



Photonic principal component analysis using an on-chip microring weight bank

PHILIP Y. MA,^{1,*} ALEXANDER N. TAIT,^{1,2} THOMAS FERREIRA DE LIMA,¹ SIAMAK ABBASLOU,¹ BHAVIN J. SHASTRI,^{1,3} AND PAUL R. PRUCNAL¹

¹*Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA*

²*Physical Measurement Laboratory, National Institute of Standards and Technology, Boulder, CO 80305, USA*

³*Department of Physics, Engineering Physics and Astronomy, Queen's University, Kingston, ON K7L 3N6, Canada*

*yechim@princeton.edu

Abstract: Photonic principal component analysis (PCA) enables high-performance dimensionality reduction in wideband analog systems. In this paper, we report a photonic PCA approach using an on-chip microring (MRR) weight bank to perform weighted addition operations on correlated wavelength-division multiplexed (WDM) inputs. We are able to configure the MRR weight bank with record-high accuracy and precision, and generate multi-channel correlated input signals in a controllable manner. We also consider the realistic scenario in which the PCA procedure remains blind to the waveforms of both the input signals and weighted addition output, and propose a novel PCA algorithm that is able to extract principal components (PCs) solely based on the statistical information of the weighted addition output. Our experimental demonstration of two-channel photonic PCA produces PCs holding consistently high correspondence to those computed by a conventional software-based PCA method. Our numerical simulation further validates that our scheme can be generalized to high-dimensional (up to but not limited to eight-channel) PCA with good convergence. The proposed technique could bring new solutions to problems in microwave communications, ultrafast control, and on-chip information processing.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

The rapid proliferation of communication networks carrying high data capacity (e.g., multiple-input multiple-output (MIMO) systems and data centers) as well as applications requiring fast-growing data volumes (e.g., video streaming and cloud services) are demanding faster, higher bandwidth information processing technologies. Conventional digital electronic approaches are facing harsh trade-offs among channel dimensionality, processing bandwidth, and energy consumption due to the bottlenecks in analog-to-digital conversion (ADC) [1, 2] and digital signal processing (DSP) [3]. In contrast, analog photonic approaches provide fascinating alternatives to break these limitations, especially in the field of microwave photonics [4–6], large reconfigurable linear systems [7–9], and neuromorphic photonics [10–13].

Advances in photonic integrated circuits (PIC) further accelerate the emergence of large-scale, high-performance, low-cost analog information processing systems such as microring (MRR) weight banks, which combine standard photonic devices for linear computation and the enormous information density made possible with wavelength-division multiplexing (WDM) [14]. MRRs are favorable processing elements due to their ubiquity in PIC platforms, small footprint, and ease of tuning through thermo-optic effect [15]. They are promising candidates for reconfigurable weighting elements in neuromorphic photonic architectures [11, 16], waveform generation [17, 18], directed logic [19], and excitable signal processing devices [20].

MRR weight banks implement a parallel matrix-vector multiplication operation called weighted

addition, where WDM signals are individually weighted by a reconfigurable spectral filter and detected together at a balanced photo-detector (BPD) whose electronic output represents their sum [14,21]. Weighted addition is the fundamental multi-input, single-output function underlying dimensionality reduction, commonly used in principal component analysis (PCA) to transform correlated variables to uncorrelated variables called principal components (PCs). These PCs are linear combinations of the original variables weighted by their contribution to the variance in a particular orthogonal dimension. Data compression with minimum information loss can be achieved by sacrificing PCs associated with smallest variances [22].

Photonic PCA is a technique of multivariate photonics: combining multivariate statistical analysis with analog photonic processing [23, 24]. Earlier implementations of photonic PCA required the inputs to be synchronized and periodic such that the complete waveform information of system inputs/outputs would be available [25, 26]. In reality, however, this information cannot be readily observed by dimension-reducing analog front-ends especially in dynamic radio frequency (RF) scenarios. A novel PCA algorithm is thus required to replace the Hebbian learning rule [27] used by these prior approaches (which needs the waveform information mentioned above). The algorithms in [23, 24] and this work are based on observations of the statistical moments of a single, reduced-dimensional output, i.e., weighted addition output.

In this work, we demonstrate a photonic PCA scheme using an on-chip MRR weight bank instead of discrete fiber-based variable optical attenuators used in [25, 26]. We adopt the feedback control procedure [28] to configure the MRR weight bank with record-breaking accuracy and precision, which is based only on electrical probing and robust to environmental fluctuations as opposed to the feedforward scheme [29, 30] adopted in [24, 31]. We consider the realistic scenario where it is impossible to directly observe the waveform information of the multi-channel correlated inputs or weighted addition output [24], and contribute a novel PCA algorithm based on optimal moment tracking that is compatible with such constrained observability. In a proof-of-concept experiment of two-channel photonic PCA, our new approach achieves PCs holding consistently high correspondence with those derived from a conventional software-based PCA method, and outperforms the results obtained by fitting a model of statistical moments [24]. Our numerical simulation further illustrates that the proposed approach can generalize to higher-dimensional PCA (up to eight-channel in this work).

2. Methods

2.1. Device fabrication and characterization

The silicon photonic chip used in this work is shown in Fig. 1(a), which is fabricated on a silicon-on-insulator (SOI) wafer at the Institute of Microelectronics (IME) A*STAR foundry [32]. Silicon thickness is 220 nm, and buried oxide thickness is 2 μm . Waveguides of 500 nm width are patterned by deep ultraviolet (DUV) lithography. The MRR weight bank (in the middle) consists of two bus waveguides that cascade four MRRs in a parallel add/drop configuration and direct light between the MRR weight bank and corresponding TE focusing grating couplers (on the right) [33]. These outputs are detected by an off-chip BPD. Metal vias and traces are deposited to connect the heater contacts of MRR weight bank to electrical metal pads (on the left). Each metal pad can be probed by a source meter to thermally tune individual MRR, while the same ground (GND) trace are shared among MRRs to reduce the electrical I/O count.

The detailed design of each MRR is shown in Fig. 1(b), where each MRR consists of a circular waveguide etched to a 90 nm thick pedestal that hosts the dopants. For the purpose of feedback control via in-ring N-doped photoconductive heaters [28, 34], a 10 μm wide N doping section is patterned to follow the MRR, outside of which heavy N⁺⁺ doping is used to make ohmic contacts. Phosphorous dopant concentrations are N: $5 \times 10^{17} \text{cm}^{-3}$ and N⁺⁺: $5 \times 10^{20} \text{cm}^{-3}$ as in [35]. The two MRRs used in our experiment have designed radii of 10.921 μm and 10.937 μm . The 0.15% difference in their radii is to avoid collision of their as-fabricated resonances. Coupling gaps

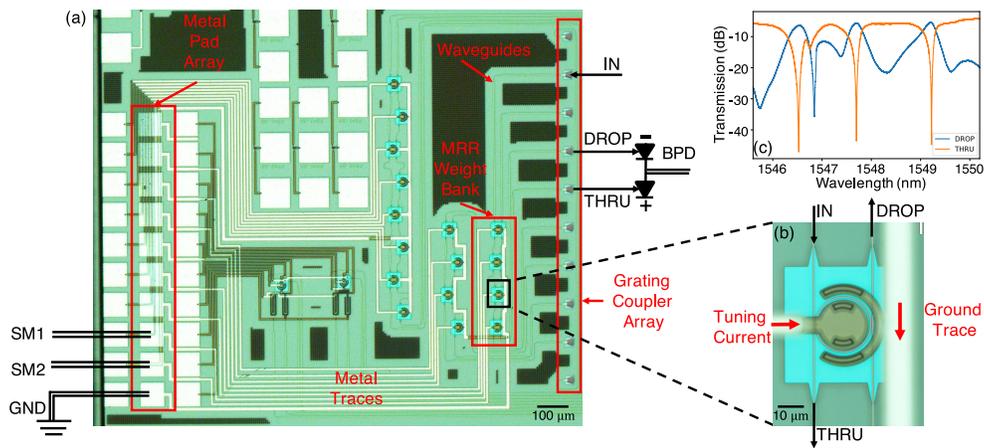


Fig. 1. (a) Overview of the fabricated silicon photonic chip containing a MRR weight bank. The metal pad array (on the left) directs the electrical signal from source meter (SM) to the MRR weight bank (in the middle), which thermally tunes the MRR optical transmission to configure its weight. The grating coupler array (on the right) guides the optical signal into and out of the MRR weight bank. Outputs are eventually directed to complementary ports of an off-chip balanced photo-detector (BPD) to produce an electrical weighted addition output. (b) Zoomed-in micrograph of a MRR with an N-doped in-ring heater. (c) Spectral response of the MRR weight bank measured from both the THRU and DROP ports. The MRRs are close to be critically-coupled with the bus waveguide.

between bus waveguides and MRR are 200 nm, and neighboring MRRs are separated by 108 μm .

We perform both optical and electrical characterizations on the MRR weight bank using an optical spectrum analyzer (Apex AP2440A) and source meter (Keithley 2400). Based on the DROP port spectral response in Fig. 1(c), the resonance peaks (λ_{peak}) of the two MRRs used in this work are at 1547.73 nm and 1549.24 nm, respectively. The full-width half maximum (FWHM) of each resonance peak is 0.28 nm (i.e., 35 GHz), and thus Q factor ($\lambda_{peak}/\text{FWHM}$) is approximately 5500. Free spectral range (FSR) is 8.62 nm for both MRRs, so the finesse (FSR/FWHM) is 30.79. Observing their resonance shift versus applied electrical power, we find the thermal tuning efficiency of MRRs to be 0.15 nm/mW (i.e., 53.3 $\mu\text{W}/\text{GHz}$ or 57.5 mW/FSR).

2.2. Experimental setup

The experimental setup is shown in Fig. 2(a). Two distributed feedback lasers (DFBs) generate optical carriers at 1548.52 nm and 1550.12 nm, respectively. Each optical carrier is modulated at a Mach-Zehnder modulator (MZM) by the data pattern from a pulse pattern generator (PPG, Anritsu MP1763B) and multiplexed together at arrayed waveguide gratings (AWG). These two distinct optical carriers then experience wavelength-dependent channel delays at a fiber Bragg grating (FBG) array to construct partially correlated inputs. The inputs enter the IN port of the MRR weight bank for weighting on the silicon photonic chip, which is mounted on a temperature-controlled fiber alignment stage. The insertion loss is about 5 dB per grating coupler, and 1 dB due to waveguide/bending loss (so round-trip loss is about 10-12 dB). The outputs at THRU and DROP ports of the MRR weight bank are summed by a 20 GHz balanced photo-detector (BPD, Discovery Semiconductors, Inc. DSC-R405ER). The resulting electrical output is recorded by a sampling oscilloscope (Oscope, Tektronix DSA8300). Each N-doped heater embedded in the MRR is driven by a source meter (SM, Keithley 2400) set in current-source, voltage-measure mode. PPG, SMs, and Oscope are all computer-controlled by *lightlab* software [36].

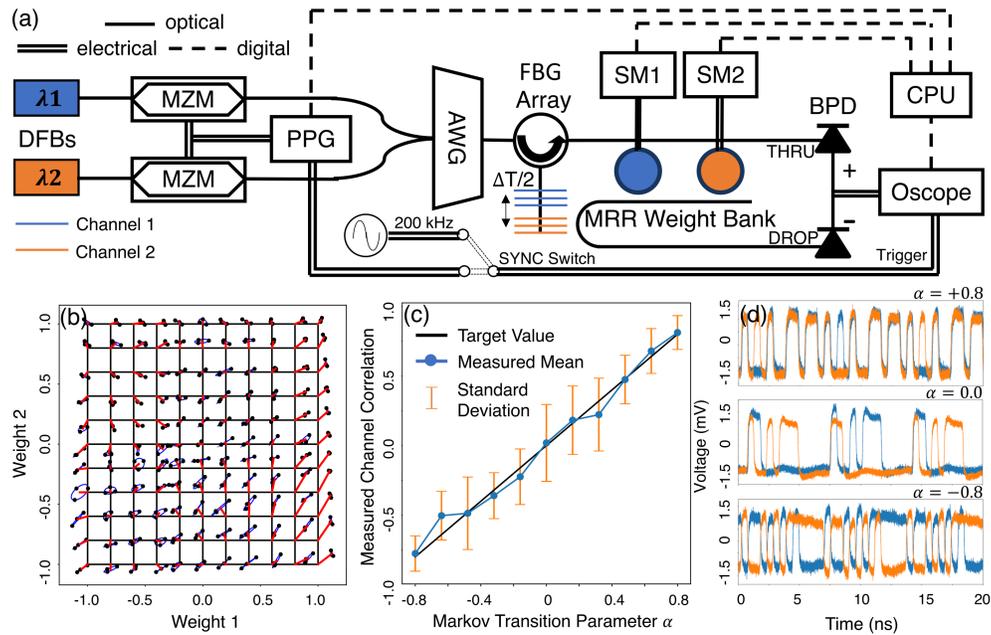


Fig. 2. (a) Schematic of the experimental setup for performing photonic PCA using an on-chip MRR weight bank. DFBs: distributed feedback lasers, MZM: Mach-Zehnder modulator, PPG: pulse pattern generator, AWG: arrayed waveguide gratings, FBG: fiber Bragg grating, SM: source meter, MRR: microring resonator, BPD: balanced photo-detector, Oscilloscope: sampling oscilloscope. The synchronization (SYNC) switch toggles the Oscilloscope triggering state between the repeating pattern from PPG and a free-running clock at 200 kHz. (b) Two-channel weight evaluation results in the same plot format of [30]. Black grid crossings are the target weights. Red lines represent the deviation between the target weights and the mean of measured weights over 3 repetitions. Blue ellipses represent the standard deviation of measured weights over 3 repetitions. (c) Measured channel cross-correlations versus Markov transition parameter α . Black line indicates the target value. Blue dots represent the mean of measured channel correlations over 6 repetitions. Orange error bars represent the standard deviation of measured channel correlations over 6 repetitions. (d) Waveform pairs of partially correlated two-channel signals associated with 3 typical α values.

2.3. MRR weight bank control

We perform MRR weight bank calibration according to the latest feedback control procedure introduced in [28]. Each probing SM sources electrical power to individual MRR to thermally shift its resonance position with respect to the paired laser channel. The resonance offset determines circulating optical power within each MRR, which is partially absorbed and results in a photoresponse. The photoresponse, in turn, affects the conductivity of the in-ring N-doped photoconductive heater, which can be measured by the probing SM at the same time [34]. As a result, sensing the photoresponse electrically provides a straightforward measure of the optical transmission of each MRR, enabling a feedback control loop for configuring MRR weights over a continuous range of $[-1, +1]$.

Weight control accuracy is assessed by comparing weight vectors measured through weight decomposition [28] to target weight vectors. Figure 2(b) shows the two-channel weight evaluation result where we calibrate the MRR weight bank with statistically independent two-channels inputs both carrying a 2 GHz pseudo-random bit sequence (PRBS) generated by PPG. We achieve 5.2 bits (2.7% error) of accuracy (red lines) and 6.3 bits (1.3% error) of precision (blue ellipses). To

the best of the authors' knowledge, this is the best record on multi-channel MRR weight control so far ([28] reported 5.1 bits of accuracy and 5.5 bits of precision). Performance improvements are owed to optimizations in the feedback control procedure and improved stabilization of the chip temperature. The demonstrated accuracy exceeds the 5 bits of resolution commonly used for weights in neuromorphic electronic systems [37].

2.4. Partially correlated two-channel inputs generation

PCA requires a way to generate pairs of signals with a controllable cross-correlation. We program the PPG to produce a data pattern satisfying a first-order Markov process according to the method of [24, 25]. The FBG array then transforms the temporal auto-correlation of the Markov pattern to instantaneous cross-correlation between the two-channel signals. In our experiment, the PPG data rate is set to be $B = 1.9259$ GHz (close to that used to calibrate the MRR weight bank before), and the FBG array introduces a relative delay of $\Delta T = 6.75$ ns between two optical carriers. Their product of $B\Delta T = 13.000$ (meaning the inter-channel delay is an integer multiple of the bit period), so the Markov pattern can be seeded by a 13-bit PRBS. The k^{th} ($k > 13$) bit has a $0.5 + \alpha(b_{k-13} - 0.5)$ probability of being the same as the $(k - 13)^{\text{th}}$ bit, where α is the Markov transition parameter. α determines the overall cross-correlation between the two-channel signals, though each bit sequence initialized for a given α is unique due to the stochastic nature of this generation process. We sweep multiple α values over the range of $[-1, +1]$ (each with 6 pattern initializations) and measure the actual channel correlations. Figure 2(c) illustrates the correspondence between the α values and resulting cross-correlations. The averages of the measured channel correlations (blue dots) are close to the target values (on black line). Among all the α values considered, the mean deviation of the measured mean from the target value is 0.04, while the mean standard deviation is 0.20. Samples of partially correlated two-channel signals associated with 3 typical α values are provided in Fig. 2(d).

2.5. Observability control

To perform and evaluate photonic PCA performance, the signal waveforms must be switched between observable and unobservable states [24]. We use a synchronization (SYNC) switch to swap between two triggering conditions on the Oscilloscope (see Fig. 2(a)). To make waveforms unobservable, the weighted addition output is not synchronized with the Oscilloscope, but with a 200 kHz clock (i.e., the internal clock on Oscilloscope). Under such circumstance, the Oscilloscope is in free-running mode such that the weighted addition output can only be sampled at a sub-Nyquist rate (i.e., 200 kS/s here). We will show that our photonic PCA is feasible even though we have no useful waveform information of the weighted addition output but solely based on its statistical properties. When waveform information is needed for the purpose of validation, the Oscilloscope is synchronized with the PPG in such way that the Oscilloscope is triggered by the repeating pattern from the PPG to synthesize the accurate waveform at a super-Nyquist sampling rate (16 GS/s here). These cases include measuring the above channel correlation or implementing the software-based PCA algorithm to obtain the ground truth.

3. PCA algorithm

Given the above constrained observability, we hereby propose a novel PCA algorithm that can be fully exploited by an MRR weight bank. We notice the weighted addition operation matches the exact intention of PCA that looks for the projection direction upon which the variance, i.e., the 2nd-order (central) moment, of the projected data is maximized. We therefore borrow the projection pursuit concept [22] to search for the direction of PC vectors by continuously updating the weight vector (set by the MRR weight bank) to the largest variance of weighted addition output in a given subspace. We describe the detailed PCA algorithm as Algorithm 1 in Appendix A, whose main body is a Nelder-Mead method [38] under the constraint that all the weight vectors

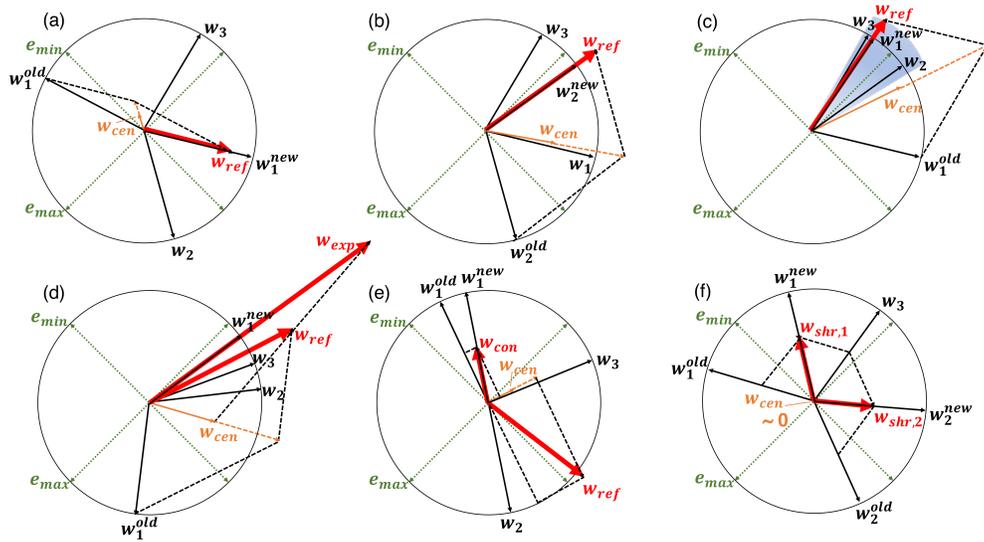


Fig. 3. Schematic illustrations of weight vector updates involved in converging onto the 1st PC vector in a 2-D subspace. Three normalized weight vectors \mathbf{w}_1 , \mathbf{w}_2 , \mathbf{w}_3 are used in this visualization. Their centroid is $\mathbf{w}_{cen} = (\mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3)/3$. Unit vectors \mathbf{e}_{max} , \mathbf{e}_{min} represent the 1st, 2nd PC vector directions, respectively. (a)-(c) show how continuous reflections $\mathbf{w}_{ref} = \mathbf{w}_{cen} + 1(\mathbf{w}_{cen} - \mathbf{w}_{min})$ have three weight vectors converge into a cone near \mathbf{e}_{max} . (d)-(f) show how other operations handle special cases of the constrained N-M to improve convergence: (d) expansion $\mathbf{w}_{exp} = \mathbf{w}_{cen} + 2(\mathbf{w}_{ref} - \mathbf{w}_{cen})$, (e) contraction $\mathbf{w}_{con} = \mathbf{w}_{cen} + 0.5(\mathbf{w}_{min} - \mathbf{w}_{cen})$, and (f) shrinkage $\mathbf{w}_i = \mathbf{w}_{max} + 0.5(\mathbf{w}_i - \mathbf{w}_{max})$ for $i = 1, 2, 3$. See Appendix A for code implementation of the PCA algorithm.

used have unit norm (hence referred to as “constrained N-M”). This constraint is essential in a sense that it is undesirable for the weight vector magnitude to influence the variance measurement of weighted addition output, and weight normalization (Line 6, 18, 23, 31, 37) is imposed to make sure this constraint is guaranteed throughout the algorithm execution.

The constrained N-M algorithm begins by initializing multiple random weight vectors distributed on a unit sphere. Then, each iteration updates the weight vector associated with the smallest weighted addition output variance \mathbf{w}_{min} , until all the weight vectors pointing in nearly the same direction at which point the algorithm is converged. There are four types of update operations: reflection, expansion, contraction, and shrinkage. We illustrate in Fig. 3 how these four operations update \mathbf{w}_{min} and control the convergence of the algorithm. In most cases, as shown in Figs. 3(a)-3(c), reflection can improve \mathbf{w}_{min} by simply reversing its direction with respect to the centroid of all weight vectors \mathbf{w}_{cen} . In some other special cases, improvements can be made beyond a simple reflection. For example, Fig. 3(d) shows a further expansion along the reflection direction may reach an even better weight vector direction; Fig. 3(e) shows a contraction of \mathbf{w}_{min} with respect to \mathbf{w}_{cen} can instead find a better alternative when reflection does not lead to a better weight vector direction than current \mathbf{w}_{min} . If none of reflection, expansion, and contraction improve the variance at \mathbf{w}_{min} , then a shrinkage operation, shown in Fig. 3(f), can be applied as a last resort (shrinkage often has an impact when \mathbf{w}_{cen} is negligible). We achieve good convergence if the initial simplex composed by weight vectors is large, and if we pick the following parameters: reflection coefficient $\beta \sim 1$, expansion coefficient $\gamma \sim 2$, contraction coefficient $\rho \sim 0.5$, and shrinkage coefficient $\sigma \sim 0.5$ (see Appendix B on how we obtain these hyperparameters). The algorithm proceeds until reaching the termination condition, which is controlled by the convergence tolerance ϵ gauging the proximity of weight vectors.

4. Results

4.1. Experimental demonstration

We first present the experimental results on two-channel photonic PCA to empirically validate the feasibility of the proposed approach. We consider 10 channel correlation values over the range of $[-1, +1]$ for the partially correlated two-channel inputs to the MRR weight bank. We first run the PCA algorithm to find the PC vectors in free-running mode where SYNC switches to 200 kHz clock, and then apply those obtained PC vectors to record the corresponding PC waveforms in synchronized mode where SYNC switches to PPG. The obtained waveforms of both the 1st and 2nd PCs associated with various α values are shown in Fig. 4. The red curves represent the measured outcome of the photonic procedure, and the black curves represent the calculated PCs via complete waveform digitization of the inputs and off-line software-based singular value decomposition (SVD). The high correspondence between these two curves evidences the effectiveness of our approach. Some short transients observed at the time scale of ps are mainly due to the small channel-dependent timing differences, which the preceding inter-channel delay (of the FBG array) does not take into account.

Table. 1 summarizes the quantitative analysis of photonic PCA performance to show the robustness of the proposed approach. Here, for the same input corresponding to a given α value, we run the algorithm 6 times (all with newly-initialized random weight vectors). We are interested in the average root-mean-squared error (RMSE) between measured and calculated PCs among all 6 runs to quantify accuracy, as well as the standard deviation (SD) of measured PCs among all 6 runs to quantify repeatability. All obtained values are normalized to the RMS intensity of the corresponding calculated PC. On one hand, the 1st PC has an average 5.36% RMSE, and the 2nd PC has an average 8.43% RMSE. On the other hand, the 1st PC has an average 2.82% SD, and the 2nd PC has an average 5.76% SD. Therefore, our photonic PCA scheme achieves consistently high accuracy and low repeatability error throughout the range of α values considered here. It is worth mentioning that we obtain PCs with higher accuracy than those reported in [24] where the principal component vectors are derived from fitting the projection-moment vs. projection-angle model. Hence, any error involved in the process of fitting

Table 1. Photonic PCA Performance Summary

α	1 st PC		2 nd PC	
	RMSE (%)	SD (%)	RMSE (%)	SD (%)
-0.80	3.54	1.31	7.36	5.89
-0.64	3.93	1.57	6.52	8.82
-0.48	3.40	1.25	9.11	3.42
-0.32	4.27	1.01	8.42	4.03
-0.16	8.03	6.34	11.86	9.57
+0.80	4.19	2.44	6.02	6.43
+0.64	5.28	4.09	7.47	5.14
+0.48	6.72	2.88	6.92	5.27
+0.32	6.81	3.97	8.97	2.45
+0.16	7.47	3.31	11.63	6.53
AVE	5.36	2.82	8.43	5.76

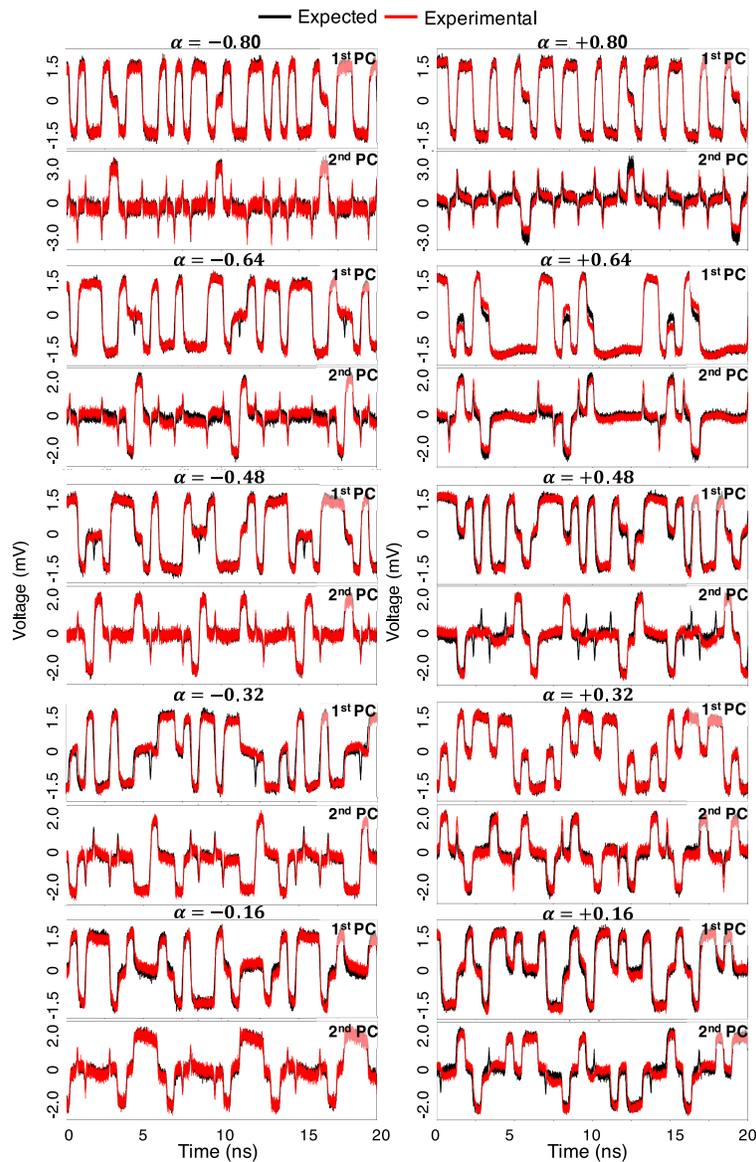


Fig. 4. Experimental PC waveforms associated with 10 typical Markov transition parameters, α , over the range of $[-1, 1]$. Those found by the proposed photonic PCA procedure (red curves) reproduce those predicted with the complete knowledge of the inputs (black curves).

the model parameters will later contribute to the error in the computed principal component vectors. In contrast, our approach is more straightforward by directly converging to the principal component vectors, which does not require estimation on any model parameters.

There are multiple error sources that contribute together to the PC accuracy error, some of which can be mitigated to provide additional improvement of the above results. One particular example is the imperfect calibration of the MRR weight bank that introduces errors to weight values used in the proposed PCA algorithm. It results mainly from inaccurate calibration model and thermal cross-talk between neighboring MRRs, and further enhancement on the weight

value accuracy may come from better MRR/chip design (e.g., higher Q factor/finesse and better thermal isolation) or taking advantage of non-thermal tuning approaches such as the one through carrier-depletion effect [39]. Some of other errors are to some extent inevitable to our scheme. Instances include the optical noise generated at multiple stages (e.g., DFBs, MZMs) to the correlated inputs, and the electrical noise produced by the BPD. Improvements on these factors would require better hardware/system implementations.

Another issue worth discussion here is that we notice an error increment from the 1st PC to the 2nd PC. We recognize this as being consistent with the nature of learning PCs in the sequential order. Prior works [25, 26, 31] did not exhibit this problem because they only show the convergence to the 1st PC. We also acknowledge the possibility of resolving this problem by reducing the tolerance parameter ϵ used in the proposed PCA algorithm. However, we need to point out a trade-off between the convergence of PCs and the error accumulations among them. A smaller ϵ results in more accurate PC results (based on our observation), but it also requires a longer convergence time. Factors like this should be taken into account when applying the proposed approach for real-time online PCA applications.

4.2. Numerical simulation

The above experimental demonstration of two-channel photonic PCA illustrates the effectiveness of the proposed algorithm for proof-of-concept purpose. We restrict ourselves to a two-channel inputs not because our method cannot be extended to higher dimensions, but because of the implementation cost that the number of required discrete DFBs, MZMs, SMs all scale linearly with the number of channels present. We can overcome these limitations via ongoing efforts that aim to integrate every components aforementioned onto the same chip [40, 41], though we do

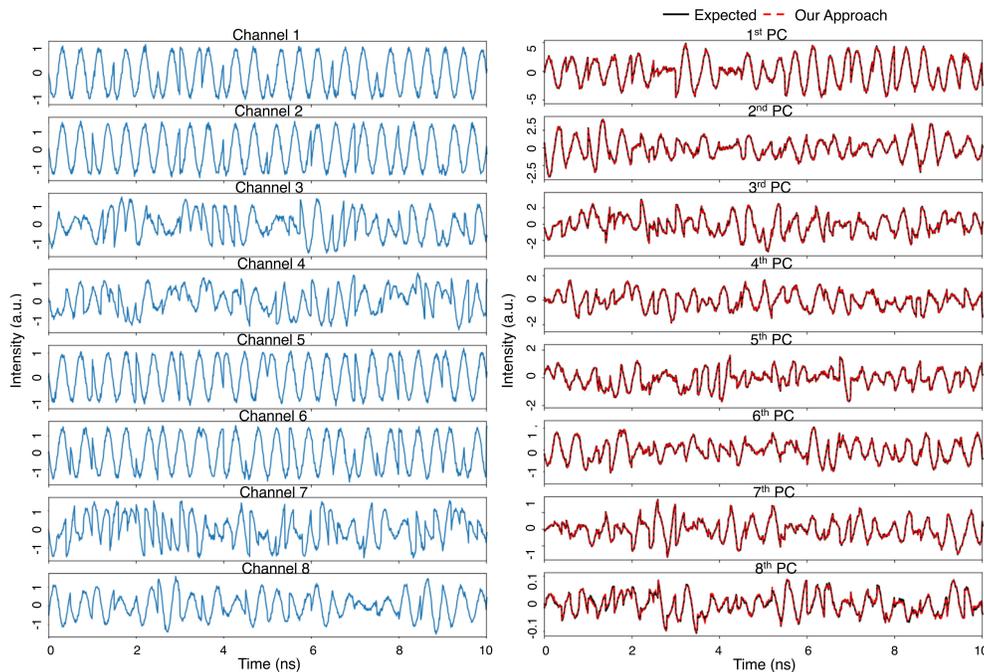


Fig. 5. Numerical demonstration of the proposed PCA algorithm in higher dimensions. Left: eight-channel inputs that are partially correlated. Right: eight PCs obtained by our approach based on the weight addition output (red curves), matching those achieved with the complete knowledge of the inputs (black curves).

face a suite of technical challenges especially the one aiming for hybrid platforms interconnecting both III-V materials and silicon [42, 43].

Here, we demonstrate through numerical simulation that the proposed photonic PCA approach can generalize up to eight-channels. For such purpose, deflationary orthogonalization is adopted to ensure PC vectors are orthogonal to each other. Since the operations in the constrained N-M algorithm are all linear, the Gram-Schmidt process (Line 4 of Algorithm 1) only needs to be performed once so that the weight vectors used to find the k^{th} PC vector \mathbf{e}_k will remain in the orthogonal subspace to all the previously found PC vectors $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$. The left column of Fig. 5 shows eight-channel correlated inputs; the right column shows the resulting eight PCs. The red curves indicate the PCs derived by our PCA algorithm running solely on the 2^{nd} -order moment information of the weighted addition output. The black curves indicate the PCs achieved by standard SVD method based on the complete knowledge of the inputs. The successful convergence on all eight PCs illustrates that the proposed algorithm can be generalized to high dimensions, and is fully compatible with the online learning objective where all the PC vectors can be found sequentially.

We make three additional observations in our numerical simulation of k-dimensional PCA. First, the larger k tends towards, the more difficult it is to converge to the lower-order PCs. The reason is intuitive since we are trying to locate one optimal direction in a k-dimensional subspace such that the effective search space grows exponentially with k. Second, the number of weight vectors needed at a given iteration of the constrained N-M algorithm is crucial to be able to ensure a good coverage of the multi-dimensional space. We recommend the number of weight vectors to be selected and optimized in the range of $[k, 2k]$. Less than k weight vectors will potentially leave certain dimensions being neglected in the initialization process, while more than 2k weight vectors will lead to too much probe cost in each iteration. Third, our approach thus presents a substantial improvement over the method proposed in [24] where the maximal weighted addition output variance is found by sweeping across the complete k-dimensional subspace for model-based fitting. Such a practice is not scalable for high-dimensional analog photonic processing simply because its number of probes scales with the power of k, in contrast to our approach where the number of probes scales linearly with k.

We acknowledge there exist limitations to generalize the proposed scheme to high dimensions. One source of scaling limitation stems from photonic weight bank hardware. Ref. [14] has identified that a) the number of channels N is limited by resonator finesse F , and a penalty $\delta w = F/N$ related to the ability of independently tuning neighboring MRR channels is determined to be 3.41-4.61 (for 1-pole MRR) and 1.2 (for 2-pole MRR in [44]); and that b) the aggregated bandwidth Nf is limited by resonator FSR , and a distortion penalty $D_{wb} = FSR/Nf$ is measured to be 4.3 in [45]. Given these limitations, for platforms containing MRRs whose Q-factors~10-20k and finesse~133, the number of channels that can be potentially supported scales up to 32 (for 1-pole MRR weight bank) and 110 (for 2-pole MRR weight bank). These limitations are imposed mainly because we construct our network via WDM, while there are other forms of multiplexing, such as polarization-division multiplexing, that can potentially relieve these limitations. Another source of scaling limitation could stem from the sequential learning aspect of the proposed photonic PCA approach. As we have demonstrated in our experiment, there is about 3% error increment from the 1^{st} PC to the 2^{nd} PC. Assuming a linear error accumulation trend in higher order learning process, we can expect the eventual error for the k^{th} PC to be approximately $(0.03 \times k)$. Depending on the exact accuracy requirement for different applications, this definitely sets a limitation on the number of PCs that can be faithfully extracted. Thus, future research directions should focus on not only improving the accuracy of individual PCs, but also reducing the error transfer from lower-order PCs to higher-order ones.

5. Conclusion

In this paper, we used an on-chip MRR weight bank to project the high-dimensional signals onto a few-dimensional representation (i.e., weighted addition operation), and set the accuracy/precision record for configuring the MRR weight bank through the feedback control procedure. Our prototype implementation produced inputs under controllable correlation conditions, and more importantly, performed the photonic PCA assuming that we are ignorant of the real-time waveform information of the correlated inputs and weighted addition output. We overcome this observability constraint via a novel PCA algorithm that works with the pure observation of the 2^{nd} -order moment of the weighted addition output. The experimental demonstration of two-channel photonic PCA achieved an average 5.36% RMSE, 2.82% SD for the 1st PC, and 8.43% RMSE, 5.76% SD for the 2nd PC. The numerical simulation on eight-channel PCA further corroborated that our scheme can be generalized to high-dimensional scenarios. Extensions to this work could include a high-dimensional experimental demonstration, larger-scale integration, and convergence acceleration.

Ultimately, this wideband multivariate technique could open new domains for high speed information processing in future communication and computing systems. In particular, it may pave the way for advanced blind source separation (BSS) in multi-antenna systems whose performance can greatly exceed current approaches based on digitizing ultra-redundant multi-dimensional signals [46]. Serving as a pre-processing technique, the proposed photonic PCA will lay a solid foundation for high-performance independent component analysis (ICA) while relieving stringent DSP requirements (e.g., channel filtering, signal digitization, data storage, bandwidth tunability, etc.) in wideband RF systems.

Appendix A PCA algorithm

See Algorithm 1.

Appendix B Hyperparameter optimization

We describe here how we optimized the hyperparameters used in the above PCA algorithm to achieve the best performance in our two-channel photonic PCA experiment. Figure 6 shows measured PC accuracy against reflection coefficient β , expansion coefficient γ , and contraction coefficient ρ . The PC accuracy for each reflection coefficient is an average over 6 runs (as reflection operation is called almost every time), while the PC accuracy for each expansion and contraction coefficient is an average over 3 runs (since expansion and contraction operations are less likely to be called compared to reflection operation). We did not obtain the result for

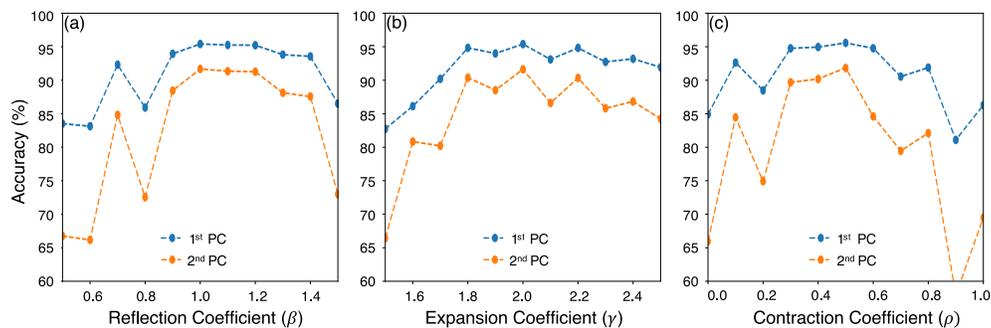


Fig. 6. Measured PC accuracy versus reflection coefficient β , expansion coefficient γ , and contraction coefficient ρ in the two-channel photonic PCA experiment.

Algorithm 1: PCA algorithm

Input: Weighted addition output \mathbf{y} , $1^{st}, \dots, (k-1)^{th}$ PC vectors $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$;
Parameters: Convergence tolerance ϵ , reflection coefficient β , expansion coefficient γ , contraction coefficient ρ , shrinkage coefficient σ ;
Output: k^{th} PC vector \mathbf{e}_k ;

- 1 Initialize n random weight vectors $\mathbf{w}_i (i = 1, \dots, n)$;
- 2 **for** $i = 1 : n$ **do**
- 3 **if** $k > 1$ **then**
- 4 Orthogonalize \mathbf{w}_i such that $\mathbf{w}_i = \mathbf{w}_i - \sum_{j=1}^{k-1} (\mathbf{w}_i^T \mathbf{e}_j) \mathbf{e}_j$;
- 5 **end if**
- 6 Normalize \mathbf{w}_i such that $|\mathbf{w}_i| = 1$;
- 7 Apply \mathbf{w}_i to the weight bank and obtain the resulted $\mathbf{y}(\mathbf{w}_i)$;
- 8 Compute the variance of \mathbf{y} : $f(\mathbf{w}_i) = E(\mathbf{y}^2)$;
- 9 **end for**
- 10 Among $f(\mathbf{w}_i) (i = 1, \dots, n)$, find $f_{max} = \max(f(\mathbf{w}_i))$, $f_{min} = \min(f(\mathbf{w}_i))$;
- 11 Set $\mathbf{w}_{max} = \operatorname{argmax}(f(\mathbf{w}_i))$, $\mathbf{w}_{min} = \operatorname{argmin}(f(\mathbf{w}_i))$;
- 12 **if** $|f_{max} - f_{min}| \leq \epsilon$ **then**
- 13 **return** $\mathbf{e}_k = \mathbf{w}_{max}$;
- 14 **end if**
- 15 **while** $|f_{max} - f_{min}| > \epsilon$ **do**
- 16 Compute the centroid of all weight vectors except \mathbf{w}_{min} : $\mathbf{w}_{cen} = E(\mathbf{w}_i)$ (for $\mathbf{w}_i \neq \mathbf{w}_{min}$);
- 17 Compute the reflected weight vector: $\mathbf{w}_{ref} = \mathbf{w}_{cen} + \beta(\mathbf{w}_{cen} - \mathbf{w}_{min})$;
- 18 Normalize \mathbf{w}_{ref} such that $|\mathbf{w}_{ref}| = 1$;
- 19 **if** $f_{min} < f(\mathbf{w}_{ref}) < f_{max}$ **then**
- 20 $\mathbf{w}_{min} = \mathbf{w}_{ref}$ and update f_{min} , \mathbf{w}_{min} ;
- 21 **else if** $f(\mathbf{w}_{ref}) \geq f_{max}$ **then**
- 22 Compute the expanded weight vector: $\mathbf{w}_{exp} = \mathbf{w}_{cen} + \gamma(\mathbf{w}_{ref} - \mathbf{w}_{cen})$;
- 23 Normalize \mathbf{w}_{exp} such that $|\mathbf{w}_{exp}| = 1$;
- 24 **if** $f(\mathbf{w}_{ref}) < f(\mathbf{w}_{exp})$ **then**
- 25 $\mathbf{w}_{max} = \mathbf{w}_{exp}$ and update $f_{max} = f(\mathbf{w}_{exp})$;
- 26 **else**
- 27 $\mathbf{w}_{max} = \mathbf{w}_{ref}$ and update $f_{max} = f(\mathbf{w}_{ref})$;
- 28 **end if**
- 29 **else if** $f(\mathbf{w}_{ref}) \leq f_{min}$ **then**
- 30 Compute the contracted weight vector: $\mathbf{w}_{con} = \mathbf{w}_{cen} + \rho(\mathbf{w}_{min} - \mathbf{w}_{cen})$;
- 31 Normalize \mathbf{w}_{con} such that $|\mathbf{w}_{con}| = 1$;
- 32 **if** $f(\mathbf{w}_{con}) > f(\mathbf{w}_{min})$ **then**
- 33 $\mathbf{w}_{min} = \mathbf{w}_{con}$ and update f_{min} , \mathbf{w}_{min} ;
- 34 **else**
- 35 **for** $i = 1 : n$ **do**
- 36 Replace all weight vectors with $\mathbf{w}_i = \mathbf{w}_{max} + \sigma(\mathbf{w}_i - \mathbf{w}_{max})$;
- 37 Normalize \mathbf{w}_i such that $|\mathbf{w}_i| = 1$;
- 38 **end for**
- 39 Update f_{max} , f_{min} , and \mathbf{w}_{max} , \mathbf{w}_{min} ;
- 40 **end if**
- 41 **end if**
- 42 **end while**
- 43 **return** $\mathbf{e}_k = \mathbf{w}_{max}$.

shrinkage operation because it is rarely called (but shrinkage coefficient $\sigma = 0.5$ generally performs well in later test). These coefficients turn out to have a range of values that can result in high PC accuracy, while those mentioned in section 3 result in the highest PC accuracy. We obtained these results with inputs whose Markov transition parameter is $\alpha = -0.32$, and they provided consistent photonic PCA performance when later applied to all Markov transition parameters considered.

Funding

National Science Foundation (NSF) (ECCS-1642962).

Acknowledgment

Fabrication support was provided via the Natural Sciences and Engineering Research Council of Canada (NSERC) Silicon Electronic-Photonic Integrated Circuits (SiEPIC) Program and the Canadian Microelectronics Corporation (CMC). Devices were fabricated at the Institute of Microelectronics (IME) A*STAR foundry in Singapore.

References

1. R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.* **17**, 539–550 (1999).
2. T. Sundstrom, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed nyquist analog-to-digital converters," *IEEE Trans. Circuits Syst. I, Reg. Papers* **56**, 509–518 (2009).
3. P. Jebashini, R. Uma, P. Dhavachelvan, and H. K. Wye, "A survey and comparative analysis of multiply-accumulate (mac) block for digital signal processing application on asic and fpga," *J. Appl. Sci.* **15**, 934–946 (2015).
4. A. J. Seeds and K. J. Williams, "Microwave photonics," *J. Lightw. Technol.* **24**, 4628–4641 (2006).
5. J. Yao, "Microwave photonics," *J. Lightw. Technol.* **27**, 314–335 (2009).
6. J. Capmany, J. Mora, I. Gasulla, J. Sancho, J. Lloret, and S. Sales, "Microwave photonic signal processing," *J. Lightw. Technol.* **31**, 571–586 (2013).
7. W. Liu, M. Li, R. S. Guzzon, E. J. Norberg, J. S. Parker, M. Lu, L. A. Coldren, and J. Yao, "A fully reconfigurable photonic integrated signal processor," *Nat. Photon.* **10**, 190 (2016).
8. N. C. Harris, D. Bunandar, M. Pant, G. R. Steinbrecher, J. Mower, M. Prabhu, T. Baehr-Jones, M. Hochberg, and D. Englund, "Large-scale quantum photonic circuits in silicon," *Nanophotonics* **5**, 456–468 (2016).
9. D. Pérez, I. Gasulla, L. Crudgington, D. J. Thomson, A. Z. Khokhar, K. Li, W. Cao, G. Z. Mashanovich, and J. Capmany, "Multipurpose silicon photonics signal processor core," *Nat. Commun.* **8**, 636 (2017).
10. P. R. Prucnal and B. J. Shastri, *Neuromorphic Photonics* (CRC Press, 2017).
11. A. N. Tait, T. F. Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Neuromorphic photonic networks using silicon photonic weight banks," *Sci. Rep.* **7**, 7430 (2017).
12. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nat. Photon.* **11**, 441 (2017).
13. B. J. Shastri, A. N. Tait, T. Ferreira de Lima, M. A. Nahmias, H.-T. Peng, and P. R. Prucnal, "Principles of neuromorphic photonics," *Unconventional Computing: A Volume in the Encyclopedia of Complexity and Systems Science, Second Edition* pp. 83–118 (2018).
14. A. N. Tait, A. X. Wu, T. F. de Lima, E. Zhou, B. J. Shastri, M. A. Nahmias, and P. R. Prucnal, "Microring weight banks," *IEEE J. Sel. Top. Quantum Electron.* **22**, 312–325 (2016).
15. J. Komma, C. Schwarz, G. Hofmann, D. Heinert, and R. Nawrodt, "Thermo-optic coefficient of silicon at 1550 nm and cryogenic temperatures," *Appl. Phys. Lett.* **101**, 041905 (2012).
16. A. N. Tait, T. Ferreira de Lima, M. A. Nahmias, H. B. Miller, H.-T. Peng, B. J. Shastri, and P. R. Prucnal, "A silicon photonic modulator neuron," *arXiv preprint arXiv:1812.11898* (2018).
17. M. H. Khan, H. Shen, Y. Xuan, L. Zhao, S. Xiao, D. E. Leaird, A. M. Weiner, and M. Qi, "Ultrabroad-bandwidth arbitrary radiofrequency waveform generation with a silicon photonic chip-based spectral shaper," *Nat. Photon.* **4**, 117–122 (2010).
18. J. Wang, H. Shen, L. Fan, R. Wu, B. Niu, L. T. Varghese, Y. Xuan, D. E. Leaird, X. Wang, F. Gan, A. M. Weiner, and M. Qi, "Reconfigurable radio-frequency arbitrary waveforms synthesized in a silicon photonic chip," *Nat. Commun.* **6**, 5957 (2015).
19. Q. Xu and R. Soref, "Reconfigurable optical directed-logic circuits using microresonator-based optical switches," *Opt. Express* **19**, 5244–5259 (2011).
20. T. V. Vaerenbergh, M. Fiers, P. Mechet, T. Spuesens, R. Kumar, G. Morthier, B. Schrauwen, J. Dambre, and P. Bienstman, "Cascadable excitability in microrings," *Opt. Express* **20**, 20292–20308 (2012).
21. A. N. Tait, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Broadcast and weight: an integrated network for scalable photonic spike processing," *J. Lightw. Technol.* **32**, 3427–3439 (2014).

22. I. T. Jolliffe, *Principal Component Analysis* (Springer-Verlag, 2011).
23. A. N. Tait, T. F. de Lima, P. Y. Ma, M. P. Chang, M. A. Nahmias, B. J. Shastri, P. Mittal, and P. R. Prucnal, "Blind source separation in the physical layer," in "52nd Annual Conference on Information Sciences and Systems (CISS)," (IEEE, 2018), pp. 1–6.
24. A. N. Tait, P. Y. Ma, T. Ferreira de Lima, E. C. Blow, M. P. Chang, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Demonstration of multivariate photonics: blind dimensionality reduction with analog integrated photonics," arXiv preprint arXiv:1903.03474 (2019).
25. A. N. Tait, J. Chang, B. J. Shastri, M. A. Nahmias, and P. R. Prucnal, "Demonstration of wdm weighted addition for principal component analysis," *Opt. Express* **23**, 12758–12765 (2015).
26. T. F. de Lima, A. N. Tait, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Scalable wideband principal component analysis via microwave photonics," *IEEE Photonics Journal* **8**, 1–9 (2016).
27. E. Oja, "Simplified neuron model as a principal component analyzer," *J. Math. Biol.* **15**, 267–273 (1982).
28. A. N. Tait, H. Jayatilleka, T. F. D. Lima, P. Y. Ma, M. A. Nahmias, B. J. Shastri, S. Shekhar, L. Chrostowski, and P. R. Prucnal, "Feedback control for microring weight banks," *Opt. Express* **26**, 26422–26443 (2018).
29. A. N. Tait, T. F. de Lima, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Continuous calibration of microring weights for analog optical networks," *IEEE Photon. Technol. Lett.* **28**, 887–890 (2016).
30. A. N. Tait, T. F. de Lima, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Multi-channel control for microring weight banks," *Opt. Express* **24**, 8895–8906 (2016).
31. A. N. Tait, T. F. de Lima, A. X. Wu, E. Zhou, M. A. Nahmias, B. J. Shastri, M. P. Chang, and P. R. Prucnal, "Silicon microring weight banks for multivariate rf photonics," in "IEEE Photonics Conference (IPC)," (2016), pp. 11–12.
32. A. E.-J. Lim, J. Song, Q. Fang, C. Li, X. Tu, N. Duan, K. K. Chen, R. P.-C. Tern, and T.-Y. Liow, "Review of silicon photonics foundry efforts," *IEEE J. Sel. Top. Quantum Electron.* **20**, 405–416 (2014).
33. Y. Wang, J. Flueckiger, C. Lin, and L. Chrostowski, "Universal grating coupler design," in "Proc. SPIE 8915, 89150Y," (2013).
34. H. Jayatilleka, K. Murray, M. Á. Guillén-Torres, M. Caverley, R. Hu, N. A. Jaeger, L. Chrostowski, and S. Shekhar, "Wavelength tuning and stabilization of microring-based filters using silicon in-resonator photoconductive heaters," *Opt. Express* **23**, 25084–25097 (2015).
35. T. Baehr-Jones, R. Ding, A. Ayazi, T. Pinguet, M. Streshinsky, N. Harris, J. Li, L. He, M. Gould, Y. Zhang, A. E.-J. Lim, T.-Y. Liow, S. H.-G. Teo, G.-Q. Lo, and M. Hochberg, "A 25 gb/s silicon photonics platform," arXiv preprint arXiv:1203.0767 (2012).
36. A. Tait, T. F. de Lima, P. Y. Ma, A. Jha, H.-T. Peng, H. Miller, and P. R. Prucnal, "lightwave-lab/lightlab: Version 1.0.5," (2018). DOI: 10.5281/zenodo.1436917.
37. F. Akopyan, J. Sawada, A. Cassidy, R. Alvarez-Icaza, J. Arthur, P. Merolla, N. Imam, Y. Nakamura, P. Datta, G.-J. Nam, B. Taba, M. Beakes, B. Brezzo, J. Kuang, R. Manohar, W. Risk, B. Jackson, and D. Modha, "Truenorth: Design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.* **34**, 1537–1557 (2015).
38. J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–313 (1965).
39. P. Dong, S. Liao, D. Feng, H. Liang, D. Zheng, R. Shafiqi, C.-C. Kung, W. Qian, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, "Low vpp, ultralow-energy, compact, high-speed silicon electro-optic modulator," *Opt. Express* **17**, 22484–22490 (2009).
40. J. S. Orcutt, B. Moss, C. Sun, J. Leu, M. Georgas, J. Shainline, E. Zraggen, H. Li, J. Sun, M. Weaver, S. Urošević, M. Popović, R. J. Ram, and V. Stojanović, "Open foundry platform for high-performance electronic-photonics integration," *Opt. Express* **20**, 12222–12232 (2012).
41. Y. Arakawa, T. Nakamura, Y. Urino, and T. Fujita, "Silicon photonics for next generation system integration platform," *IEEE Commun. Mag.* **51**, 72–77 (2013).
42. G. Roelkens, L. Liu, D. Liang, R. Jones, A. Fang, B. Koch, and J. Bowers, "Iii-v/silicon photonics for on-chip and intra-chip optical interconnects," *Laser Photonics Rev.* **4**, 751–779 (2010).
43. M. J. Heck, J. F. Bauters, M. L. Davenport, J. K. Doylend, S. Jain, G. Kurczveil, S. Srinivasan, Y. Tang, and J. E. Bowers, "Hybrid silicon photonic integrated circuit technology," *IEEE J. Sel. Top. Quantum Electron.* **19**, 6100117 (2013).
44. A. N. Tait, A. X. Wu, T. F. de Lima, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Two-pole microring weight banks," *Opt. Lett.* **43**, 2276–2279 (2018).
45. A. N. Tait, T. Ferreira de Lima, M. P. Chang, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Application regime and distortion metric for multivariate rf photonics," in "IEEE Optical Interconnects Conference (OI)," (2017), pp. 25–26.
46. Tuan-Do-Hong and P. Russer, "Signal processing for wideband smart antenna array applications," *IEEE Microw. Mag.* **5**, 57–67 (2004).