Experimental Demonstration of an Optical Neural PDE Solver via On-Chip PINN Training

Yequan Zhao,[1,2] Xian Xiao,[1] Antoine Descos,[1] Yuan Yuan,[1] Xinling Yu,[1,2] Geza Kurczveil,[1] Marco Fiorentino,[1] Zheng Zhang[2] and Raymond G. Beausoleil[1]

[1] *Hewlett Packard Labs, Hewlett Packard Enterprise, 820 N. McCarthy Blvd., Milpitas, California 95305, USA*
[2] *Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA*

yequan_zhao@ucsb.edu, xian.xiao@hpe.com

Abstract: Partial differential equation (PDE) is an important math tool in science and engineering. This paper experimentally demonstrates an optical neural PDE solver by leveraging the back-propagation-free on-photonic-chip training of physics-informed neural networks. © 2025 The Author(s)

1. Introduction

Partial differential equation (PDE) is one of the most important math tools in science and engineering. Examples include electromagnetic modeling and thermal analysis of IC chips [1], medical imaging [2], safety verification of autonomous systems [3]. Discretization-based solvers (e.g., finite-difference and finite-element methods) convert a PDE to a large-scale algebraic equation via spatial and temporal discretization. Solving the resulting algebraic equation often requires massive digital resources and run times. Physics-informed neural network (PINN) is a promising discretization-free and unsupervised approach to solve PDEs [4]. PINN uses the residuals of a PDE operator and the boundary/initial conditions to set up a loss function, then minimizes the loss to train a neural network as a global approximation of the PDE solution. However, current PINN training typically needs several to dozens of hours on a powerful GPU, hindering the deployment of an real-time neural PDE solver on edge devices.

Optical neural networks (ONNs) provide a promising high-throughput, low-energy-consumption, low-latency, and high-parallelism solution. However, training PINNs on a photonic chip is very challenging. It is hard to realize training on photonic chips with back propagation (BP) due to 1) BP needs extra memory and photonics hardware to implement backward computation graphs and 2) BP needs a fully differentiable neural network, requiring a calibration process to accurately model the on-chip noises and imperfections. Several BP-free [5] and in-situ BP methods [6] are proposed, but they all have poor scalability. This problem becomes more severe in PINN training since the loss function also includes derivative terms, requiring multiple BPs.

In this paper, we experimentally demonstrate an optical neural PDE solver by calibration-free and BP-free on-chip training of PINNs. We leverage a photonics-friendly back-propagation-free training algorithm [7] and
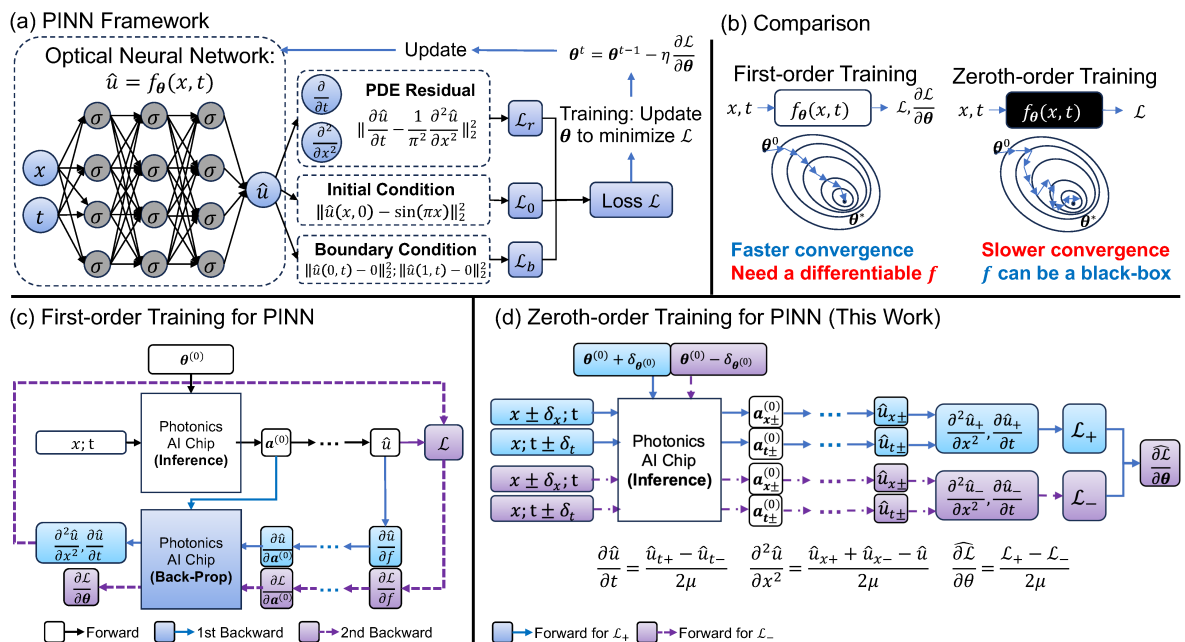

Fig. 1: (a) PINN framework. (b) Comparison between first-order training and zeroth-order training. (c) First-order training framework of PINN. (d) Zeroth-order training framework of PINN proposed in this work.
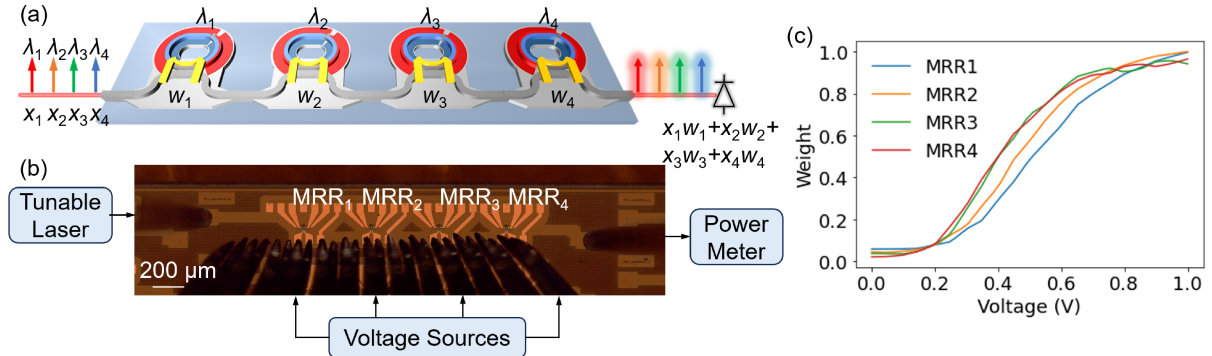
Fig. 2: (a) Schematic of the 1×4 MRR weight bank. (b) Microscope image of the weight bank and the schematic of the experimental setup for on-chip training. (c) Weight curves of the four MRRs.

implement it on a photonics AI chip (a 1×4 micro-ring resonator (MRR) weight bank). The on-chip training experiment learns a solution for a one-dimensional heat equation with 5E-3 mean ℓ_2 error after 1000 iterations of update, which experimentally validates our concepts of training PINN on an optical chip to solve a PDE.

2. Principle and Architecture

Fig. 1(a) shows the PINN architecture. A PINN is a feed-forward neural network parameterized by $\boldsymbol{\theta}$ that uses the output $\hat{u}(x,t) = f_{\boldsymbol{\theta}}(x,t)$ to approximate the PDE solution $u(x,t)$, where x and t repersent the spatial and temporal dimensions, respectively. We give an example of training a PINN to solve a heat equation with one (1) spatial dimension (x) and one (1) temporal dimension (t): $u_t = \frac{1}{\pi^2} u_{xx}$. The initial condition is $u(x,0) = \sin(\pi x)$, and boundary conditions are $u(0,t) = 0$; $u(1,t) = 0$. During training, $\boldsymbol{\theta}$ are updated to minimize a physics-informed loss function $\mathscr{L} = \mathscr{L}_r + \mathscr{L}_0 + \mathscr{L}_b$. $\mathscr{L}_r, \mathscr{L}_0, \mathscr{L}_b$ measure how well the approximation $\hat{u}$ complies with the PDE operator, the initial conditions, and boundary conditions that describe the physics laws of a system, respectively. Fig. 1(b) gives a comparison between first-order training and zeroth-order training. First-order training has a faster convergence rate since it utilizes the gradient information to guide the parameter update direction. However, as shown in Fig. 1(c), implementing first-order training on photonics hardware to train a PINN requires 1) additional photonics chip design for back propagation computation, which poses expensive hardware overhead and 2) an accurate and differentiable model of photonics devices, which requires an exhaustive calibration process to model the fabrication error. Moreover, since unknown environmental noises always exist, such modeling is not ideally accurate. To address the challenge of implementing PINN training on photonics hardware, we design our BP-free training with zeroth-order (ZO) optimization (Fig. 1(d)). The training uses forward propagation only by repeatedly calling a photonics inference accelerator. Since zeroth-order training does not require a differentiable neural network model, we can directly optimize the tunable parameters of photonic devices (*e.g.,* voltage values of a micro ring resonators, phase shifter values of Mach-Zehnder interferometers) to minimize the training loss. The fabrication errors and other unknown environmental noises can be mitigated on-the-run, thus a pre-calibration process is not required.

3. Experimental Demonstration

We experimentally demonstrated the optical neural PDE solver for the heat equation by training a small-scale PINN. The PINN contains a feed-forward neural network with two hidden layers, as shown in Fig. 1(a). The scale of the weight matrices are 2×4, 4×4, 4×4, and 4×1, respectively. The weight matrices are computed by 1×4 tiles in different clock cycles with time multiplexing. The 1×4 tile is implemented by a 1×4 microring resonator (MRR) weight bank, as shown in Fig. 2(a). Using the wavelength division multiplexing (WDM) technology, by encoding the input data at $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, the data can multiply with each of the MRR weights w_1, w_2, w_3, w_4, and a photodetector at output can compute $\lambda_1 w_1 + \lambda_2 w_2 + \lambda_3 w_3 + \lambda_4 w_4$. Fig. 2(b) shows the experimental setup and the microscope image of the MRR weight bank chip fabricated at AMF foundry. Fig. 2(c) shows the weight curves with different thermal tuning voltages, indicating the tuning range of the weight values from 0 to 1. The variances of the weight curves among different MRRs caused by fabrication errors can be mitigated by our BP-free on-chip training without the need for pre-calibration.

Simulation Results. We first numerically simulate the BP-free on-chip training under different bit precisions. We use an experimentally measured voltage-weight look-up table [as shown in Fig. 2(c)] to model the MRR weight banks. The learned PDE solution is compared with the ground-truth solution by the mean ℓ_2 error on a hold-out test point set unseen during training. The solutions are visualized with heatmaps. Each pixel indicates the value of the PDE solution $u(x,t)$. Fig. 3(a) visualizes the ground truth solution and learned solutions. With a low bit accuracy (8-bit), the neural network cannot learn the governing physics due to the restricted expressive power. The results indicate the affect of bit accuracy on the performance of on-chip PINN training.
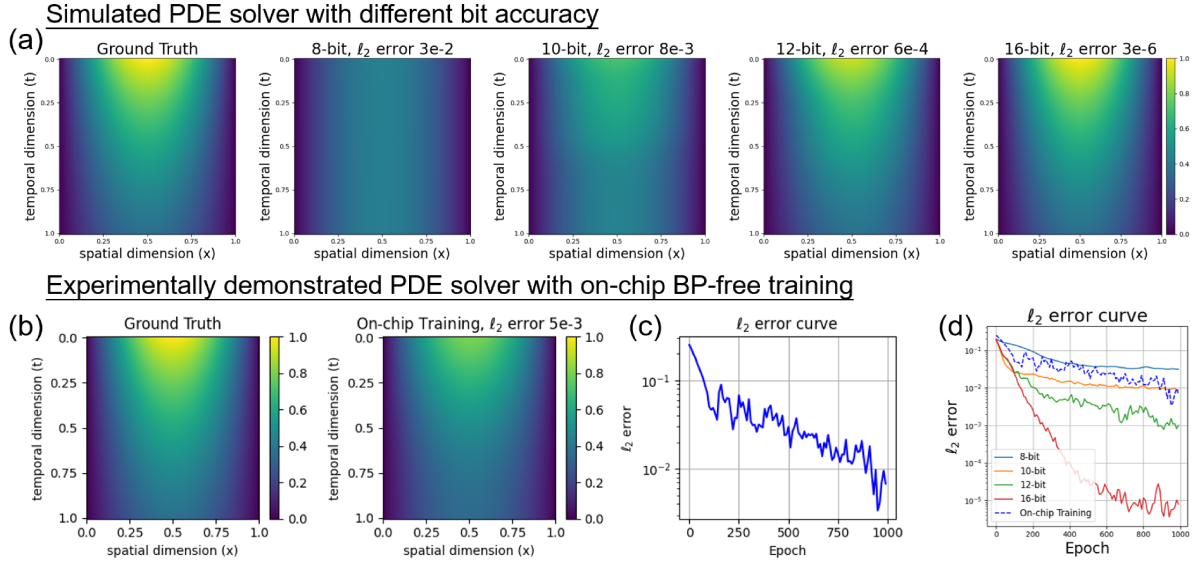
Fig. 3: Comparison between the ground truth solution and learned solutions of (a) BP-free training simulation and (b) experimentally demonstrated BP-free on-chip training. (c) The ℓ_2 error curve of on-chip training. (d) The ℓ_2 error curve comparison between on-chip training and different simulation setups.

Hardware Demo. Then, we experimentally demonstrated our BP-free training on the fabricated photonic chip. A tunable laser and photodetector are used to measure the spectrum of the MRR weight bank at different tuning voltages, as shown in Fig. 2(b). Four voltage sources are used for thermally tuning the MRRs. The weight values are read out at the four resonances of the MRRs. Fig. 3(b) visualizes the learned solution after on-chip training, and Fig. 3(c) shows the ℓ_2 error curve. The decreasing ℓ_2 error shows that on-chip training gradually captures the governing physics laws described by PDE even under fabrication errors, environmental noise, etc. However, the limited bit accuracy of analog computation limits the accuracy of the final learned solution. Fig. 3(d) shows the ℓ_2 error curve comparison between on-chip training and simulation results with different bit accuracy setups. The experimentally demonstrated on-chip training showcases the bit accuracy between 8-bits and 10-bits.

4. Conclusion

In this paper, we have experimentally demonstrated an optical neural PDE solver for solving heat equations, using BP-free on-chip training of PINNs. The experiment results learn the solution with 5E-3 mean ℓ_2 error. To scale up the optical on-chip training framework for real-size PINN (*e.g.*, 1000 neurons per layer), the tensor-train decomposed PINN (TT-PINN) [8, 9] implemented on an energy-efficient tensorized ONN (TONN) [10] inference accelerator is a promising approach.

References

1. Li, P., Pileggi, L. T., Asheghi, M., and Chandra, R. (2004, November). Efficient full-chip thermal modeling and analysis. In IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004. (pp. 319-326). IEEE.
2. Villena, J. F., Polimeridis, A. G., Wald, L. L., Adalsteinsson, E., White, J. K., and Daniel, L. (2015). MARIE–a MATLAB-based open source software for the fast electromagnetic analysis of MRI systems. In Proceedings of the 23rd Annual Meeting of ISMRM, Toronto, Canada (p. 709).
3. Bansal, S., and Tomlin, C. J. (2021, May). Deepreach: A deep learning approach to high-dimensional reachability. In 2021 IEEE International Conference on Robotics and Automation (ICRA) (pp. 1817-1824). IEEE.
4. Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 378, 686-707.
5. Gu, J., Zhao, Z., Feng, C., Li, W., Chen, R. T., and Pan, D. Z. (2020, July). FLOPS: Efficient on-chip learning for optical neural networks through stochastic zeroth-order optimization. Design Automation Conference (pp. 1-6).
6. Pai, S., Sun, Z., Hughes, T. W., Park, T., Bartlett, B., Williamson, I. A., ... and Miller, D. A. (2023). Experimentally realized in situ backpropagation for deep learning in photonic neural networks. Science, 380(6643), 398-404.
7. Zhao, Y., Xiao, X., Yu, X., Liu, Z., Chen, Z., Kurczveil, G., ... and Zhang, Z. (2023). Real-Time FJ/MAC PDE Solvers via Tensorized, Back-Propagation-Free Optical PINN Training. arXiv preprint arXiv:2401.00413.
8. Liu, Z., Yu, X., and Zhang, Z. (2022). TT-PINN: a tensor-compressed neural PDE solver for edge computing. arXiv preprint arXiv:2207.01751.
9. Zhao, Y., Yu, X., Chen, Z., Liu, z., Liu, Z., and Zhang, Z. (2023). Tensor-compressed back-propagation-free training for (physics-informed) neural networks. arXiv preprint arXiv:2308.09858.
10. Xiao, X., On, M. B., Van Vaerenbergh, T., Liang, D., Beausoleil, R. G., and Yoo, S. J. (2021). Large-scale and energy-efficient tensorized optical neural networks on III–V-on-silicon MOSCAP platform. Apl Photonics, 6(12).