Check for updates

Laser & Photonics Reviews

LASER & PHOTONICS REVIEWS

www.lpr-journal.org

RESEARCH ARTICLE

Optical Arbitrary Vector-Matrix Multiplication in a Single Time-Varying Ring

Zhengnuo Chen[1] | Yuhan Wang[1] | Yixiang Zhang[1] | Guangzhen Li[1] | Bo Wang[1] | Xianfeng Chen[1,2] | Luqi Yuan[1]

[1]State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China | [2]Collaborative Innovation Center of Light Manipulation and Applications, Shandong Normal University, Jinan, China

Correspondence: Luqi Yuan (yuanluqi@sjtu.edu.cn)

Received: 17 January 2026 | **Revised:** 16 April 2026 | **Accepted:** 24 April 2026

Keywords: Optical computing | Vector-matrix multiplication | Synthetic frequency dimension

ABSTRACT

Optical computing is increasingly seen as a promising route to fast machine learning by utilizing the ultrafast speed of light and its intrinsic parallelism. A recent challenge is to obtain a scalable and programmable vector-matrix multiplication module with compact integrated photonic technology. Here, we propose an architecture based on a time-varying optical ring to realize arbitrary vector-matrix multiplication with $O(1)$ spatial complexity and $O(N)$ evolution time. The non-Hermitian synthetic frequency lattice is built where we prove its universality, and further confirming its scalability with high precision and full programmability remaining. We also highlight the strong expressivity within a fixed physical footprint by demonstrating no intrinsically hard-to-learn matrix in our architecture. Our proposal is naturally suitable for chip-scale realization by overcoming the limitation from spatial complexity in conventional optical computing, which paves the way toward integrated, large-scale, and scalable photonic computing systems.

1 | Introduction

Optical computing has been well recognized as a promising approach to accelerate a range of machine learning applications due to its ultrafast propagation speed and intrinsic parallelism capability [1–7]. Among different fundamental operations in optical computing, arbitrary vector-matrix multiplication plays an essential role underpinning a variety of tasks, including machine learning [2–7], linear transforms and convolutions for optical signal processing [8–10], and programmable optical filtering [11]. With the growing desires for the general purpose and programmable functionality, compact and reconfigurable photonic modules capable of performing high-dimensional linear matrix operations with low latency are therefore needed [12]. However, existing implementations using architectures such as Mach–Zehnder interferometer meshes [13, 14], diffractive optical elements [8, 9, 15, 16] and phase-change memory arrays [17] usually suffer from large physical footprints, complex calibration procedures, and limited scalability which relies heavily on spatial configurations and hence imposes severe constraints on the potential integration density for on-chip applications.

It has been recently noted that synthetic dimensions may provide an alternative paradigm to overcome the spatial limitations for optical computing [18–22]. Internal degrees of freedom of light can be used to replace the linear transformation of states over spatial distributions, which may enable programmable functionality beyond conventional spatial-domain implementations [23]. For example, recent proposals suggest cascaded modulated rings for frequency-space operations [20], photonic time circuits for programmable unitary transformation [21], and time-multiplexed designs [24, 25]. Nevertheless, all these proposals require multiple optical components, which limit their further on-chip implementation. Moreover, to achieve arbitrary vector-matrix

Zhengnuo Chen and Yuhan Wang contributed equally to this work.

© 2026 Wiley-VCH GmbH

multiplication in these systems, it usually demands cascaded modulation stages or complex optical routing networks, resulting in non-negligible $O(N)$ spatial complexity. Therefore it is still imperative to seek photonic proposal that can simultaneously offer programmability and compactness for implementing large-scale optical computing with on-chip devices.

In this work, we demonstrate a time-varying synthetic frequency lattice based on a modulated ring can realize arbitrary vector–matrix multiplication. By using resonant frequency modes as the computational basis vector and designing time-dependent inter-mode couplings, we confirm this architecture capable for realizing arbitrary multiplications with the universality proof and inverse verification in simulations. The scalability of our proposal is further demonstrated where the evolution time scales linearly with the matrix size N. We exhibit that there is no intrinsically hard-to-learn matrix after applying a re-study approach, highlighting the strong expressivity of our design. Our work verifies a programmable architecture can be established for realizing high-accuracy arbitrary vector–matrix operations within a compact optical module.

2 | Results

2.1 | Model

We propose arbitrary vector-matrix multiplications based on a synthetic frequency-time architecture in a single ring resonator (see Figure 1a). By ignoring the group-velocity dispersion, we can take the set of resonant modes at frequencies $\omega_n = \omega_0 + n\Omega_R$ as basis vector. Here ω_0 is a reference frequency, n is a positive integer that we take from 1 to N (the lowest and highest modes under our consideration), and Ω_R is the free spectral range (FSR) of the ring [26]. We then introduce both phase modulation (PM) and amplitude modulation (AM) inside the ring with modulation signal composed of a sequence of the sinusoid function at modulation frequency $n\Omega_R (n \in [1, N-1])$, namely $f_{\mathrm{PM}}(t) = \sum_n \alpha_n \cos(n\Omega_R t + \phi_n)$ and $f_{\mathrm{AM}}(t) = \sum_n \beta_n \cos(n\Omega_R t + \psi_n)$ respectively. Here α_n (β_n) and ϕ_n (ψ_n) are corresponding modulation strength and modulation phase for PM (AM). For an injection of light pulse $\tilde{\boldsymbol{a}}$ labeled as $\tilde{\boldsymbol{a}} = (a_1 e^{i\omega_1 t}, a_2 e^{i\omega_2 t}, \ldots, a_N e^{i\omega_N t})$, where a_n is the field amplitude for the n-th mode, the output light after it passes both PM and AM $\tilde{\boldsymbol{a}}'$ is given by (see Note S1 for details)

$$\tilde{\boldsymbol{a}}' = \sum_k a_k e^{i\omega_k t}\left[1 + \sum_{n'=1}^{N-1}\left(\frac{\beta_{n'}}{2i}e^{i\psi_{n'}+in'\Omega_R t} - \frac{\beta_{n'}}{2i}e^{-i\psi_{n'}-in'\Omega_R t}\right)\right] \times \prod_{n=1}^{N-1}\left[\sum_{h=-\infty}^{\infty} i^h J_h(\alpha_n) e^{ih(\phi_n + n\Omega_R t)}\right]. \tag{1}$$

We take the weak-modulation condition ($\alpha_n < 0.7$, see Note S2 for discussion) to apply the first-order approximation, and take the sharp frequency cutoff to exclude frequency components with frequencies higher than ω_n or lower than ω_1 [27, 28] (see Note S3 for the justification). Therefore, the transfer matrix $\boldsymbol{T_m}$ for the amplitude vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_N)$ after each round-trip propagation along the ring is obtained as

$$\boldsymbol{T_m} = \begin{pmatrix} 1 & \frac{\beta_{1,m}}{2i}e^{i\psi_{1,m}} & \frac{\beta_{2,m}}{2i}e^{i\psi_{2,m}} & \cdots & \frac{\beta_{N-1,m}}{2i}e^{i\psi_{N-1,m}} \\ -\frac{\beta_{1,m}}{2i}e^{-i\psi_{1,m}} & 1 & \frac{\beta_{1,m}}{2i}e^{i\psi_{1,m}} & & \frac{\beta_{N-2,m}}{2i}e^{i\psi_{N-2,m}} \\ -\frac{\beta_{2,m}}{2i}e^{-i\psi_{2,m}} & -\frac{\beta_{1,m}}{2i}e^{-i\psi_{1,m}} & 1 & & \vdots \\ \vdots & & & \ddots & \\ -\frac{\beta_{N-1,m}}{2i}e^{-i\psi_{N-1,m}} & -\frac{\beta_{N-2,m}}{2i}e^{-i\psi_{N-2,m}} & \cdots & -\frac{\beta_{1,m}}{2i}e^{-i\psi_{1,m}} & 1 \end{pmatrix} \times \prod_{n=1}^{N-1} J_0(\alpha_{n,m}) \begin{pmatrix} 1 & i\frac{J_1(\alpha_{1,m})}{J_0(\alpha_{1,m})}e^{i\phi_{1,m}} & i\frac{J_1(\alpha_{2,m})}{J_0(\alpha_{2,m})}e^{i\phi_{2,m}} & \cdots & i\frac{J_1(\alpha_{N-1,m})}{J_0(\alpha_{N-1,m})}e^{i\phi_{N-1,m}} \\ i\frac{J_1(\alpha_{1,m})}{J_0(\alpha_{1,m})}e^{-i\phi_{1,m}} & 1 & i\frac{J_1(\alpha_{1,m})}{J_0(\alpha_{1,m})}e^{i\phi_{1,m}} & & i\frac{J_1(\alpha_{N-2,m})}{J_0(\alpha_{N-2,m})}e^{i\phi_{N-2,m}} \\ i\frac{J_1(\alpha_{2,m})}{J_0(\alpha_{2,m})}e^{-i\phi_{2,m}} & i\frac{J_1(\alpha_{1,m})}{J_0(\alpha_{1,m})}e^{-i\phi_{1,m}} & 1 & & \vdots \\ \vdots & & & \ddots & \\ i\frac{J_1(\alpha_{N-1,m})}{J_0(\alpha_{N-1,m})}e^{-i\phi_{N-1,m}} & i\frac{J_1(\alpha_{N-2,m})}{J_0(\alpha_{N-2,m})}e^{-i\phi_{N-2,m}} & \cdots & & 1 \end{pmatrix}. \tag{2}$$

For external modulation signals, we consider time-varying modulations, where signals are composed of a series of modulation segments (layers). Each layer of modulations is labeled by a positive index $m \in [1, M]$, and is applied for a total of Q_m round

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

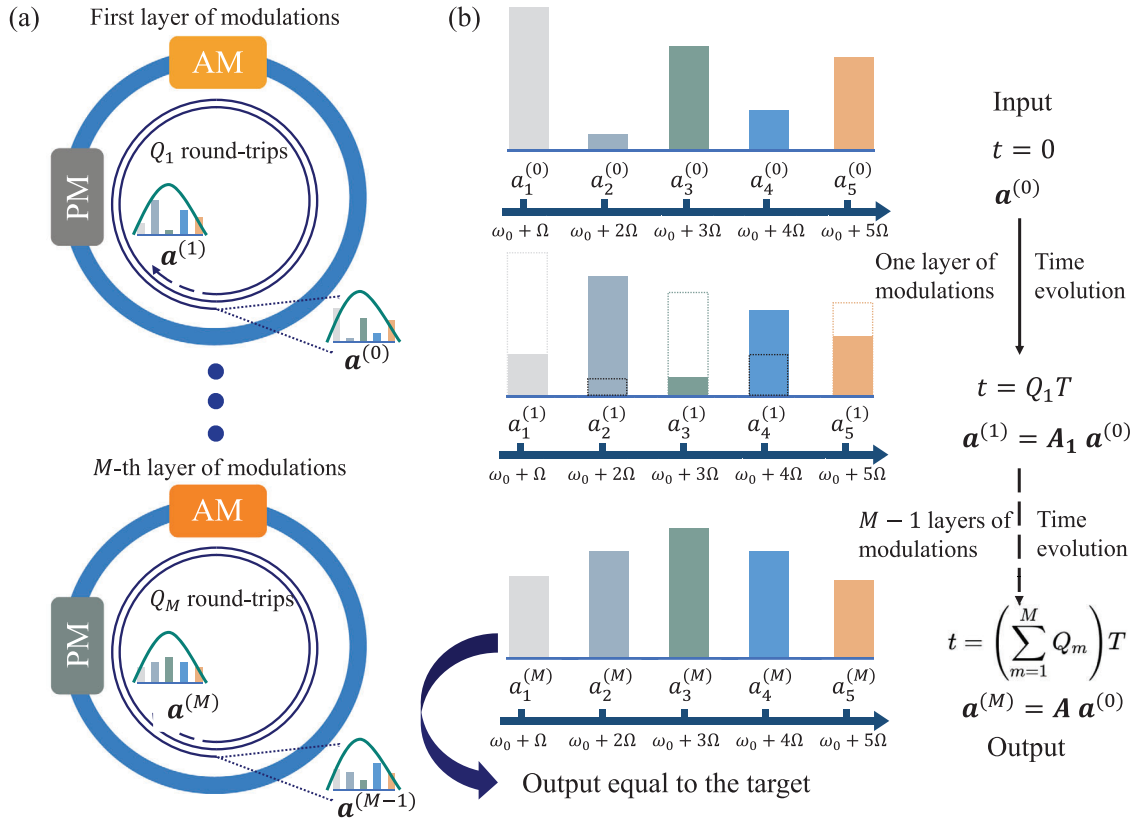


FIGURE 1 | Architecture for implementing vector-matrix multiplication in a time-varying synthetic frequency lattice. (a) A dynamically modulated ring resonator (blue) incorporating both an amplitude modulator (orange) and a phase modulator (gray). Different shades of the same color denote different modulation parameters. The circulating light field in the single ring evolves from $\boldsymbol{a}^{(0)}$ to $\boldsymbol{a}^{(1)}$ after the first layer of modulations, i.e. after Q_1 round-trips. After the series of M layers of modulations, the input light field is transformed into the output $\boldsymbol{a}^{(M)}$. (b) Schematic of the vector-matrix multiplication scheme based on the synthetic frequency dimension with time-varying modulation. Each modulation layer $\boldsymbol{A_m}$ corresponds to light circulating within the ring for Q_m round-trips under a fixed modulation $\boldsymbol{T_m}$, such that $\boldsymbol{A_m}=\boldsymbol{T_m}^{(Q_m)}$. The modulation is changed every Q_mT evolution time, where $T=2\pi/\Omega_R$ is one round-trip time. Consequently, the light field evolves from $\boldsymbol{a}^{(0)}$ to $\boldsymbol{A_1a}^{(0)}$, and so on. By sequentially applying appropriate modulation layers obtained by inverse design, the input light field is progressively transformed into an output field that approximates the desired target waveform.

trips (see Figure 1a for the illustration). During these Q_m round trips, the modulation parameters remain fixed, which relaxes the requirement on the modulation varying rate. By modulating the ring layers by layers, we finally get a $N\times M$ frequency-time vector-matrix multiplication as shown in Figure 1b. In each layer of modulations, the one-layer tranfer matrix $\boldsymbol{A_m}$ is given as $\boldsymbol{A_m}=(\boldsymbol{T_m})^{Q_m}$. After M layers of modulations, the light is transformed from its initial state $\boldsymbol{a}^{(0)}$ to $\boldsymbol{a}^{(M)}$:

$$\boldsymbol{a}^{(M)}=\boldsymbol{A_M A_{M-1}}\cdots\boldsymbol{A_1a}^{(0)}\equiv\boldsymbol{Aa}^{(0)}. \tag{3}$$

In this manner, we successfully construct a frequency-time vector-matrix multiplication within a single ring resonator, whose total transfer matrix $\boldsymbol{A}$ emulates (or generally speaking, proportional to) a target matrix.

To find the minimum required number of layers of modulations, we analyze from degrees of freedom. For each layer of modulations, if we fix the total round-trip number Q_m for simplicity, the intrinsic degrees of freedom of $\boldsymbol{A_m}$ has a total of $4(N-1)$ degrees of freedom, corresponding to the modulation intensities and phases of both PM and AM at frequencies $\Omega_R, 2\Omega_R, \ldots, (N-1)\Omega_R$. Thus the total degree of freedom for M layers of modulations is $4(N-1)M$. In order to have the total transfer matrix $\boldsymbol{A}$ to emulate arbitrary matrix transformation, one expects the minimum required number of degrees of freedom is $2N^2$, where the factor of 2 accounts for real and imaginary parts. Therefore, we desire the minimum number of modulation layers

$$M_{\min}=\left\lceil\frac{N^2}{2(N-1)}\right\rceil, \tag{4}$$

where $\lceil\rceil$ denotes the ceiling function. Once we have enough degrees of freedom for parameters, the next important question is whether the operation proposed in our time-varying modulated ring architecture can perform the arbitrary vector-matrix multiplication, and we here provide a proof to demonstrate its universality.

To discuss the universality of our system in a transparent way, we begin with the simplest case with $N=3$, $Q_m=1$, and $M=3$. For the m-th layer, the transfer matrix of PM or AM is a 3×3 complex matrix. We normalize the PM matrix elements by setting

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the diagonal elements as 1 for simplicity, and then write

$$\boldsymbol{T}_{\mathrm{PM},m}=\begin{pmatrix}1 & p_m & t_m\\ -p_m^* & 1 & p_m\\ -t_m^* & -p_m^* & 1\end{pmatrix},\qquad \boldsymbol{T}_{\mathrm{AM},m}=\begin{pmatrix}1 & r_m & s_m\\ r_m^* & 1 & r_m\\ s_m^* & r_m^* & 1\end{pmatrix},\qquad m=1,2,3. \tag{5}$$

Here $p_m, t_m, r_m, s_m \in \mathbb{C}$, each of which supplies two real degrees of freedom. The target matrix is 3×3, so there are 18 free parameters to be determined. We then assume that there is no AM operation in the third layer and hence $T_{\mathrm{AM},m=3}$ is the identity. Therefore, there are 20 real parameters in total in this case, which leads to the parameter vector

$$\theta=\left(\theta_{\mathrm{PM},1},\,\theta_{\mathrm{AM},1},\,\theta_{\mathrm{PM},2},\,\theta_{\mathrm{AM},2},\,\theta_{\mathrm{PM},3}\right)^{\top}\in\mathbb{R}^{20}, \tag{6}$$

$$\text{where}\quad \theta_{\mathrm{PM},m}=\left(\mathrm{Re}(p_m),\,\mathrm{Im}(p_m),\,\mathrm{Re}(t_m),\,\mathrm{Im}(t_m)\right),$$

$$\theta_{\mathrm{AM},m}=\left(\mathrm{Re}(r_m),\,\mathrm{Im}(r_m),\,\mathrm{Re}(s_m),\,\mathrm{Im}(s_m)\right),$$

and the total transfer matrix $\boldsymbol{A}$

$$\boldsymbol{A}(\theta)=\boldsymbol{T}_{\mathrm{PM},3}\,\boldsymbol{T}_{\mathrm{AM},2}\,\boldsymbol{T}_{\mathrm{PM},2}\,\boldsymbol{T}_{\mathrm{AM},1}\,\boldsymbol{T}_{\mathrm{PM},1},\qquad \boldsymbol{A}\in\mathbb{C}^{3\times 3}. \tag{7}$$

To examine whether all directions in the 18-dimensional real space associated with 3×3 complex matrices are accessible under the mapping $\theta\mapsto\boldsymbol{A}(\theta)$, we first evaluate the Jacobian matrix

$$J(\theta)=\left[\frac{\partial\,\mathrm{Re}(A_{ij})}{\partial\theta_k},\ \frac{\partial\,\mathrm{Im}(A_{ij})}{\partial\theta_k}\right]\in\mathbb{R}^{18\times 20}, \tag{8}$$

from simulations with extensive Monte–Carlo sampling (over 2×10^7 random parameter sets for θ). In all cases (see Note S4 for details), the Jacobian matrix exhibits full row rank $r=18$, which means the local surjectivity of the mapping is guaranteed. In other words, such evaluation indicates the existence of a local inverse mapping around every sample point [29], i.e., every small perturbation in the 18-dimensional image space around the target matrix generated by the sample parameters can be obtained by a suitable variation in the parameters.

Second, we further prove that the set of attainable directions in the 18-dimensional image space is dense from the mathematics perspective. In details, we carry out an inverse design experiment in simulations: we randomly sample 50 000 target matrices with different directions and, for each target, solve the inverse problem of finding parameters θ such that $\boldsymbol{A}(\theta)$ approximates the target matrix within accuracy as 10^{-2} (see Note S5 for details). In most cases, suitable parameters are successfully obtained, indicating that the reachable set is dense in all directions. Together with the local su rjectivity ensured by full row rank Jabian matrix (which is also checked for these 50 000 parameter sets), our simulation experiment proves that, in the simple case $N=3$, our architecture is universal, in the sense that any target transfer matrix can be obtained via this mapping from the parameter space.

For more general cases with larger N, the universality of our proposed architecture can be naturally generalized if all phase and amplitude modulations are nonzero and Q_m are not necessarily equal to 1, i.e., any target transfer matrix can still be represented by a suitable choice of parameters. However, it is important to note that in the proof above, we have not imposed any constraint on the parameter space, and have not considered the feasibility or efficiency of retrieving these inverse mappings in practice. In the following, we perform an inverse verification procedure to demonstrate its expressivity, i.e., one can obtain arbitrary target matrix.

2.2 | Inverse Verification of Expressivity

To validate our proposed architecture can reach arbitrary target matrices under practical constraints (i.e. the weak-modulation condition $\alpha_n < 0.7$), we perform an inverse verification procedure. Specifically, we randomly generate a complex target matrix $\boldsymbol{U}\in\mathbb{C}^{N\times N}$, and recover it in our architecture by optimizing modulation parameters following realistic experimental conditions [30]. The resulting total transfer matrix $\boldsymbol{A}$ is expected to approximate $\boldsymbol{U}$ up to a real scaling factor, i.e., $\boldsymbol{U}\approx\lambda\boldsymbol{A}$ for some $\lambda\in\mathbb{R}$.

We first introduce our learning process using our architecture with $N=5$. We fix the layers of modulation to $M=4$, which is the minimum number of layers required given Equation (4), and take $Q_m=32$. To calibrate the accuracy of the generated $\boldsymbol{A}$ with $\boldsymbol{U}$, we define the loss function:

$$\mathcal{L}=\left\|\frac{\boldsymbol{U}}{\|\boldsymbol{U}\|_{\max}}-\frac{\boldsymbol{A}}{\|\boldsymbol{A}\|_{\max}}\right\|_F^2. \tag{9}$$

Here, $\|\cdot\|_F$ denotes the Frobenius norm, defined as $\|\boldsymbol{V}\|_F=\sqrt{\sum_{i,j}|V_{ij}|^2}$, which measures the overall element-wise error. $\|\cdot\|_{\max}$ labels the largest modulus from all matrix elements. Different from deterministic frameworks [31], we retrieve the modulation parameters by formulating the target-matrix implementation as an optimization-based inverse design problem [32] over the modulation parameter space. Specifically, in the learning process with finding modulation parameters to minimize $\mathcal{L}$, we adopt gradient descent strategy [33] with a batch including a large number of sets in the modulation parameter space.

In simulations, we randomly initialize 65 536 sets of modulation parameters and iteratively perform matrix calculations with Equation (3) and optimize parameters with Equation (9) using the gradient over this batch. For a typical batch, Figure 2a shows the distribution of $\mathcal{L}$ vs. iteration number up to 50 000. The figure is presented in a heatmap format by plotting the number of parameter sets whose $\mathcal{L}$ falls within the small-value regime at a particular iteration number. For each set in this batch, $\mathcal{L}$ experiences an initial exponential decay, followed by reaching convergence. For most sets, $\mathcal{L}$ quickly converges to a relatively high value, while a small portion of sets reach much lower values of $\mathcal{L}$. In all these 65 536 sets of calculations, the best one gives $\mathcal{L}$ as 1.42×10^{-12}.

We further apply 300 batches, where each batch includes a newly generated random target matrix $\boldsymbol{U}$ and the corresponding learning process to obtain a total transfer matrix $\boldsymbol{A}$. To evaluate the expressive capability of our architecture, we introduce two metrics to quantify their consistency between $\boldsymbol{U}$ and $\boldsymbol{A}$: the root

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

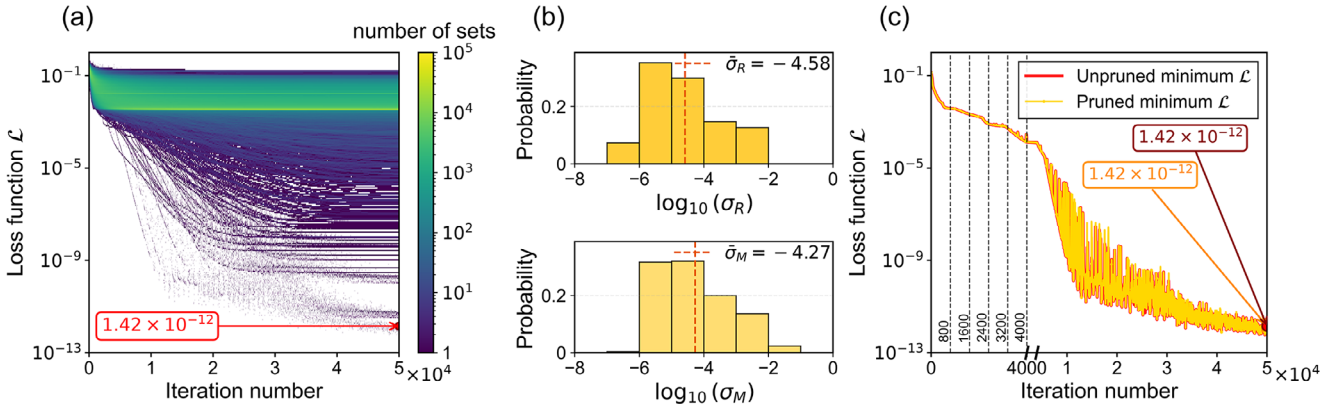


FIGURE 2 | Inverse verification of expressivity in our architecture and learning behavior with pruning strategy. (a) 65 536 randomly initialized modulation parameter sets are optimized using gradient-based algorithms. The color bar denotes the number of parameter sets. The loss $\mathcal{L}$ exhibits a near-exponential decline followed by gradual saturation, reaching a final minimum of 1.42×10^{-12}. (b) 300 arbitrary complex target matrices are generated and optimized through the proposed learning process. The distributions of σ_R and σ_M are statistically analyzed, with red dashed lines indicating their logarithmic mean values $\bar{\sigma}_R$ and $\bar{\sigma}_M$. (c) Every 800 iterations (vertical dashed lines), half of the poorly performing sets are pruned until the batch size reaches 2048. To improve readability, the horizontal axis is plotted on a nonlinear scale. The thick red curve represents the lowest $\mathcal{L}$ without pruning, while the thin yellow curves show those with pruning. The orange–yellow numbers indicate the final results with pruning, while the dark-red numbers indicate the final results without pruning. Both two curves achieve the same minimum of 1.42×10^{-12} and nearly overlap throughout the learning process, demonstrating that pruning maintains convergence performance. Modulation parameters for (a) are provided in Note S6.

mean square error (RMSE) σ_R and the maximum element error (MEE) σ_M. In Figure 2b, we plot distributions of σ_R and σ_M using 300 batches of learning results, respectively. One sees that among all results, 72.67% σ_R is smaller than 10^{-4}, while 64% σ_M is smaller than 10^{-4}, indicating our learning process can output a matrix $\boldsymbol{A}$ very close to the target one $\boldsymbol{U}$. The logarithmic mean values of RMSE and MEE are $\bar{\sigma}_R = -4.58$ and $\bar{\sigma}_M = -4.27$, respectively. Here, the logarithmic mean value is defined as $\bar{\sigma} = \frac{1}{N_s}\sum_{j=1}^{N_s} \log_{10}(\sigma_j)$, where N_s denotes the total number of batches. Even in worst case, σ_R gives a value of 7.98×10^{-3} with σ_M being 1.27×10^{-2}. We also perform eigenvalue spectrum analyses of the discrepancy between $\boldsymbol{U}$ and $\boldsymbol{A}$, which gives the reasonable consistence with the results from RMSE and MEE (see Note S7 for detailed analysis). These results confirm that our architecture can perform high-quality arbitrary vector-matrix multiplication functioning as the target matrix.

In order to reduce computational cost in the later simulations, we will apply a pruning strategy [34]. Pruning strategies have been proposed in optical systems for improving hardware performance [35, 36]. In our work, we employ pruning strategy in a different role, namely in the inverse optimization to progressively disregard lower-performing parameter sets, thereby improving the search efficiency in the modulation parameter space. As illustrations, for the same batch of Figure 2a, we disregard the lower-performing half of sets every 800 iterations until the set number is reduced to 2,048. In Figure 2c, we compare the lowest values of $\mathcal{L}$ at each iteration numbers for the unpruned calculation and the calculation using the pruning strategy. For this batch, both calculations give approximately good optimization, highlighting the efficiency for using the pruning strategy in our later calculations. However, we clarify that the performance of such pruning strategy is not always as good as the unpruned calculation, but is acceptable considering its reduction of computational cost (see the Note S8 for details).

2.3 | The Linear Relation Between Evolution Time and Matrix Size N

In this time-varying single ring architecture, we can implement arbitrary vector-matrix multiplication with any matrix size N in principle. It is therefore important to understand the relationship between the time of light evolving through this architecture and N. For vector-matrix multiplication, it has a computational complexity of $O(N^2)$ [37], and for a given spatial cost of $O(1)$, the total processing time typically scales as $O(N^2)$. In our architecture, however, each round-trip time for light field is $T = 2\pi/\Omega_R$. The total evolution time is thus determined by the cumulative numbers of round-trips across all modulation layers (also see in Figure 1b), i.e.,

$$t_{\text{tot}} = \left(\sum_{m=1}^{M} Q_m\right) T. \tag{10}$$

which is simplified to $t_{\text{tot}} = M \times Q \times T$, if Q_m is chosen to be equal to Q. To achieve an arbitrary operation of a $N \times N$ complex target matrix, the minimal number of modulation layers required to ensure full expressivity is given by $M_{\min}$ in Equation (4), scaling asymptotically as $O(N)$. Consequently, the total evolution time in our architecture—proportional to the number of modulation layers—also scales linearly with N, i.e., $t_{\text{tot}} \sim O(N)$.

To validate such linear scaling dependence of the evolution time in our architecture, we perform several batches of simulations with N chosen from 6 to 10. In each batch, we fix $Q_m = 16$ and set a loose value of $M = N$. Such choices of parameters make the total round-trip number increase linearly with N. For each N, we perform a series of batches, where each batch involves a learning process using the pruning strategy for a randomly generated target matrix, and then perform statistical analysis on the RMSE and MEE distributions between $\boldsymbol{U}$ and $\boldsymbol{A}$ in each batch.

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

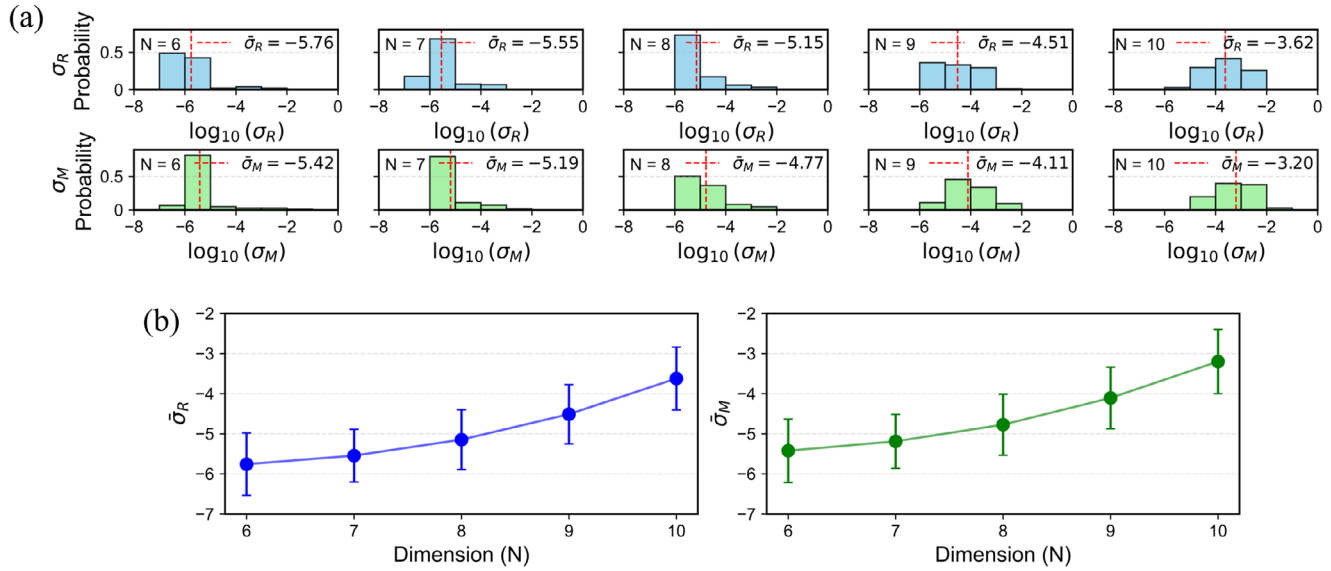


FIGURE 3 | Linearity test of modulation layer number M with respect to matrix dimension N. (a) Statistical distributions of σ_R and σ_M for matrix sizes $N = 6$–**10** under the condition $M = N$ and $Q = 16$. All cases use the pruning strategy during the learning process. Red dashed lines indicate the logarithmic mean values $\bar{\sigma}_R$ and $\bar{\sigma}_M$. Batch numbers are 194 ($N = 6$), 176 ($N = 7$), 155 ($N = 8$), 136 ($N = 9$), and 102 ($N = 10$). (b) The logarithmic means $\bar{\sigma}_R$ and $\bar{\sigma}_M$ are plotted as functions of N, with line segments connecting data points to indicate the trend. Error bars denote the standard deviations. Although a slight performance degradation appears with increasing N, all results remain within acceptable accuracy.

We plot statistical results in Figure 3a. First of all, one can see all σ_R is smaller than 10^{-2}. For results with $N = 6 - 9$, more than 90% σ_R and σ_M is smaller than 10^{-3}, respectively. Notably, for $N = 10$, the portion of σ_R and σ_M smaller than 10^{-3} reaches 74.26% and 59.41%, respectively. Moreover, there are only three batches having the resulting σ_M larger than 10^{-2}, but they give the maximum value being 0.0243 (eigenvalue spectrum analyses see Note S7 for detail). These results demonstrate the high-precision scalability of our architecture and its robustness in representing arbitrary matrices across increasing matrix dimension N.

In Figure 3b, we further plot the logarithmic mean values $\bar{\sigma}_R$ and $\bar{\sigma}_M$ as functions of N. One can see that both $\bar{\sigma}_R$ and $\bar{\sigma}_M$ gradually become larger with N, due to the increasing N and parameter space. Nevertheless, all $\bar{\sigma}_R$ and $\bar{\sigma}_M$ is smaller than -3.62 and -3.20. Moreover, the standard deviations in Figure 3b are smaller than 0.8 in all cases, indicating the applied pruning strategy can yield results with high accuracy. From all batches of simulations performed in Figure 3, we prove that our architecture exhibits reliable linear scalability while maintaining high computational accuracy.

2.4 | No Intrinsically Hard-to-Learn Matrices

After seeing that our proposed architecture can perform arbitrary vector-matrix multiplication in a single ring, a natural question remaining to be answered is whether there exists any target matrix intrinsically hard to learn, i.e., inherently untrainable target matrix that limits the efficiency of our architecture.

To answer this question, we take a 6×6 target matrix $\boldsymbol{U_w}$ that gives one of the worst $\sigma_R = 0.00769093$ from previous simulations as an example (see Note S6 for the detailed matrix $\boldsymbol{U_w}$). In Figure 4a, we plot the distribution of the absolute error $\boldsymbol{\sigma}$ between elements in $\boldsymbol{U_w}$ and $\boldsymbol{A}$, where $\boldsymbol{A}$ is obtained using the learning process with pruning strategy. In this particular example, the $\sigma_M = 0.0168$, which hints the inefficient computing by applying our architecture. However, here we use a re-study approach to find more appropriate modulation parameters that can make our optical computing more accuracy to reach the target matrix $\boldsymbol{U_w}$.

Here we briefly introduce this re-study approach. We apply a 5% random perturbation to each modulation parameter obtained from the previous learning in Figure 4a, and generate 65 536 random initial parameter sets. All these sets are further studied by using the learning process and the best set with the smallest $\sigma_R = 0.00000190$ is obtained using the same pruning strategy. Figure 4b plots $\boldsymbol{\sigma}$, which shows a better performance with 3 orders of magnitude decreasing in errors compared to Figure 4a. This approach takes advantage of previous learning results, preventing the learning process trapped in a local minima of the loss function $\mathcal{L}$.

Therefore, there is no intrinsically hard-to-learn target matrix once we apply the re-study approach on those cases with worse performance to obtain better modulation parameters, giving a proof of the strong expressivity of our architecture.

3 | Discussions and Conclusion

We discuss a few experimental aspects for on-chip implementation of our architecture. A practical constraint on the accessible matrix size is affected by the cavity lifetime, which is determined by the loaded quality factor $\boldsymbol{Q_L}$. In general, $\boldsymbol{Q_L^{-1}} = \boldsymbol{Q_i^{-1}} + \boldsymbol{Q_c^{-1}}$, where $\boldsymbol{Q_i}$ and $\boldsymbol{Q_c}$ are the intrinsic and coupling quality factors, respectively. Under the critical-coupling condition, $\boldsymbol{Q_c} = \boldsymbol{Q_i}$. Using a state-of-the-art thin-film lithium niobate microring with intrinsic quality factor $\boldsymbol{Q_i} \sim 10^7$ [38], one obtains $\boldsymbol{Q_L} \sim 10^7$ near

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

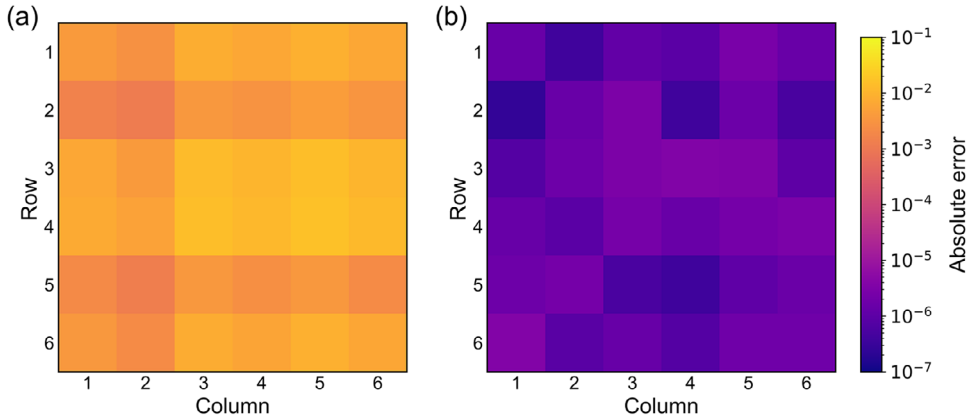


FIGURE 4 | Learning performance before and after the re-study approach. (a) Distribution of the absolute error $\boldsymbol{\sigma}$ between $\boldsymbol{U_w}$ and $\boldsymbol{A}$ obtained from the previous learning process using the pruning strategy. The result yields $\sigma_R = 7.69 \times 10^{-3}$ and $\sigma_M = 1.68 \times 10^{-2}$, indicating limited computational accuracy in this case. (b) Results obtained with the re-study approach, in which each modulation parameter is first perturbed by 5% and then re-optimized from 65,536 random initializations using the same pruning strategy. One can see that figure (a) and (b) are localized in different ranges of absolute error. In particular, optimized result achieves $\sigma_R = 1.90 \times 10^{-6}$ in (b), representing a 99.975% reduction in RMSE and a three-order-of-magnitude improvement in accuracy compared to (a). Modulation parameters for (a) and (b) are provided in Note S6.

the critical coupling. The corresponding photon lifetime is $\tau_{\text{ph}} = \frac{Q_L}{\omega_0} \sim 10$ ns, for $\lambda = 1550$ nm. We further consider a modulation frequency on resonance with the FSR of a thin-film lithium niobate electro–optic microring at $\Omega_R \sim 10$ GHz [38]. Assuming a classical input power of $P_0 = 1$ mW, and a detectable signal level of $P_{\text{det}} = 1$ nW, one can estimate the maximum accessible round-trip number $Q_{\max}$ as $Q_{\max} = \frac{Q_L \lambda}{2\pi c} \Omega_R \ln\left(\frac{P_0}{P_{\text{det}}}\right) \sim 10^3$, which could support vector-matrix multiplication with the matrix size $N \sim 100$ in our architecture. By further increasing the quality factor of the ring, the modulation frequency, and the signal-to-noise ratio of the detector, one may prompt the improvement of the accessible matrix size and hence reach to larger-scale vector-matrix multiplication in the future implementations.

We propose an architecture that performs inverse-designed arbitrary $O(N^2)$ vector-matrix multiplication. This architecture is based on a single ring as $O(1)$ spatial complexity, and supports a time-varying synthetic frequency lattice that makes the multiplication operation within $O(N)$ evolution time. The principle of our work aligns closely with the previous study based on the waveguide-based design of coupled ring resonators [20]. However, different from it and other implementations based on ring arrays [21, 39, 40], ours realizes the matrix operation in a single time-varying ring resonator using a single ring that occupies the fixed spatial area. This design then features flexible scalability in footprint, reduced device count as well as simplified device calibration, and is therefore suitable for on-chip implementation [39–43], without scaling of the spatial footprint with matrix size N. Moreover, compared with the optical neural network using time-multiplexed networks [44], our proposal here leverages frequency dimension of freedom and offers stronger modal orthogonality and thus alleviates stringent requirements on temporal synchronization, making it more stable and feasible for large-scale expansion once this approach is implemented in on-chip device. Very recent, a microwave experiment exhibits the possibility of the matrix multiplication using several frequency components [45], and our work extends this concept to show a time-varying architecture can be harnessed as a programmable and scalable engine for large-scale vector-matrix operations, which reinforces its feasibility as a practical device for optical computing. Our proposal can be potentially realized across different experimental platforms from fiber loops to integrated photonics [39–43, 46–48], both of which have many experimental progresses supporting necessary state-of-art technologies for the future experimental demonstration. In particular, recent efforts have been greatly applied on lithium niobate on insulator (LNOI) that forms the chip-based ring resonators under modulations [41, 43, 46–48], so our work can be a timely proposal for achieving optical neural networks in these LNOI chips. We also note a recent work on proposing high-dimensional topological photonics in a single ring resonator [49], which relies on mode-selective modulations and mode conversions. In our architecture, the programmability is achieved through universal time-varying phase and amplitude modulations over all frequency modes, so such additional mode-dependent degrees of freedom may also be incorporated for higher-dimensional matrix computations in the future.

In summary, we propose that arbitrary vector-matrix multiplication can be realized in a time-varying synthetic frequency lattice on a single ring, achieving $O(1)$ spatial complexity and $O(N)$ evolution time with excellent precision, universality, programmability, and scalability. We confirm the universality and scalability of our system, and show that there exhibits no intrinsically hard-to-learn matrices in our proposed architecture. Our work hence provides an efficient solution to its spatial counterpart implementations, substantially reducing the spatial footprint while preserving scalability and precision, and offering a compact and programmable route toward large-scale integrated photonic computing.

Acknowledgements

The research was supported by National Key R&D Program of China (No. 2023YFA1407200), National Natural Science Foundation of China

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(12122407, 12192252 and 12574417) and Natural Science Foundation of Shanghai (25ZR1402221).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. P. L. McMahon, "The Physics of Optical Computing," *Nature Reviews Physics* 5, no. 12 (2023): 717–734.

2. X. Lin, Y. Rivenson, N. T. Yardimci, et al., "All-Optical Machine Learning Using Diffractive Deep Neural Networks," *Science* 361, no. 6406 (2018): 1004–1008.

3. J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. Pernice, "All-Optical Spiking Neurosynaptic Networks with Self-Learning Capabilities," *Nature* 569, no. 7755 (2019): 208–214.

4. F. Ashtiani, A. J. Geers, and F. Aflatouni, "An On-Chip Photonic Deep Neural Network for Image Classification," *Nature* 606, no. 7914 (2022): 501–506.

5. S. Pai, Z. Sun, T. W. Hughes, et al., "Experimentally Realized In Situ Backpropagation for Deep Learning in Photonic Neural Networks," *Science* 380, no. 6643 (2023): 398–404.

6. T. Zhou, X. Lin, J. Wu, et al., "Large-Scale Neuromorphic Optoelectronic Computing with a Reconfigurable Diffractive Processing Unit," *Nature Photonics* 15, no. 5 (2021): 367–373.

7. J. Li, D. Mengu, N. T. Yardimci, et al., "Spectrally Encoded Single-Pixel Machine Vision Using Diffractive Networks," *Science Advances* 7, no. 13 (2021): eabd7690.

8. J. Hu, D. Mengu, D. C. Tzarouchis, B. Edwards, N. Engheta, and A. Ozcan, "Diffractive Optical Computing in Free Space," *Nature Communications* 15, no. 1 (2024): 1525.

9. H. Zhu, J. Zou, H. Zhang, et al., "Space-Efficient Optical Computing with an Integrated Chip Diffractive Neural Network," *Nature Communications* 13, no. 1 (2022): 1044.

10. L. Fan, K. Wang, H. Wang, A. Dutt, and S. Fan, "Experimental Realization of Convolution Processing in Photonic Synthetic Frequency Dimensions," *Science Advances* 9, no. 32 (2023): eadi4956.

11. D. A. Miller, C. Roques-Carmes, C. G. Valdez, et al., "Universal Programmable and Self-Configuring Optical Filter," *Optica* 12, no. 9 (2025): 1417–1426.

12. W. Bogaerts, D. Pérez, J. Capmany, et al., "Programmable Photonic Circuits," *Nature* 586, no. 7828 (2020): 207–216.

13. Y. Shen, N. C. Harris, S. Skirlo, et al., "Deep Learning with Coherent Nanophotonic Circuits," *Nature Photonics* 11, no. 7 (2017): 441–446.

14. M. Moralis-Pegios, G. Giamougiannis, A. Tsakyridis, D. Lazovsky, and N. Pleros, "Perfect Linear Optics Using Silicon Photonics," *Nature Communications* 15, no. 1 (2024): 5468.

15. M. Camacho, B. Edwards, and N. Engheta, "A Single Inverse-Designed Photonic Structure That Performs Parallel Computing," *Nature Communications* 12, no. 1 (2021): 1466.

16. V. Nikkhah, A. Pirmoradi, F. Ashtiani, B. Edwards, F. Aflatouni, and N. Engheta, "Inverse-Designed Low-Index-Contrast Structures on a Silicon Photonics Platform for Vector–Matrix Multiplication," *Nature Photonics* 18, no. 5 (2024): 501–508.

17. W. Zhou, B. Dong, N. Farmakidis, et al., "In-Memory Photonic Dot-Product Engine with Electrically Programmable Weight Banks," *Nature Communications* 14, no. 1 (2023): 2887.

18. L. Yuan, Q. Lin, M. Xiao, and S. Fan, "Synthetic Dimension in Photonics," *Optica* 5, no. 11 (2018): 1396–1405.

19. D. Yu, W. Song, L. Wang, et al., "Comprehensive Review on Developments of Synthetic Dimensions," *Photonics Insights* 4, no. 2 (2025): R06–R06.

20. S. Buddhiraju, A. Dutt, M. Minkov, I. A. Williamson, and S. Fan, "Arbitrary Linear Transformations for Photons in the Frequency Synthetic Dimension," *Nature Communications* 12, no. 1 (2021): 2401.

21. X. Piao, S. Yu, and N. Park, "Programmable Photonic Time Circuits for Highly Scalable Universal Unitaries," *Physical Review Letters* 132, no. 10 (2024): 103801.

22. X. Fang, X. Hu, B. Li, et al., "Orbital Angular Momentum-Mediated Machine Learning for High-Accuracy Mode-Feature Encoding," *Light: Science & Applications* 13, no. 1 (2024): 49.

23. K. Kim, K. Park, H. Park, S. Yu, N. Park, and X. Piao, "Programmable Photonic Unitary Circuits for Light Computing," *Nanophotonics* 14, no. 10 (2025): 1429–1449.

24. A. V. Pankov, I. D. Vatnik, and A. A. Sukhorukov, "Optical Neural Network Based on Synthetic Nonlinear Photonic Lattices," *Physical Review Applied* 17, no. 2 (2022): 024011.

25. B. Peng, S. Yan, D. Cheng, et al., "Optical Neural Network Architecture for Deep Learning with Temporal Synthetic Dimension," *Chinese Physics Letters* 40, no. 3 (2023): 034201.

26. P. Del'Haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, and T. J. Kippenberg, "Optical Frequency Comb Generation from a Monolithic Microresonator," *Nature* 450, no. 7173 (2007): 1214–1217.

27. A. Dutt, L. Yuan, K. Y. Yang, et al., "Creating Boundaries Along a Synthetic Frequency Dimension," *Nature Communications* 13, no. 1 (2022): 3377.

28. Y. Hu, M. Yu, N. Sinclair, et al., "Mirror-Induced Reflection in the Frequency Domain," *Nature Communications* 13, no. 1 (2022): 6293.

29. A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings* (Springer, 2009), vol. 543, see Theorem 1F.6 (Inverse Selections When $m \leq n$).

30. K. Wang, A. Dutt, K. Y. Yang, C. C. Wojcik, J. Vučković, and S. Fan, "Generating Arbitrary Topological Windings of a Non-Hermitian Band," *Science* 371, no. 6535 (2021): 1240–1245.

31. M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental Realization of Any Discrete Unitary Operator," *Physical Review Letters* 73, no. 1 (1994): 58–61.

32. M. Markowitz, K. Zelaya, and M. A. Miri, "Learning Arbitrary Complex Matrices by Interlacing Amplitude and Phase Masks with Fixed Unitary Operations," *Physical Review A* 110, no. 3 (2024): 033501.

33. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, 2006).

34. S. Han, J. Pool, J. Tran, and W. Dally, "Learning Both Weights and Connections for Efficient Neural Networks," *Advances in Neural Information Processing Systems* 28 (2015): 1135–1143.

35. S. Yu and N. Park, "Heavy Tails and Pruning in Programmable Photonic Circuits for Universal Unitaries," *Nature Communications* 14, no. 1 (2023): 1853.

36. T. Xu, W. Zhang, J. Zhang, et al., "Control-Free and Efficient Integrated Photonic Neural Networks via Hardware-Aware Training and Pruning," *Optica* 11, no. 8 (2024): 1039–1049.

37. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, 2013).

38. X. Zhu, Y. Hu, S. Lu, et al., "Twenty-Nine Million Intrinsic Q-Factor Monolithic Microresonators on Thin-Film Lithium Niobate," *Photonics Research* 12, no. 8 (2024): A63–A68.

39. S. Ohno, R. Tang, K. Toprasertpong, S. Takagi, and M. Takenaka, "Silicon Microring Resonator Crossbar Array for On-Chip Inference and

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Training of the Optical Neural Network," *ACS Photonics* 9, no. 8 (2022): 2614–2622.

40. R. Tang, S. Ohno, K. Tanizawa, et al., "Symmetric Silicon Microring Resonator Optical Crossbar Array for Accelerated Inference and Training in Deep Learning," *Photonics Research* 12, no. 8 (2024): 1681–1688.

41. B. Desiatov, A. Shams-Ansari, M. Zhang, C. Wang, and M. Lončar, "Ultra-Low-Loss Integrated Visible Photonics Using Thin-Film Lithium Niobate," *Optica* 6, no. 3 (2019): 380–384.

42. A. Balčytis, T. Ozawa, Y. Ota, S. Iwamoto, J. Maeda, and T. Baba, "Synthetic Dimension Band Structures on a Silicon Complementary Metal-Oxide-Semiconductor Photonic Platform," *Science Advances* 8, no. 4 (2022): eabk0468.

43. H. X. Dinh, A. Balčytis, T. Ozawa, et al., "Reconfigurable Synthetic Dimension Frequency Lattices in an Integrated Lithium Niobate Ring Cavity," *Communications Physics* 7, no. 1 (2024): 185.

44. B. Wu, Y. Ren, and R. Zhao, "A Scalable and Programmable Optical Neural Network in a Time-Synthetic Dimension," *arXiv Preprint arXiv:2507.02297* (2025), https://arxiv.org/abs/2507.02297.

45. Y. Guan and J. Yao, "A Microwave Photonic Neural Network in the Frequency Synthetic Dimension Using Multi-Tone Single-Sideband Modulation in a Fiber Loop," *Journal of Lightwave Technology* 43 (2025): 9934–9940.

46. R. Ye, G. Li, S. Wan, et al., "Construction of Various Time-Varying Hamiltonians on Thin-Film Lithium Niobate Chip," *Physical Review Letters* 134, no. 16 (2025): 163802.

47. W. Liu, X. Su, C. Li, et al., "Reconfigurable Chiral Edge States in Synthetic Dimensions on an Integrated Photonic Chip," *Physical Review Letters* 134, no. 14 (2025): 143801.

48. Z. A. Wang, X. D. Zeng, Y. T. Wang, et al., "Versatile Photonic Frequency Synthetic Dimensions Using a Single Programmable On-Chip Device," *Nature Communications* 16, no. 1 (2025): 7780.

49. D. Cheng, H. Wang, C. Roques-Carmes, J. Zhong, and S. Fan, "Creating High-Dimensional Topological Physics Using a Single Ring Resonator," *Newton* 1, no. 7 (2025): 100163.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting File: lpor71242-sup-0001-SuppMat.pdf.

18638899, 2026, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.71242 by Shanghai Jiaotong University, Wiley Online Library on [17/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License