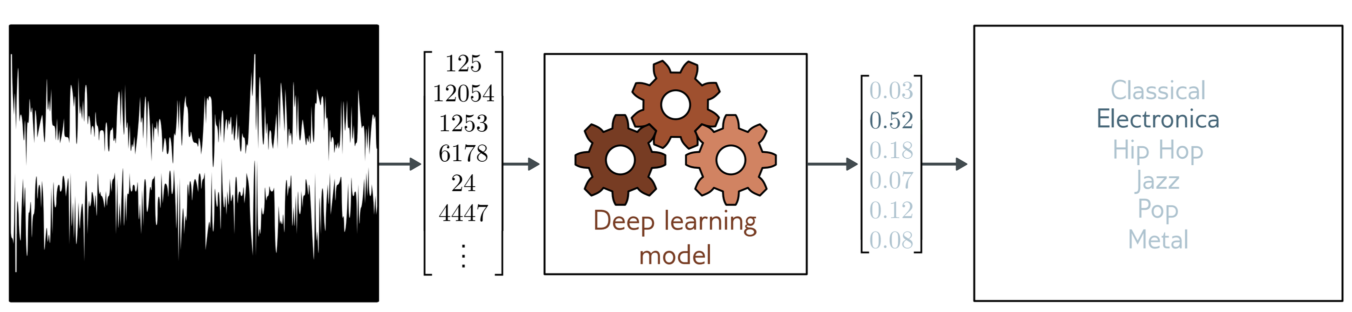
Artificial neural networks (ANNs) have achieved remarkable success at a variety of machine learning tasks, and have thus become the foundation of modern artificial intelligence. Their success has motivated the study of physical neural networks (PNNs), which replace the heavy digital computational load of ANNs with a physical system capable of performing nonlinear operations at a fraction of the energy and time. Here, we realize a PNN using a compact graph of interconnected coaxial cables that produces wave-chaotic scattering. The rapidly varying frequency dependence of the scattering matrix provides a large set of effectively uncorrelated nonlinear responses within a narrow bandwidth, enabling computational complexity to scale through frequency diversity rather than hardware size. Using probe frequencies as physical trainable parameters, our four-vertex, fully connected network achieves higher accuracy on the MNIST digit-classification task than other PNN proposals. We also demonstrate that the Universal Approximation Theorem holds for our system, meaning it is expressive enough for any machine learning task.

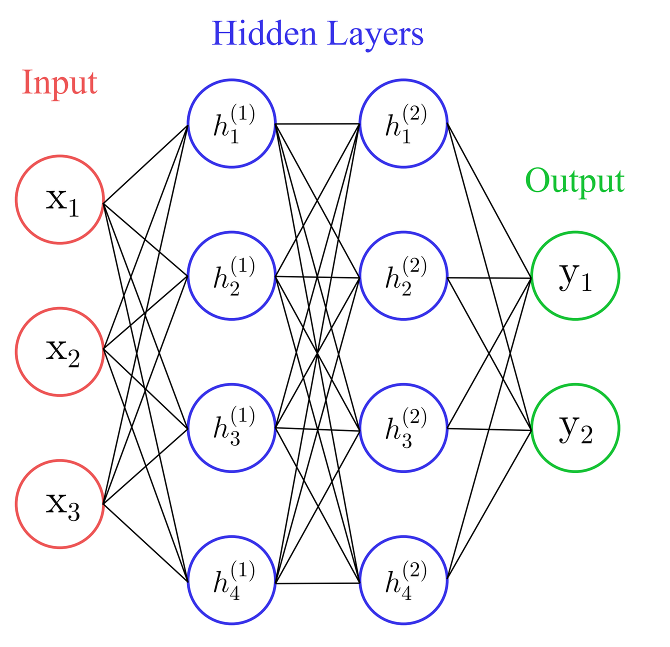
Motivation



- AI is increasing demand for data centers, consuming massive amounts of energy
- A machine learning model that relies on fewer digital computions would require fewer data centers and would consume less energy

Digital Neural Networks

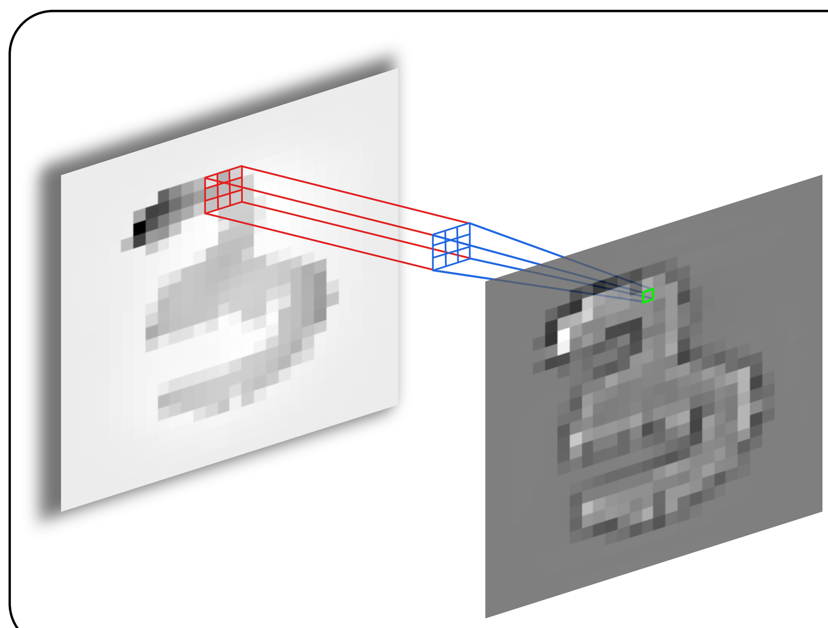




$$\mathbf{h}^{(1)} = \sigma(A_1\mathbf{x} + \mathbf{b}_1)$$
$$\mathbf{h}^{(2)} = \sigma(A_2\mathbf{h}^{(1)} + \mathbf{b}_2)$$
$$\mathbf{y} = A_3\mathbf{h}^{(2)} + \mathbf{b}_3$$

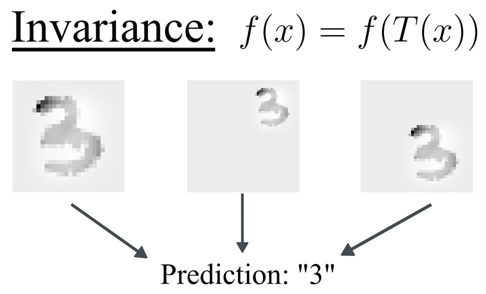
$\mathbf{x}$: Input vector
$\mathbf{h}$: Hidden unit vectors
σ: Nonlinear activation function
A: Trainable weights matrices
$\mathbf{b}$: Trainable bias vectors
$\mathbf{y}$: Output vector

Convolutional Neural Networks

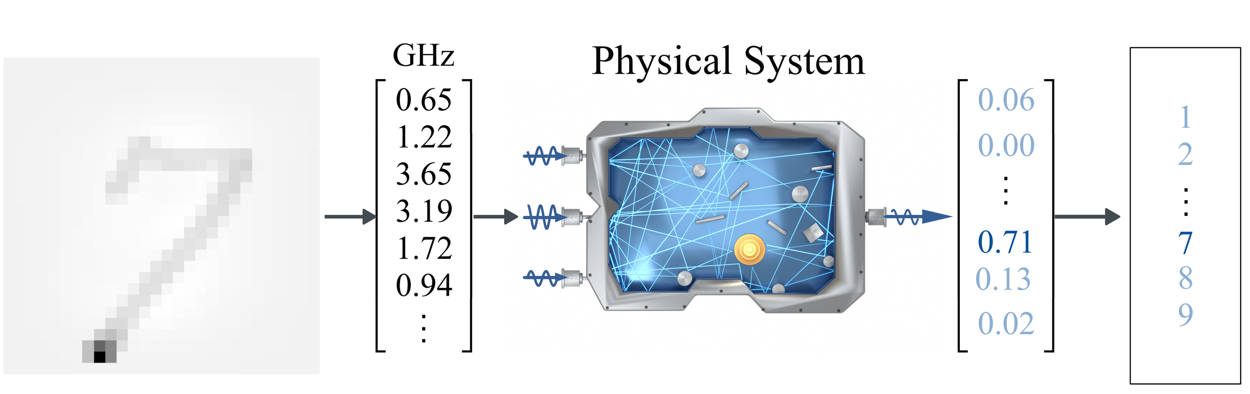


$$\overbrace{Y_{ij}}^{\text{Output pixel}} = \sum_{m,n} \overbrace{K_{mn}}^{\text{Kernel}} \underbrace{X_{i+m,j+n}}_{\text{Input pixel}}$$

Invariance: $f(x) = f(T(x))$



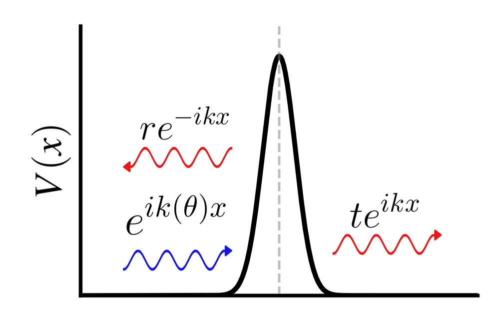
Physical Neural Networks



- Use natural dynamics of a physical system as the nonlinear activation function
- Tunable parameters are physical rather than digital

Wave Scattering

One-dimensional scattering:

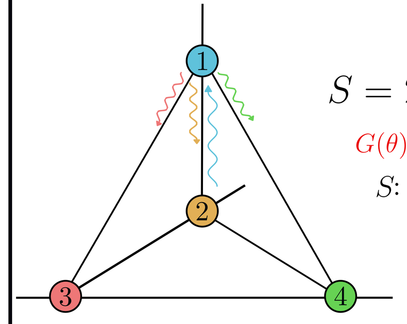


$\theta = \{A,\ \mathbf{b}\}$

Output amplitudes are **nonlinearly** related to input parameters

Network scattering:

Scattering in a graph is just a series of one-dimensional scattering events

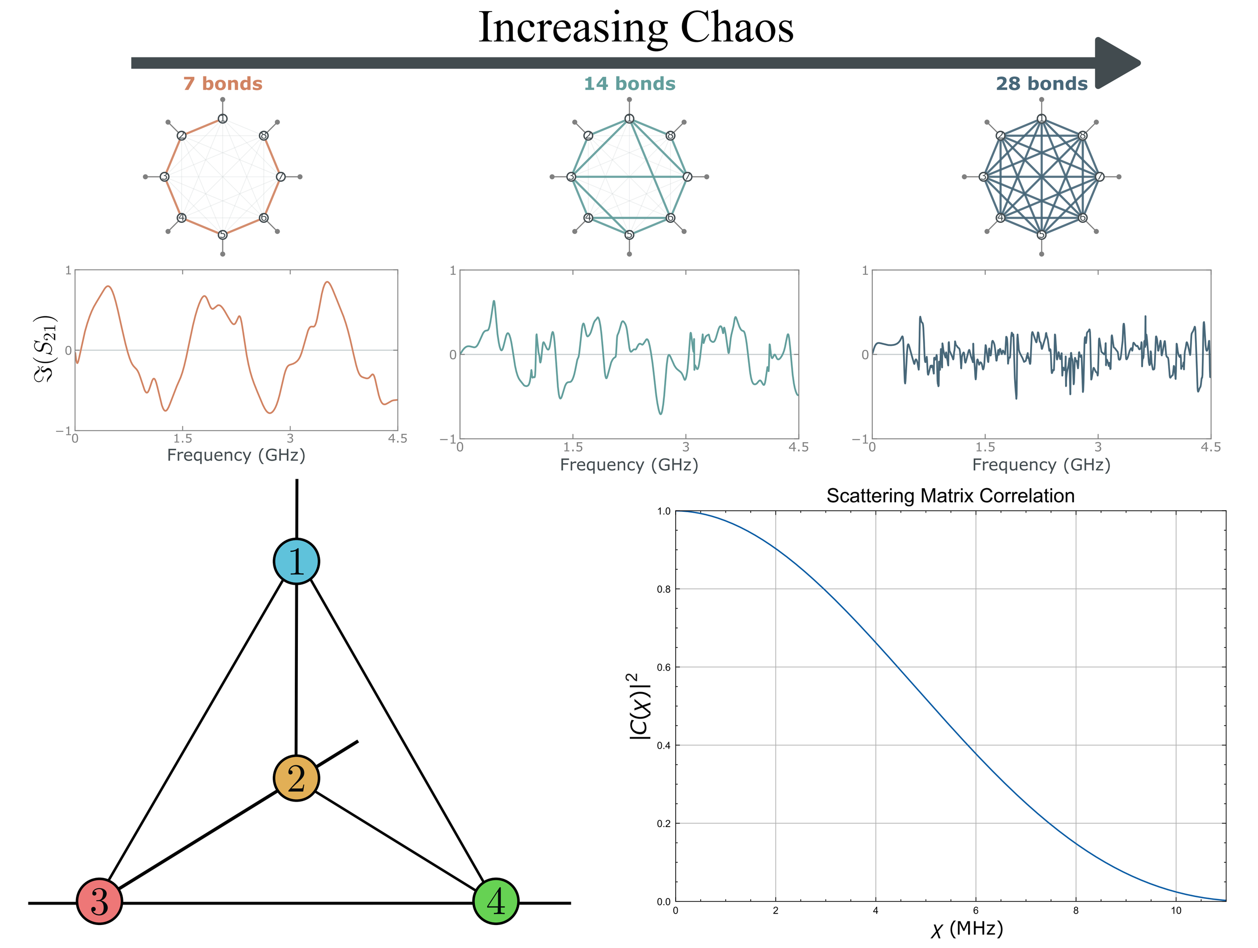


$$O = SI$$
$$S = 2iWG(\theta)W^T - \mathbb{I}$$
$$G(\theta) = [iW^TW + H(\theta)]^{-1}$$

S: Scattering Matrix
O: Output
I: Input

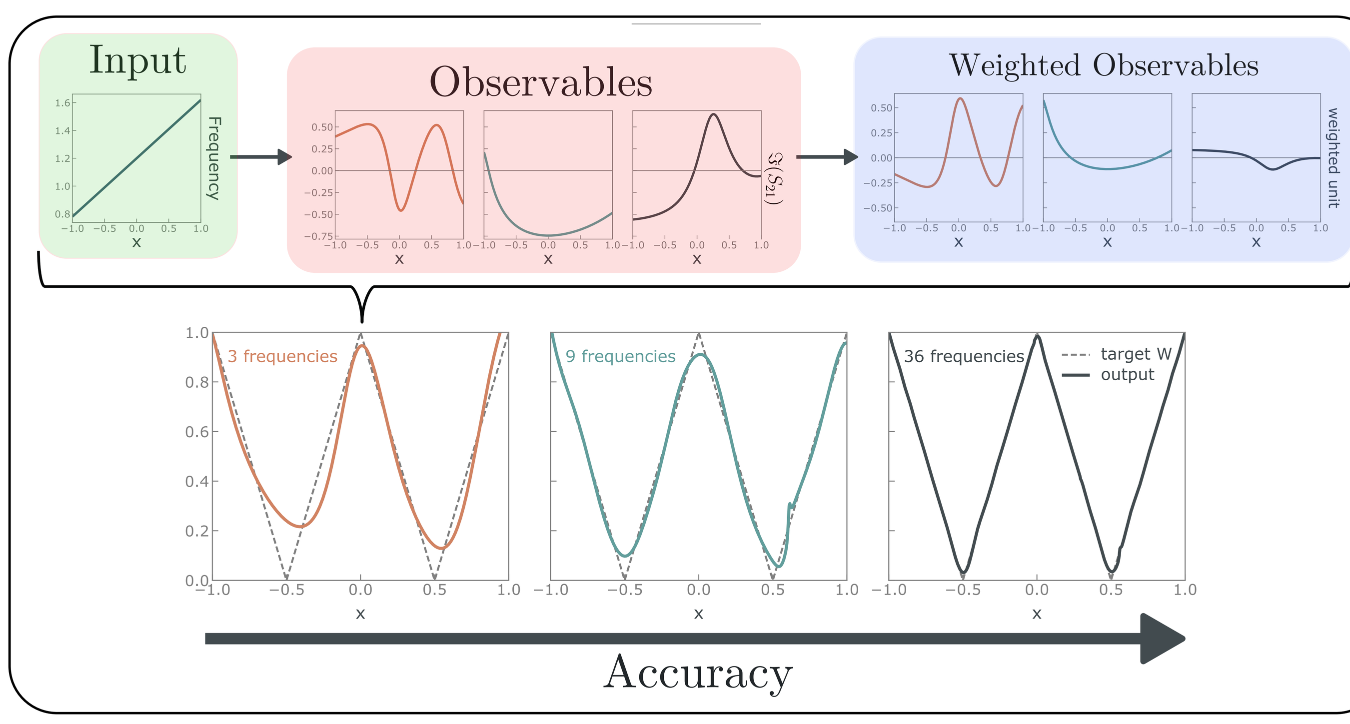
Matrix inversion provides a **nonlinear** transformation from parameters to output

Chaotic Frequency Response



A sufficiently complicated graph yields **high decorrelation** between scattering matrices at different frequencies, minimizing the amount of redundant information we extract from the system.

Function Approximation



Optimization

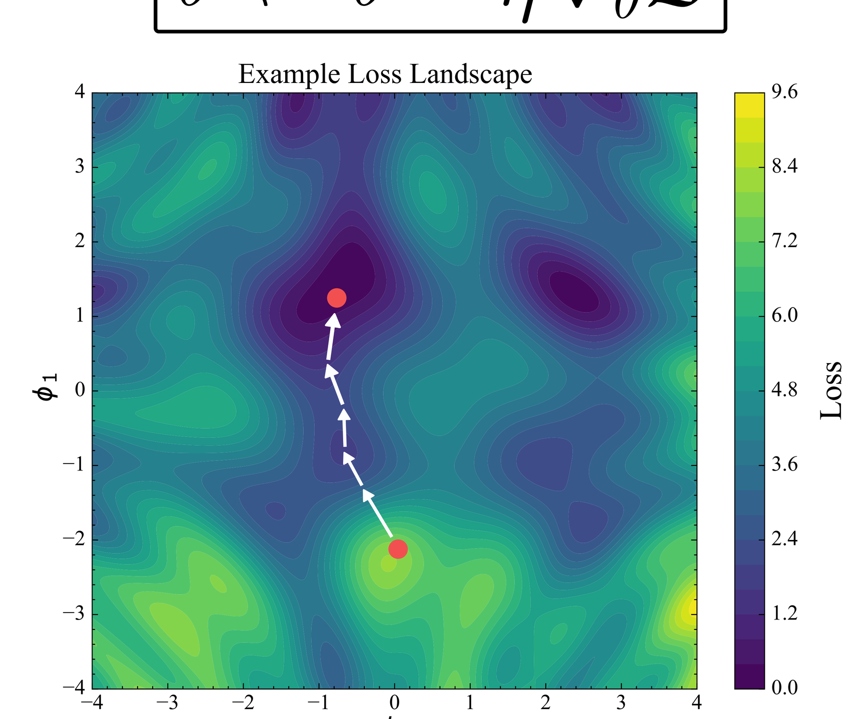
Cross-entropy loss quantifies the model's prediction error:

$$\boxed{\mathcal{L} = -log(p_y)}$$

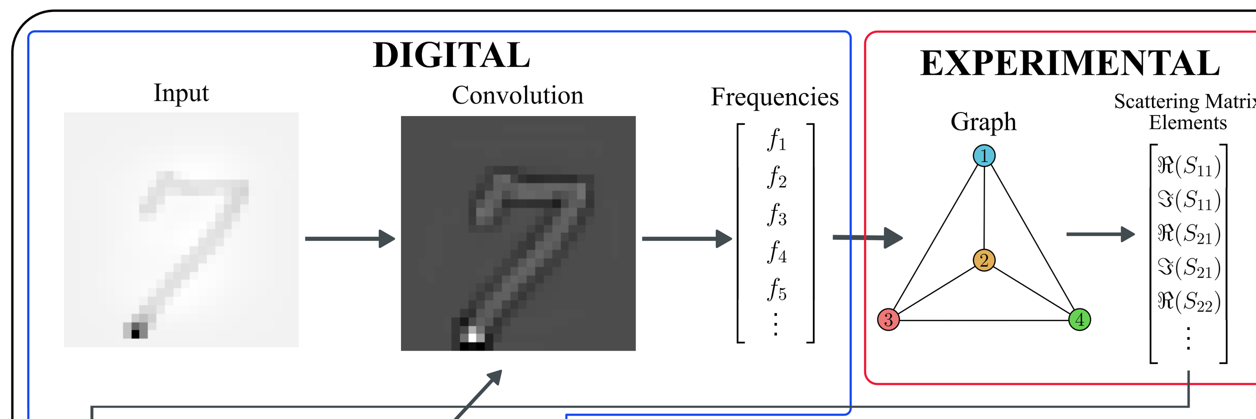
p_y: probability assigned to the correct digit

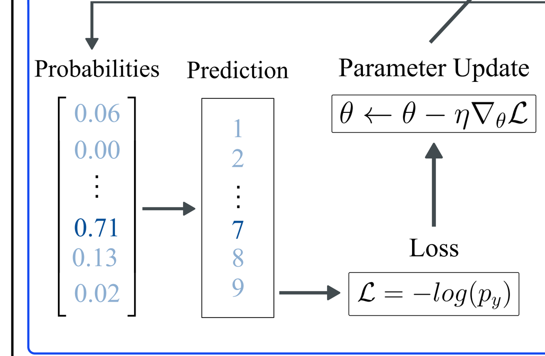
Parameter update with learning rate η:

$$\boxed{\theta \leftarrow \theta - \eta\nabla_\theta\mathcal{L}}$$

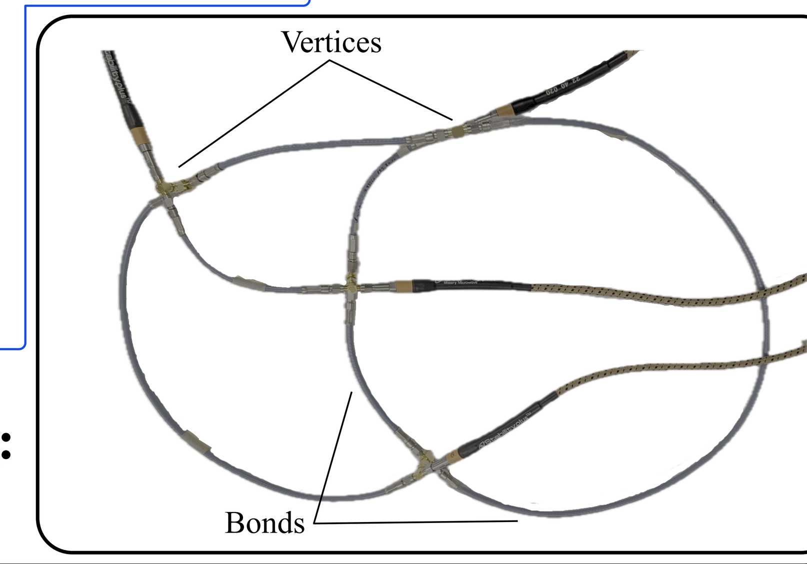


Experimental Platform

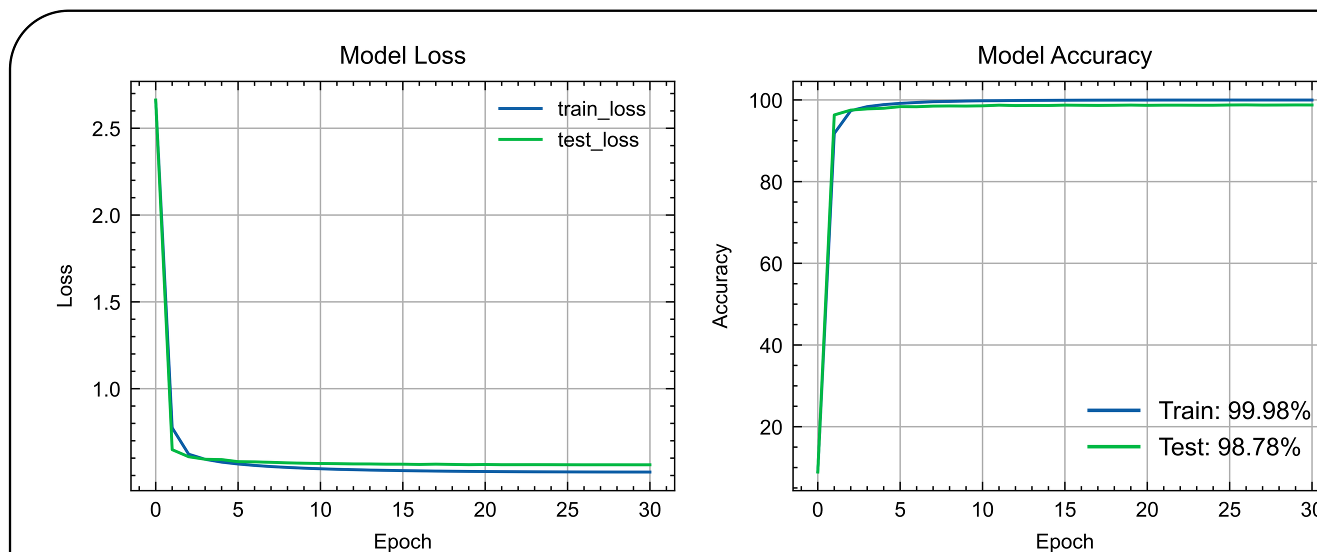


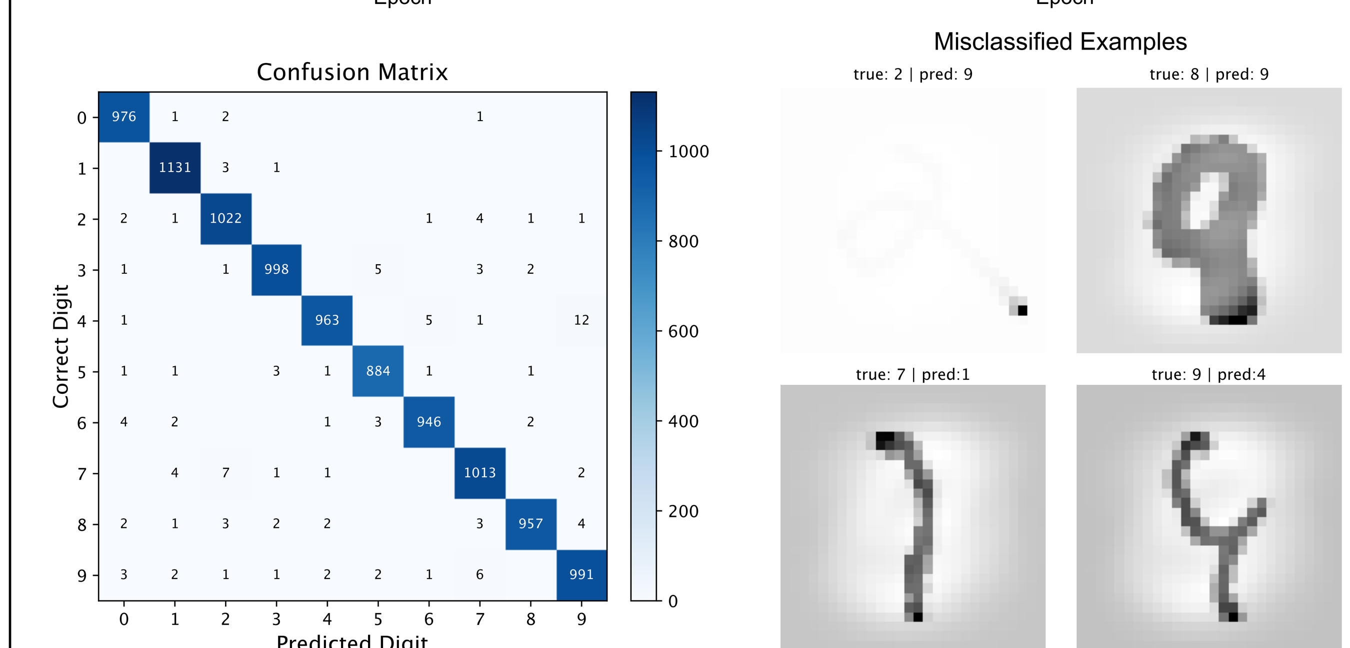


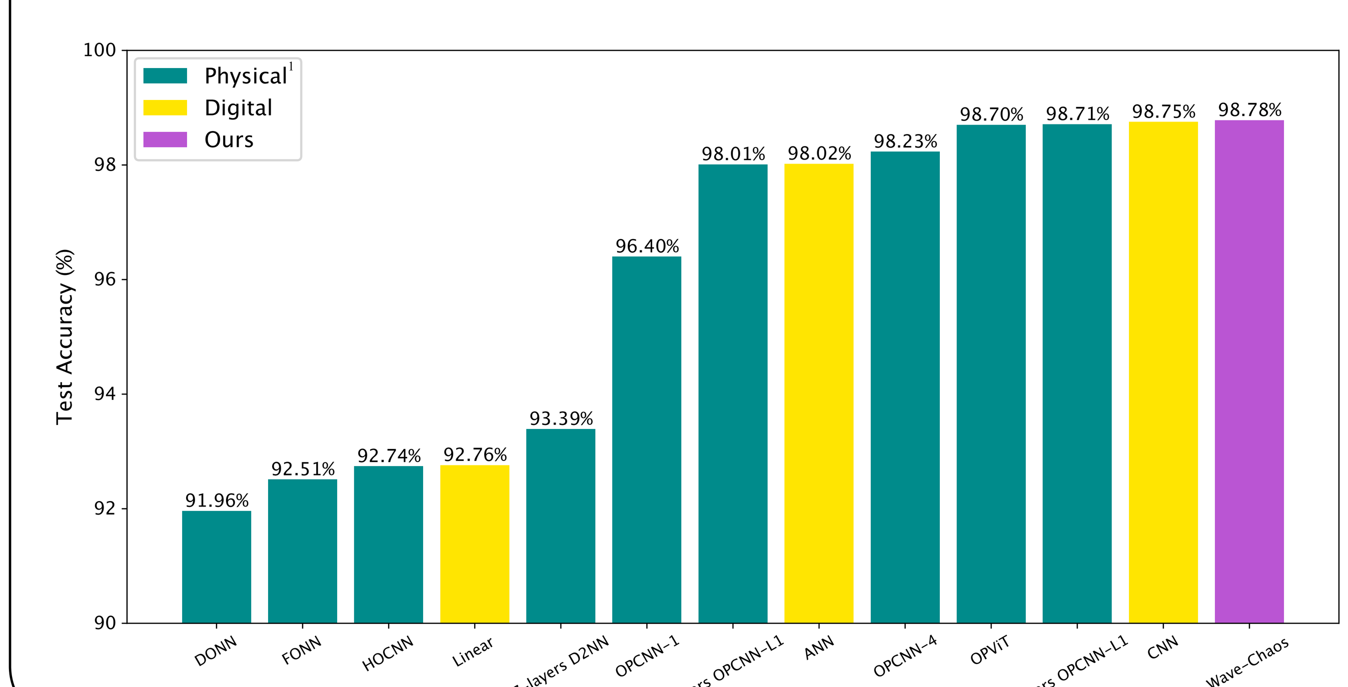
Experimental Graph:



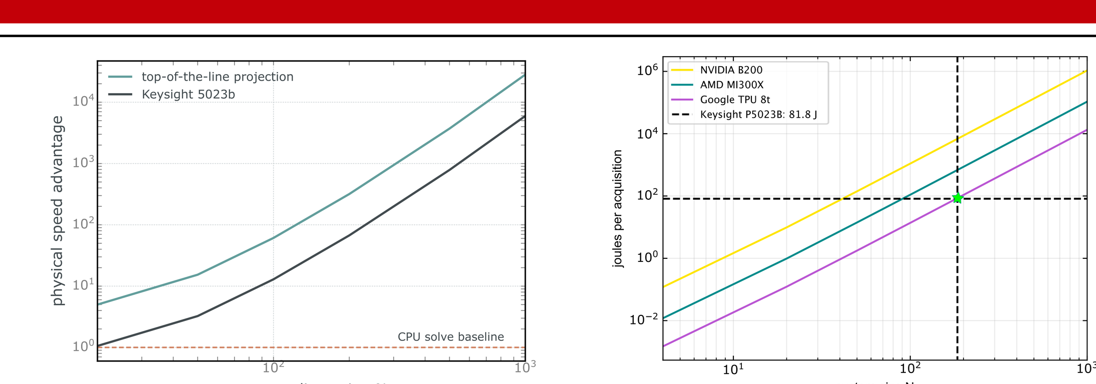
Results

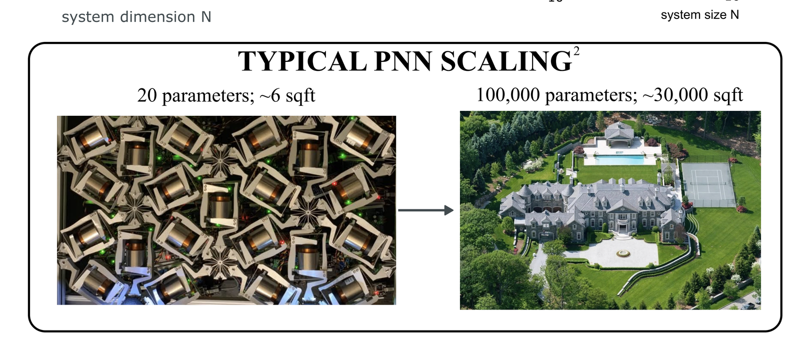






Discussion





References

1. Chen Xu, Xiubao Sui, Jia Liu, Yuhang Fei, Liping Wang, Qian Chen, Transformer in optronic neural networks for image classification, Optics & Laser Technology, Volume 165, 2023, 109627, ISSN 0030-3992, https://doi.org/10.1016/j.optlastec.2023.109627.
2. Lee RH, Mulder EAB, Hopkins JB. Mechanical neural networks: Architected materials that learn behaviors. Sci Robot. 2022 Oct 26;7(71):eabq7278. doi: 10.1126/scirobotics.abq7278. Epub 2022 Oct 19. PMID: 36260698.