ThermoPhoton: Fast 3D Thermal Simulation of Photonic Integrated Circuits via Operator Learning

Weilong Guan†, **Li Huang†,** Yuxuan Lin, Yuchao Wu, Yeyu Tong
Yuzhe Ma*

The Hong Kong University of Science and Technology (Guangzhou)

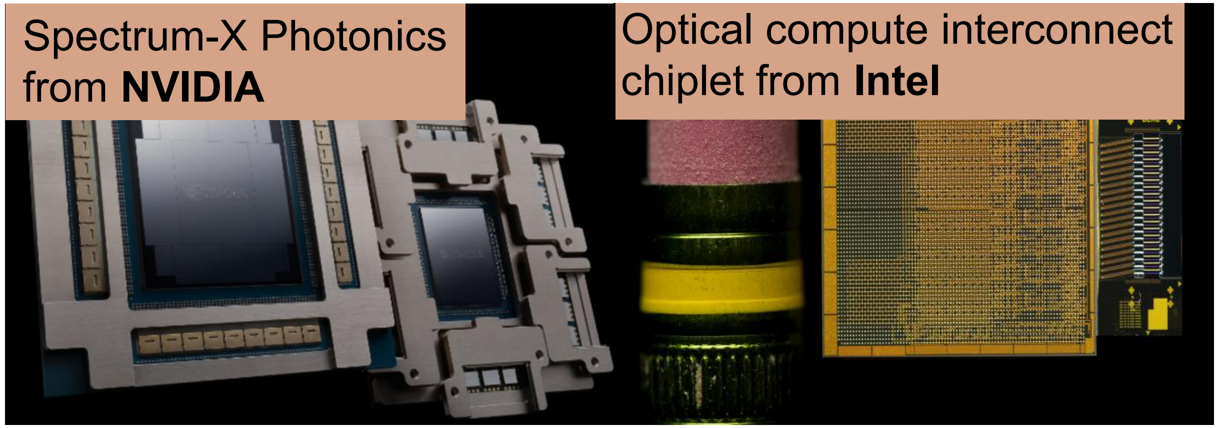
IEEE/ACM ICCAD 2025 INTERNATIONAL CONFERENCE ON COMPUTER-AIDED DESIGN 44th Edition

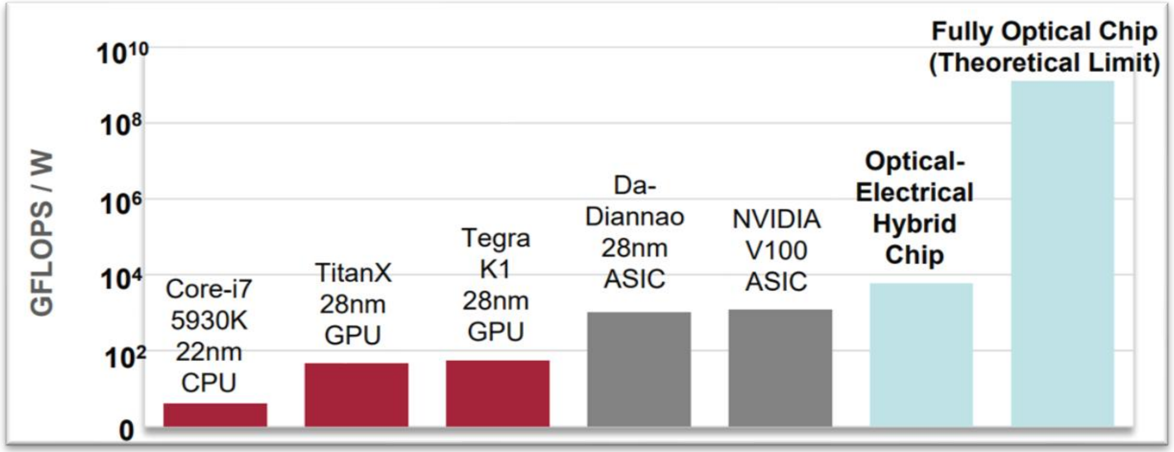
1/4/2026

CONTENTS

Photonic Integrated Circuits

- ➢ Moore's Law is slowing down
- ➢ Photonic integrated circuits offers a solution
  - Low transmission loss
  - No electrical shorts and ground loops
  - Low cost and abundant material sources





[1]Xu Z et al. Large-scale photonic chiplet Taichi empowers 160-TOPS/W artificial general intelligence[J]. Science, 2024.

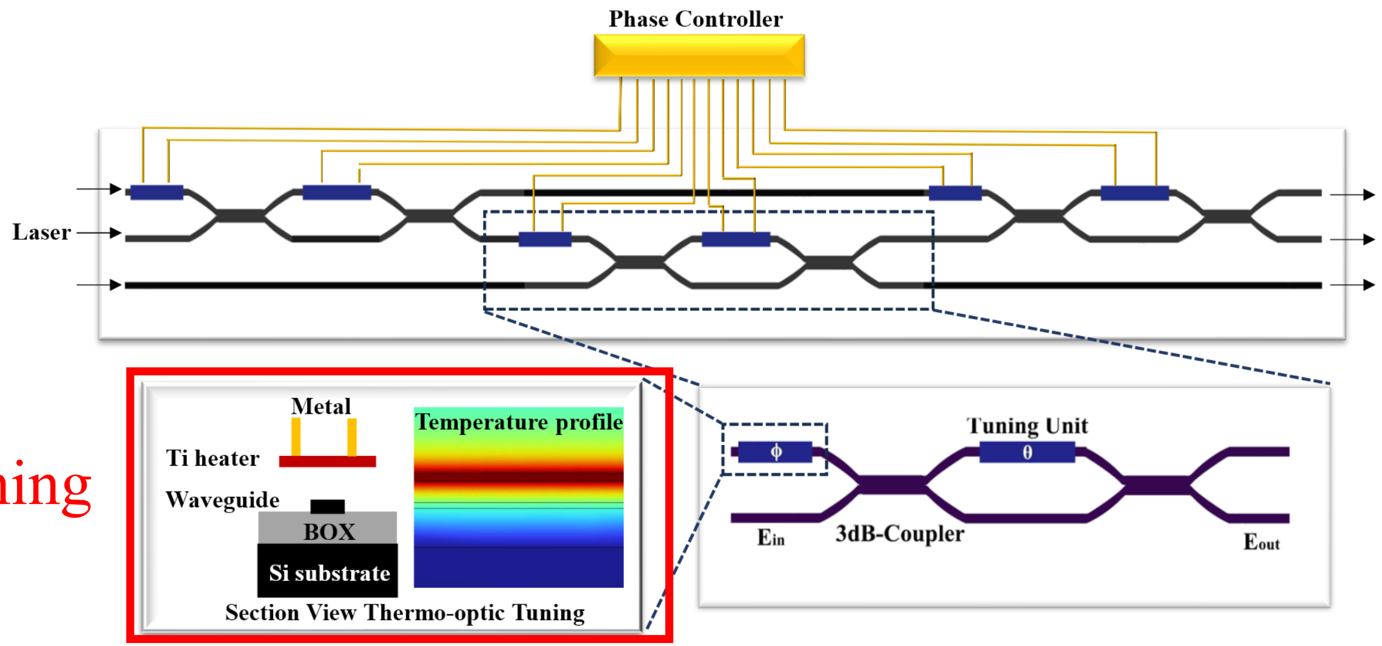
Photonic Integrated Circuits (PICs) have emerged as a promising solution

Key Components of PICs

➢ Thermal Modulation in PICs:

- **Heater** increases nearby **waveguide** temperature
- Temperature change causes a linear phase shift
- Enables precise tuning of the chip's behavior

Thermo-optic Tuning



Need fast and accurate thermal analysis

Thermal Analysis of Photonic Integrated Circuits

➢ PICs structure:

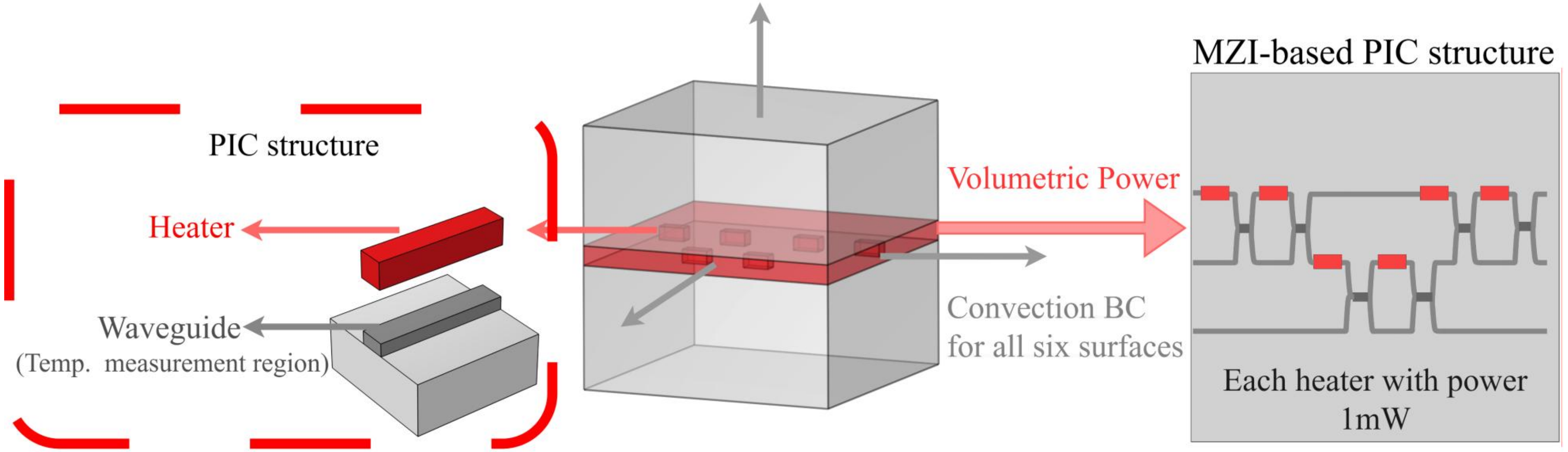
- Small-scale heaters
- The heaters and waveguides are in distinct layers

⇨ High spatial resolution

Full three-dimensional thermal simulation

➢ Traditional method:

- COMSOL Multiphysics (time consuming)

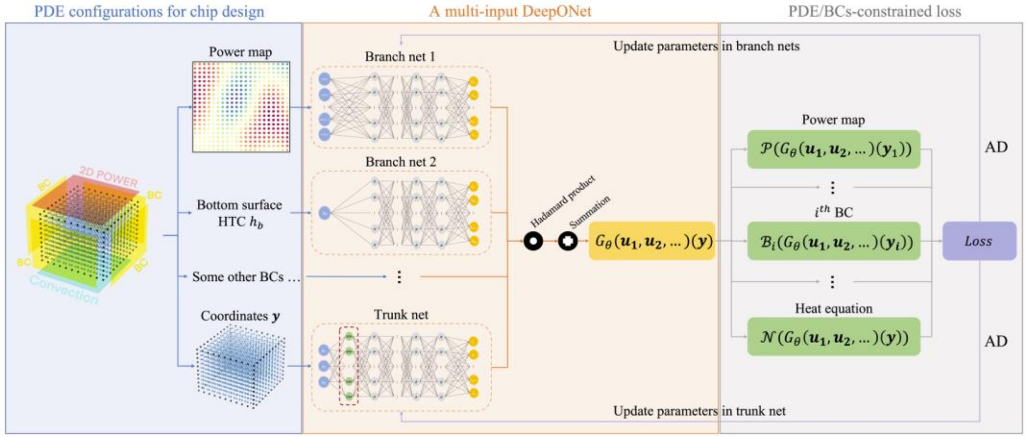


Need fast and accurate thermal prediction for PICs

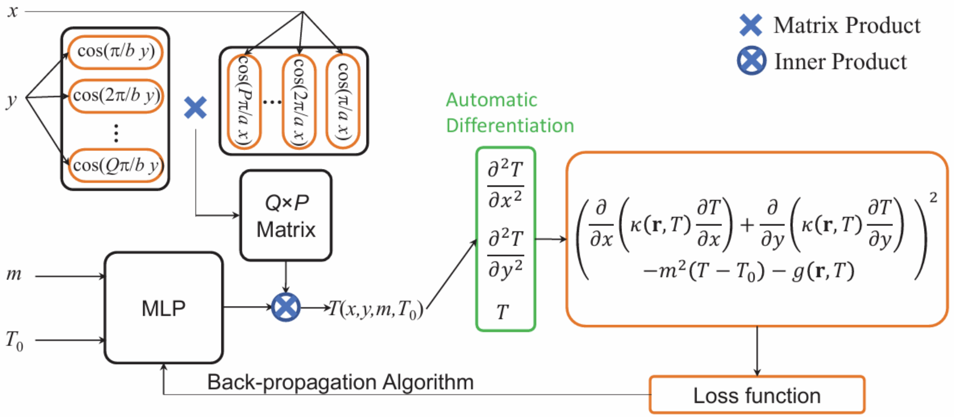
SOTAs

➢ ML-based methods:

- DeepOHeat[1] (operator learning)
- ThermPINN[2] (physics-informed neural networks)
- ARO[3] (multi fidelity fusion)



DeepOHeat[1]



ThermPINN[2]

[1]Liu Z, et al. DeepOHeat: Operator Learning-based Ultra-fast Thermal Simulation in 3D-IC Design[C]. DAC, 2023.
[2]Chen L, et al. Fast Full-Chip Parametric Thermal Analysis Based on Enhanced Physics Enforced Neural Networks[C]. ICCAD, 2023.
[3]Wang M, et al. ARO: Autoregressive Operator Learning for Transferable and Multi-fidelity 3D-IC Thermal Analysis With Active Learning[C]. ICCAD, 2024.

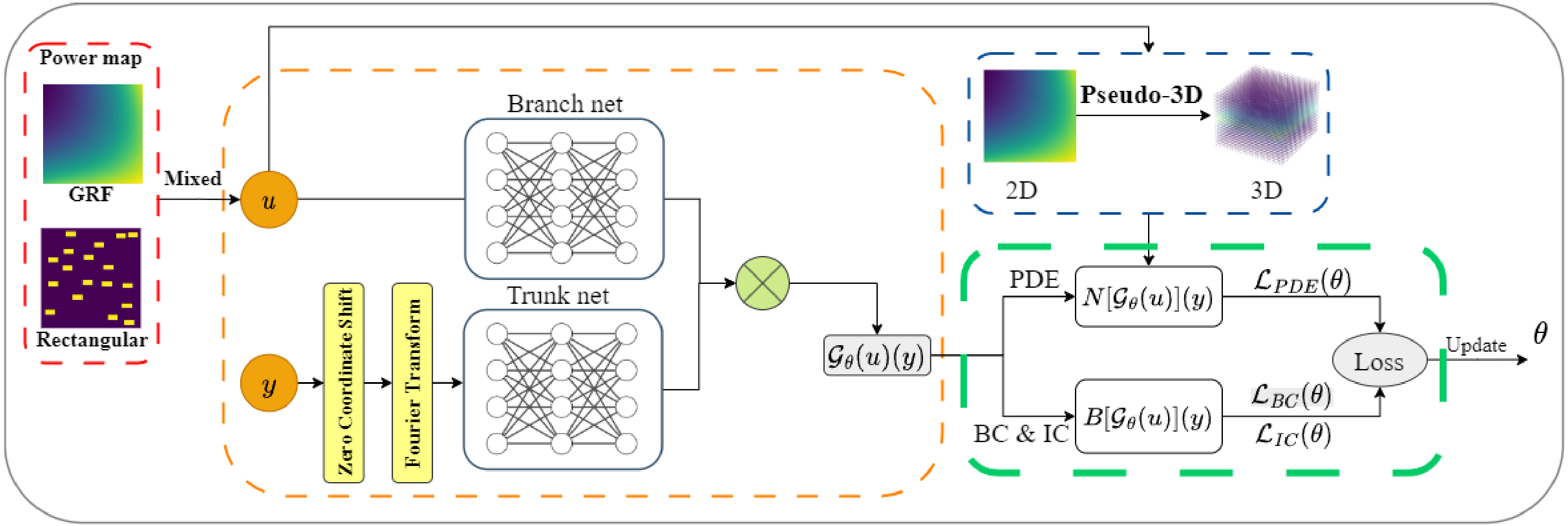
ThermoPhoton Framework

➢ Operator learning model

- Branch net encodes 2D power map
- Trunk net encodes spatial coordinates

➢ Physics-informed training

- Minimizing a loss function that penalizes residuals of the PDE and boundary conditions



Computational Complexity Reduction

➤ Pseudo-3D heat source representation

$$Q(x, y, z) = C \cdot \mathbf{u}(x, y) \cdot w(z),$$

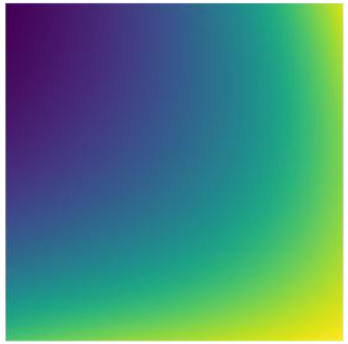
u(x, y): normalized lateral 2D power map

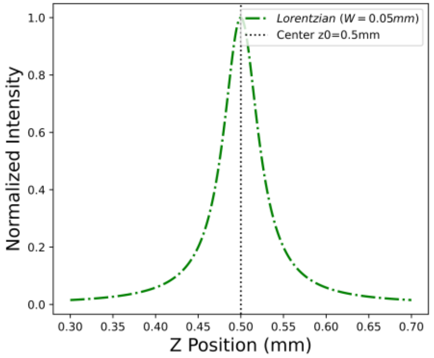
$w(z)$: vertical profile, here we use:

$$w_{\text{Lorentz}}(z) = \frac{1}{1 + \alpha(z - z_0)^2}$$

C: normalization constant ensuring the total power equals P_{total}

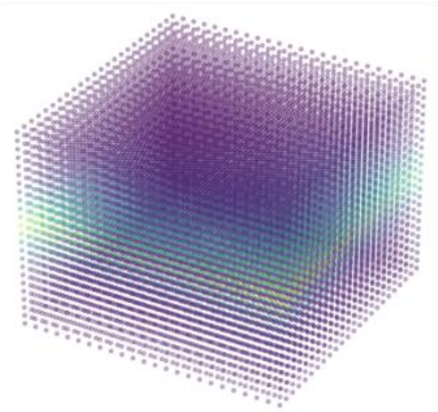
2D





Pseudo-3D

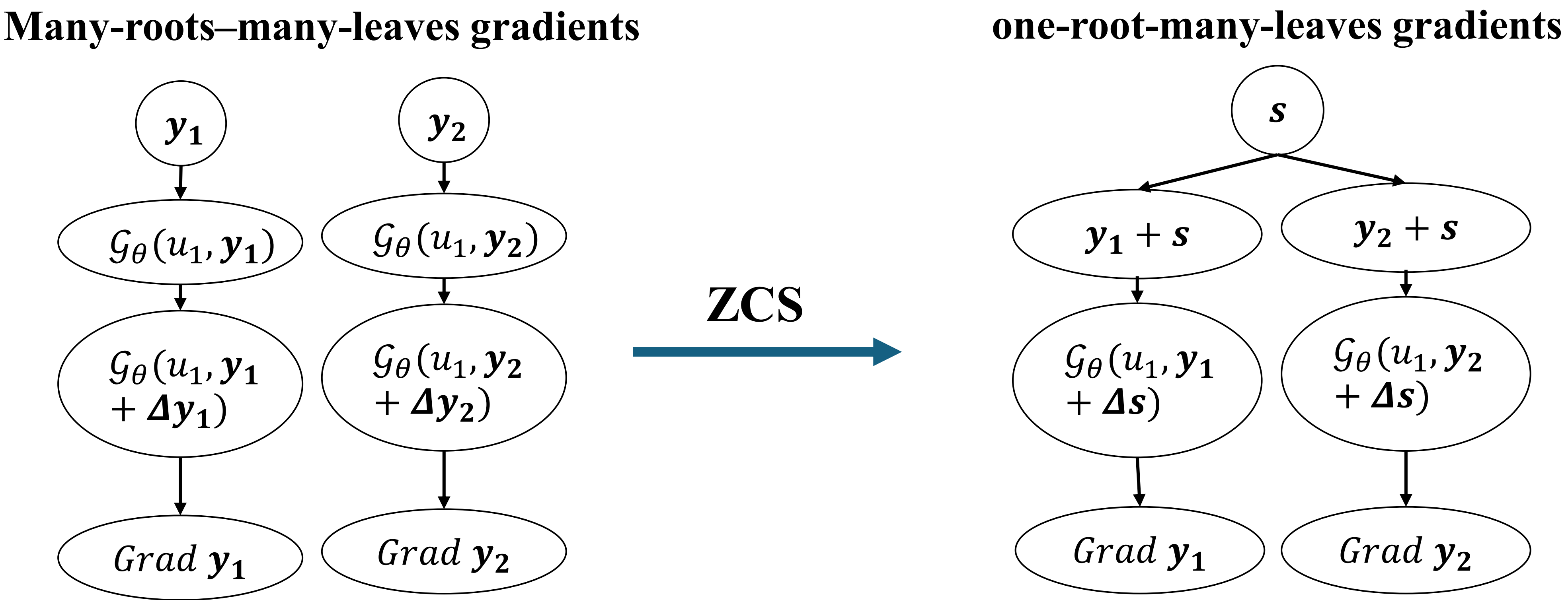
3D



Efficient Physics-Informed Loss Evaluation

➢ Zero coordinate shift (ZCS) gradient computation

- Improves both **memory** and **computational efficiency** in automatic differentiation
- Instead of calculating the gradient directly for each coordinate individually, ZCS introduces a shared dummy shift variable $\boldsymbol{s}$
- Avoids explicit loops and preserves a compact computation graph



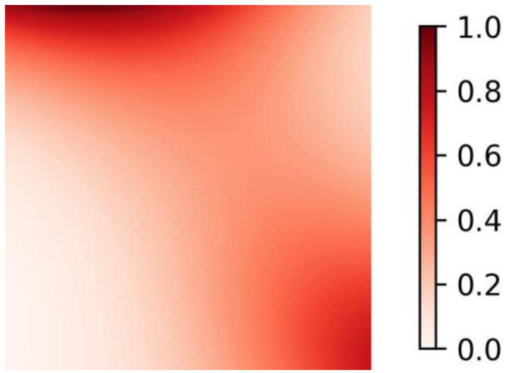
Enhanced Generalization

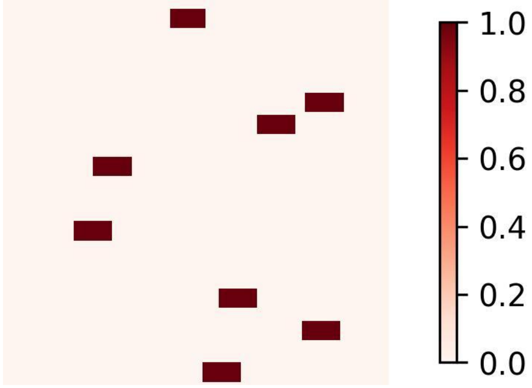
➢ Training dataset design

$$\mathbf{u}^{(i)}(x,y) = (1-\alpha)\cdot\mathbf{u}_{\mathrm{GRF}}^{(i)}(x,y) + \alpha\cdot\mathbf{u}_{\mathrm{rect}}^{(i)}(x,y),$$

Hybrid gaussian random field (GRF) and rectangular component

$$\mathbf{u}_{\mathrm{GRF}}^{(i)} \sim \mathcal{GP}(0, K((x,y),(x',y'))) \qquad \mathbf{u}_{\mathrm{rect}}^{(i)}(x,y) = \sum_{r=1}^{R_i} A_r \cdot \mathbb{1}_{\mathcal{R}_r}(x,y)$$





Experiment Setup

➢ Test cases of PICs:

- Mach-Zehnder interferometers (MZI)-based PICs (3×3, 4×4)
- Microring resonators (MRR)-based PICs (2×2, 3×3)
- Random blocks

➢ Input:

- 2D-power map (Grid size: 120 × 120)
- Spatial coordinates (x, y, z)

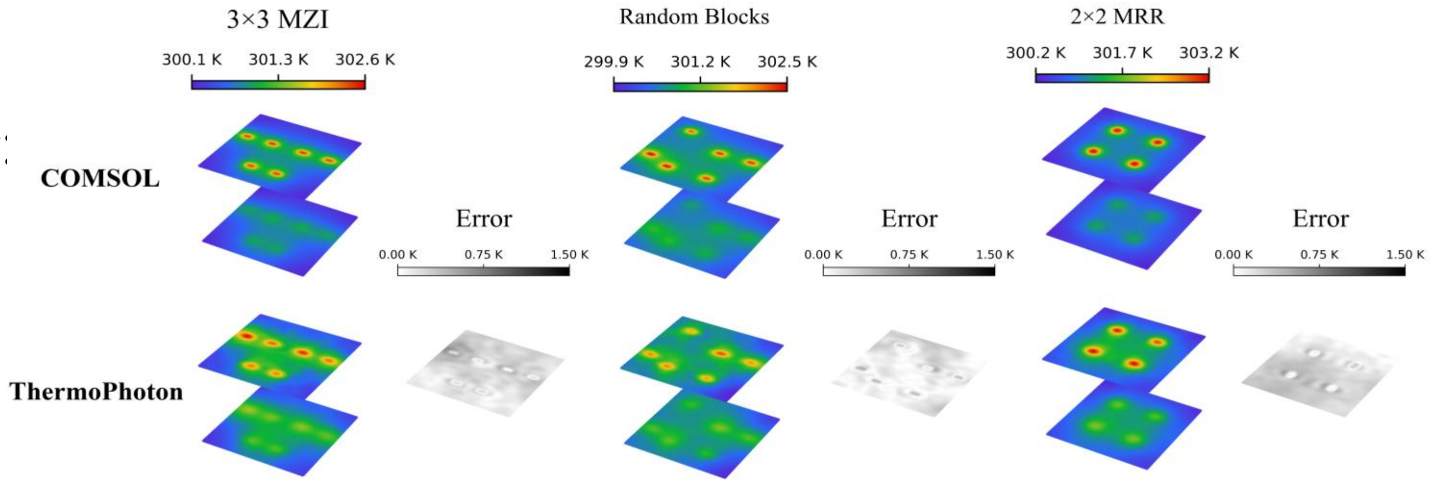
➢ Output:

- Spatial-temporal temperature profile(Grid size: 120 × 120 × 120)

Prediction Accuracy Comparison

Compared to COMSOL:

- Mean absolute percentage error:
  0.07%
- Average temperature error:
  0.2K



Temperature distribution of ThermoPhoton and COMSOL

TABLE I Comparison of ThermoPhoton and DeepOHeat across all test cases in terms of thermal prediction accuracy.

Pattern	**Heaters Num**	**Avg. T_{err} (K)**			**MAPE (%)**			**COMSOL Temp. T (K)**
		DeepOHeat	ResNet	Transformer	DeepOHeat	ResNet	Transformer	
Random Blocks	6	146.084	0.898	0.214	48.37	0.30	0.07	302.014
3×3 MZI	6	146.374	0.716	0.109	48.45	0.24	0.04	302.113
4×4 MZI	12	147.127	0.778	0.150	48.49	0.26	0.05	303.417
2×2 MRR	4	125.013	0.338	0.084	41.30	0.11	0.03	302.694
3×3 MRR	9	137.925	1.037	0.445	45.11	0.34	0.15	305.753
Overall	–	140.505	0.753	**0.200**	46.7	0.25	**0.07**	–

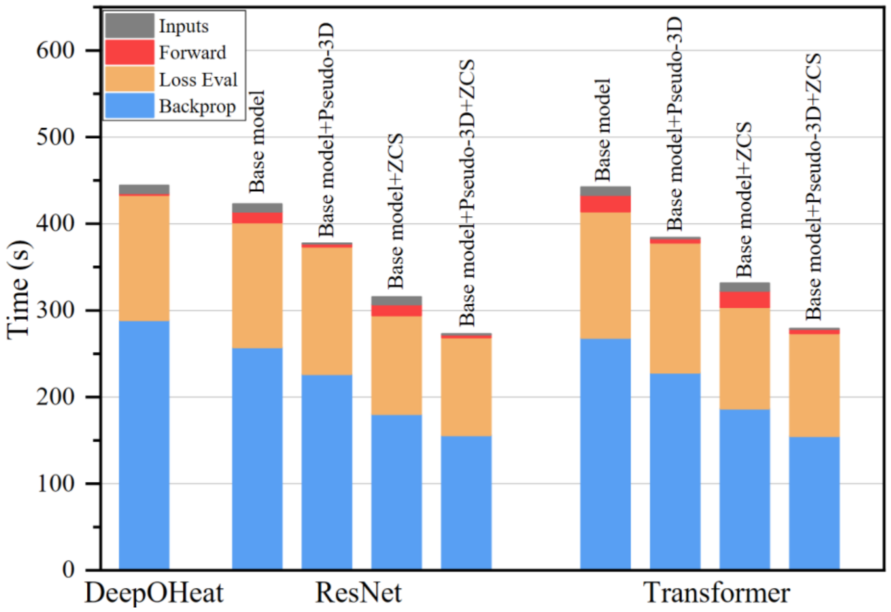
Overhead Evaluation

Vs. DeepOHeat:

- Model Training Time:
  37.9% Reduction
- Memory Usage:
  67.1% Reduction

TABLE II Peak GPU memory and thermal prediction error.

Method	GPU Mem. (GB)	Avg. T_{err} (K)	MAPE (%)
DeepOHeat	74.1	144.417	47.82
ResNet-based branch network			
Base model	44.1	7.787	2.58
Base model + Pseudo-3D	40.5	0.226	0.08
Base model + ZCS	27.9	2.649	0.88
Base model + Pseudo-3D + ZCS	**24.3**	0.230	0.08
Transformer-based branch network			
Base model	46.0	8.035	2.65
Base model + Pseudo-3D	40.6	0.327	0.11
Base model + ZCS	29.9	0.555	0.18
Base model + Pseudo-3D + ZCS	24.4	**0.212**	**0.07**



Training time breakdown

Observation:

- Pseudo-3D mainly improves **prediction accuracy**.
- ZCS is primarily used for reducing **memory usage**

Conclusion and Future Work

➢ Key contributions:

- Novel operator learning-based neural architecture for 3D thermal simulation of PICs
- Two key techniques, **Pseudo-3D** source representation and **Zero Coordinate Shift**
- Mean absolute percentage error **0.07%** (vs. COMSOL)
- Up to **67.1%** memory usage and **37.9%** training time reduction (vs. DeepOHeat)

➢ Future directions:

- Embed into border PIC design tasks (e.g., PIC placement)
- Extension to large scale deployment scenarios
- Explore more advanced training strategies

Thank you & Questions

Contact: lhuang913@connect.hkust-gz.edu.cn