



APPLIED PHYSICS

Experimental realization of convolution processing in photonic synthetic frequency dimensions

Lingling Fan¹, Kai Wang^{1,2}, Heming Wang¹, Avik Dutt³, Shanhui Fan^{1*}

Convolution is an essential operation in signal and image processing and consumes most of the computing power in convolutional neural networks. Photonic convolution has the promise of addressing computational bottlenecks and outperforming electronic implementations. Performing photonic convolution in the synthetic frequency dimension, which harnesses the dynamics of light in the spectral degrees of freedom for photons, can lead to highly compact devices. Here, we experimentally realize convolution operations in the synthetic frequency dimension. Using a modulated ring resonator, we synthesize arbitrary convolution kernels using a pre-determined modulation waveform with high accuracy. We demonstrate the convolution computation between input frequency combs and synthesized kernels. We also introduce the idea of an additive offset to broaden the kinds of kernels that can be implemented experimentally when the modulation strength is limited. Our work demonstrates the use of synthetic frequency dimension to efficiently encode data and implement computation tasks, leading to a compact and scalable photonic computation architecture.

INTRODUCTION

Neural networks (1) have been ubiquitously used in machine learning tasks such as computer vision, speech, audio, and language comprehension. Among these networks, convolutional neural networks (CNNs) (2) play a critical role in recognizing features embedded in complex input data. With this capacity of feature extraction, CNNs achieve superior accuracy in predicting unseen data, with a reduced number of parameters compared with dense neural networks (3).

Convolution, as a central operation for spatiotemporal perception in CNNs (4), is particularly energy- and memory-intensive using conventional electronic architecture that is limited by the data movement bottleneck (5). As a promising substitute, optical neural networks (ONNs) (6) process information by propagating a light signal through an optical structure (7, 8). ONNs have the potential for improved computing performance, with parallel input processing (9), high computing speed (10–12), broad information bandwidth (13), and low energy consumption (14). ONNs have been implemented in a variety of schemes ranging from Mach-Zehnder interferometers for matrix-vector multiplication (15–17) to microring resonators for reservoir computing (18, 19). In addition, diffractive layers (20, 21) and scattering media (22) are used for image and vowel detection. However, most of the ONNs are limited to using only the spatial degrees of freedom of photons (23). Frequency degree of freedom is seldom used for kernel formation (24), and most of the previous works using frequency degree of freedom of photons for computing did so without mixing the frequencies as light propagates through the device (11, 12). With the need to scale up ONNs to meet the demands of various applications, it is desirable to use other degrees of freedom of photons, e.g., frequency, to further enhance the scalability of ONNs.

In this work, we experimentally demonstrate the use of a synthetic frequency dimension (25, 26) as formed by a dynamically modulated ring resonator to enable convolution operations. Specifically, we synthesize a wide range of convolution kernels with analytically predetermined modulation waveforms. We achieve various intended convolution kernels with good agreement with theory. We demonstrate the convolution computation by generating different frequency-mode inputs. The output frequency comb obtained from the ring agrees well with the target output as processed by convolution. We also introduce a pathway to broaden the kinds of kernels that can be implemented experimentally when the modulation strength is limited.

The concept of synthetic frequency dimension has been previously used to demonstrate topological physics (27–30) and matrix-vector multiplication (24, 31). However, the use of synthetic frequency dimension for convolution (32) has rarely been demonstrated experimentally. While convolution can be viewed as a special case of matrix-vector multiplication, the matrix that corresponds to a convolution operation has a translational symmetry. Exploiting this translational symmetry enables a simpler implementation as compared with the implementation of a general matrix-vector multiplication operation. Frequency combs have been previously used for optical convolution purposes (10–12). Previous studies (10–12), however, do not use the dynamics of light along the frequency dimension, i.e., these works do not use the possibility of frequency mixing and conversion as offered by a dynamically modulated system, which is at the heart of the concept of synthetic frequency dimension. Our work introduces a new physics mechanism for achieving optical convolution and is important for the quest to achieve large-scale parallel optical computation with compact devices.

RESULTS

Modulation waveform design

The experimental setup in this work consists of an optical ring resonator undergoing electro-optical modulation as schematically

¹Department of Electrical Engineering, Ginzton Laboratory, Stanford University, Stanford, CA 94305, USA. ²Department of Physics, McGill University, 3600 Rue University, Montreal, Quebec H3A 2T8, Canada. ³Department of Mechanical Engineering and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA.

*Corresponding author. Email: shanhui@stanford.edu.

shown in Fig. 1A. Assuming that the waveguide forming the ring resonator as well as all other waveguides that provide input and output coupling to the ring all support a single mode and the group velocity dispersion is negligible, $\Omega_R = 2\pi c/n_g \ell$ corresponds to the free spectral range (FSR) of the ring resonator. Here, c , n_g , and ℓ represent light speed in vacuum, group refractive index, and ring circumference, respectively. $t_R = 2\pi/\Omega_R$ denotes the round-trip time of the ring. Specifically, here, we consider the case that the modulator exclusively modulates the amplitude of light, which can be described by a temporal transmission factor (30)

$$T_{AM}(t) = \exp \left\{ \left[\sum_{m \geq 1} B_m \sin(m\Omega_R t + \beta_m) - \gamma t_R \right] \right\} \quad (1)$$

B_m and β_m correspond to the magnitude and phase angle of the waveforms in the amplitude modulators for the m th-order resonant modulation component, respectively. γ corresponds to time-averaged loss as induced by the amplitude modulator. In using Eq. 1 to describe a passive amplitude modulator that has no gain, γ is positive and needs to be sufficiently large so that $T_{AM}(t) < 1$ for all t . The ring resonator is coupled to an input and an output waveguide. Since $T(t) = T(t + t_R)$, the frequency components of the modulation waveform are located at integer multiples of the FSR of the ring resonator. Therefore, with modulations, the resonant modes of the ring at different frequencies can resonantly couple with each other.

In Fig. 1A, there is an input waveguide that couples to the ring with a coupling coefficient γ_{e1} , as well as a drop-port waveguide that

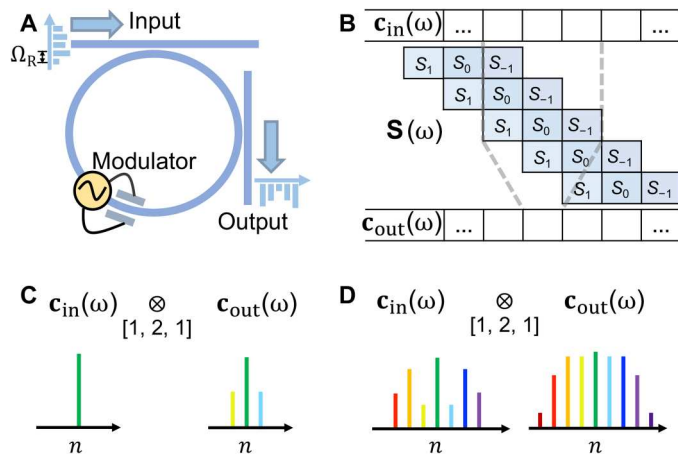


Fig. 1. Schematic illustration of the convolution experiment. (A) The experimental setup, where the convolution operation is performed by a ring resonator modulated by an electro-optical modulator. The modulation has its frequency components located at the free-spectral range Ω_R of the ring and its integer multiples. An input optical frequency comb is injected into the modulated ring resonator. The output frequency comb is detected at the drop-port optical waveguide. (B) A translationally symmetric scattering matrix \mathbf{S} transforms the input \mathbf{c}_{in} to the output \mathbf{c}_{out} . This transformation is equivalent to a 1D convolution operation with a kernel. Here, we show a three-element kernel $[s_{-1}, s_0, s_{+1}]$ for illustration purposes. (C) A CW laser is injected into a modulated ring resonator implementing a smoothing kernel $[1, 2, 1]$. The output frequency comb manifests the kernel shape. (D) A multifrequency comb is injected into the same modulated ring resonator as in (C) that implements the same smoothing kernel $[1, 2, 1]$. The output frequency comb is smoother as compared with the input. In (C) and (D), the height of a comb line represents the electric field amplitude at the corresponding frequency site.

couples to the ring with a coupling coefficient γ_{e2} . We inject a frequency comb in the input waveguide to generate an output frequency comb in the drop waveguide. The modulation waveform as described above can be used to implement a convolution kernel in the frequency dimension. The ring resonator supports $2N + 1$ equally spaced resonant modes with frequencies $\omega_n = \omega_0 + n\Omega_R$ ($-N \leq n \leq N$), with ω_0 corresponding to the central resonant frequency. We assume an input wave with a form $c_{in} = \sum_n c_{in,n} \exp(j\omega_n t + j\Delta\omega t)$, with $\Delta\omega$ being the detuning. The wave inside the modulated ring then has the form $a(t) = \sum_n a_n(t) \exp(j\omega_n t + j\Delta\omega t)$. The modal amplitudes a_n can

be determined by the temporal coupled-mode theory (32). Defining the input and output wave amplitude vectors $\mathbf{c}_{in} = [\dots, c_{in,m}, \dots]$ and $\mathbf{c}_{out} = [\dots, c_{out,m}, \dots]$, we obtain the scattering matrix that connects \mathbf{c}_{in} and \mathbf{c}_{out} by $\mathbf{c}_{out} = \mathbf{S}\mathbf{c}_{in}$ and \mathbf{S} is given by

$$\mathbf{S} = j2\gamma_{e1} \gamma_{e2} \{ \mathbf{K} + [j(\gamma + \gamma_{cst}) - \Delta\omega] \mathbf{I} \}^{-1} \quad (2)$$

Here, γ_{cst} is the total rate of loss in the resonator from mechanisms other than the amplitude modulator. These mechanisms can include, for example, the propagation loss of light in the fiber, as well as input and output coupling, as characterized by the input and output coupling rate of γ_{e1} and γ_{e2} , respectively. Here, we assume that such a loss rate is the same for every resonant mode in the system. This assumption of uniform loss rates holds within a spectral window of a few nanometers, as the variations of coupling ratios for the coupler and the gain profile of amplifiers could be negligible. \mathbf{I} is an identity matrix. The matrix elements of \mathbf{K} satisfy the translational symmetry, i.e., $K_{m,n} = \kappa_{m-n}$, where m and n are the indices of the modes. $\kappa_{\pm p}$, with $p > 0$, is the coupling constant between two modes m and n satisfying $|m - n| = p$ and is related to the modulation parameters by $\kappa_{\pm p} = \pm \frac{1}{2t_R} B_p e^{\pm j\beta_p}$. To simplify the representation, we denote $\kappa_0 = j(\gamma + \gamma_{cst}) - \Delta\omega$ to combine the loss and detuning factors into \mathbf{K} matrix. \mathbf{S} , consequently, is a matrix with elements $S_{m,n} = s_{m-n}$ so it has a translational symmetry along the frequency axis (32).

Because of the translational symmetry along the frequency axis, the scattering matrix in Eq. 2 implements a one-dimensional (1D) convolution operation

$$\begin{aligned} c_{out,m} &= \sum_n S_{m,n} c_{in,n} = \sum_n s_{m-n} c_{in,n} \\ &= \sum_n s_n c_{in,m-n} \end{aligned} \quad (3)$$

Here, s_n is the n th element of a kernel for the convolution operation (33). Figure 1B illustrates the convolution operation of a kernel consisting of three elements s_{-1} , s_0 , and s_{+1} . Here, for compactness of notation, we often represent a kernel as a row vector with an odd number of elements. The three-element kernel here, for example, is denoted as $[s_{-1}, s_0, s_{+1}]$. Figure 1 (C and D) illustrates the operation of such kernel when $s_0 = 2$ and $s_{\pm 1} = 1$. Figure 1C corresponds to a case with the input consisting only of a single frequency, namely, the continuous-wave (CW) laser. Figure 1D corresponds to a case with the input being a multiple-frequency comb. 1D convolution is essential for feature extraction in sequential data processing such as audio and speech comprehension (34).

For experimental design, it would be desirable to generate a prescribed kernel with an analytical modulation waveform. Assuming zero detuning, $\Delta\omega = 0$, for a given kernel with elements s_n , the corresponding modulation parameters B_m and β_m are given by (32)

$$B_m \exp(j\beta_m) = t_R (\kappa_{+m} - \kappa_{-m}^*) \quad (4)$$

$$\gamma + \gamma_{\text{cst}} = -j\kappa_0 \quad (5)$$

where coupling constants κ_m are related to the kernel elements s_n via

$$K_m = \frac{j\gamma_{e1} \gamma_{e2}}{\pi} \int_0^{2\pi} \frac{e^{-jmk}}{\sum_n e^{jnk} s_n} dk \quad (6)$$

Kernel synthesis

Our experiments use a fiber ring resonator modulated by an electro-optic modulator as shown in Fig. 1A. The ring has an FSR of $\Omega_R = 2\pi \cdot 5.99$ MHz, corresponding to a circumference of $\ell = 34.3$ m. From the input waveguide, we launch a CW laser into the ring resonator through a fiber coupler. The laser's frequency is scanned across a resonance of the unmodulated ring. Within the cavity, we use an Er-doped fiber amplifier to compensate for part of the roundtrip loss. At each detuning $\Delta\omega$, we measure the time-resolved output power $I(\Delta\omega, t)$ at the drop port, using a fast photodiode with a bandwidth over 5 GHz and an oscilloscope of 1 GHz analog bandwidth.

We experimentally construct various convolution kernels based on the theory as discussed in the previous section. Here, the input-output transformation is as described in Fig. 1C where we launch a single frequency, and, therefore, the output manifests the kernel. In the first example (Fig. 2, A to D), we demonstrate the high boost kernel, which consists of three nonzero elements of $s_0 = 6$ and $s_{\pm 1}$

$= -1$. This kernel is widely applied in image processing to sharpen the high-frequency edge information and enhance the low-frequency feature information in the image (35).

To generate this kernel, we first calibrate the loss rate $\gamma + \gamma_{\text{cst}} = 0.027\Omega_R$. This calibration is described in more detail in Materials and Methods. With this $\gamma + \gamma_{\text{cst}}$ and Eqs. 4 and 5, we obtain the modulation waveform. For the amplitude modulation as described by Eq. 1, the magnitudes are $B_1 = 5.858 \times 10^{-2}$, $B_2 = 9.994 \times 10^{-3}$, $B_3 = 1.679 \times 10^{-3}$, $B_4 = 2.676 \times 10^{-4}$, and $B_5 = 3.323 \times 10^{-5}$, the phase angles are $\beta_1 = 1.576$, $\beta_2 = 1.580$, $\beta_3 = 1.585$, $\beta_4 = 1.590$, and $\beta_5 = 1.594$. At any given time, the instantaneous loss rate of the cavity is defined as

$$\gamma(t) = \gamma + \gamma_{\text{cst}} + \sum_{m \geq 1} \frac{B_m}{t_R} \cos(m\Omega_R t) \quad (7)$$

$\gamma(t)$ is above zero as shown in Fig. 2A. Therefore, the modulation as designed in this way satisfies the passivity constraint (36), and the system is always dissipative.

We apply the modulation waveform, as designed above, to the ring resonator. In the experiment, we vary the detuning $\Delta\omega$ by adjusting the input laser frequency. At each detuning $\Delta\omega$, we record the intensity at the drop port $I(\Delta\omega, t)$ as a function of time. The resulting 2D plot of $I(\Delta\omega, t)$ is plotted in Fig. 2B. We observe that the linewidth of the resonance is the smallest at about $t = \pi/\Omega_R$ in the horizontal axis defined in Fig. 2B. This is consistent with Fig. 2A, where the instantaneous loss rate is lowest at the same t .

To determine the kernel from the output intensity measurement $I(\Delta\omega, t)$, we recall that $I(\Delta\omega, t) = |S(\Delta\omega, t)|^2$, with $S(\Delta\omega, t)$ being the time-domain scattering factor of Eq. 2. Since the modulation in Fig. 2A is designed for the kernel at $\Delta\omega = 0$, we plot $I(\Delta\omega = 0, t)$ as shown in Fig. 2C. Throughout the paper, all the kernel generation and convolution are based on this $\Delta\omega = 0$ line only. As the high-boost kernel demonstrated here is real-valued and symmetric, the time-domain scattering factor $S(0, t) = \sum_n s_n \exp(jn\Omega_R t)$ should be real-valued

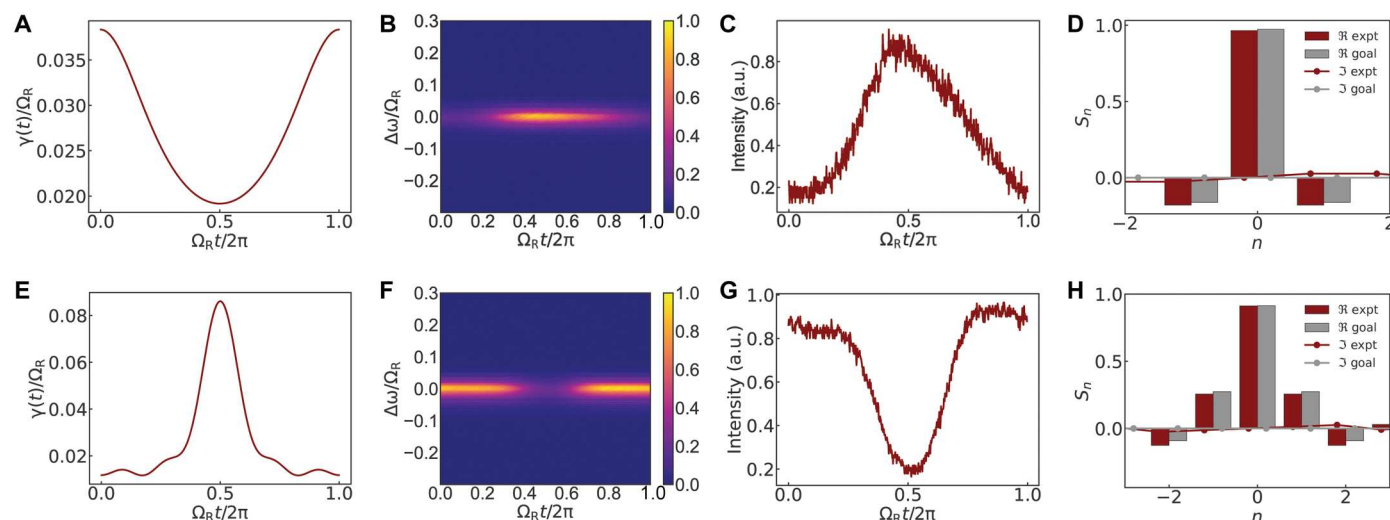


Fig. 2. Experimental synthesis of convolution kernels. A high-boost kernel $[-1, 6, -1]$ in (A to D) and a Laplacian of Gaussian kernel $[-1, 3, 10, 3, -1]$ in (E to H). (A) and (E) Calculated instantaneous loss rate $\gamma(t)$ as a function of time in a roundtrip. (B) and (F) Measured time- and frequency-detuning resolved output intensity $I(\Delta\omega, t)$. This is measured at the drop port from a dynamically modulated ring resonator. (C) and (G) Measured $I(\Delta\omega = 0, t)$ in (B) and (F), respectively. (D) and (H) Comparison of the synthesized kernel and target kernel. The red bar/line corresponds to the real/imaginary part of the experimental kernel. The gray bar/line corresponds to the real/imaginary parts of the target kernel. a.u., arbitrary units.

as well, with s_n defined in Eq. 3. In the Supplementary Materials, we use an example single cosine modulation to prove that the modulation waveform only results in a change of $\gamma(t)$ in Eq. 7. We confirm that the amplitude modulation waveform that is obtained from the experiment agrees well with what is implemented on the modulator. This proves that $S(0, t)$ purely results from amplitude modulation, so $S(0, t)$ is real-valued, and the phase variation in a round-trip is negligible. From $I(0, t)$ as shown in Fig. 2C and $I(0, t) = |S(0, t)|^2$, we obtain $S(0, t) = \sqrt{I(0, t)}$. We then perform a Fourier transform of $S(0, t)$ to determine the kernel s_n that is obtained in the experiment.

In Fig. 2D, we compare the kernel obtained from experiments and target designs. The s_n from the experimental measurement is shown next to the s_n from the target design. Both kernels are normalized such that $\sum_n |s_n|^2 = 1$. These two kernels agree well and verify that the high-boost kernel is synthesized successfully.

As one more example of kernel synthesis, in Fig. 2 (E to H), we synthesize a quantized Laplacian of Gaussian kernel with its nonzero elements being $s_0 = 10$, $s_{\pm 1} = 3$, and $s_{\pm 2} = -1$. This quantized kernel is suitable for compressing features and tracking the machine learning process (37). The modulation waveform is designed in a similar way as above using Eqs. 4 and 5. The magnitudes of the modulation waveform are $B_1 = 0.1539$, $B_2 = 0.1014$, $B_3 = 0.05854$, $B_4 = 0.03525$, and $B_5 = 0.02094$. The corresponding phase angles are $\beta_1 = -1.566$, $\beta_2 = 1.580$, $\beta_3 = -1.557$, $\beta_4 = 1.590$, and $\beta_5 = -1.547$. Using these parameters, the instantaneous loss rate given by Eq. 7 is plotted in Fig. 2E. Contrary to the prior example, the instantaneous loss rate is highest in the middle of the roundtrip in this case. In Fig. 2F, we present the measured time- and frequency-detuning resolved output intensity $I(\Delta\omega, t)$. $I(\Delta\omega = 0, t)$ is

plotted in Fig. 2G. Using a similar method as in the previous example, we extract the experimental kernel s_n values from $I(\Delta\omega = 0, t)$. As shown in Fig. 2H, the experimental kernel agrees well with the target kernel, which verifies that our analytically designed modulation waveform can faithfully synthesize a multielement quantized Laplacian of Gaussian kernel.

Convolution kernel construction with an additive offset

As seen in the two examples provided in the previous section, the implemented kernel typically has a strong s_0 component in our modulated ring setup. This arises because of the high internal loss factor γ_{cst} and the limited modulator strength in our setup. In this section, we implement the convolution kernel with an additive offset, as described in the form of

$$c_{\text{out},n} = bc_{\text{in},n} + \sum_n s_{n-m} c_{\text{in},m} \quad (8)$$

where $b < 0$ is the additive offset. Alternatively, we consider the implementations of Eq. 8 to broaden the kinds of kernels that can be implemented in a fiber experimental system.

In our setup, the operation of Eq. 8 can be implemented by synthesizing a kernel $\{\tilde{s}_n\}$ where $\tilde{s}_0 = s_0 + b$ and $\tilde{s}_n = s_n$ for $n \neq 0$, in the same way as we described in the previous section. On the other hand, in scenarios where the strength of the modulation is insufficient to directly achieve Eq. 8 using the procedure as described in the previous section, we note that Eq. 8 can be implemented in an alternative all-optical implementation (38). In this alternative implementation, one passes the input light through a beam splitter to separate it into two paths. In the first path, one implements the operation of the first term in Eq. 8 using a π phase shifter and an

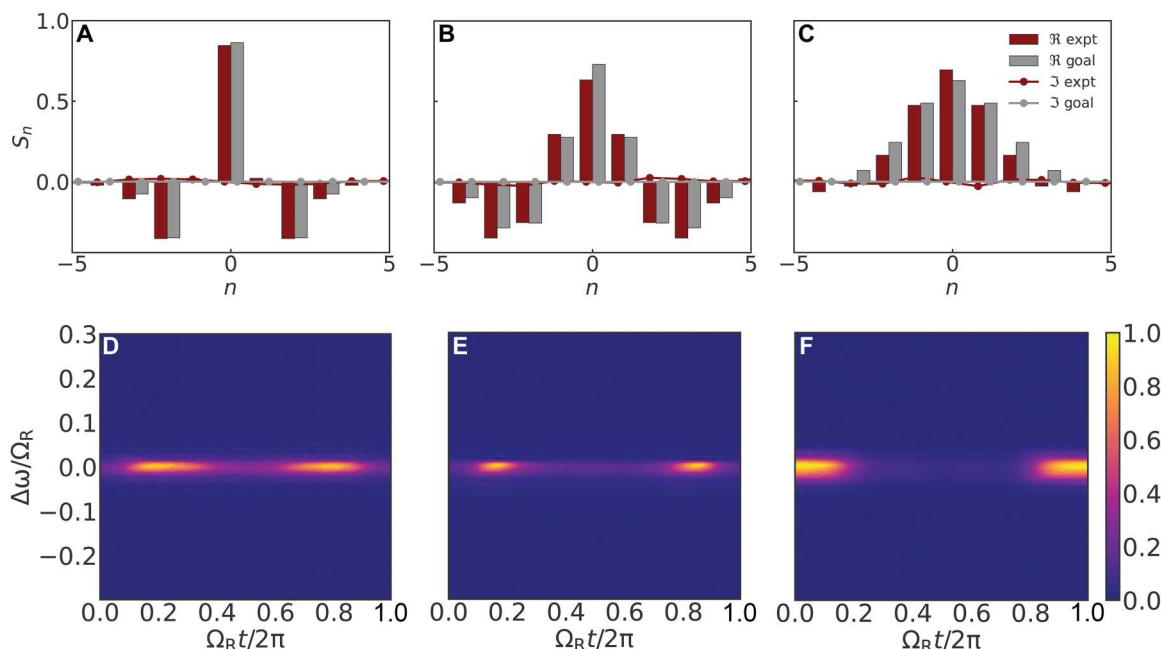


Fig. 3. Construction of convolution kernels with multiple examples of various kernels. (A and D) A standard Laplacian of Gaussian kernel $[-1, -4.56, 0.028, 11.304, 0.028, -4.56, -1]$ with $b = -20$, (B and E) another standard Laplacian of Gaussian kernel $[-1, -2.9, -2.6, 2.8, 7.4, 2.8, -2.6, -2.9, -1]$ with $b = -20$, (C and F) a Gaussian kernel $[1, 3.5, 7, 9, 7, 3.5, 1]$ with $b = -8$. The top panels of (A) to (C) correspond to the synthesized kernel measured (in red) and target (in gray) kernels with the real and imaginary parts plotted in bar and lines, respectively. The bottom panels of (D) to (F) correspond to the time- and frequency-detuning resolved output intensity measurements. The experimentally synthesized kernel in (A) to (C) is obtained from (D) to (F), respectively.

attenuator or amplifier. In the second path, one implements the second term in Eq. 8 using our modulated fiber ring setup. The transmitted lights from these two paths are then combined to realize Eq. 8. A schematic of this realization is provided in the Supplementary Materials.

Here, as an illustration of Eq. 8 and for simplicity, instead of the all-optical implementation as discussed above, we present results from a hybrid implementation. In the hybrid implementation, for a prescribed target kernel $\{\tilde{s}_n\}$, we separate it into two terms in Eq. 8 such that the second term can be implemented using our modulated ring setup. We then present the end results assuming that the first term and the summation operation in Eq. 8 have been carried out digitally.

In Fig. 3, we present the implementations of various kernels using this hybrid approach. Figure 3 (A and B) demonstrates a standard Laplacian of Gaussian kernel with different parameters. In both cases, the kernel elements are summed to zero. This Laplacian of Gaussian kernel is widely applied in noise-robust spatial filtering and edge detection (39). Figure 3A corresponds to a seven-element kernel with a standard deviation $\sigma = 1.0$. Figure 3B corresponds to a nine-element kernel with an SD $\sigma = 1.4$. Figure 3C presents a Gaussian kernel with an SD $\sigma = 1.4$. Such a Gaussian kernel is useful for suppressing high-frequency noise in a limited spatial spread area, which is essential for digital telecommunications (40). The details of the modulation waveform parameters can be found in the Supplementary Materials.

Figure 3 (D to F) corresponds to the time- and frequency-detuning resolved output intensity measurement. The experimentally synthesized kernel in Fig. 3 (A to C) is obtained from Fig. 3 (D to F), respectively, using the same method discussed in the previous section. All of the kernels are normalized such that $\sum_n |s_n|^2 = 1$.

In Fig. 3 (A to C), the measured kernels agree very well with the target kernels in both real and imaginary parts. This verifies that we can synthesize a broad range of kernels at high accuracy with the approach as described by Eq. 8.

Convolution processing

In the previous sections, we demonstrated the synthesis of several convolution kernels. In these demonstrations, we performed convolution operations with an input vector that consists of only a single element. In this section, we provide an experimental demonstration of the convolution operation of the kernels with various input vectors that consist of multiple frequency comb lines.

To start with, we first synthesize a modified Laplacian kernel $s_0 = 3$ and $s_{\pm 1} = -1$. This functions in a similar way as a high boost kernel introduced before, but the reduced s_0 term enables an improved edge detection property. We follow the same procedure of applying a predetermined modulation waveform, as introduced in previous sections. In Fig. 4A, we compare the kernel obtained from experiments and target designs. The s_n from the experimental measurement is shown next to the s_n from the target design. Both kernels are normalized such that $\sum_n |s_n|^2 = 1$. These two kernels agree

well and verify that the modified Laplacian kernel is synthesized successfully. The slice of $\Delta\omega = 0$ in the time- and frequency-detuning resolved drop-port intensity measurement is shown in Fig. 4D, which shows a consistent lineshape as in the high-boost kernel case.

We emphasize that in this kernel synthesis example, there is no additive offset term involved.

To generate the input vector, we use a CW laser operating at a swept frequency across the resonant frequency of the ring and pass the output of the CW laser through an electro-optic amplitude modulator. The modulator is driven by an arbitrary waveform generator (AWG), which has frequency components of the FSR and its integer multiples. This modulation is periodic with a periodicity equal to the round-trip time. Such a modulation results in a comb of discrete frequencies equally separated by FSR, which is injected into the ring.

The input vector thus generated can be characterized by measuring the time-dependent intensity $I_{in}(t)$ that is transmitted through the modulator. For an amplitude modulator, the amplitude of the transmitted light, up to a global phase that is unimportant, can be determined as $A_{in}(t) = \sqrt{I_{in}(t)}$. A Fourier transform of $A_{in}(t)$ then determines the input vector, i.e., the complex amplitudes of the input light at various frequencies.

Figure 4 (B and C) shows two different input vectors thus generated by applying multiple sinusoidal bands and a sharp pulse, respectively. We choose these two modulations to generate as broadband frequency combs as possible. For each of these input vectors, we send it through the setup corresponding to the kernel shown in Fig. 4 (A and D). To determine the generated output vector, we measure the output intensity $I_{out}(t)$ as a function of time. Since only the amplitude modulator is used in synthesizing the kernels, we determine the output amplitude $A_{out}(t) = \sqrt{I_{out}(t)}$, we then Fourier-transform $A_{out}(t)$ to obtain the output vector. The experimentally determined output vector agrees very well with the direct calculation of the convolution operation of the kernels on the input vectors using the output signal from Fig. 4D, as shown in Fig. 4 (E and F). We have thus demonstrated that our setup can indeed achieve convolution operation in the synthetic frequency dimension.

DISCUSSION

In summary, we experimentally demonstrate convolution operation in the synthetic frequency space. We show that the prescribed kernel can be implemented by an analytically determined modulation waveform applied to the electro-optic modulator. Our work demonstrates the promise of using frequency to encode data and implement convolution tasks. We anticipate that our demonstration of convolution operation via frequency synthetic dimensions may lead to new types of scalable photonic computation architecture.

We note that throughout the paper, we only use amplitude modulators, both for the generation of the input signals and for kernel synthesis. As a proof-of-principle experiment, this suffices to demonstrate a wide range of convolution. Although we demonstrated some symmetric convolution kernels in our paper, our work can be readily generalized to arbitrary convolution kernels within the capabilities of current experimental setups. For example, an asymmetric kernel can be encoded using an amplitude modulator and a phase modulator in the same ring resonator, as demonstrated in our previous theory work (32). For kernels with an even number of nonzero elements, we can pad a zero to either side of the kernel and implement it as an asymmetric kernel with an odd number of elements. In addition, to realize kernels with a small $|s_0|$, we

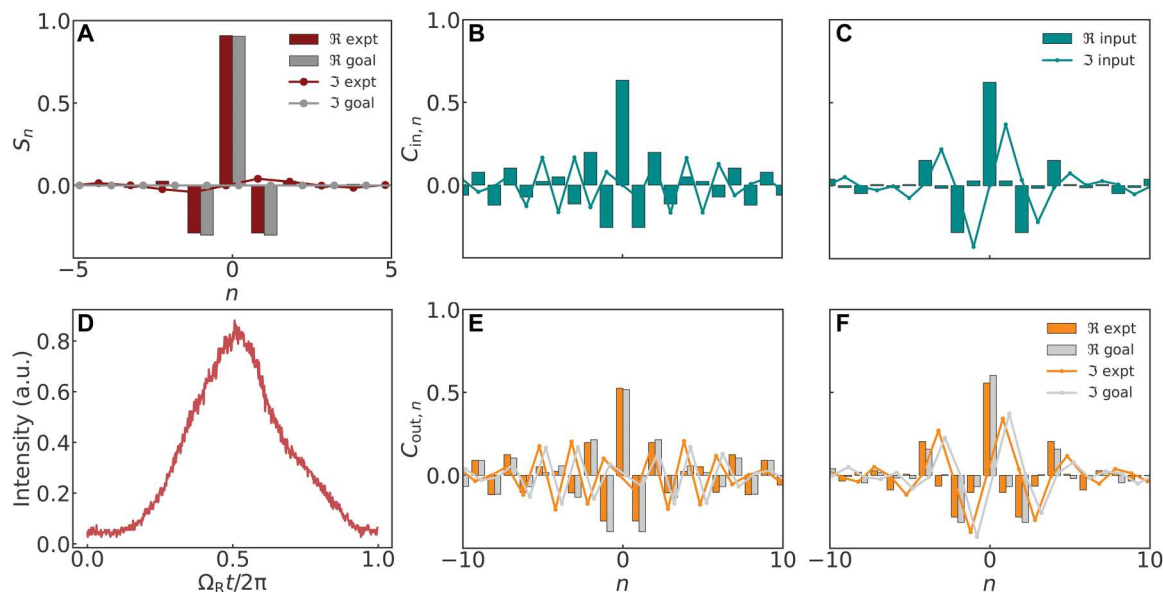


Fig. 4. Convolution processing of the kernels generated from a modulated ring resonator with an input frequency comb consisting of multiple nonzero frequency comb lines. (A) Comparison of the synthesized kernel and target kernel. The red bar/line corresponds to the real/imaginary part of the experimental kernel. The gray bar/line corresponds to the real/imaginary parts of the target kernel. (B and C) correspond to the input frequency comb measured from experiments. (D) Measured time-resolved intensity from the drop port of the modulated ring resonator $I(\Delta\omega = 0, t)$ for the kernel synthesis in (A). (E and F) correspond to the output frequency comb measured (in orange) and expected (in gray) outputs with the real and imaginary parts plotted in bar and lines, respectively.

discuss in the Supplementary Materials an all-optical approach using interference to provide an offset term to the kernel. A deterministic modulation waveform can be obtained in all cases, and we numerically demonstrate two asymmetric kernels in the Supplementary Materials. In the case of a general asymmetric kernel, decoding the output signal would require the retrieval of the phase information in the output light, which can be done with the use of heterodyne detection (41).

The size of the kernel matrix is constrained by the modulation speed bandwidth and FSR, with larger matrices requiring higher modulation speed bandwidth and lower FSR allowing for higher modulation orders. To estimate the performance of our scheme implemented in an on-chip integrated platform, we assume a device with a pump power of 2 mW, 200 input comb lines (12), a modulation speed of 10 GHz, and a modulator power cost of 100 mW for a 1 mm² area chip (42). The computation density for this device is about 4 tera "trillion" operations per second (TOPS) mm⁻², four orders of magnitudes higher than graphics processing unit (43), and four times faster than the previous state-of-the-art photonic convolution unit (12). In terms of power efficiency, we achieve 40 TOPS W⁻¹, 6× more power efficiency than Nvidia's A100X (43). Our platform is limited only by the photodetector bandwidth. In anticipation, future advances in fabricating high-speed and high-confinement modulators, as well as high-speed photodetectors, will improve our estimations further.

Our system based on electro-optic modulation complements the previously reported acousto-optic modulation approach (24) and can be easily integrated with existing electronic circuitry while allowing a wider range of operating frequencies. All components of the convolution setup demonstrated here can be integrated on a chip. Potential benefits of moving to an integrated platform include lower energy consumption with integrated modulators

and lasers, higher computation density with chip-scale compact areas, and more robust and portable edge computing platforms. A major limitation of integration is the large FSR for integrated resonators, which limits the number of modes the modulators and photodetectors can cover at the same time. With the advances in on-chip low-loss waveguides and modulators, it is possible to integrate the entire system without amplifiers and achieve energy-efficient on-chip convolution processes.

MATERIALS AND METHODS

Calibration of the loss rate

In this section, we describe the experimental calibration process of $\gamma + \gamma_{\text{cst}}$. Without any modulation from the electro-optical modulator (JDSU model 10020476), we measure the output intensity $I(\Delta\omega)$ from the drop port of the ring resonator, in the same way as described in the "Kernel synthesis" section. $I(\Delta\omega)$ is related to $\gamma + \gamma_{\text{cst}}$ by

$$I(\Delta\omega) = \left| \frac{2j\gamma_{e1}\gamma_{e2}}{j\gamma + j\gamma_{\text{cst}} - \Delta\omega} \right|^2 \quad (9)$$

We then perform the least square fitting of $I(\Delta\omega)$ to obtain the optimal parameters of $\gamma + \gamma_{\text{cst}}$. In our system, the calibrated loss factor is $\gamma + \gamma_{\text{cst}} = 0.027\Omega_R$. We provide more details of extracting the loss factor in the Supplementary Materials.

Data processing and time sequence acquisition

In our experiments, we use a narrow-linewidth laser with tunable lasing frequency as input (ORION 1550 nm laser module) under an amplitude modulator (JDSU, model 10020476) controlled by the radio frequency signal from an AWG (Agilent 33250A-U 80 MHz function). We use an erbium-doped amplifier (IRE-Polus,

model EAU-2 M) to amplify the optical signal. We use a radio frequency amplifier (Mini-Circuits, model ZHL-3A+) to amplify the modulation signal.

To measure the time-dependent output intensity $I(\Delta\omega, t)$ at the drop port, we use a photodiode (Thorlabs DET08CFC) with a 5 GHz bandwidth to detect the output signal and use an oscilloscope (LeCroy LC584AL) with a bandwidth of 1 GHz to obtain 1-ms time-sequence data. The 1-ms-long time-sequence data were then reshaped into multiple time sequences, one for a round-trip time of the ring $[1/(5.99 \text{ MHz}) = 167 \text{ ns}]$.

We determine the starting time of one round-trip sequence by comparing the intensity peak of the theoretical design peak location. We shift one sequence so that the experimental resonant peak is aligned with the designed peak. The entire measured time sequence is shifted by the same amount of time. We then unflatten the 1D data sequences along the vertical axis to obtain the 2D intensity measurement in Figs. 2 (B and E) and 3 (E to H). The details of the experimental setup can be found in the Supplementary Materials.

Supplementary Materials

This PDF file includes:

Figs. S1 to S5
Tables S1 and S2

REFERENCES AND NOTES

1. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
2. A. Krizhevsky, I. Sutskever, G. E. Hinton, *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger, Eds. (Curran Associates Inc., 2012), vol. 25.
3. D. Ciregan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, in *2012 IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2012)*, pp. 3642–3649.
4. C. Choy, J. Gwak, S. Savarese, 4D spatio-temporal ConvNets: Minkowski convolutional neural networks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR, 2019)*, pp. 3075–3084.
5. Y.-H. Chen, J. Emer, V. Sze, in *Proceedings of the 43rd International Symposium on Computer Architecture, ISCA '16 (IEEE Press, 2016)*, pp. 367–379.
6. G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljčić, C. Denz, D. A. B. Miller, D. Psaltis, Inference in artificial intelligence with deep optics and photonics. *Nature* **588**, 39–47 (2020).
7. V. Bangari, B. A. Marquez, H. Miller, A. N. Tait, M. A. Nahmias, T. F. de Lima, H.-T. Peng, P. R. Prucnal, B. J. Shastri, Digital electronics and analog photonics for convolutional neural networks (deap-cnns). *IEEE J. Sel. Top. Quantum Electron.* **26**, 1–13 (2020).
8. B. J. Shastri, A. N. Tait, T. Ferreira de Lima, W. H. P. Pernice, H. Bhaskaran, C. D. Wright, P. R. Prucnal, Photonics for artificial intelligence and neuromorphic computing. *Nat. Photonics* **15**, 102–114 (2021).
9. J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, W. H. P. Pernice, All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **569**, 208–214 (2019).
10. X. Xu, M. Tan, B. Corcoran, J. Wu, T. G. Nguyen, A. Boes, S. T. Chu, B. E. Little, R. Morandotti, A. Mitchell, D. G. Hicks, D. J. Moss, Photonic perception based on a Kerr microcomb for high-speed, scalable, optical neural networks. *Laser Photonics Rev.* **14**, 2000070 (2020).
11. X. Xu, M. Tan, B. Corcoran, J. Wu, A. Boes, T. G. Nguyen, S. T. Chu, B. E. Little, D. G. Hicks, R. Morandotti, A. Mitchell, D. J. Moss, 11 tops photonic convolutional accelerator for optical neural networks. *Nature* **589**, 44–51 (2021).
12. J. Feldmann, N. Youngblood, M. Karpov, H. Gehring, X. Li, M. Stappers, M. Le Gallo, X. Fu, A. Lukashchuk, A. S. Raja, J. Liu, C. D. Wright, A. Sebastian, T. J. Kippenberg, W. H. P. Pernice, H. Bhaskaran, Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021).
13. T. Alexoudi, G. T. Kanellos, N. Pleros, Optical ram and integrated optical memories: A survey. *Light Sci. Appl.* **9**, 91 (2020).
14. H. J. Caulfield, S. Dolev, Why future supercomputing requires optics. *Nat. Photonics* **4**, 261–263 (2010).
15. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Laroche, D. Englund, M. Soljčić, Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017).
16. T. W. Hughes, M. Minkov, Y. Shi, S. Fan, Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **5**, 864–871 (2018).
17. I. A. D. Williamson, T. W. Hughes, M. Minkov, B. Bartlett, S. Pai, S. Fan, Reprogrammable electro-optic nonlinear activation functions for optical neural networks. *IEEE J. Sel. Top. Quantum Electron.* **26**, 1–12 (2020).
18. A. N. Tait, T. F. de Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, P. R. Prucnal, Neuromorphic photonic networks using silicon photonic weight banks. *Sci. Rep.* **7**, 7430 (2017).
19. A. N. Tait, T. Ferreira de Lima, M. A. Nahmias, H. B. Miller, H.-T. Peng, B. J. Shastri, P. R. Prucnal, Silicon photonic modulator neuron. *Phys. Rev. Appl.* **11**, 064043 (2019).
20. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, A. Ozcan, All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
21. Y. Luo, Y. Zhao, J. Li, E. Çetintaş, Y. Rivenson, M. Jarrahi, A. Ozcan, Computational imaging without a computer: Seeing through random diffusers at the speed of light. *eLight* **2**, 4 (2022).
22. T. W. Hughes, I. A. D. Williamson, M. Minkov, S. Fan, Wave physics as an analog recurrent neural network. *Sci. Advances* **5**, eaay6946 (2019).
23. M. J. Collins, C. Xiong, I. H. Rey, T. D. Vo, J. He, S. Shahnia, C. Reardon, T. F. Krauss, M. J. Steel, A. S. Clark, B. J. Eggleton, Integrated spatial multiplexing of heralded single-photon sources. *Nat. Commun.* **4**, 2582 (2013).
24. H. Zhao, B. Li, H. Li, M. Li, Enabling scalable optical computing in synthetic frequency dimension using integrated cavity acousto-optics. *Nat. Commun.* **13**, 5426 (2022).
25. L. Yuan, Q. Lin, M. Xiao, S. Fan, Synthetic dimension in photonics. *Optica* **5**, 1396–1405 (2018).
26. L. Yuan, A. Dutt, S. Fan, Synthetic frequency dimensions in dynamically modulated ring resonators. *APL Photonics* **6**, 071102 (2021).
27. A. Dutt, M. Minkov, Q. Lin, L. Yuan, D. A. B. Miller, S. Fan, Experimental band structure spectroscopy along a synthetic dimension. *Nat. Commun.* **10**, 3122 (2019).
28. A. Dutt, Q. Lin, L. Yuan, M. Minkov, M. Xiao, S. Fan, A single photonic cavity with two independent physical synthetic dimensions. *Science* **367**, 59–64 (2020).
29. Y. Hu, C. Reimer, A. Shams-Ansari, M. Zhang, M. Loncar, Realization of high-dimensional frequency crystals in electro-optic microcombs. *Optica* **7**, 1189–1194 (2020).
30. K. Wang, A. Dutt, K. Y. Yang, C. C. Wojcik, J. Vuković, S. Fan, Generating arbitrary topological windings of a non-hermitian band. *Science* **371**, 1240–1245 (2021).
31. S. Buddhiraju, A. Dutt, M. Minkov, I. A. D. Williamson, S. Fan, Arbitrary linear transformations for photons in the frequency synthetic dimension. *Nat. Commun.* **12**, 2401 (2021).
32. L. Fan, Z. Zhao, K. Wang, A. Dutt, J. Wang, S. Buddhiraju, C. C. Wojcik, S. Fan, Multidimensional convolution operation with synthetic frequency dimensions in photonics. *Phys. Rev. Appl.* **18**, 034088 (2022).
33. B. Jähne, *Neighborhood Operations* (Springer Berlin Heidelberg, 2002), pp. 99–124.
34. S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, D. J. Inman, 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **151**, 107398 (2021).
35. A. Satapathy, L. M. J. Livingston, Optimized openCL™ kernels for frequency domain image high-boost filters using image vectorization technique. *SN Applied Sciences* **1**, 1424 (2019).
36. S. Sandhu, S. Fan, Lossless intensity modulation in integrated photonics. *Opt. Express* **20**, 4280–4290 (2012).
37. D. Qin, X. Chen, M. Guillaumin, L. V. Gool, Quantized kernel learning for feature matching, in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, K. Weinberger, Eds. (Curran Associates Inc., 2014), vol. 27.
38. P. Minzioni, C. Lacava, T. Tanabe, J. Dong, X. Hu, G. Csaba, W. Porod, G. Singh, A. E. Willner, A. Almailan, V. Torres-Company, J. Schröder, A. C. Peacock, M. J. Strain, F. Parmigiani, G. Contestabile, D. Marpaung, Z. Liu, J. E. Bowers, L. Chang, S. Fabbri, M. R. Vázquez, V. Bharadwaj, S. M. Eaton, P. Lodahl, X. Zhang, B. J. Eggleton, W. J. Munro, K. Nemoto, O. Morin, J. Laurat, J. Nunn, Roadmap on all-optical processing. *J. Opt.* **21**, 063001 (2019).
39. J. S. Chen, A. Huertas, G. Medioni, Fast convolution with Laplacian-of-Gaussian masks. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**, 584–590 (1987).
40. S. Rapuano, G. Truglia, An improved image processing-based method for disturbance classification in telecommunication networks. *IEEE Trans. Instrum. Meas.* **54**, 2068–2074 (2005).
41. A. Dutt, M. Minkov, Q. Lin, L. Yuan, D. A. B. Miller, S. Fan, Experimental demonstration of dynamical input isolation in nonadiabatically modulated photonic cavities. *ACS Photonics* **6**, 162–169 (2019).
42. M. Zhang, B. Buscaino, C. Wang, A. Shams-Ansari, C. Reimer, R. Zhu, J. M. Kahn, M. Loñar, Broadband electro-optic frequency comb generation in a lithium niobate microring resonator. *Nature* **568**, 373–377 (2019).

43. Nvidia, "Converged accelerators" (2022); <https://resources.nvidia.com/en-us-accelerated-networking-resource-library/converged-accelerators>.

Acknowledgments: We thank D. A. B. Miller for providing the laboratory space. **Funding:** This work is supported by a MURI project from the U.S. Air Force Office of Scientific Research (AFOSR) (grant no. FA9550-22-1-0339). **Author contributions:** Conceptualization: L.F. and S.F. Methodology: L.F., K.W., and S.F. Investigation: L.F., K.W., H.W., A.D., and S.F. Visualization: L.F. and S.F. Supervision: S.F. Writing—original draft: L.F. Writing—review and editing: L.F., K.W., H.W., A.D., and S.F. **Competing interests:** L.F., K.W., and S.F. have filed a patent (U.S. Provisional Patent

Application 63/349413, filed 6/6/2023) based on this work. The authors declare that they have no other competing interests. **Data and materials availability:** All data are available in the main text and/or the Supplementary Materials.

Submitted 29 April 2023

Accepted 12 July 2023

Published 11 August 2023

10.1126/sciadv.adi4956

Experimental realization of convolution processing in photonic synthetic frequency dimensions

Lingling Fan, Kai Wang, Heming Wang, Avik Dutt, and Shanhui Fan

Sci. Adv., **9** (32), eadi4956.
DOI: 10.1126/sciadv.adi4956

View the article online

<https://www.science.org/doi/10.1126/sciadv.adi4956>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)