nature photonics

Article

https://doi.org/10.1038/s41566-025-01660-x

Field-programmable photonic nonlinearity

Received: 4 June 2024

Accepted: 14 March 2025

Published online: 15 April 2025

Check for updates

Tianwei Wu [1], Yankun Li [2], Li Ge [3,4] & Liang Feng [1,2]

Propelled by advancements in artificial intelligence, the demand for field-programmable devices has grown rapidly in the last decade. Among various state-of-the-art platforms, programmable integrated photonics emerges as a promising candidate, offering a new strategy to drastically enhance computational power for data-intensive tasks. However, intrinsic weak nonlinear responses of dielectric materials have limited traditional photonic programmability to the linear domain, leaving out the most common and complex activation functions used in artificial intelligence. Here we push the capabilities of photonic field-programmability into the nonlinear realm by meticulous spatial control of distributed carrier excitations and their dynamics within an active semiconductor. Leveraging the architecture of photonic nonlinear computing through polynomial building blocks, our field-programmable photonic nonlinear microprocessor demonstrates in situ training of photonic polynomial networks with dynamically reconfigured nonlinear connections. Our results offer a new paradigm to revolutionize photonic reconfigurable computing, enabling the handling of intricate tasks using a polynomial network with unparalleled simplicity and efficiency.

Reconfigurable computing[1,2], enabled by highly flexible and reprogrammable hardware such as field-programmable gate array (FPGA) chips, offers a promising solution to a persistent challenge in traditional computational architectures, that is, the von Neumann bottleneck associated with large, non-local datasets with high throughputs. Compared with the general-purpose CPUs and GPUs, reconfigurable computing-based accelerators can execute computation-intensive operations—for example, matrix-vector multiplication or convolutions that are the fundamental operations used in artificial neural networks[3,4]—orders of magnitude faster. Integrated photonic demonstrations of reconfigurable computing[5–16] can further improve the performance of this specialized hardware by processing a large dataset at an unprecedented speed with minimal energy consumption by exploiting intrinsic parallelism, elevated frequency rates and large bandwidths that inherently come with working in the optical domain[17–21]. Despite a number of benefits intrinsically associated with computing using light, photonics has been so far constrained to linear operations.

Extensive research into nonlinear optics—although widely deemed as a promising approach—has offered limited assistance. Fundamentally, nonlinear optics focuses on effective light–light interactions that can be described as a power series of light amplitude(s) involved through the high order of the susceptibility of materials. For example, the Kerr effect—due to the $\chi^{(3)}$ optical nonlinearity—modifies the refractive index and causes light to self-focus, but the overall power of the beam remains relatively unchanged between input and output. This is a completely different nonlinear relation than the nonlinear function envisioned in reconfiguration computing, which must dynamically map inputs to outputs in a nonlinear fashion, such as through a tunable polynomial function. Although structural nonlinearity[22–24] in linear optical systems can facilitate signal processing through nonlinear mapping, due to the natural linearity of light, it is still difficult to achieve direct programmable nonlinear optical interconnects. The demonstrated nonlinear function in this context is between two different types of light field, where the measured field (that is, the probe) still behaves linearly: an increase in the probe amplitude results in a proportional

[1]Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA. [3]Department of Physics and Astronomy, College of Staten Island, CUNY, Staten Island, NY, USA. [4]The Graduate Center, CUNY, New York, NY, USA. e-mail: fenglia@seas.upenn.edu

increase in the output, as a manifestation of the natural linearity of light discussed above. This challenge lies in materials limitations that nonlinear interactions between light and available carriers in materials are intrinsically weak, even in commonly used active semiconductors with saturable absorption and amplification[25,26]. These fundamental issues in nonlinear optics have posed a considerable challenge to leveraging light-based computing for complex nonlinear operations.

Instead of pursuing the conventional approach of developing new materials through composition and bandgap engineering, here we start from a completely distinguished perspective: all-optical control of spatially varying carrier dynamics to strategically tailor nonlinear light–matter interactions in existing active semiconductors (such as InGaAsP), directly enabling field-programmable photonic nonlinearity for the first time, without the need to cascade multiple linear multiply–accumulate units as typically done in digital computing. Field-programmable photonic nonlinearity facilitates the configuration of polynomials of various orders, forming a polynomial nonlinear network capable of in situ photonic learning. Our field-programmable photonic nonlinear computing network targets complex tasks that linear networks fail to handle, surpassing conventional architectures composed of reconfigurable linear components and fixed nonlinear activations.

Figure 1 illustrates our disruptive concept of field-programmable photonic nonlinearity demonstrated on an active III–V semiconductor platform of InGaAsP multiple quantum wells[27–31]. The carrier dynamics greatly impacts the electronic states and recombination processes within the semiconductor, thereby offering a powerful toolbox to tailor the light–matter interaction. Efficient manipulation of carrier excitations to the valence band and their recombination with holes in the conduction band therefore becomes crucial for shaping the transmission of signal light on demand. Dependent on the pump power, this enables the transition of the transmission curve from saturable absorption with a transmission curve as a convex function, to optical transparency with a linear transmission curve and, ultimately, to saturable amplification featuring a concave function in transmission (Fig. 1a). Although nonlinear transmission functions naturally arise from saturable absorption or amplification, their responses are weak, only slightly deviating from a linear function regardless of the scale of the device platform. This is due to the fact that sole absorption or amplification rapidly alters the intensity of signal light away from the saturation intensity, leading to inefficient interaction between signal light and available carriers. To overcome this fundamental limitation, strategic control over spatial carrier dynamics within a two-dimensional (2D) space, specifically, the spatial distribution of excited carrier density, is essential for effectively manipulating nonlinear light–matter interactions (Fig. 1b), synthesizing high-order polynomials P^N of a degree of N in nonlinear transmission curves. Here, the distributed carrier density is precisely controlled by optimized external optical pumping fields, allowing for convenient reconfiguration of both the polynomial degree (N) and the coefficients (α_l for order l) (Fig. 1c). This enables the creation of reconfigurable on-chip nonlinear port-to-port connections with polynomial functions in real time, leading to the construction of a field-programmable photonic polynomial network (Fig. 1d). In contrast to the conventional deep neural networks, which typically involve the stacking of linear matrices and nonlinear activation units, the realization of a polynomial network[32,33] can drastically simplify the design of computational hardware, facilitating a reduction in the required depth of the computing network. Especially since our scheme is compatible with the recently developed lithography-free integrated photonics paradigm[34], field-programmable photonic nonlinearity enables direct excitation and reconfigurable manipulation of nonlinear connections, denoted as $P_{ij}^{N_{ij}}(x_j)$, which represent the transmission of signal light traveling from input port j to output port i, within a fully connected scheme. Taking a global perspective for optimization, the patterned pumping generation algorithm necessitates fewer parameters to simultaneously optimize all polynomial input–output (I/O) connections of varying degrees N_{ij}, which enables in situ training to be performed with high accuracy.

To validate field-programmable photonic nonlinearity and its ability to reconfigure polynomial functions, we fabricated the sample (see Methods) on a 250-nm-thick thin film of InGaAsP (Fig. 2a). The signal light with a wavelength of approximately 1,550 nm is generated by a micro-ring laser on the input side, with its transmission monitored using a grating coupler on the output side for free space detection. The I/O connection, that is, the InGaAsP thin film in between, remains lithographically unpatterned, such that carrier excitations within InGaAsP are purely modulated by 2D distributed external optical pumping, at a wavelength of 1,064 nm, programmed by a spatial light modulator (see Methods). In this work we focus on the reconfiguration of enhanced saturable absorption which, compared with saturable amplification, offers convex transmission curves with higher-N polynomials commonly used in data-intensive applications. To facilitate a strong interaction between signal light and local carriers within the semiconductor, it is important to maintain the intensity of signal light near the transparency intensity along the entire pathway of an I/O connection. Note that the available local carrier density regulates the imaginary part of the susceptibility of the semiconductor, consequently influencing its associated optical gain/loss coefficient (see Supplementary Section 1). Therefore, the spatial modulation between intense and weak external pumping emerges as efficient in managing local light intensity and its nonlinear interaction with carriers, effectively reshaping the transmission curve. In this scenario, a periodic modulation of pump light in space between the input and the output is employed under the overall envelope of the elliptically shaped pumping profile derived from an adjoint method[34], which is conveniently programmed with two parameters: the duty cycle β, defined by the ratio of the pumped area in a spatial period ($\beta = a/b$ in Fig. 2a), and the total pumping power P_l along this single I/O connection. With our focus on the convex functions, the nonlinearity in transmission curves is dominated by saturable absorption; therefore, the decrease in the value of β provides more unexcited carriers, leading to the reconfigurable synthesis of the nonlinear transmission functions continuously from nearly linear to strongly nonlinear (Fig. 2b). Simultaneously, by elevating P_l to counterbalance absorption stemming from the expanded unpumped region, the signal intensity can be maintained at a level for efficient carrier dynamics of saturable absorption throughout the I/O connection, allowing the strategy to be deployed for networks with varying scales.

The minimum of β is set by the maximum gain provided by the material and in this work is positioned at 0.2. The dominant polynomial order of the corresponding transmission curve is mainly determined by the value of β, as can be seen clearly from the slope in the logarithmic scale (Fig. 2c). In principle, the obtained nonlinear function should take the form of a polynomial comprising multiple orders, but the dominant order can serve as an effective indicator for quantifying the nonlinearity of the function. Evidently, the predominant polynomial term can be effectively tuned from low to high orders with decreasing β, culminating in a polynomial order as high as $N \approx 14$ with $\beta = 0.3$. Furthermore, it is crucial to highlight that concave-shaped nonlinear functions with $N < 1$—as well as more complex polynomials with precisely tailored combinations of different orders—can also be achieved (see Supplementary Section 2), with the ability to reconfigure their polynomial orders in real-time. By leveraging the convenient programmability in polynomial orders, we can extend our capabilities further to synthesize even more complex nonlinear functions through cascading the photonic polynomial connections, tackling increasingly intricate computing tasks. Also note that the result shown in Fig. 2b closely resembles the widely used rectified linear unit (ReLU)[35] function in deep neural networks, thereby being able to serve as fixed activation functions in conventional neural networks (see Supplementary Section 3).

https://doi.org/10.1038/s41566-025-01660-x

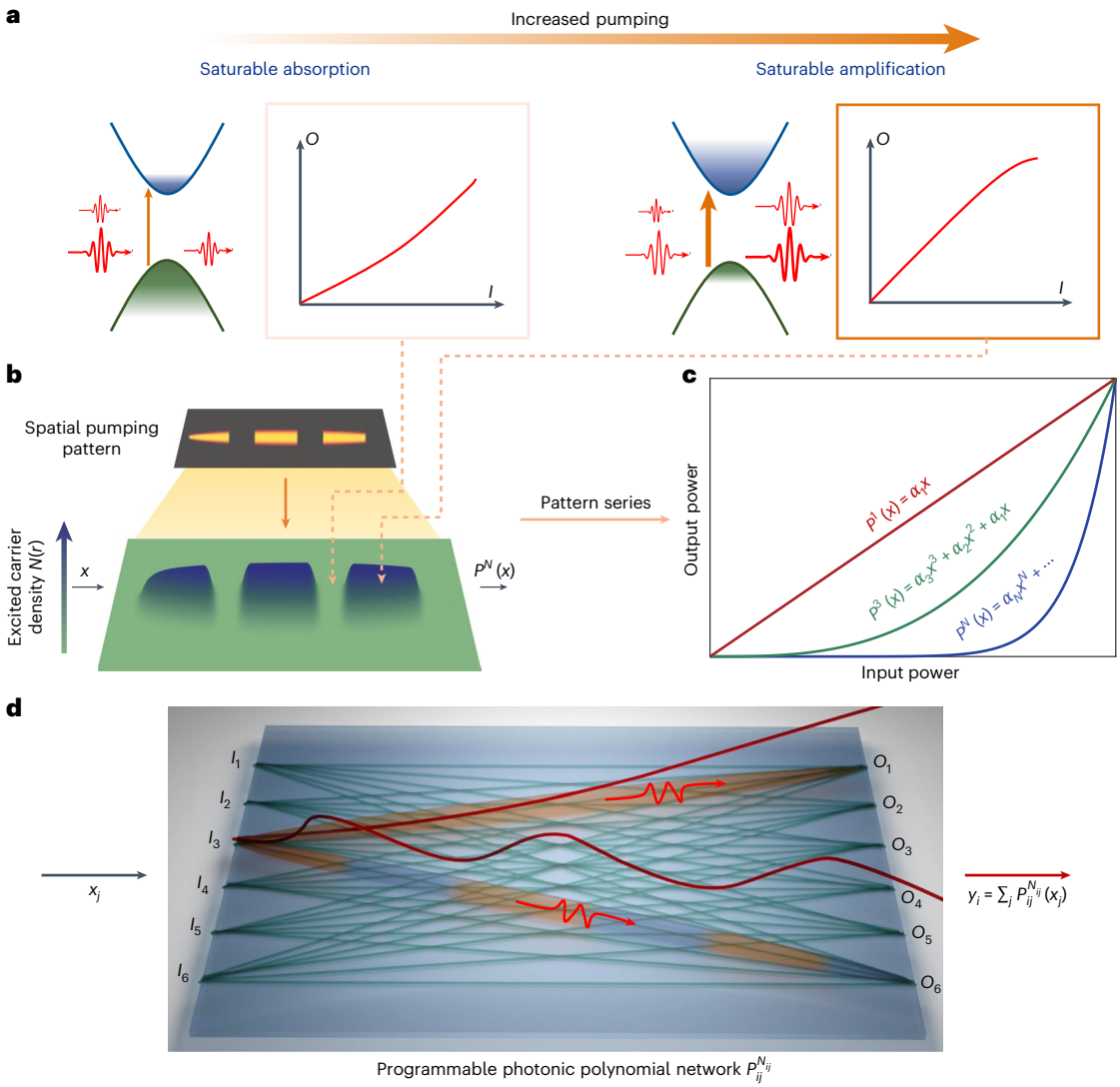


Fig. 1 | Field-programmable photonic nonlinearity. a, Pump-dependent photonic nonlinearity in an active semiconductor such as InGaAsP. When InGaAsP is unpumped or weakly pumped, a substantial portion of electrons resides in the valence band. Their transition to the conduction band, excited by the signal light, results in intrinsic absorption. Due to finite carrier populations, the transmission of the signal light is intensity-dependent, reflecting saturable absorption. As the pump power increases, more carriers become excited in the conduction band. Their recombination with holes in the valence band provides optical gain to the signal light, leading to saturable amplification. **b**, Spatial distribution of excited carriers controlled by the external pumping pattern, incident perpendicular to the sample plane, as indicated by the orange solid arrow. The combination of spatially varying pumping levels across the 2D space yields strong nonlinear light–matter interactions, enabling the reconfiguration of transmission curves programmed into different polynomial functions for the signal propagating from left to right. **c**, Programmable polynomial functions, showing that both the degree N and the l-order coefficients α_l of the synthesized polynomial functions can be tuned by spatial pumping patterns. **d**, Field-programmable photonic polynomial microprocessor for on-chip information processing and in situ learning. Nonlinear connections between arbitrary input and output ports can be established to create a fully connected integrated photonic network with complete programmability (indicated by the green lines). Selective spatial pumping allows for individual tuning of polynomial connections. For example, two different photonic connections (I_3 to O_6 and I_3 to O_1) with different spatial pumping strategies are demonstrated: segmented pumping (I_3 to O_6) alternates between amplification (marked in orange) and absorption (marked in blue), causing oscillations in signal power along the propagation direction (illustrated by the light trail above the connection), which favours high-order polynomials with enhanced nonlinearity. In contrast, continuous pumping (I_3 to O_1) leads to a monotonic increase in light power along propagation (also indicated by the light trail above), yielding low-order polynomials close to a linear function.

Its integration with linear networks holds considerable promise if heterogeneously integrated with a variety of platforms, particularly with silicon photonics through mature III–V/silicon integration[36].

Instead of conventional neural network models with elementwise nonlinear activation functions typically sandwiched between adjacent linear networks, our focus is to field-program optical polynomial functions in real-time, where polynomial functions are directly employed to connect neurons for deploying polynomial networks. Polynomial networks, which leverage the richer information provided by polynomial functions, can outperform purely linear matrix operations in various contexts, as demonstrated below. To deploy the field-programmable photonic polynomial network, we developed an iterative training algorithm (see Supplementary Section 4) to optimize the pumping pattern controlling the spatial modulation of carrier excitations within the 2D InGaAsP thin film. In our experiments, each I/O connection is individually controlled by the two parameters mentioned previously, that is, the duty cycle $\beta_{i,j}$ and the pump power $P_{I,ij}$ (for the connection from input j to output i), and the iterative algorithm

https://doi.org/10.1038/s41566-025-01660-x

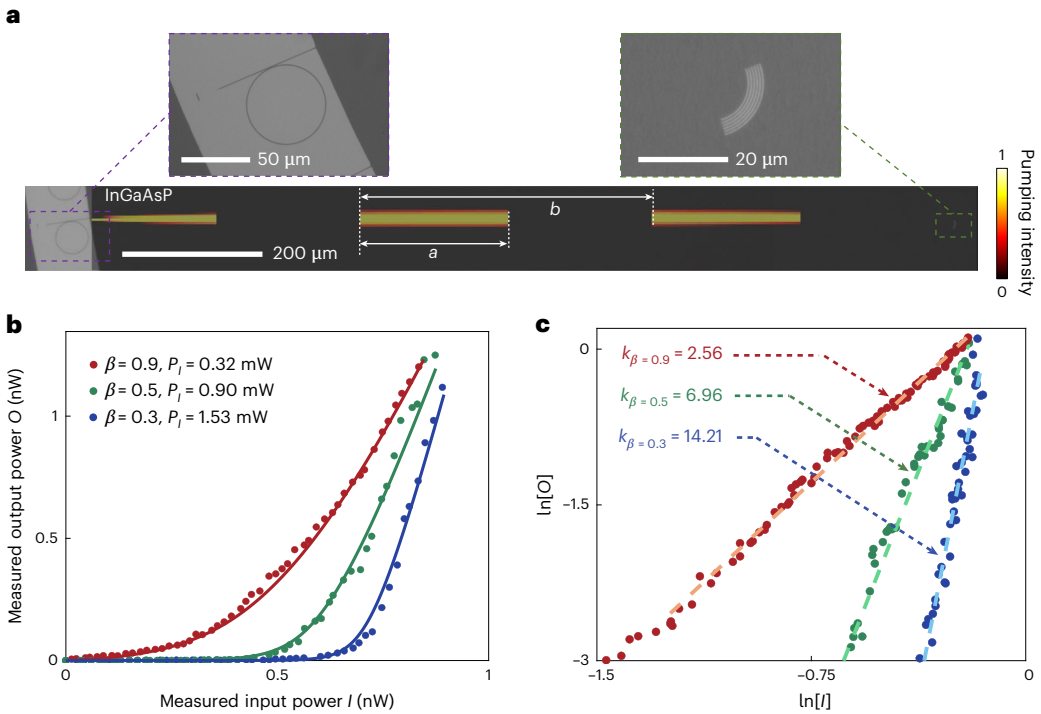


Fig. 2 | Reconfigurable polynomial I/O connection. a, Optical microscope image of the fabricated sample used for characterizing the reconfigurable polynomial I/O connection, alongside the overlay of the spatial modulation of optical pumping applied in experiments. Here, *a* denotes the length of a pumping segment, whereas *b* represents the length of a spatial period. The two upper insets show the input unit comprising a micro-ring laser coupled to a straight waveguide connected to the unpatterned InGaAsP region and a grating coupler in InGaAsP designed for output detection. **b**, The relationship between output and input power under different pumping scenarios, highlighting the reconfigurable polynomial I/O connection controlled by the duty cycle β. The overall pumping power P_I is adjusted to maintain a consistent output signal intensity across different scenarios. The solid lines are the simulation results from the travelling-wave model, showing remarkable consistency with experimental measurements. **c**, The relationship between output and input power in a logarithmic scale, directly showing the dominant polynomial order of the synthesized polynomial function (estimated via slopes). The fitting is applied to $O > 0.05$ nW ($\ln[O] > -3$) to avoid the influence of noise at low signal levels.

simultaneously updates both parameters for the optimization of all connections in the network.

To exemplify the superior performance of field-programmable photonic nonlinearity, we start with a benchmark two-class classification problem involving two input variables, where a synthetic dataset consists of two features in the input layer, labelled as x_1 and x_2 (Fig. 3a). The boundary between the two classes forms a curve within the feature vector space, showing an intrinsic nonlinear characteristic that can only be addressed using a nonlinear network. In experiments, we deployed a field-programmable 2 × 2 photonic network to execute a nonlinear transformation, directing the processed signal to the two output ports corresponding to the two classes (Fig. 3b). The feature vector (x_1, x_2) was encoded with normalized emission power from two micro-ring lasers at inputs. The predicted class was directly indicated by the higher intensity observed between the two output channels. Initially, a symmetric pumping pattern was configured with a uniform duty cycle and total pumping power for all four nonlinear I/O connections. In each training epoch, following the measurements of all 300 data points, a loss function $L = \frac{1}{2}\sum_{k=1}^{300}|(O_{\mathrm{tgt}} - O_{\mathrm{mea}})|^2$ was calculated, where O_{tgt} and O_{mea} are the target and measured output vectors defined as the power of each output port, respectively. The polynomial function of each connection ($P_{ij}^{N_{ij}}(x_j)$) was then updated. After 40 epochs, the classification accuracy reached 97.3%. In the feature vector space of the outputs, the two classes were predominantly separated, demonstrating effective distinguishability (Fig. 3c). When transforming back to the input vector space, only a few misclassifications near the boundary were observed (Fig. 3d). It is important to note that, for the intrinsic nonlinearity of this benchmark classification problem, these two classes cannot be effectively distinguished by any linear network. For comparison, a control experiment with a linear network was also conducted, resulting in a final accuracy of approximately 64.5% (see Supplementary Section 5).

The same task was also tested using digital conventional neural networks with one intermediate layer, where fixed nonlinear ReLU functions (f_α) were applied to each neuron between two linear networks (Fig. 3e). Two fully connected matrices w_{ij} and v_{ij} are digitally trained using a computer, resulting in $4M$ trainable weights with M neurons in the middle layer. The final accuracy achieved by the digital networks largely depends on the total number of connections (Fig. 3f). Despite using a much simpler architecture with only four connections, the photonic polynomial network can achieve a classification accuracy that is comparable with that attained through digital computation, which uses a conventional neural network with 16 to 20 linear connections and applies nonlinear activation to four neurons. This highlights the intrinsic higher information density provided by the direct programming of polynomial functions.

To harness the dynamic reconfiguration capabilities of field-programmable photonic nonlinearity for accelerating computing tasks, we performed in situ training for the Iris dataset[37] (see Supplementary Section 6), which is one of the most recognized datasets in the field of machine learning. The dataset consists of 150 samples across three iris flower species: *Iris setosa* (*I. setosa*), *Iris versicolor* (*I. versicolor*) and *Iris virginica* (*I. virginica*), each characterized by four features: sepal length, sepal width, petal length and petal width, which form the input feature vector. Notably, *I. versicolor* and *I. virginica* are challenging to distinguish using a linear network due to their similar feature distributions. A fully connected 4 × 3 photonic polynomial

https://doi.org/10.1038/s41566-025-01660-x

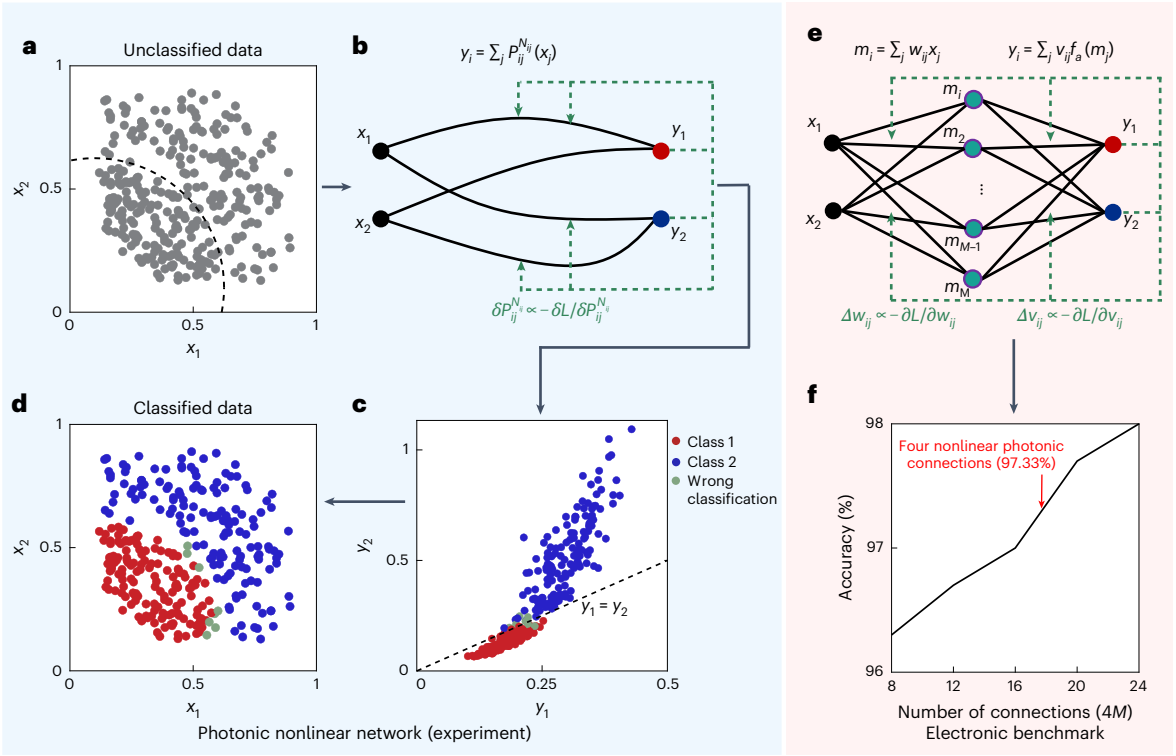


Fig. 3 | In situ training with field-programmable photonic nonlinearity for benchmark problem solving. a, Feature vector space of input data where two distinct datasets reside in two domains demarcated by a curved boundary. **b**, Training schematic for the photonic polynomial network illustrating the process of in situ training using field-programmable photonic nonlinearity. Each nonlinear connection (solid black lines) takes individual programming in a parallel manner according to feedback loops (green dashed lines) derived from measurement results. **c**, Post-training classification results where the classification outcomes are delineated by a straight line of $y_1 = y_2$. **d**, Input vector space post-training, where the clear separation of two classes represents effective performance. **e**, Training schematic for a computer-trained digital neural network outlining the training process. Nonlinear activation functions are applied to M neurons in the intermediate layer. The network is trained by adjusting the weights (w_{ij} and v_{ij}) in the linear transformations within each layer, as depicted by the green dashed lines. **f**, Classification accuracy as a function of computer-trained network complexity. Remarkably, it reveals that 20 trainable linear connections are required to barely surpass the experimental performance of our single-layer photonic polynomial network with only four connections.

network is conducted on our device (Fig. 4a), where four micro-ring lasers are excited to form the input vector, with the power of each micro-ring laser corresponding to the respective feature vector element (see Methods). In the output layer, the three classes are categorized as three output channels, representing *I. setosa* (O_1), *I. versicolor* (O_2) and *I. virginica* (O_3). The predicted class is determined by identifying the output channel with the highest intensity. At the core of the system is the unpatterned InGaAsP thin film that can host field-programmable photonic nonlinearity, where distributed carrier excitations enable direct programming of the polynomial network and real-time learning in a fully connected architecture. Two training strategies are employed to validate the importance of nonlinear connections: one using a nonlinear polynomial network trained on both duty cycle β and power P_l for each individual I/O connection, and the other employing a linear network with a fixed $\beta = 1$ for all connections (see Supplementary Section 7), while keeping all other parameters constant. Evidently, the nonlinear network exhibits a much lower training loss upon convergence, indicating a substantial performance enhancement (Fig. 4b). Upon the completion of the training, the nonlinear network has achieved a final classification accuracy of 96.7% for both the training and testing datasets (Fig. 4c,d), outperforming the linear network, which reaches only 85.8% accuracy for the training dataset and 86.7% for the testing dataset (Fig. 4e,f). The experimental results are consistent with the digital network simulations (see Supplementary Section 6). In addition to accuracy, another important performance indicator is confidence, which reflects how strongly the model believes that an input belongs to a given class. For the nonlinear network (Fig. 4g), the output power corresponding to the correct class is much higher than that of the other two classes, as shown by the power distribution across the output ports. By contrast, for the linear network (Fig. 4h), the output powers for *I. versicolor* (green) and *I. virginica* (blue) are comparable across output ports O_2 and O_3, indicating difficulty in clearly distinguishing between the two classes. To quantify the distinguishability, we define a parameter D for a sample belonging to class l as $D = p_l - \max\{p_m, m \neq l\}$, where $p_l = O_l/(O_1 + O_2 + O_3)$ is the normalized power ratio of the output port l. For correctly classified samples, D represents the difference in normalized powers between the correct port and the most misleading incorrect port. The average distinguishability for *I. versicolor* and *I. virginica* is $D_{nl} = 0.12$ for the nonlinear network, in stark contrast to $D_l = 0.02$ for the linear network.

Achieving high distinguishability is essential for ensuring robust and scalable performance in analogue photonic computing systems. As the network size and complexity grow, this capability becomes even more critical, enabling photonic nonlinear polynomial networks to handle intricate computational tasks with greater reliability and precision. To demonstrate this, we performed in situ training for speech command recognition[38] using a 16 × 4 fully connected polynomial network. The data were pre-processed, and a subset was created to match the scale of the network (see Supplementary Section 8). A dataset of 480 audio samples was randomly divided into 360 samples for training and 120 samples for validation, representing four distinct words: 'Bird', 'Happy', 'House' and 'Tree'. To reduce redundancy in the raw audio data, the intensity distribution across 16 frequency bands was extracted as the feature vector for each sample (q_s for data s). These feature vectors were then encoded as input signals and sequentially fed into the polynomial network. After training, the nonlinear

https://doi.org/10.1038/s41566-025-01660-x

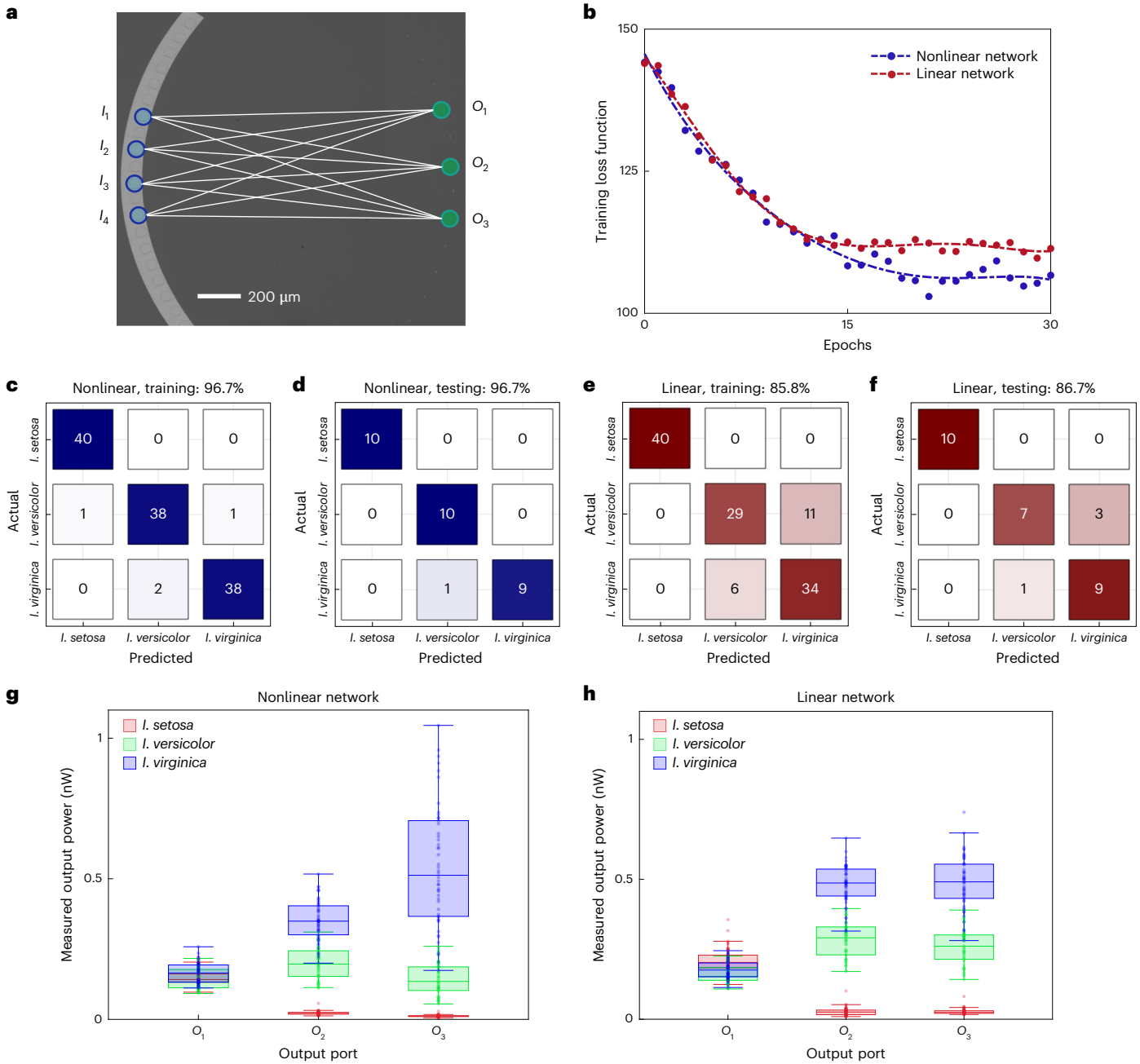


Fig. 4 | Photonic polynomial networks for iris flower classification. a, The schematic depicts a single layer of a 4 × 3 network designed for the iris flower classification task. The four-element feature vectors are encoded by the four micro-ring lasers on the left side (marked with blue circles). The output signal is measured through three grating couplers, each corresponding to a different iris class. **b**, Evolution of the loss function during the training of nonlinear (blue) and linear (red) networks. In the nonlinear network (blue), both the duty cycle β and power P_i are trainable across all 12 connections, whereas in the linear network (red), only the power of individual connections is trainable, with the duty cycle β fixed at $\beta = 1$. Both strategies start with the same initial pumping pattern; slight differences in the loss function at the beginning of training are attributed to fluctuations in the pumping pattern generation and camera detection during extensive dataset measurements. The dashed lines are polynomial fittings of the experimental data (denoted by dots). **c**–**f**, Confusion matrices of the training (**c**,**e**) and testing (**d**,**f**) datasets for the nonlinear (**c**,**d**) and linear (**e**,**f**) networks. The value in each box represents the number of samples. **g**,**h**, Box charts of the measured power at the output ports for all data after training for nonlinear (**g**) and linear (**h**) networks. For each class, the boxes indicate the interquartile range with the central line indicating the median. The coloured whiskers extend from the upper (lower) quartile to the non-outlier maximum (minimum), marked by the horizontal lines. Outliers are defined as data points more than 1.5 times the interquartile range from the edges of the boxes.

network achieved a classification accuracy of 92.5% and 91.7% on the training and testing datasets, respectively (Fig. 5c,d), also outperforming the linear network (see Supplementary Section 9), which attained accuracies of 89.2% and 86.7% on the training and testing datasets, respectively (Fig. 5e,f). These results are consistent with our digital simulations (see Supplementary Section 8) and highlight the advantages of incorporating nonlinearity into the network. This result marks the first demonstration of intricate computations using a photonic polynomial network, made possible by the direct programming of photonic nonlinear I/O connections. By unlocking new levels of computational power and versatility in integrated photonics through reconfigurable nonlinear light–matter interactions, our approach

https://doi.org/10.1038/s41566-025-01660-x

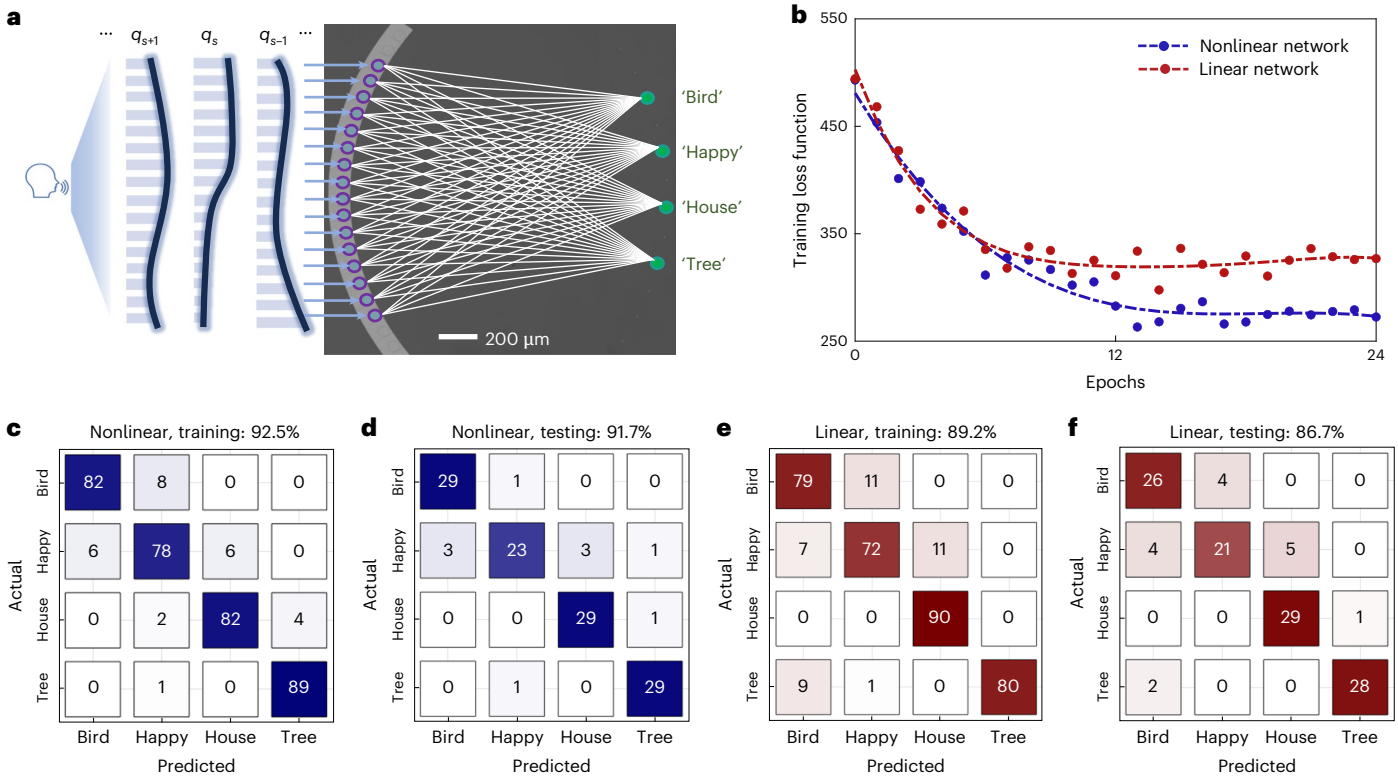


Fig. 5 | Photonic polynomial networks for speech command recognition. **a**, Speech command recognition task in a 16 × 4 photonic polynomial network. Fourier transform is applied to each audio data (s) to encode it as intensity distribution across 16 frequency bands (depicted by the blue bars) generated from 16 micro-ring lasers on the left side of the sample, forming the input feature vector. The output signal is measured through four grating couplers corresponding to 'Bird', 'Happy', 'House' and 'Tree', respectively. **b**, Evolution of the loss function during training nonlinear (blue) and linear (red) networks. Dashed lines represent polynomial fitting of the experimental data denoted by dots. **c**–**f**, Confusion matrices of the training (**c**,**e**) and testing (**d**,**f**) datasets for the nonlinear (**c**,**d**) and linear (**e**,**f**) networks.

paves the way for advancing more complex, real-world applications (see Supplementary Section 10).

We have demonstrated a groundbreaking advancement in field-programmable photonic nonlinearity through quantum control of carrier dynamics within an active semiconductor thin film, executing polynomial functions with a simpler architecture and faster processing speed than the electronic counterpart (see Supplementary Section 11). The precise manipulation of spatially excited carrier densities enables the realization of highly reconfigurable polynomial photonic connections. Distinguished from the replications of well-developed electronic architectures on photonic platforms, our work serves as a pioneer in the exploration of photonic paradigms specifically tailored for computing with light, accounting for the intrinsic disparities between photons and electrons. By facilitating direct, real-time programming of polynomial networks, our demonstrated field-programmable photonic nonlinearity represents a significant leap forwards in photonic computing and its integration with reconfigurable computing architectures.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41566-025-01660-x.

References

1. Compton, K. & Hauck, S. Reconfigurable computing: a survey of systems and software. *ACM Comput. Surv.* **34**, 171–210 (2002).
2. Bondalapati, K. & Prasanna, V. K. Reconfigurable computing systems. *Proc. IEEE* **90**, 1201–1217 (2002).
3. Reagen, B. et al. Minerva: enabling low-power, highly-accurate deep neural network accelerators. In *2016 ACM/IEEE 43rd Annual International Symposium on Computer Architecture (ISCA)* 267–278 (ACM/IEEE, 2016).
4. Chen, Y.-H., Krishna, T., Emer, J. S. & Sze, V. Eyeriss: an energy-efficient reconfigurable accelerator for deep convolutional neural networks. *IEEE J. Solid-State Circuits* **52**, 127–138 (2017).
5. Bogaerts, W. et al. Programmable photonic circuits. *Nature* **586**, 207–216 (2020).
6. Wang, J., Sciarrino, F., Laing, A. & Thompson, M. G. Integrated photonic quantum technologies. *Nat. Photon.* **14**, 273–284 (2019).
7. Shen, Y. et al. Deep learning with coherent nanophotonic circuits. *Nat. Photon.* **11**, 441–446 (2017).
8. Liu, W. et al. A fully reconfigurable photonic integrated signal processor. *Nat. Photon.* **10**, 190–195 (2016).
9. Pai, S. et al. Experimentally realized in situ backpropagation for deep learning in photonic neural networks. *Science* **380**, 398–404 (2023).
10. Feldmann, J. et al. Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021).
11. Shastri, B. J. et al. Photonics for artificial intelligence and neuromorphic computing. *Nat. Photon.* **15**, 102–114 (2021).
12. Ashtiani, F., Geers, A. J. & Aflatouni, F. An on-chip photonic deep neural network for image classification. *Nature* **606**, 501–506 (2022).

13. Ying, Z. et al. Electronic-photonic arithmetic logic unit for high-speed computing. *Nat. Commun.* **11**, 2154 (2020).
14. Xu, Z. et al. Large-scale photonic chiplet Taichi empowers 160-TOPS/W artificial general intelligence. *Science* **384**, 202–209 (2024).
15. Pérez-López, D., López, A., DasMahapatra, P. & Capmany, J. Multipurpose self-configuration of programmable photonic circuits. *Nat. Commun.* **11**, 6359 (2020).
16. Feldmann, J., Youngblood, N., Wright, C. D., Bhaskaran, H. & Pernice, W. H. P. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **569**, 208–214 (2019).
17. Wang, T. et al. An optical neural network using less than 1 photon per multiplication. *Nat. Commun.* **13**, 123 (2022).
18. Sludds, A. et al. Delocalized photonic deep learning on the internet's edge. *Science* **378**, 270–276 (2022).
19. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
20. Zhou, H. et al. Photonic matrix multiplication lights up photonic accelerator and beyond. *Light: Sci. Appl.* **11**, 30 (2022).
21. Camacho, M., Edwards, B. & Engheta, N. A single inverse-designed photonic structure that performs parallel computing. *Nat. Commun.* **12**, 1466 (2021).
22. Wanjura, C. C. & Marquardt, F. Fully non-linear neuromorphic computing with linear wave scattering. *Nat. Phys.* **20**, 1434–1440 (2024).
23. Xia, F. et al. Nonlinear optical encoding enabled by recurrent linear scattering. *Nat. Photon.* **18**, 1067–1075 (2024).
24. Yildirim, M., Dinç, N. U., Oguz, I., Psaltis, D. & Moser, C. Nonlinear processing with linear optics. *Nat. Photon.* **18**, 1076–1082 (2024).
25. Dubovitsky, S., Steier, W. H., Mathur, A. & Dapkus, P. D. Gain saturation properties of a semiconductor gain medium with tensile and compressive strain quantum wells. *IEEE J. Quantum Electron.* **30**, 380–391 (1994).
26. Haiml, M., Grange, R. & Keller, U. Optical characterization of semiconductor saturable absorbers. *Appl. Phys. B* **79**, 331–339 (2004).
27. Zhao, H. et al. Non-Hermitian topological light steering. *Science* **365**, 1163–1166 (2019).
28. Zhang, Z. et al. Spin-orbit microlaser emitting in a four-dimensional Hilbert space. *Nature* **612**, 246–251 (2022).
29. Qiao, X. et al. Higher-dimensional supersymmetric microlaser arrays. *Science* **372**, 403–408 (2021).
30. Zhang, Z. et al. Tunable topological charge vortex microlaser. *Science* **368**, 760–763 (2020).
31. Zhang, Y. et al. High-dimensional quantum key distribution by a spin–orbit microlaser. *Phys. Rev. X.* **15**, 011024 (2025).
32. Chrysos, G. G. et al. Deep polynomial neural networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 4021–4034 (2022).
33. Chrysos, G. G. et al. P–nets: Deep polynomial neural networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 7325–7335 (IEEE, 2020).
34. Wu, T., Menarini, M., Gao, Z. & Feng, L. Lithography-free reconfigurable integrated photonic processor. *Nat. Photon.* **17**, 710–716 (2023).
35. Nair, V. & Hinton, G. E. Rectified linear units improve restricted boltzmann machines. In *Proc. 27th International Conference on International Conference on Machine Learning* 807–814 (Omnipress, 2010).
36. Feng, X. et al. Non-Hermitian hybrid silicon photonic switching. *Nat. Photon.* **19**, 264–270 (2025).
37. Fisher, R. *Iris* (UCI Machine Learning Repository, 1936); https://doi.org/10.24432/C56C76
38. Warden, P. *Speech Commands: A Public Dataset for Single-Word Speech Recognition* (2017); https://storage.googleapis.com/download.tensorflow.org/data/speech_commands_v0.01.tar.gz

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Methods

Sample fabrication

To fabricate the field-programmable nonlinear photonic devices, we used a III–V semiconductor wafer composed of 250-nm-thick InGaAsP multiple quantum wells on an InP substrate. Electron-beam lithography was used exclusively to pattern the micro-ring lasers, input waveguides and grating couplers, whereas the bulk nonlinear processing area remained lithography free. A layer of SiO_2 was initially deposited on top via plasma-enhanced chemical vapour deposition. ZEP520A was then spin-coated onto the sample as a positive resist. Following exposure, the wafer underwent development using *ortho*-xylene and immersion in isopropyl alcohol. The SiO_2 layer was then etched in a reactive-ion etching process using CHF_3, with the remaining ZEP removed via oxygen plasma. The patterned SiO_2 layer served as a hard mask for inductively coupled plasma reactive-ion etching using BCl_3:Ar plasma. After dry etching of InGaAsP, the remaining SiO_2 layer was removed using buffered oxide etchant. A 3-µm-thick cladding layer of silicon nitride (Si_3N_4) was subsequently deposited as a cladding layer using plasma-enhanced chemical vapour deposition. Finally, the sample was bonded to a piece of glass slide, and the InP substrate was selectively removed via wet etching using a mixture of hydrochloric and phosphoric acid.

Optical measurements

The optical set-up (Extended Data Fig. 1a) uses a nanosecond pulse laser (Spectra-Physics, EXPL-1064) operating at a wavelength of 1,064 nm, with a pulse width of approximately 10 ns and a repetition rate of 10 kHz, as the pumping beam source. The pumping beam is divided into two paths by a beam splitter. The first path, modulated by SLM 1 (Hamamatsu, X15213), generates the structured patterns required for programming photonic nonlinearity. A structured-light toolbox[39] implements holographic pumping sequences on the basis of the Gerchberg–Saxton algorithm. The incident 4× objective lens, with a numerical aperture of 0.2, produces an Airy disk radius of ~3.25 µm. The pixel resolution of the pumping pattern is set at 10 µm × 10 µm, well within the system's diffraction limit. Although the generation of the pumping pattern is realized by SLM 1 containing 1,272 × 1,024 pixels, the number of the independent variables defining an interconnect between any input and output is only two (that is, duty cycle β and the pumping power P_l), making the in situ training convenient to perform. The other path, modulated by SLM 2 (Meadowlark Optics, HSP-1K-500-1200), is designated for micro-ring laser excitation and input signal encoding. The maximum operation speed of SLM 2 is around 300 Hz. The micro-ring lasers operate around 1,550 nm, with slight detuning for each individual ring. The emitted signal from the photonic chip is collected using a 5× objective lens with a numerical arpeture of 0.14. A band-pass filter, centred at 1,550 nm with a 30 nm bandwidth, removes unwanted photoluminescence. The intensities of the output ports are captured by an infrared camera (Goldeye, G-032 Cool TEC2) operating at a frame rate of 50 Hz. To fully eliminate the influence of photoluminescence, a reference image (with only the pumping pattern from SLM 1 applied, no microlaser excitation) is subtracted from the captured image. Extended Data Fig. 1b illustrates an optical measurement example for the integrated photonic network. A 16-element input feature vector is encoded by the emission power of 16 micro-ring lasers on the left, corresponding to the 16 input ports ($I_1 - I_{16}$). The input power of each microlaser is monitored by the power emitted by the scatterer (shown by each white dashed box). The output power of four output ports ($O_1 - O_4$) is measured by the power emitted by the grating coupler on the right (shown by each yellow dashed box). An integrating sphere photodiode is used to monitor the power of the pump light. The transmission between the power measuring position and the sample plane is calibrated and all of the pumping power values shown in the manuscript indicate the incident power on the sample plane. The digital readout of the camera is also calibrated by the integrating sphere photodiode with a large incident laser light spot covering an area over 90,000 pixels of the camera. During the measurements, the signal power and camera exposure are monitored to prevent sensor saturation. As shown in Extended Data Fig. 1c, the camera's digital readout exhibits a linear relationship with the actual incident power within the working range.

Signal encoding

The signal is encoded through the emission power of the micro-ring lasers, which is monitored and optimized using real-time images captured by the camera. SLM 2, which controls the pumping of the micro-ring resonators, is iteratively reconfigured until the measured input feature vector (I_e) matches the target (I_t), guided by a gradient-descent algorithm. Once optimized, the corresponding hologram is saved. During in situ training, the saved hologram is reloaded to SLM 2 in each training epoch. The average accuracy of the encoding (defined by $\frac{I_e \cdot I_t}{|I_e||I_t|}$) is above 99.95%. The stability of the encoded feature vector is also tested by evaluating the similarity of the input feature vector of the same sample in the dataset across different epochs (defined by $\frac{I_{m1} \cdot I_{m2}}{|I_{m1}||I_{m2}|}$ for epochs m_1 and m_2). The average similarity value is 99.97%, confirming a highly stable input signal throughout the training process.

Training process

During each training epoch for the Iris and Speech datasets, encoded samples are sequentially excited, and an image of each sample is captured by the camera. Once measurements for all samples are completed, the input and output powers are calculated. A cross-entropy loss function is then computed, and a new desired pumping pattern is generated based on the algorithm, guided by the measurement results from the training dataset. The hologram corresponding to this new pumping pattern is then created using the Gerchberg–Saxton algorithm. However, due to residual errors in the algorithm and imperfect wavefront correction, the actual generated pumping pattern may deviate from the desired pattern. This limitation becomes more pronounced near the end of training, where precise power balancing between interconnects is crucial. This has been the primary cause of the fluctuation in the loss function in our experiments.

To mitigate these effects and ensure a more stable training process, the hologram from the current epoch is used as an initial guess for the next epoch. Furthermore, the Gerchberg–Saxton algorithm is restricted to only two optimization iterations per epoch to ensure a smooth and continuous evolution of the hologram. The primary bottleneck in training speed arises from the low frame rate of the camera and the data transfer speed between the camera and the desktop. For stable operation, the experiments are conducted at 20 measurements per second, resulting in an approximate epoch duration of 8 s for the Iris dataset (150 samples). It is important to note that only the power information from the input and output ports is needed for training, rendering most of the data recorded by the camera redundant. With faster detectors, the training speed could be greatly improved.

Data availability

The main data that support this study are available within the Article and its Supplementary Information. The additional raw and processed data are available at https://doi.org/10.6084/m9.figshare.28560848.v1.

Code availability

The computer codes that support the plots within this paper and other findings of this study are available from the corresponding author on reasonable request.

References

39. Lenton, I. C. D., Stilgoe, A. B., Nieminen, T. A. & Rubinsztein-Dunlop, H. OTSLM toolbox for structured light methods. *Comput. Phys. Commun.* **253**, 107199 (2020).

Acknowledgements

We acknowledge the support from Defense Advanced Research Projects Agency (DARPA) (grant no. W911NF-21-1-0340), Office of Naval Research (ONR) (grant no. N00014-23-1-2882) and National Science Foundation (NSF) (grant nos. ECCS-2023780, ECCS-2425529 and DMR-2326699). L.G. acknowledges the support from the National Science Foundation (NSF) (grant nos. PHYS-1847240 and DMR-2326698).

Author contributions

T.W. and L.F. conceived the project. T.W. developed the in situ training algorithm for nonlinear networks, fabricated the devices, and performed optical measurements and digital network simulations. Y.L., T.W. and L.G. developed theoretical models for nonlinear photonic transmission. L.F. guided the research. All authors participated in discussions and contributed to the writing of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41566-025-01660-x.

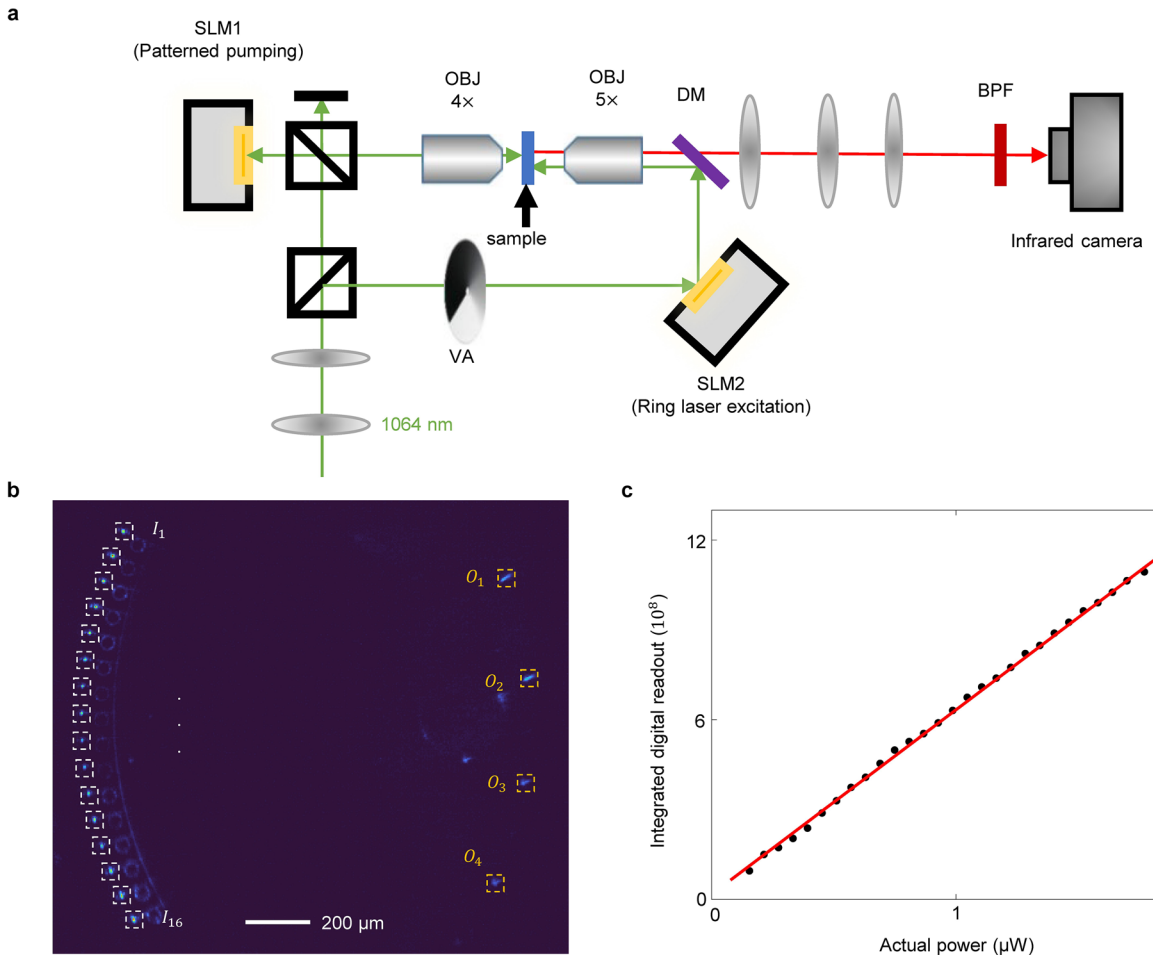
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41566-025-01660-x.

Correspondence and requests for materials should be addressed to Liang Feng.

Peer review information *Nature Photonics* thanks Fei Xia and Jianping Yao for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41566-025-01660-x



Extended Data Fig. 1 | Optical measurement. a, Set-up for the measurement. VA: variable attenuator; OBJ: objective; DM: dichroic mirror; BPF: band-pass filter. **b**, An image captured by the camera, with the influence of the PL removed by subtracting a reference image taken without input signals. The white dashed boxes on the left indicate the integration area for input port powers, while the yellow dashed boxes on the right highlight the integration area for output port powers. **c**, Calibration of the digital readout of the camera, performed using a large external laser light spot that covers most of the camera's working area. The results show a convincing linear input–output power response within the working range.