

# Supplementary Information

for

## Passive All-Optical Nonlinear Neuron Activation via PPLN Nanophotonic Waveguides

Wujie Fu<sup>1,2,†</sup>, Xiaodong Shi<sup>3,4,†</sup>, Sakthi Sanjeev Mohanraj<sup>5</sup>, Lei Shi<sup>1</sup>, Yuan Gao<sup>1</sup>, Zexian Wang<sup>1</sup>, Jianing Wang<sup>1</sup>, Xu Chen<sup>5</sup>, Luo Qi<sup>1</sup>, Pragati Aashna<sup>1</sup>, Guanyu Chen<sup>6,7</sup>, Di Zhu<sup>1,3,4,5,\*</sup>, Aaron Danner<sup>1,2,\*</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, National University of Singapore, 117583, Singapore

<sup>2</sup> Graduate School for Integrative Sciences and Engineering, National University of Singapore, 119077, Singapore

<sup>3</sup> A\*STAR Quantum Innovation Centre (Q.InC), Agency for Science, Technology and Research (A\*STAR), 138634, Singapore

<sup>4</sup> Institute of Materials Research and Engineering (IMRE), Agency for Science, Technology and Research (A\*STAR), 138634, Singapore

<sup>5</sup> Department of Materials Science and Engineering, National University of Singapore, 117575, Singapore

<sup>6</sup> Key Laboratory of Optoelectronic Technology & Systems, Ministry of Education, Chongqing University, Chongqing 400044, China

<sup>7</sup> College of Optoelectronic Engineering, Chongqing University, Chongqing 400044, China

<sup>†</sup> These authors contributed equally.

\* Correspondence: [adanner@nus.edu.sg](mailto:adanner@nus.edu.sg); [dizhu@nus.edu.sg](mailto:dizhu@nus.edu.sg)

---

## Supplementary Note I: Digital-twin SHG activation function

For physical neural network training, a differentiable activation function is required to enable backpropagation in the network's digital twin. We analyzed  $\chi^{(2)}$ -based nonlinear input–output response of the fundamental harmonic in PPLN nanophotonic waveguides, for the all-optical neuron implementation that concatenates the PPLN and MZI chips. We solved the coupled-wave equations describing the underlying parametric interaction to construct a physics-informed surrogate model. This solution yields a differentiable, physics-based baseline curve that is then fitted to the experimental data and used as the digital-twin activation function, explicitly accounting for pump depletion as well as linear propagation losses at both FH and SH frequencies. The governing equations are given by [1]:

$$\frac{dE_1}{dz} = -\alpha_1 E_1 + i\kappa E_1^* E_2 e^{i\Delta k z}, \quad (1)$$

$$\frac{dE_2}{dz} = -\alpha_2 E_2 + i\kappa E_1^2 e^{-i\Delta k z}, \quad (2)$$

where  $E_1$  and  $E_2$  denote the complex field amplitudes of the FH and SH waves, respectively, and  $\alpha_1$  and  $\alpha_2$  are their amplitude-loss coefficients, estimated as  $\alpha_i = (\text{loss}_{\text{FH,SH}} [\text{dB/cm}]) \times (\ln 10/10) \times 100$ , corresponding to propagation losses of 0.2 dB/cm and 0.5 dB/cm for the FH and SH waves, respectively. The effective nonlinear coefficient was set as  $d_{\text{eff}} \approx 20$  pm/V, and the refractive indices were  $n_1 = 1.93$  and  $n_2 = 2.10$  for the FH and SH wavelengths, as obtained from COMSOL simulations. The nonlinear coupling coefficient is given by  $\kappa = \omega d_{\text{eff}} / (n_1 c)$ , where  $\Delta k$  denotes the phase mismatch between the interacting waves. Under the experimental wavelength and temperature conditions,  $\Delta k$  is estimated to lie in the range of 0–60  $\text{m}^{-1}$ , and it is used as an optimization parameter for fitting the experimental data. The coupled-wave equations were numerically integrated along the propagation direction  $z$  using a fourth-order Runge–Kutta (RK4) algorithm with a step size of  $\Delta z = 10^{-6}$  m and a total device length of  $L = 1.1$  cm, consistent with the physical geometry of the TFLN device.

The input pump power ( $P_{\text{in}}$ ) was varied from 0 to 208 mW, covering the power range for both CW-pumped and pulse-pumped (peak power) SHG experiments (Fig. S1). The corresponding FH and SH output powers were then calculated as follows:

$$P_{\text{FH,out}} = \frac{1}{2} n_1 \epsilon_0 c A_{\text{eff,FH}} |E_1(L)|^2, \quad (3)$$

$$P_{\text{SH,out}} = \frac{1}{2} n_2 \epsilon_0 c A_{\text{eff,SH}} |E_2(L)|^2, \quad (4)$$

where  $A_{\text{eff}}$  is the effective mode area, simulated to be 1.22  $\mu\text{m}^2$  and 0.83  $\mu\text{m}^2$  for the fundamental TE mode profile at FH and SH wavelengths. The simulated pump-depletion curve  $P_{\text{FH,out}}(P_{\text{in}})$  exhibits a sigmoid-like transition that directly characterizes the nonlinear transfer from input power to transmitted FH power, corresponding to the optical neuron activation (Fig. S2).

To align the simulated transfer function with the experimentally measured response under pulsed excitation, the simulated curve was first normalized with respect to its linear low-power regime. Specifically, we define the normalized transmission as

$$T_{\text{sim}}(P_{\text{in}}) = \frac{P_{\text{FH,out}}(P_{\text{in}})}{P_{\text{in}}}. \quad (5)$$

This transmission curve was then scaled along both the input (power) and output (transmission) axes to match the experimentally recorded FH intensity trace,  $I_{\text{FH,exp}}(t)$ , from the detector. The temporal input pulse intensity,  $I_{\text{in}}(t)$ , was mapped to a normalized input-power sequence,  $P_{\text{in}}(t)$ . A least-squares fitting was performed between  $T_{\text{sim}}(P_{\text{in}})$  and the experimental ratio,  $I_{\text{FH,exp}}(t)/I_{\text{in}}(t)$ , to determine two scaling coefficients  $a$  and  $b$  in  $I_{\text{fit}}(t) = a T_{\text{sim}}(b I_{\text{in}}(t))$ , where  $a$  adjusts the output normalization (vertical scaling) and  $b$  defines the input normalization (horizontal scaling). The optimal pair  $(a, b)$  was obtained by minimizing the mean-squared error between  $I_{\text{fit}}(t)$  and the experimentally measured trace. The resulting scaled continuous functions in Fig. 2c and Fig. 2f were then implemented as a differentiable nonlinear activation within the digital twin of the ONNs.

Therefore, the RK4-based simulation provides a physics-informed analytical representation of the

PPLN activation function, facilitating further ONN training in the digital-twin differentiable models.

## Supplementary Note II: Physics-aware system calibration and training

To enable gradient-based optimization of the all-optical neuron-based ONN for the binary and Iris classification tasks, as well as multilayer ONNs that use the experimentally measured activation curve for the medical-image and airfoil tasks, we employ a hybrid physics-aware training (PAT) framework [2]. This framework combines in-situ optical measurements for the forward pass with in-silico differentiation of a differentiable digital twin, ensuring accurate gradient estimation while remaining robust to device drift and measurement noise.

The procedure consists of three main stages: independent calibration of the MZI and PPLN modules, hybrid alignment of their joint transfer characteristics, and synchronized digital-domain training of the optical weight matrix and ONN parameters. First, each MZI on the silicon photonic chip is characterized by sweeping the drive voltage  $V_k$  applied to its thermo-optic phase shifters and measuring the corresponding transmission  $T_k(V_k)$ . The half-wave voltage  $V_{\pi,k}$  is extracted by fitting the transfer relation:

$$T_k(V_k) = \cos^2\left(\frac{\pi V_k}{2V_{\pi,k}} + \phi_{0,k}\right), \quad (6)$$

where  $V_k$  denotes the applied heater voltage,  $V_{\pi,k}$  the voltage required to induce a  $\pi$  phase shift, and  $\phi_{0,k}$  a fabrication-induced static phase offset. To establish an accurate voltage–phase mapping for the Clements-type interferometer mesh, we first perform a cross-bar transmission/extinction calibration to obtain each phase shifter’s  $V_{\pi,k}$ . Before calibration, the dark-current responses of all photodetectors (PD1–PD5) are recorded for bias correction during both calibration and practical operation. The calibration sequence follows cross-bar condition PD response maximization and minimization for  $V_{\pi,k}$  extraction, and adheres to the singular value decomposition (SVD) of the target matrix and its corresponding input vector, consisting of four main steps:

1. **Amplitude calibration:** Amplitude modulators (AM1 and AM2) are first swept with a triangular-wave DC signal to roughly identify their  $V_{\pi}$ , followed by a fine static voltage sweep with a 0.01 V resolution to locate the zero-transmission (minimum) and full-transmission (maximum) points, corresponding to the minima and maxima of the summed response currents from PD2 and PD5, respectively. These two modulators correspond to the diagonal matrix (singular values) and are calibrated to set amplitude scaling.
2. **Unitary calibration:** With AM1 set to the all-pass state and AM2 to the zero-pass state, MZI1 is swept with a triangular-wave DC signal to estimate its  $V_{\pi}$ , then fine-tuned via static voltage sweeps under the *bar* condition (PD2 maximum) and *cross* condition (PD5 maximum). MZI2 is calibrated in the same manner. These two MZIs constitute the unitary-matrix section of the decomposition.
3. **Input encoding calibration:** Input-side shifters (AM3–AM6) are then calibrated for data encoding. With AM1 in the all-pass state, AM2 in the zero-pass state, and both MZI1 and MZI2 set to the bar condition, AM4 is calibrated by maximizing and minimizing PD2’s response. The same procedure applies to AM3, AM5, AM6, and the phase shifters PM1–PM4, using voltage sweeps under bar–cross conditions of the monitored PD signals.
4. **Iterative fine tuning:** In each iteration, a single input channel or singular-value component is encoded while the corresponding MZI unitary block is adjusted. The measured input–output discrepancy is iteratively minimized through fine tuning around the identified  $V_{\pi}$  operating points.

This iterative calibration effectively minimizes inter-heater thermal cross-talk and layout-dependent phase bias, yielding a well-tuned linear transformation matrix  $W(\phi)$  for subsequent optical computation. The resulting configuration produces a series of verified linear transformation matrices  $W(\phi)$  with computation fidelities of approximately 97.7% and 96.2% under continuous-wave and pulsed operation

(Fig. 3), respectively. The complete voltage–phase calibration dataset, after dark-current subtraction and normalization of both voltage and transmission amplitudes to the range of  $[0, 1]$ , constitutes a physics-constrained digital twin  $\hat{W}(\phi)$ . Such normalization ensures consistent scaling between the experimental measurements and the differentiable model, enabling stable gradient back-propagation and mitigating the simulation–reality gap during hybrid physics-aware training of the ONNs.

Second, the PPLN nonlinear activation is calibrated by continuously sweeping the input optical power from its minimum to maximum operating level (normalized to 0–1) and recording the corresponding output FH intensity. The measured transfer curve is then fitted with differentiable fourth-order Runge–Kutta solutions of the coupled-wave equations under the optimized phase-mismatch condition (Supplementary Note I), yielding a physics-informed surrogate activation function compatible with gradient-based training. This model remains fixed throughout training and serves as the nonlinear activation in the digital twins for both CW and pulsed working situations.

To achieve joint alignment between the physical systems and their digital representation for ONN inference, the PPLN chip is optically cascaded to the MZI chip (Fig. S4). In this setup, a programmed optical signal generated by arbitrary wave generator through a fiber-based intensity modulator with normalized amplitude between 0 and 1 is launched into the PPLN. The modulated FH output is injected into the on-chip MZI array, and the detected optical intensity at the output port is compared against the numerically computed response:

$$\hat{y} = \hat{W} f_{\text{PPLN}}(x). \quad (7)$$

During the alignment experiments, as the output FH signal increases to a relatively high carrier level, it indicates the upper limit of the PPLN activation’s effective power window. Within this region, saturation behavior characterized by a inflection point in the first derivative of the carrier signal curve is observed (Fig. S5), accompanied by a monotonic growth of the generated SH power. This response is highly consistent with the nonlinear activation behavior obtained from the independent characterization of the PPLN device (Fig. 2), confirming that the  $\chi^{(2)}$ -based nonlinear transformation is successfully encoded into the ONN input signal. For improved consistency between the nonlinear and linear transfer stages, the MZI responses are fine-tuned via the same least-squares phase-retrieval algorithm, which iteratively adjusts each phase-shifter voltage to minimize the difference between the measured optical outputs and the hybrid nonlinear predictions, thereby ensuring precise alignment between the physical and virtual system responses.

During training of the ONN, the physical PPLN–MZI responses are used for the forward pass, while gradient computation is performed entirely in the digital twin. Specifically, the measured optical outputs  $y_{\text{meas}}$  are substituted into the differentiable surrogate model:

$$\hat{y} = \hat{W}(\phi) f_{\text{PPLN}}(x), \quad (8)$$

from which gradients  $\partial \mathcal{L} / \partial \phi_k$  of the loss function  $\mathcal{L}$  are obtained via standard back-propagation. Here,  $\mathcal{L}$  is defined as the mean-squared error between the measured and predicted optical intensities, optionally regularized by a small phase-smoothness term to suppress overfitting, and all measured signals are normalized to  $[0, 1]$  after dark-current subtraction. These gradients are mapped to phase-shifter voltages using the local sensitivity  $d\phi_k/dV_k$ , yielding the update rule:

$$V_k^{(t+1)} = V_k^{(t)} - \eta \frac{\partial \mathcal{L}}{\partial \phi_k} \left( \frac{d\phi_k}{dV_k} \right)^{-1}, \quad (9)$$

where  $\eta$  denotes the learning rate, and  $d\phi_k/dV_k$  is experimentally obtained by small-signal voltage perturbations around each heater’s operating point.

In this way, the optimization of the linear weights and ONN tunable parameters is performed entirely within the digital twin. Convergence is determined when the relative loss reduction falls below a preset threshold for each task or when  $\|\nabla_{\phi} \mathcal{L}\|$  becomes sufficiently small, ensuring stable alignment between the experimental and simulated responses. Once training converges, the final optimized network parameters can be obtained and used to time-multiplex the programming of the MZI network and the intensity

modulator, enabling all-optical neuron inference driven by the  $\chi^{(2)}$ -based nonlinear activations.

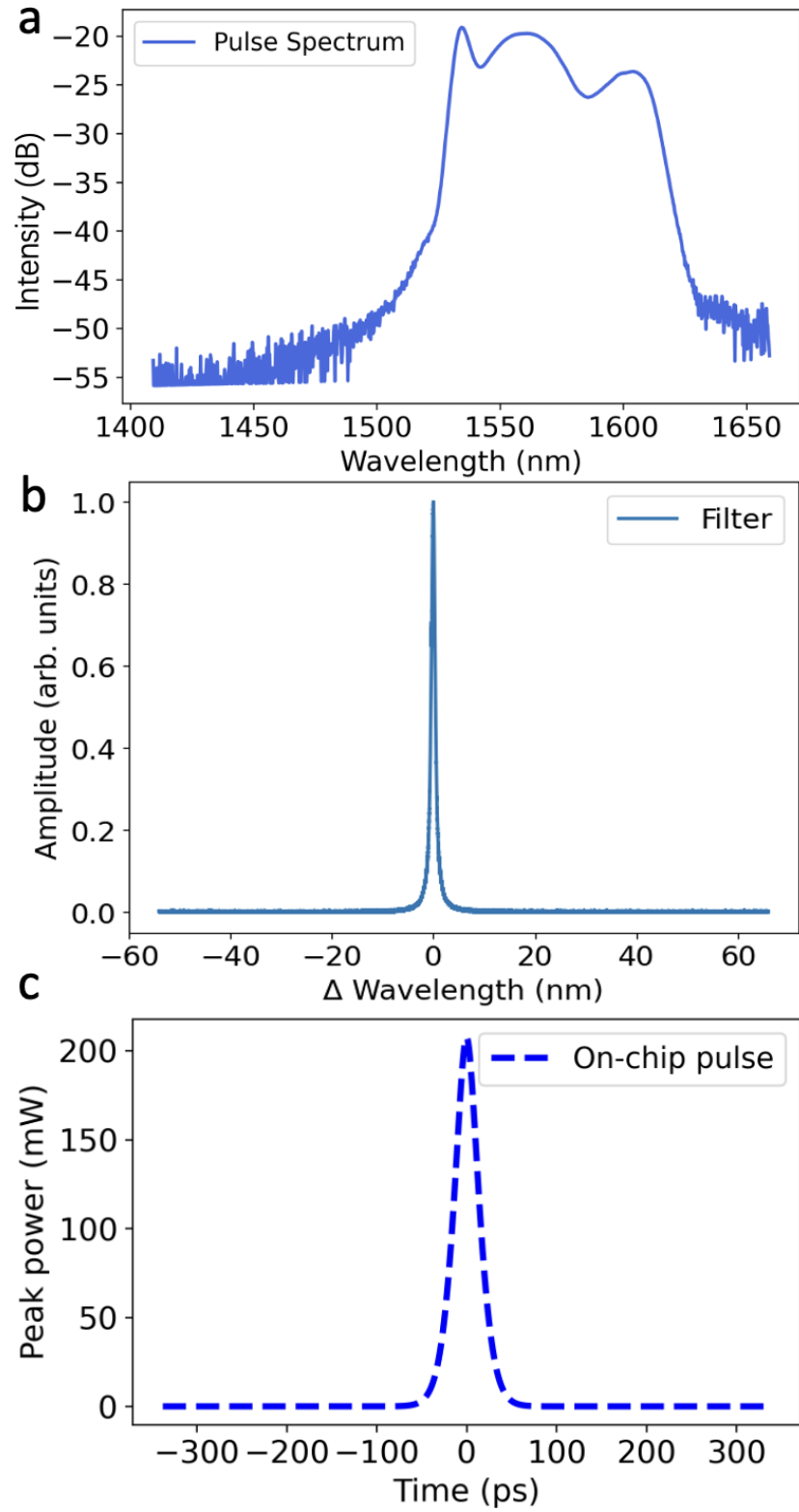
## Supplementary Note III: Analysis of Casadability in Deep Optical Networks

Figures 2c and 2f of the main text show that the proposed PPLN activation exhibits a nonlinear onset at pump levels below 10 mW under continuous-wave (CW) excitation and below 0.05 nJ per pulse under pulsed excitation. This behavior reflects the threshold-free nature of  $\chi^{(2)}$  parametric interactions: even weak pump fields induce electronic polarization in lithium niobate and generate a second-harmonic radiation source. Thus, from a physical standpoint,  $\chi^{(2)}$ -driven activation does not impose a strict energy threshold.

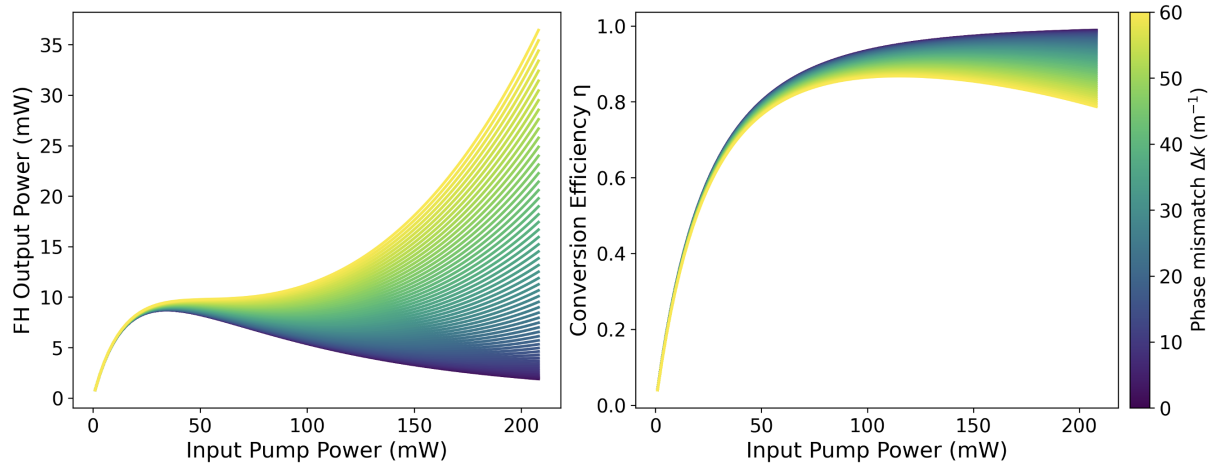
From a neural-network perspective, however, practical operation requires a nonlinear response with sufficient curvature—such as a sigmoid-like saturation trend—to enable expressive function approximation beyond linear models. To quantify this regime, we characterize the nonlinear loss at pump levels approximately twice the nonlinearity onset region, where the activation exhibits clear saturated curvature useful for ONN inference. As summarized in Fig. S6, the measured nonlinear loss at 20 mW CW pump power is 1.88 dB, while at 0.04 nJ under pulsed excitation the loss is 2.2 dB. These values indicate that the device provides neuron-effective nonlinearity while maintaining modest optical loss in both operating conditions.

To assess scalability toward multilayer networks, we consider a representative coherent-computing architecture supplying 50 mW per channel at the first layer. Assuming a typical TFLN waveguide propagation loss of 0.2 dB/cm and a linear-network path length of 1 cm per layer, the linear propagation loss is approximately 0.2 dB per layer. The PPLN activation contributes a major  $\sim 4$  dB of loss (Fig. S6). After the first activation stage, approximately 20 mW remains to drive the second layer, which remains safely above the nonlinear-onset region of below 10 mW. After the second activation, the available power is  $\sim 13$  mW, still exceeding the required nonlinear-operating threshold. These estimates suggest that at least three to four cascaded nonlinear layers are feasible within a monolithic TFLN implementation using present device performance. Under pulsed excitation, the much higher peak-to-average power ratio enables stronger activation at approximately one order of magnitude lower average power than in the CW case, providing additional headroom for multilayer architectures with improved energy efficiency.

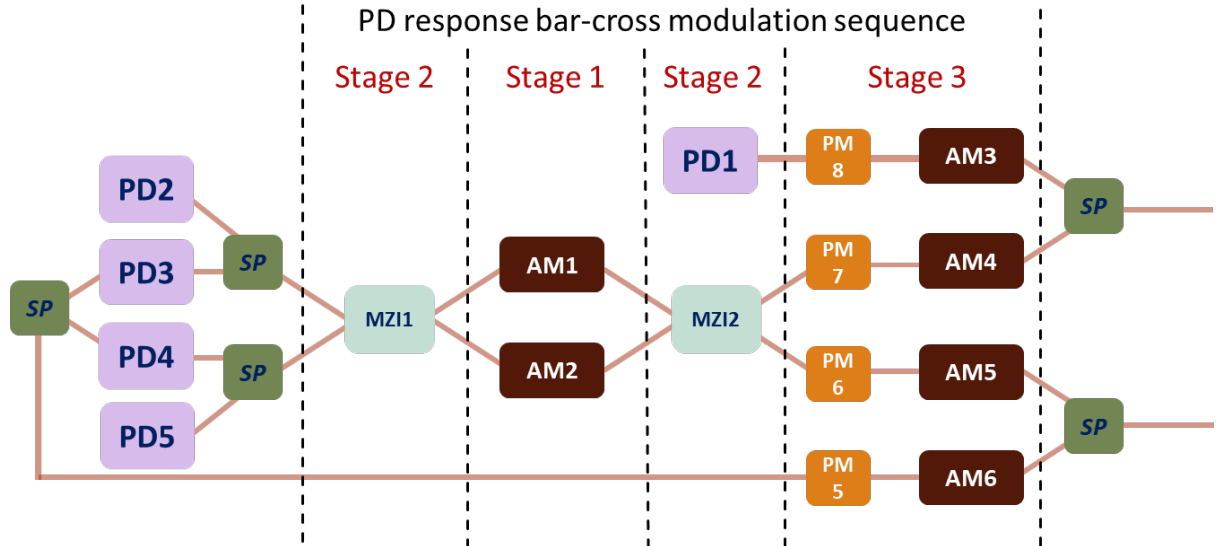
We also note that existing demonstrations of fully optical deep neural networks typically implement only two to three nonlinear layers [3, 4, 5, 6]. The cascability estimates above indicate that the proposed PPLN-based activation is fully compatible with such depths and, with further optimization, offers a promising path toward deeper ( $> 3$ -layer) integrated ONN systems on TFLN platforms.



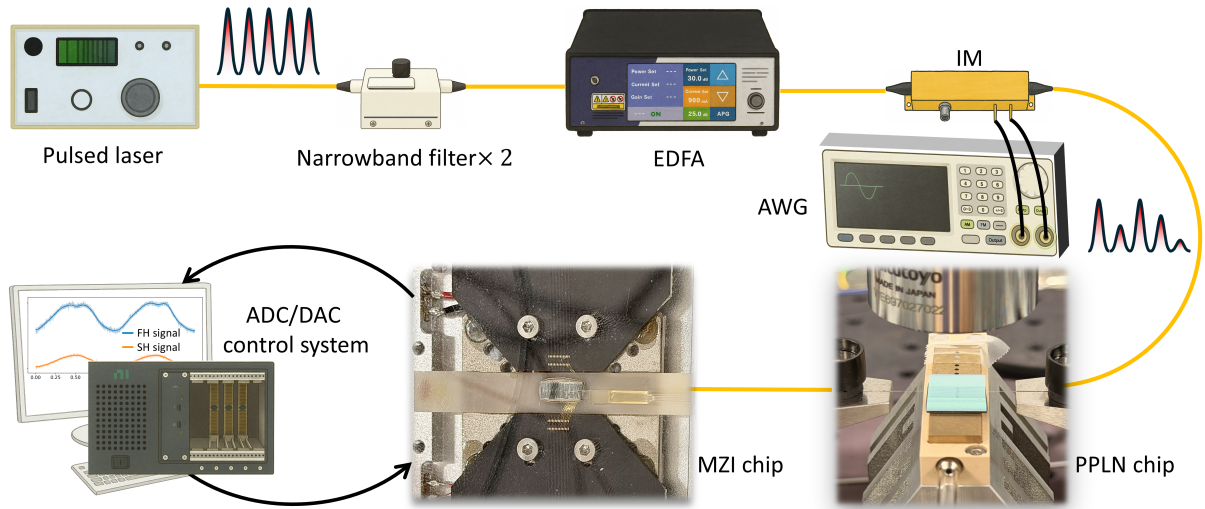
**Supplementary Figure S1.** (a) Measured spectrum of the original pulsed laser. (b) Measured transmission characteristics of the narrow-band filters defining the operational wavelength window. (c) Simulated temporal waveform of the pulse following spectral filtering, amplification, and on-chip coupling, representing the effective excitation delivered to the PPLN waveguide.



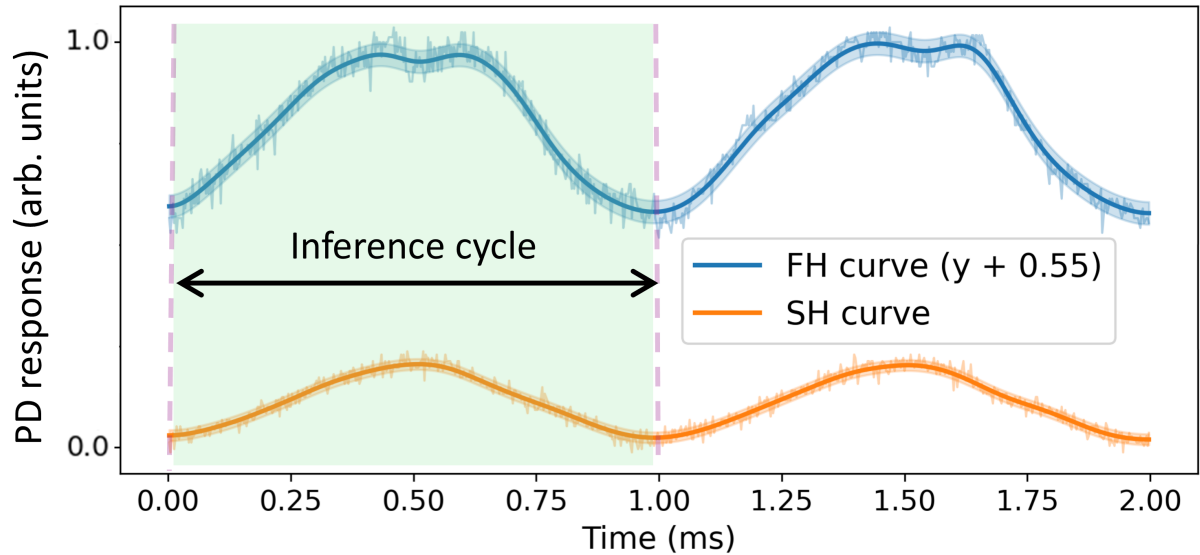
**Supplementary Figure S2.** Simulated fundamental-harmonic input-output characteristics and the corresponding absolute conversion efficiency, obtained using a fourth-order Runge-Kutta solution with the fabricated PPLN parameters, as a function of phase mismatch. Three distinct saturation behaviors emerge: a monotonic increase under imperfect phase matching; an increase-saturation-decrease-increase profile near perfect matching; and an increase followed by a stable decrease under almost perfect matching. The measured PPLN activation operates in the second regime, exhibiting the characteristic increase-saturation-decrease-increase behavior mimicking the sigmoid activation function.



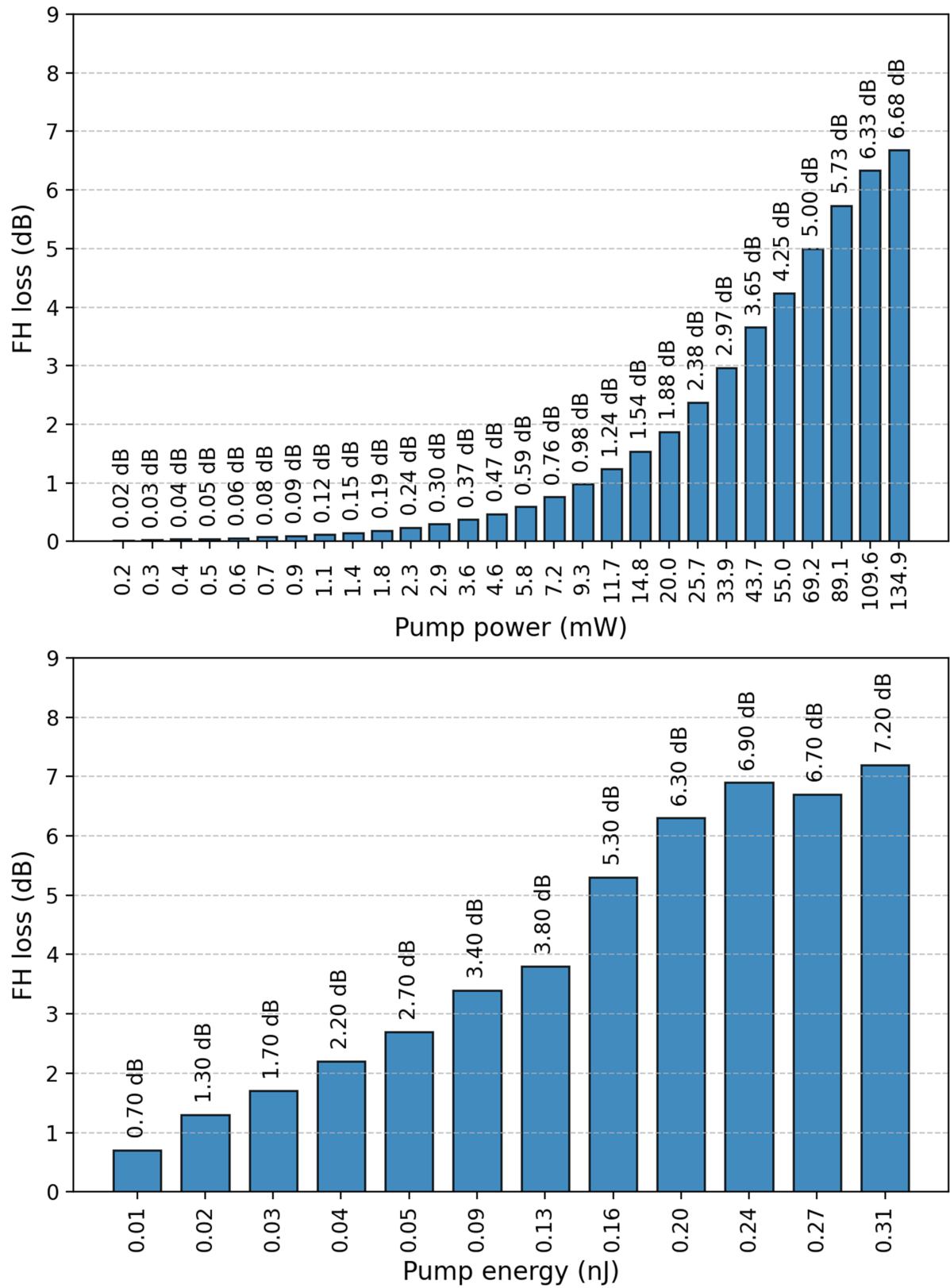
**Supplementary Figure S3.** Calibration procedure of the MZI linear product module with splitters (SPs). It consists of three stages of cross-bar calibration, in which the voltages applied to each amplitude modulator (AM), Mach-Zehnder interferometer (MZI), and phase modulator (PM) are sequentially swept to maximize the corresponding photodetector (PD) responses.



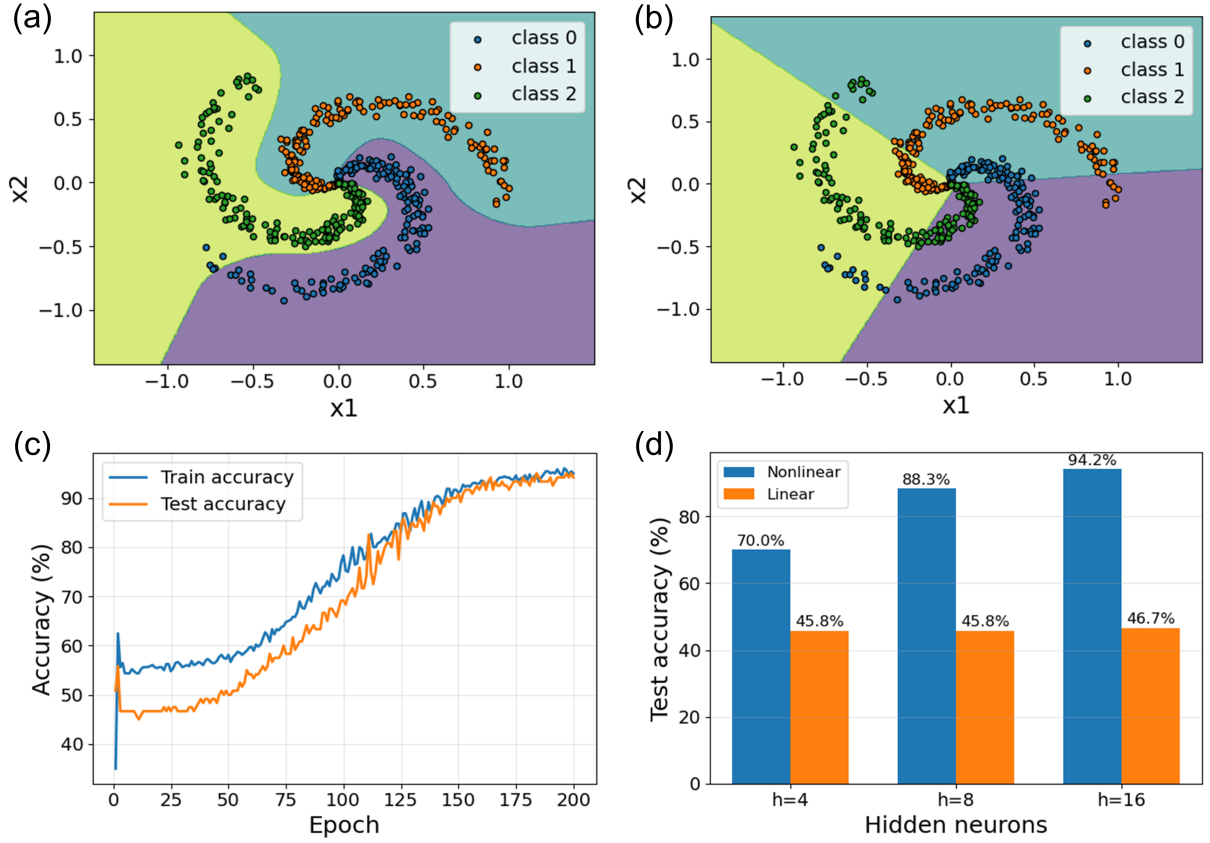
**Supplementary Figure S4.** Experimental setup for all-optical neuron inference based on an activation-first and weighted-sum-later architecture. A pulsed light source is connected to bandpass filters followed by an EDFA for power amplification at the phase-matching wavelength of 1552 nm. The carrier for data loading is then prepared and sent to a fiber-based intensity modulator, where both the neural-network inputs  $x_j$  and its weighting parameter  $k_j$  parameters are encoded using an arbitrary waveform generator operating at 1 kHz. The intensity-encoded pulses are coupled into the PPLN chip to undergo SHG-induced nonlinear activation. The activated optical pulses are subsequently routed to a silicon MZI system to perform the weighted linear operation. The output results of the ONN is finally detected using on-chip photodetectors.



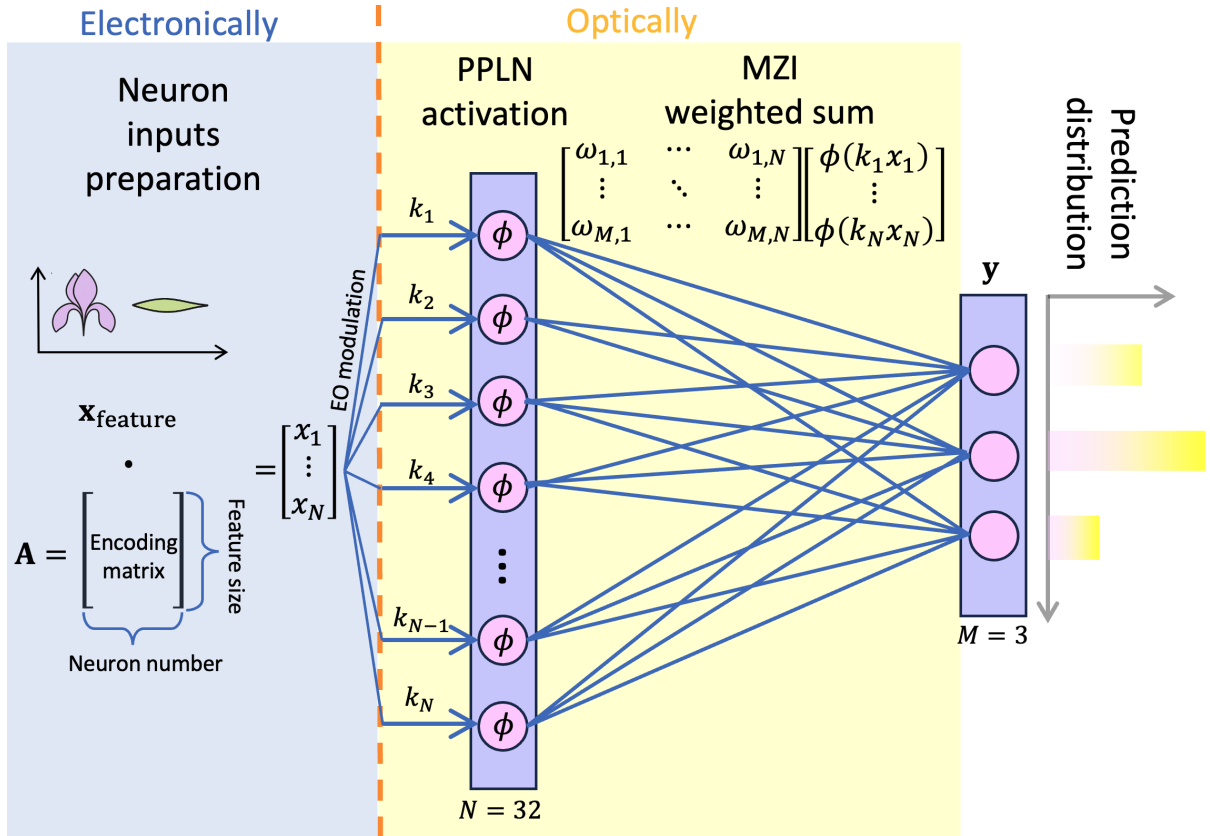
**Supplementary Figure S5.** Experimental on-chip response of the periodic SHG-based nonlinear activation under a 1-kHz data-encoding modulation in pulsed operation. In each cycle, the pump at the fundamental harmonic is intensity-modulated to encode the input signal and subsequently routed through the PPLN nanophotonic waveguide to undergo SHG-based nonlinear activation. The optical energy is monotonically converted to the harmonic (tapped out for detection via the wavelength-division multiplexer), and a sigmoid-like saturation emerges when the input reaches its maximum level (normalized data-encoding value of 1).



**Supplementary Figure S6.** FH power reduction introduced by the nonlinear activation process acting on the FH light that encodes the neural signal, evaluated under both CW and pulsed operating conditions.



**Supplementary Figure S7.** (a) Decision boundary generated by an 8-neuron single-hidden-layer MLP using the PPLN-based activation function characterized in the  $< 20$  mW CW regime, evaluated on a 3-class spiral dataset. The spiral dataset contains 200 points per class and is synthetically generated by sampling a radius  $r \in [0, 1]$  and an angle  $\theta = \theta_0 + 4r + \epsilon$ , where  $\theta_0 \in \{0, 4, 8\}$  for the three classes,  $\epsilon \sim \mathcal{N}(0, 0.2^2)$ , and mapping to Cartesian coordinates  $x = r \sin \theta$  and  $y = r \cos \theta$ . (b) Decision boundary generated by an 8-neuron single-hidden-layer linear model (no activation). (c) Training and testing accuracies of the 8-hidden-neuron MLP over 200 epochs, obtained digitally using the Adam optimizer with a learning rate of 0.001. (d) Comparison of classification performance between SHG-activated MLPs and linear models with different numbers of hidden neurons.



**Supplementary Figure S8.** Computing pipeline for the  $\chi^{(2)}$ -driven ONN, including electronic input preparation, electro-optic (EO) modulation, and an activation-first-weighting-later all-optical neuron-level operation stage for hidden-layer inference.

## References

- [1] Robert W. Boyd. *Nonlinear Optics*. Academic Press, 4th edition, 2020.
- [2] Ali Momeni, Babak Rahmani, Benjamin Scellier, Logan G Wright, Peter L McMahon, Clara C Wanjura, Yuhang Li, Anas Skalli, Natalia G Berloff, Tatsuhiro Onodera, et al. Training of physical neural networks. *Nature*, 645(8079):53–61, 2025.
- [3] Farshid Ashtiani, Alexander J Geers, and Firooz Aflatouni. An on-chip photonic deep neural network for image classification. *Nature*, 606(7914):501–506, 2022.
- [4] Saumil Bandyopadhyay, Alexander Sluuds, Stefan Krastanov, Ryan Hamerly, Nicholas Harris, Darius Bunandar, Matthew Streshinsky, Michael Hochberg, and Dirk Englund. Single-chip photonic deep neural network with forward-only training. *Nature Photonics*, 18(12):1335–1343, 2024.
- [5] Tao Yan, Yanchen Guo, Tiankuang Zhou, Guocheng Shao, Shanglong Li, Ruqi Huang, Qionghai Dai, and Lu Fang. A complete photonic integrated neuron for nonlinear all-optical computing. *Nature Computational Science*, pages 1–12, 2025.
- [6] Fuhao Yu, Kangjian Di, Wenjun Chen, Sen Yan, Yuanyuan Yao, Silin Chen, Xuping Zhang, Yixin Zhang, Ningmu Zou, and Wei Jiang. Time-wavelength multiplexed photonic neural network accelerator for distributed acoustic sensing systems. *Advanced Photonics*, 7(2):026008–026008, 2025.