Improved Spectral Sensing in Cognitive Radios Using Photonic-Based Principal Component Analysis

Thomas Ferreira de Lima*, Alexander N. Tait, Mitchell A. Nahmias,
Bhavin J. Shastri and Paul R. Prucnal
Department of Electrical Engineering
Princeton University
Princeton NJ 08544, USA
*tlima@princeton.edu

***Abstract*—We propose and experimentally demonstrate a microwave photonic system that iteratively performs principal component analysis on partially correlated, 8-channel, 13 Gbaud signals. The system that is presented is able to adapt to oscillations in interchannel correlations and follow changing principal components. The system provides advantages in bandwidth performance and fan-in scalability that are far superior to electronic counterparts. Wideband, multidimensional techniques are relevant to >10 GHz cognitive radio systems and could bring solutions for intelligent radio communications and information sensing, including spectral sensing.**

Keywords—Microwave Photonics; Analog Signal Processing; RF Photonics

I. Introduction

Principal component analysis (PCA) is a well-known technique for pattern recognition and dimensionality reduction for multidimensional random variables. It is the basis of many statistical analyses that rely on multivariate correlations [1]. We interest ourselves in the spatial correlation of multiple time series, particularly RF signals. PCA and its cousin statistical procedures, such as independent component analysis (ICA), are a fundamental part of the solution to the blind source separation (BSS) problem in RF signal processing [2], which is a useful technique for spectrum sensing in cognitive radio systems [3]. The multi-antenna system design used to solve BSS bears a fatal challenge to digital signal processing: the marriage between high bandwidth and large fan-in. The more antennas and the higher the bandwidth requirement in the system, the higher the requirements for clock speed and memory of digital processors. Microwave photonics (MWP) is the enabling technology for the processing of radio frequency (RF) signals, in particular the part of the spectrum belonging to SHF (3–30 GHz) and EHF (30–300 GHz) as classified by the International Telecommunications Union (ITU). A MWP implementation of PCA significantly relaxes the hardware constraints of digital signal processing systems.

Furthermore, microwave photonics brings additional advantages relative to electronic microwave filters, such as low loss, high bandwidth, tunability and immunity to electromagnetic interference. Because fast response times are inherent to photonics, MWP has found fruitful applications in arbitrary waveform generation, chirped microwave pulse generation, microwave differentiators and real-time operations over microwave signals [4], but few present the advantages of tunability or wavelength-division multiplexing (WDM). Recent work [5] demonstrated iterative principal component extraction of partially correlated 8-channel 1 Gbaud input. The present work extends the design in [5] to handle electronic signals up to 13 Gbaud, bringing signal processing capabilities to the SHF radio spectrum, which include wireless LAN (local area network), radar, communication satellites and television broadcasting. More importantly, WDM allows the fan-in to be as scalable as the number of channels that can be multiplexed into an optical fiber.

II. Methods

PCA linearly transforms a set of N variables $x_i(t)$, with $i = 1,...,N$, into a set of principal components (PCs) devoid of second-order correlations, i.e. $\langle \tilde{x_i} \cdot \tilde{x_j} \rangle_t = 0$ where $\langle \, . \, \rangle_t$ is a time average. The present work extends the experiment design in [5] to allow for 7-dimensional control of interchannel correlations. It is very similar to a finite impulse response (FIR) microwave photonic filter (cf. [4]), but with a wavelength-dependent tunable weight bank (Fig. 1). An electric non-return-to-zero (NRZ) signal with several GHz bandwidth is constructed using a single pulse-pattern generator (PPG) carrying a programmable 8192-long bit pattern.

This signal is imprinted in 16 wavelength carriers via a Mach Zehnder modulator (MZM), producing complimentary modulations at each arm. 16 wavelengths are necessary to encode 8 channels with two polarities (positive and negative) (see Fig. 2(a)). After modulation, an equally-spaced fiber Bragg grating (FBG) array adds wavelength-dependent time-of-flight delays so that each successive channel is one nanosecond delayed with respect to each other (see Fig. 2(a)). Another function of the FBG array is to select 8 wavelengths corresponding to positive or negative modulations of respective channels (Fig. 1(a): (1) and (2)). A tunable wavelength-dependent weight bank weights each channel independently and a photodetector (PD) performs the desired *weighted addition* by carrying out optical-to-electrical conversion (Fig. 1(b): (3) and (4)).

978-1-4673-8118-5/15/$31.00/©2015 IEEE

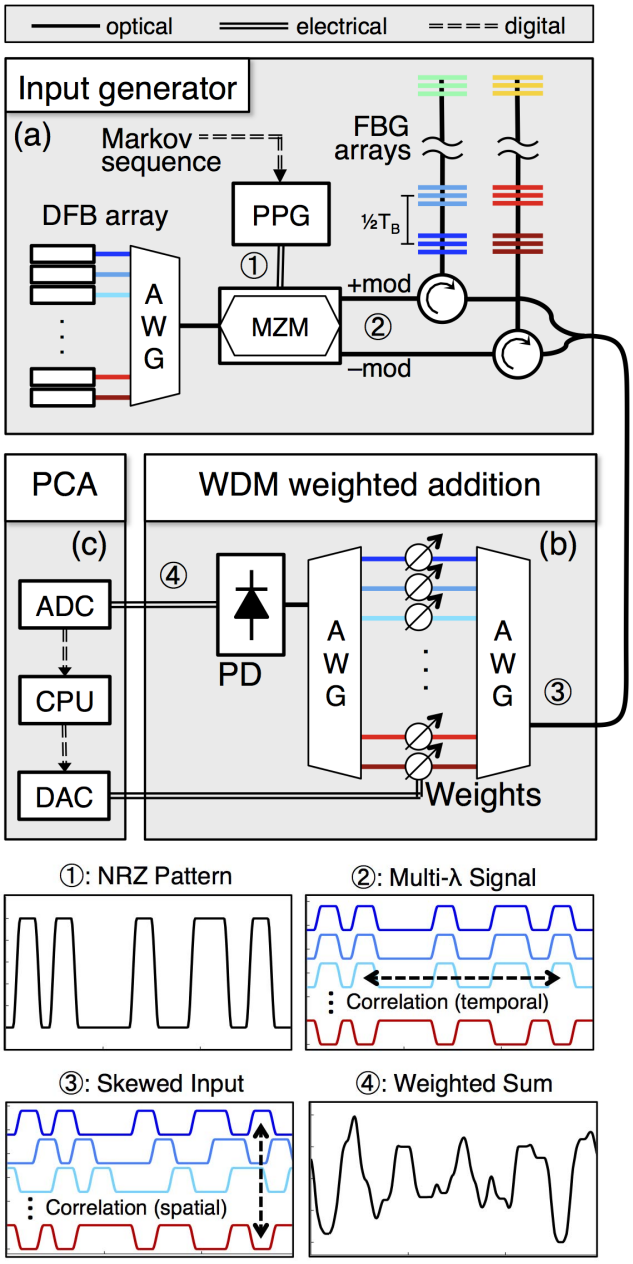

Fig. 1. Experimental setup, adapted from [5]. (a) Input generator, where DFB: distributed feedback laser; AWG: arrayed-waveguide grating multiplexer; PPG: pulse pattern generator; MZM: Mach-Zehnder modulator; FBG: fiber Bragg grating. (b) WDM weighted addition where circles are variable optical attenuators, and PD: photodiode, and (c) PCA algorithm, where ADC: analog-digital converter; CPU: central processor; DAC: digital-analog converter. Illustrations show signals at various points in the circuit: (1) RF input to the system that exhibits temporal autocorrelation, (2) Modulated WDM optical signals—complementary modulations are provided by the MZM, (3) λ-dependent delays transform initial temporal correlation into instantaneous spatial correlations, (4) electrical output of the PD representing the weighted sum of correlated channels.

In order to study the performance of a wideband PCA algorithm, one needs to reliably generate partially correlated input signals with continuous control. We explain next how one can use the programmability of the NRZ signal and the fixed interchannel delays to continuously control the correlation between channels. Assume that the bit sequence takes the form $X(n) \in \{-1,1\}$. Each channel carries a copy $X_k(n)$ of the bit pattern $X(n)$, but shifted by 1 ns ± 5 ps slots of time (see eq. (1)). In order to lay an integer number of bits per slot, the PPG bit rate is set to a multiple of 1 Gbaud, say d Gbaud, i.e. d bits per nanosecond. The interchannel correlation can be calculated as follows.

$$X_{k+1}(n + kd) = X_1(n) \equiv X(n) \tag{1}$$

$$\begin{aligned}\langle X_i(n), X_j(n)\rangle &= \langle X(n-(i-1)d), X(n-(j-1)d)\rangle \\ &= \langle X(n'), X(n'+(i-j)d)\rangle\end{aligned} \tag{2}$$

$$\begin{aligned}\Sigma_{ij} &= \mathbb{E}\big(\langle X_i(n), X_j(n)\rangle\big) = \\ &\mathbb{E}(\langle X(n'), X(n'+|i-j|d)\rangle) \equiv K(|i-j|)\end{aligned} \tag{3}$$

Eq. (2) demonstrates that a temporal correlation translates into an interchannel correlation, also called spatial correlation. In addition, eq. (3) indicates that the correlation matrix element Σ_{ij} only depends on the distance $|i-j|$: such matrices are called symmetric Toeplitz matrices. It is necessary to program the bit sequence such that we can control the correlation $K(|i-j|)$ in eq. (3) for $|i-j| \in \{1,\ldots,m\}$—evidently, $K(0) = 1$. To achieve this, the NRZ bit pattern is defined by an additive Markov-chain (MC) sequence of order m and dimension d (see eq. 4).

$$\begin{aligned}\mathbb{P}\big(X(n) = 1 \big| X(n-d), \ldots, X(n-(m-1)d)\big) = \\ \textstyle\sum_{k=1}^{m} f(X(n-kd,k)\end{aligned} \tag{4}$$

$$\begin{aligned}\Longrightarrow K(r) &= \textstyle\sum_{r'=1}^{m} F(r')K(|r-r'|)\ \ \forall r \geq 1, \\ \text{where } F(r) &= f(1,r) - f(-1,r)\end{aligned} \tag{5}$$

The additive property of the MC sequence can be understood as a probability of the current number being linearly dependent on the set of samples in the past, hence the sum in eq. (5). The result in eq. (5) is demonstrated in [6], where $F(r)$ is referred to as "memory function". Its recursive definition stems from the fact that different channels are correlated by different memories. For example, correlation between channels 3 and 1 depends directly on $F(2)$, but since the correlation between channels 3 and 2 (and between 2 and 1) is related by $F(1)$, then correlation between channels 3 and 1 also have 'indirect' dependence on $F(1)^2$.

For this study, we chose $m = 7$ and bit rates $d = 1,\ldots,13$ Gbaud—the PPG is limited to 13 Gbaud. The symmetric Toeplitz matrix Σ has the property of having 8 positive eigenvalues (E_i; $i = 1,\ldots,8$), corresponding to the variances of the PCs; and 8 symmetric or antisymmetric eigenvectors (see Fig. 3(b)), corresponding to the respective 8 weight vectors μ_i—also called principal component coefficients—that transform signals $x(t)$ into the i-th PC $\tilde{x_i}(t) = \mu_i \cdot x(t)$.

These PCs, especially the first, can be approximated by power iteration methods, which explore the fact that the first eigenvalue is greater than all others. The ratio between the first and second eigenvalues of Σ—$R \equiv E_1/E_2$—is thus a natural figure of merit for the analysis of convergence rates; i.e. the greater R, the more "well-defined" the first PC is with respect to all others. In [5], we parametrized the memory function with a single scalar variable α, but in this work we discovered that $m = 7$ parameters provide access to a wider range of PC coefficients.

Having described the partially correlated, multichannel NRZ sequence as input, we proceed to describe the PCA algorithm. Its goal is to control the weight bank so that it outputs the first principal component of this multidimensional signal. In this setup, the CPU is responsible for all correlation computations, and therefore limits the control loop latency. The algorithm is divided in two steps.

First, the CPU digitizes and stores in memory each of the 8 channels waveforms by individually selecting them using the filter bank.

$$\Delta\boldsymbol{\mu}_1 = \gamma\langle \boldsymbol{x}(t)\boldsymbol{\mu}_1 \cdot \boldsymbol{x}(t)\rangle_t = \gamma\langle \boldsymbol{x}^{\boldsymbol{T}}(t)\boldsymbol{x}(t)\rangle_t \cdot \boldsymbol{\mu}_1 \approx \gamma\boldsymbol{\Sigma}\cdot\boldsymbol{\mu}_1$$

$$\boldsymbol{\mu}_1 \leftarrow \frac{\boldsymbol{\mu}_1 + \Delta\boldsymbol{\mu}_1}{\|\boldsymbol{\mu}_1 + \Delta\boldsymbol{\mu}_1\|} \tag{6}$$

In the second step, we simulate an iterative weight control algorithm called the normalized Hebbian learning rule, described in eq. (6) [7, 5]. The algorithm only requires a sample of the output waveform after weighted addition—which is carried out by the ADC with a time window corresponding to 2000 bits—and its correlations with the stored waveforms samples. After a finite number of iterations, μ_1 in eq. (6) converges successfully to the first PC. In this experiment, the iteration count was limited to 40, when the converged weight value is compared with the PC computed by the SVD method. The extension of this algorithm to the computation of all PCs is described in [5].

In [5], electronic non-idealities caused the accuracy of the PCA algorithm to suffer under certain conditions. As an example, one RF amplifier (not shown in Fig. 1) was band limited to 1.3 GHz. Impedance mismatch was also present in the circuit, introducing overshoot and ringing. These obstacles were overcome by either using faster electronic components or removing them.

Injecting high bandwidth signals into the photonic pathway becomes increasingly difficult for bandwidths higher than 10 Gbaud. Two effects can cause the experimental waveform to deviate from the MC bit sequence (see Fig. 2(a)): low pass filtering in electronic circuitry of the PPG; and unwanted delay mismatch between channels across the optical circuit. We found that, although the first effect distorts the waveforms significantly, it does not ultimately destroy the interchannel partial correlations. The second effect, more critical, was minimized in this setup by finetuning delay lines across all channels.

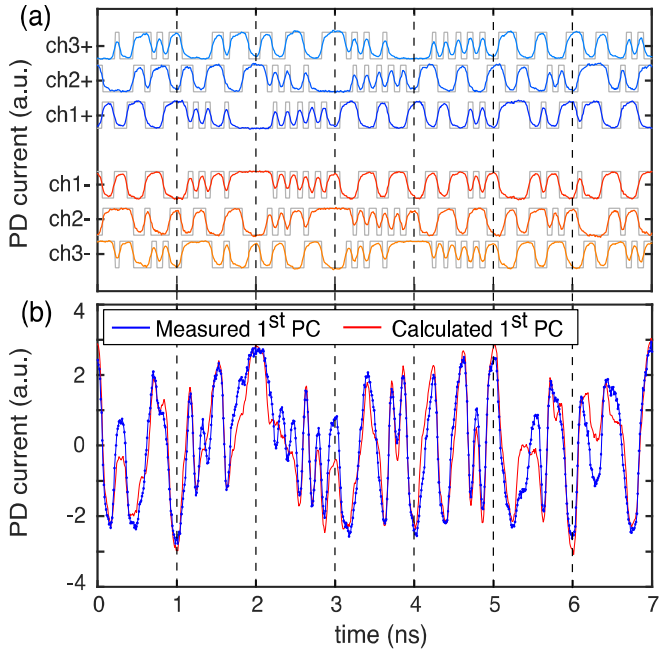

Fig. 2. Example of a PCA task on a 13 Gbaud bit pattern yielding 93% accuracy (cf. caption of Fig. 3). (a) The modulated waveforms before the weight bank (Fig. 1.(3)). Note the delay of 1 ns between subsequent channels. The distortion from the MC bit pattern (gray) arises from the bandwidth limitation in the PPG circuit. (b) After completion of PCA task, the measured first PC waveform (Fig. 1.(4)) is compared to the SVD-calculated one.

III. Results

First, it is necessary to evaluate the characteristics of the input signal. In Fig. 2(a), we present the intended NRZ bit pattern at a 13 Gbaud bit rate overlapped by the waveform generated by the PPG. Since the PPG is operating at maximum speed, one can observe the effects of low pass filtering at the measured signals. However, as previously mentioned, the signals are still strongly correlated with principal component vectors close to the originally calculated by the Markov chain method.

After a multichannel signal with partial spatial correlations is generated, the PCA iterative algorithm consistently converges to the well-defined first PC. Fig. 2(b) indicates a good match between the converged signal and the one solved by SVD. Note that in this case the converged output signal and the SVD-predicted follow each other without visible low pass filtering. This illustrates the fact that the MWP circuit is designed to handle a bandwidth of 200GHz, much higher than the currently tested 13Gbaud.

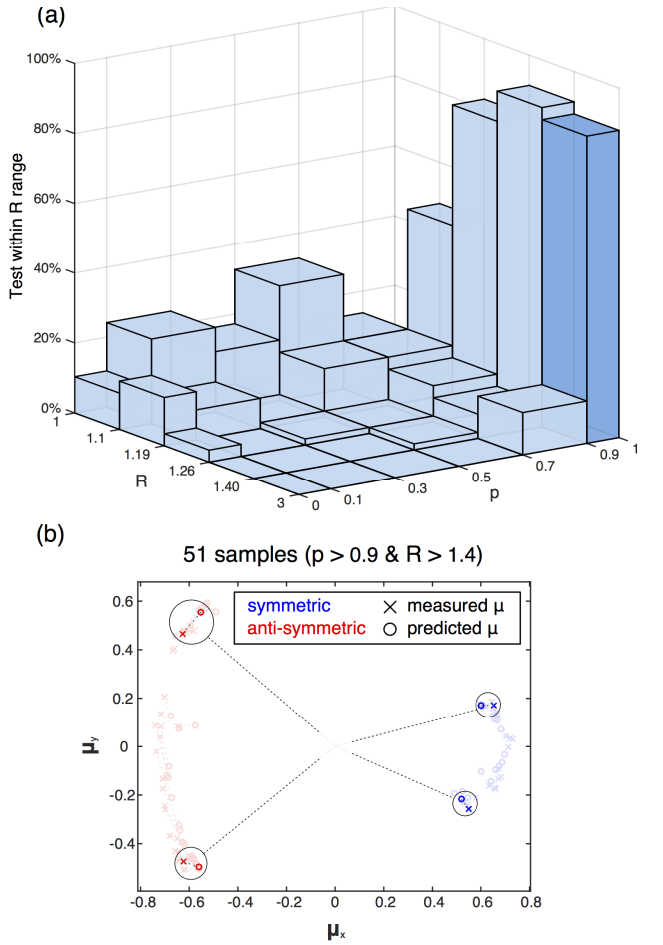

Fig. 3. Accuracy analysis of 290 runs of the experimental PCA task. (a) Histogram relating the accuracy p (square of dot product between predicted and measured PC weight vector) and R ratio between the first and second eigenvalues. The histogram was binned according to quintiles of R—each bin contains many different Markov memory functions mapping to the same range of R. The experimental PCA task accurately ($p > 0.9$) converges to the correct weight for $R > 1.2$, which translates to the first PC being 20% more "well-defined" than the second PC. Low accuracy can be explained by lack of convergence or convergence to the wrong PC. (b) 51 weight vectors contained in the rightmost bin (darker blue) of the histogram in (a). Four examples were highlighted for clarity. The weight vectors were plotted onto a convenient plane that highlights the spatial diversity of the explored weight vectors generated by different Markov parameters.

The necessary key to reveal each of the 8 PCs of an 8-dimensional RF signal is a weight vector of dimension 8. Finding this key is the fundamental objective of PCA: if the correlation changes, the key changes. In this study, we generated an unbiased binary additive Markov sequences of seventh order (m = 7), as described in eq. (4). From the memory function $F(r)$ used to create such a sequence, we can mathematically predict the interchannel correlation values $K(r)$ and the matrix Σ, and therefore the expected "PC" eigenvectors μ_i. Fig. 3(b) shows that the weight vectors found by the iterative algorithm match the predicted ones to great accuracy. A great number of memory functions were used to generate a continuous trace of PC coefficients in order to demonstrate that the experiment can handle arbitrary weight vectors. The setup can therefore reverse engineer the key to the first PC of the input signals in less than 40 iterations. That key is only dependent on the partial correlation of input signals—not dependent, notably, on the bit rates of the individual signals themselves. Thus, system performs equally well for NRZ patterns of 1 Gbaud and 13 Gbaud. Moreover, we observe that performance does not depend on the orientation of the PC coefficients. Its only depends on the magnitude of R (see Fig. 3(a)).

The presence of noise and the inexact command of effective weight in the weight bank prevent the system from achieving perfect close-to-100% in all cases. The former stems mostly from the electronic components, since we observed that thermal noise overwhelms amplified spontaneous emission. The latter, more important, can cause convergence to a different PC or to a linear combination of PCs in cases where $R < 1.2$ (see Fig. 3). Both effects can be minimized by integrating the system onto an optoelectronic chip.

IV. Application to Spectral Sensing in Cognitive Radios

Previous work in MWP filtering [4], beamforming [8], and cognitive radio [2] has included high bandwidth analog signal processing. However, they lack control algorithms that could adapt to changing environments, necessary for online analysis. In particular, one can expect the partial correlations of signals coming from an antenna array to drift over time. In the cognitive radio BSS context, for example, this corresponds to moving sources. Thus, the circuit has to learn and self-adjust to these new conditions—a task performed by the PCA module shown in Fig. 1. Such is the reason why we opted to solve the PCA task using an iterative method rather than using singular value decomposition (SVD) to calculate all principal components at once.

If the weight control feedback loop featured a custom analog controller, instead of a CPU, this process would take at most $40 \cdot 2000/10$ Gbaud = 8 μs to complete for a 10 Gbaud signal of the sort here presented. This considers a conservative statistical confidence of $40 \cdot 2000$ "bits" of information, over which we assumed that the cross-correlation stayed constant. This figure can be further reduced if R is kept large enough ($R > 1.2$). A short convergence time demonstrates that optoelectronic circuits can be used to follow non-stationary PCs in real-time.

In contrast to standard DSP, the WDM approach offer the possibility to scale fan-in without severely jeopardizing bandwidth in the weighted addition operation—because a single PD detects all channels simultaneously. The fundamental limit of fan-in lies not in the transmission medium, but rather in the power loss while multiplexing different inputs [9].

MWP can be used with cognitive radios to help solve an important problem in traditional RF spectrum management, in which static frequency bands are allocated to an exclusive user, resulting in large swaths of the spectrum being significantly underutilized during certain times and across geographical regions. Cognitive radio, using dynamic spectrum access, has evolved to enable secondary users to intelligently recycle the underutilized spectrum by sensing the RF environment and exploiting available spectrum holes, or white spaces. Upon identifying a spectrum white space, a cognitive radio may

temporarily access the spectrum until a licensed user is detected or until network quality degrades below a application threshold, at which point the radio must relocate to a new spectrum white space. To provide continuous service, while avoiding interference with other users, a cognitive radio must efficiently deal with constantly fluctuating spectrum availability and network quality. To do so, the radio must be able to quantify spectrum characteristics at all times and across a wide bandwidth.

Currently, radios cannot sense the RF environment while transmitting, because their own strong transmission overwhelms weak signals arriving from distant nodes, a problem known as self-interference. Thus, to perform spectrum sensing, the radio must stop transmitting, decreasing spectrum efficiency and adding overhead. In addition, while transmitting, the radio may not be aware of the new presence of a licensed user, resulting in undesirable interference. By reducing the self-interference to the noise floor using PCA, cognitive radios can continuously scan the spectrum, even while transmitting, to prevent interference and locate new spectrum white spaces.

This approach employs several well known advantages of optics, such as wideband performance and high precision that are highly advantageous for self-interference cancellation in cognitive radio. The wideband performance enables a cognitive radio to operate across a broad bandwidth without the need to reconfigure internal hardware. Unlike typical RF front ends, the optical system does not require a separate RF bandpass filter for each frequency band of operation, reducing size, weight, and cost. The high precision of optics allows the system to cancel the self-interference signal to very low levels so that the radio can detect very weak user signals in the band. Finally, because the system is completely analog, it is compatible with all modulation formats.

V. CONCLUSION

We have demonstrated how a WDM-based MWP system can be used to find and follow principal components of wideband, multidimensional RF signals. This can be applied to spectral sensing in cognitive radios, allowing a user to transmit while its receiver simultaneously scans across the entire radio spectrum. The present work extended the novel methodology of using additive Markov chains and fixed wavelength-dependent delays to generate partially correlated wavelength-carried signals for processing. Demonstration with higher-bandwidth RF waveforms was made possible by removing bandlimited electronic amplifiers. Finally, we showed that the iterative normalized Hebbian rule allows for a type of unsupervised learning in real time—in this case, learning the first principal component of a multichannel signal. Although a more thorough analysis will requires faster electronic sources, we showed that high frequency components could be recovered from the photonic circuit in Fig. 1. Further work will focus on the integration; on the acceleration of control loop latency to allow for faster conve rgence; on the generation of faster aperiodic partially correlated analog signals; and on the fundamental limitations imposed by inexact weight command and noise. This work may have application to PCA-based learning for blind source separation—ultimately spectrum sensing—and other areas of RF processing.

REFERENCES

[1] I.T.Jolliffe, *Principal Component Analysis* (Springer-Verlag NewYork, 2002), 2nd ed.

[2] M. Baylor, D. Z. Anderson, and Z. Popovic, "Holographic Blind Source Separation at Radio Frequencies," in "Controlling Light with Light: Photorefractive Effects, Photosensitivity, Fiber Gratings, Photonic Materials and More," (Optical Society of America, 2007), p. TuB1.

[3] N.T.Khajavi, S.Sadeghi, and S.M.-S.Sadough, "An Improved Blind Spectrum Sensing Technique for Cognitive Radio Systems," 2010 5th International Symposium on Telecommunications pp. 13–17 (2010).

[4] J. Capmany, J. Mora, I. Gasulla, J. Sancho, J. Lloret, and S. Sales, "Microwave Photonic Signal Processing," Journal of Lightwave Technology **31**, 571–586 (2012).

[5] A. N. Tait, J. Chang, B. J. Shastri, M. A. Nahmias, and P. R. Prucnal, "Demonstration of WDM weighted addition for principal component analysis," Optics Express **23**, 12758 (2015).

[6] S. S. Melnyk, O. V. Usatenko, V. A. Yampol'skii, S. S. Apostolov, and Z. A. Maiselis, "Memory functions and Correlations in Additive Binary Markov Chains," Journal of Physics A: Mathematical and General **39**, 14289 (2006).

[7] E. Oja, "A Simplified Neuron Model as a Principal Component Analyzer." Journal of mathematical biology **15**, 267–273 (1982).

[8] J. Chang, M. P. Fok, R. M. Corey, J. Meister, and P. R. Prucnal, "Highly Scalable Adaptive Photonic Beamformer Using a Single Mode to Multimode Optical Combiner," IEEE Microwave and Wireless Components Letters **23**, 563–565 (2013).

[9] J. W. Goodman, "Fan-in and Fan-out with Optical Interconnections," Optica Acta: International Journal of Optics **32**, 1489–1496 (1985).

[10] J. Hasler and B. Marr, "Finding a roadmap to achieve large neuromorphic hardware systems," Frontiers in Neuroscience **7**, 118 (2013).

[11] A. Tait, M. Nahmias, B. Shastri, and P. Prucnal, "Broadcast and Weight: an Integrated Network for Scalable Photonic Spike Processing," Journal of Lightwave Technology **32**, 4029–4041 (2014).

[12] M. A. Nahmias, B. J. Shastri, A. N. Tait, and P. R. Prucnal, "A Leaky Integrate-and-Fire Laser Neuron for Ultrafast Cognitive Computing," IEEE Journal of Selected Topics in Quantum Electronics **19** (2013).