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

PHILIP Y. MA,[1,*] ALEXANDER N. TAIT,[1,2] THOMAS FERREIRA DE LIMA,[1] CHAORAN HUANG,[1] BHAVIN

frequency dependent microwave electronics that exhibit poor reconfigurability and tunability across wideband RF spectrum [17].

Lately, a photonic architecture was proposed for networking and computation on wavelength-division multiplexing (WDM) signals using standard photonic integrated circuit (PIC) components [18]. This architecture enables a parallel matrix-vector multiplication operation called weighted addition, where an array of microrings (MRRs) performs channel-wise weighting (i.e., spectral filtering) on individual WDM signals while a balanced photo-detector (BPD) outputs the sum of weighted signals on all channels [19]. The MRRs carry at least two advantages: a) they are compact passive devices with small footprint (100 m^2) but holding high operational bandwidth (>10 GHz), which enables unparalleled information density if instantiated into a dense weight bank; b) they have wide tunability (over THz bandwidth) achieved by thermal tuning [20], which equips them with the capability of performing frequency independent processing on arbitrary RF signals. Therefore, MRRs have been adopted as the key reconfigurable elements in directed-logic circuits [21], waveform generation [22], reservoir computing [23], and neuromorphic photonic architectures [24,25].

In this paper, we present the first demonstration of photonic ICA using an on-chip MRR weight bank to identify the underlying sources that form the basis of the observed data. Source separation is a critical extension to our prior work on photonic principal component analysis (PCA) that only decomposes the data to uncorrelated components [26]. We follow the methodology of multivariate photonics assuming the waveform information of received mixtures may not be readily available in realistic field scenarios [13,27], and demonstrate that such information is unnecessary for ICA task but requires sophisticated front-end ADCs and additional data storage. Instead, we propose a novel photonic ICA algorithm that is able to extract independent components (ICs) solely based on the higher-order statistics of a single reduced-dimensional weighted addition output. We deploy a prototype implementation consisting of an RF subsystem that can dynamically reconfigure the mixing process of two independent sources, and a photonic subsystem that accomplishes record-high accuracy and precision on MRR weight control. Our proof-of-concept experiment of two-channel photonic ICA achieves performance whose accuracy and repeatability are close to that of conventional software-based ICA method (i.e., FastICA [28]) under di□erent channel separability and near-far conditions. In addition, our numerical simulation further studies how the Gaussian noise a□ects the performance of photonic ICA procedure, and identifies the operating regimes where our approach holds its fidelity. Overall, the proposed photonic ICA scheme paves the way for future research on integrated photonic systems that could enable advanced BSS pipelines for more complicated information mixing process.

2. Theory

In this section, we provide a succinct exposition of the basic ICA theory. Our photonic ICA scheme follows the same two-step procedure as most ICA algorithms: a preliminary whitening and the actual ICA estimation [29].

2.1. Definition

Consider we have source matrix **S** where each row represents an independent time-series signal $s_i{}^1t^{o1}i = 1, : : : , n^o$. These signals are mixed and result in the observed mixtures **X** = **AS** with the mixing coe□cients collectively placed in the mixing matrix **A**. ICA aims to figure out the corresponding inverse matrix of **A**, called the demixing matrix $\mathbf{A}^{-1}$, such that we can retrieve the independent sources by multiplying it with the received mixtures $\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} = \mathbf{A}^{-1}\mathbf{AS}$.

2.2. Whitening using PCA

Whitening applies a whitening matrix **V** to transforms the mixtures such that **Z** = **VX** = **VAS** is white: expectation

straightforward solutions from PCA: $\mathbf{V} = \mathbf{U}\boldsymbol{\Sigma}^{-1/2}\mathbf{U}^{T}$, where $\mathbf{U}$ has principal component (PC) vectors as its columns while diagonal matrix $\boldsymbol{\Sigma}$ has PC variances as its diagonal entries. Since

Fig. 1. (a) Schematic of the experimental setup for performing photonic ICA using an on-chip MRR weight bank. DFB: distributed feedback laser, MZM: Mach-Zehnder modulator, ATT: RF attenuator, AWG: arbitrary wave generator, MUX: WDM multiplexer, SM: source meter, BPD: balanced photo-detector, Oscope: sampling oscilloscope. The synchronization (SYNC) switch toggles the Oscope triggering state between a repeating pattern from AWG and a free-running clock at 200 kHz. (b) Micrograph of the fabricated MRR weight bank. MRRs are coupled with two bus waveguides that input the WDM signals at the IN port, and output the weighted WDM signals at the THRU and DROP ports. Metal traces are deposited to deliver the tuning current to the MRR weight bank to thermally tune the optical transmission of MRRs to configure their weights. (c) Two-channel weight evaluation results in the same format of [26,33]. 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. (d) Example of IC sources generated by two AWGs; top: square wave with kurtosis of 2, and bottom: sinusoidal wave with kurtosis of 1.5.

We list here some basic characterizations on the two MRRs used in this work, with more details described in [26] as we use the same chip for both photonic PCA and ICA. The two MRRs have radii of 10.921 m and 10.937 m, respectively. Their bare resonance peaks are at 1547.73 nm and 1549.24 nm. For both MRRs, the free spectral range (FSR) is 8.62 nm and finesse F is 30.79. The coupling gap between the bus waveguides and ring waveguide is 200 nm, and the Q factor is approximately 5500. To enable feedback control on the MRR weight bank through in-ring photoconductive heaters [33,35], an N-doped section of 10 m width is patterned to follow MRRs, outside of which heavy N++ doping is used to make ohmic contacts. Phosphorous dopant concentrations are N: 5 $10^{17} cm^{-3}$ and N++: 5 $10^{20} cm^{-3}$ as in [36]. The thermal tuning e□ciency of MRRs is measured to be 0.15 nm/mW. We perform MRR weight bank calibration using the feedback control procedure [33], and achieve 5.7 bits of accuracy (1.9% error, red lines) and 6.9 bits of precision (0.84% error, blue ellipses) as shown in Fig. 1(c). Such accomplishment improves upon the latest weight accuracy/precision record (5.2 bits of accuracy and 6.3 bits of precision) set by our recent photonic PCA work [26].

We employ two AWGs as the independent sources in our two-channel photonic ICA experiment. One AWG generates square wave (10 MHz) whose kurtosis is 2. The other AWG generates sinusoidal wave (20 MHz) whose kurtosis is 1.5. Examples of IC sources are shown in Fig. 1(d), which will be used as the ground-truth in later performance evaluation phase (with some temporal shift for alignment purpose). These two signals are mixed by a dedicated mixer consisting of four tunable RF attenuators (ATTs). As shown in Fig. 1(a), these four ATTs are inter-connected in such a way that a) ATT1 and ATT3 receive the source signal from AWG1, while ATT2 and ATT4 receive the source signal from AWG2; and that b) the outputs from ATT1 and ATT2 are combined as the input to MZM1, while the outputs from ATT3 and ATT4 are combined as the input to MZM2. As a result, we have an effective mixing matrix

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \tag{2}$$

where each mixing coefficient $a_i (i = 1, 2, 3, 4)$ represents the attenuation of corresponding ATT. In our prototype implementation, the attenuation of each ATT is in the range of 0–12 dB, so a_i is in the range of [0.25, 1].

We also use a synchronization (SYNC) switch to swap between two triggering conditions for the Oscope [26]. The motivation is that the input mixtures and weighted addition output are generally not synchronized with the Oscope in field (i.e., real life) scenarios. The only way to obtain an accurate digital waveform is to take samples in a real-time order at a real-time rate

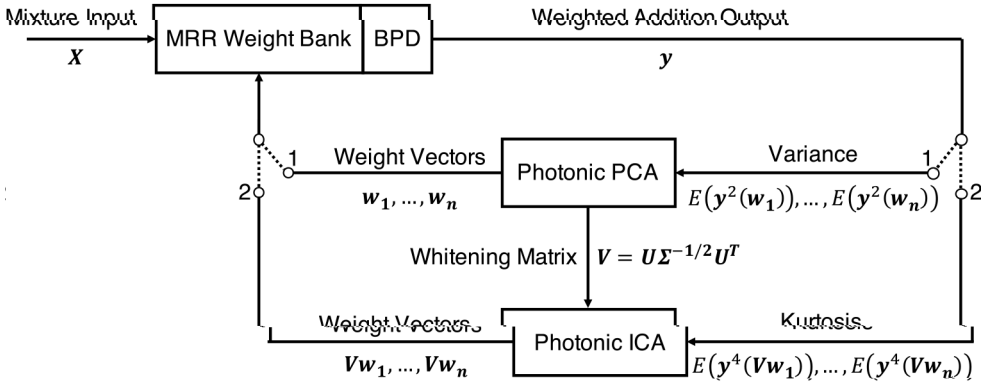

Fig. 2. Flowchart of the complete photonic ICA procedure exploiting the interaction between the photonic hardware (MRR weight bank plus BPD) and software (photonic PCA/ICA algorithms). First, the photonic hardware produces the weighted addition output **y** of the input mixtures **X**. Next, the photonic PCA algorithm updates the weight vectors $\mathbf{w_1}, :::, \mathbf{w_n}$ (set at the MRR weight bank) to be the target PC vectors by maximizing the variance of **y** (when switch is on 1). Finally, the photonic ICA algorithm takes the whitening matrix **V** computed from photonic PCA results, and updates the weight vectors in the whitened subspace $\mathbf{Vw_1}, :::, \mathbf{Vw_n}$ to be the target IC vectors by maximizing the kurtosis of **y** (when switch is on 2). Both photonic PCA and ICA algorithms require multiple iterations for convergence. Detailed code implementations of photonic PCA can be found in [26], while code implementations of photonic ICA can be found in Appendix A of this manuscript.

the input to the subsequent ICA procedure. Instead, we introduce **V** as an input argument to the photonic ICA algorithm, so as to directly process the input mixtures rather than having to sample, store, and replay **VX**. This is consistent with our e□orts of performing photonic ICA without redundant waveform information, not to mention **VX** expands the size of **V** by $T \bullet 2$ times here where T is the number of samples taken at super-Nyquist rate.

4.1. ICA algorithm

We describe ICA algorithm as Algorithm 1 in Appendix A, which proceeds by first initializing multiple weight vectors $\mathbf{w_1}, .., \mathbf{w_n}$. However, the actual weight vectors applied at the MRR weight bank now become $\mathbf{Vw_1}, .., \mathbf{Vw_n}$, which will equivalently transform our ICA procedure to the whitened subspace of the input mixtures. More importantly, these weight vectors are constrained to have unit norm because a) the weight vector magnitude is not supposed to a□ect the kurtosis measurement of the weighted addition output, and b) we are only di□ered from the actual ICs by an orthogonal transformation after whitening (i.e., only the weight vector direction matters). Therefore, weight normalization is imposed throughout the algorithm execution (where “normalized" is used).

Our ICA pursuit algorithm borrows inspirations from the well-known Nelder-Mead method [37], aiming to converge all the weight vectors to the target IC vector by updating the weight vector associated with the smallest weighted addition output kurtosis at each iteration. Please refer to Algorithm 1 for the detailed code implementations on four types of update operations available: reflection, expansion, contraction, and shrinkage, which contribute to the final convergence collectively. The algorithm proceeds until reaching the termination condition, which is controlled by a convergence tolerance gauging the proximity of weight vectors. The algorithm can extract ICs sequentially by enforcing the Gram-Schimit process (i.e., deflationary orthogonalization) given the fact that IC vectors are orthogonal against each another (see Line 40 of Algorithm 1).

4.2. Ill-conditioning

The difficulty of ICA problem largely depends on the mixing condition where so-called ill-conditioning can increase the chance of non-separable ICs. In this work, we consider the following two basic forms of ill-conditioning:

a) Low channel separability: the transmission paths for ICs are close $a_1 \bullet a_3 \quad a_2 \bullet a_4$ (e.g., when the sources/receivers are close to each other), so that the received mixtures $x_1 = a_1 s_1 + a_2 s_2$, $x_2 = a_3 s_3 + a_4 s_4$ are similar (could be differed approximately by a scaling factor). We quantify the channel separability using the condition number of the mixing matrix $\mathbf{A}$ [38]

$$^{1}\mathbf{A}^{o} = jj\mathbf{A}jj \quad jj\mathbf{A}^{-1}jj \tag{3}$$

where $jj \quad jj$ stands for the maximum absolute column sum. The channels are most separable when $^{1}\mathbf{A}^{o}_{min} = 1$ (e.g., if $a_1 \bullet a_3 = a_4 \bullet a_2 = 1$), and least separable when $^{1}\mathbf{A}^{o}_{max} = 1$ (e.g., if $a_1 \bullet a_3 = a_2 \bullet a_4$).

b) Near-far problem: the ratio of received power levels of ICs is high (e.g., when the desired signal is far away while an undesired interferer is nearby). Suppose s_1 is the signal of interest, and s_2 is the signal of interference. Then, we can fix the received power level of s_2 while varying the received power level of s_1 by considering the following simple but illustrative mixing matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 6 & 1 \quad 7 \\ 6 & 7 \\ 6 & 7 \\ 4 & 0 \quad 1 \quad 7 \\ & 5 \end{bmatrix} \tag{4}$$

where $\quad$ controls the intensity ratio between two ICs in mixture $x_1 = \quad s_1 + s_2$. If both ICs have equal power enter the mixer, then $20 \quad log_{10}{}^{1} \; ^{o}$ quantifies their signal-to-interference ratio (SIR) within mixture x_1.

5. Experimental results

In this section, we present the empirical results obtained from the two-channel photonic ICA experiment, demonstrating the feasibility of the proposed algorithm and the robustness under ill-conditioning cases above.

5.1. Channel separability case

We first present the experimental results of two-channel photonic ICA under various channel

Table 1. ICA Performance Under Channel Separability Conditions

Mixing Matrix (**A**)	Condition Number ( )	SNR (dB)	1st IC				2nd IC			
			FastICA		PhotonicICA		FastICA		PhotonicICA	
			RMSE	SD (%)	RMSE	SD (%)	RMSE	SD (%)	RMSE	SD (%)
" 1 0 / 0 1 #	1	22.70	8.04	0.06	8.72	0.23	7.80	0.08	7.97	0.13
" 0.8 0.4 / 0.4 0.9 #	3	21.03	14.38	0.09	14.65	1.46	15.93	0.10	17.88	1.17
" 0.8 0.5 / 0.6 0.9 #	5	21.88	21.04	0.11	22.63	1.51	22.45	0.24	24.02	0.99
" 0.8 0.6 / 0.5 0.7 #	7	21.22	26.46	0.15	28.03	2.32	28.52	0.16	29.07	2.11
" 0.7 0.5 / 0.7 0.8 #	10	21.38	30.61	0.17	34.57	2.36	32.60	0.39	35.89	2.07
" 0.7 0.8 / 0.4 0.6 #	21	22.19	36.43	0.24	40.78	3.27	38.13	0.81	39.58	3.16
" 0.8 0.7 / 0.5 0.4 #	65	22.14	44.71	0.39	47.38	3.66	46.88	1.29	50.76	3.91
0.8 0.7 / 0.6 0.5	105	22.12	54.60	0.64	58.58	4.14	58.07	1.81	61.35	4.05

recovered by FastICA are plotted as black curves in parallel to demonstrate such di□culty results from the worsening of channel separability rather than the algorithm itself.

Table 1 summarizes quantitative photonic ICA performance under various channel separability conditions. For each value, we run the algorithm 6 times (all with newly-initialized random weight vectors) to study the robustness of the proposed approach. We are interested in a) the average root-mean-squared error (RMSE) between measured and ground-truth ICs among all 6 runs to quantify accuracy, and b) the standard deviation (SD) of measured ICs among all 6 runs to quantify repeatability. All values are obtained when both the measured and ground-truth ICs are temporally aligned and normalized to unit power (the scale of ICs does not matter in ICA problem). We also report the results of FastICA algorithm in the table.

While the accuracy of both FastICA and photonic ICA degrades with the increase of , their performance di□erence is limited throughout the values considered here. For the best case of = 1, photonic ICA achieves almost the same performance as FastICA, both claiming about 8% RMSE. For the worst case of = 105, both their RMSEs are over 50% with photonic ICA having about 4% more error than FastICA. The low accuracy of ICs at large stems directly from the inherent di□culty of decomposing two almost the same linear combinations into two di□erent components. This di□culty holds even if the demixer has the full access to the mixing matrix, and the lack of knowledge on the mixing matrix in ICA settings only makes it worse.

The repeatability of both FastICA and photonic ICA worsens with respect to the increase of as well. However, the SDs of FastICA are mostly small (less than 1% for IC1, less than 2% for IC2), due to the comparatively invariant input mixtures sampled under super-Nyquist rate. In contrast, the SDs of photonic ICA express less stability. From the best case of = 1 to the worst case of = 105, the SD increases by about 4% for both ICs. We do not attribute this uncertainty to our statistics-based photonic ICA approach itself as [13] has demonstrated the histogram of weighted addition output are identical regardless of whether it is sampled under super-Nyquist rate or sub-Nyquist rate. Instead, the errors originate from our prototype photonic subsystem, particularly the instability of MRR weight bank that oversees the drifting of weights over time (we will have more discussions later about the reasons behind this phenomenon).

Fig. 3. Experimental waveforms of the received mixtures (left column) and corresponding ICs (right column) associated with 4 typical condition numbers of the mixing matrix: (a) = 1, (b) = 5, (c) = 10, (d) = 21. The degradation of channel separability conditions (from top to bottom) makes it more di□cult for FastICA (black curves) and PhotonicICA (red curves) to retrieve correct ICs.

5.2. Near-far case

We next present the experimental results of two-channel photonic ICA under various near-far conditions. Table 2 lists the e□ective mixing matrix (**A**) configured by the mixer in the experimental setup, with ranging from +1.5 to -0.25. Here, we pick the square wave as signal of interest s_1 (i.e., far source), a.693 -11.95 Td[(Reu20(uare)u20siTJ-167.693 vide)-22ei407-3e7.6ignaluaree as signal of

Fig. 4. Experimental waveforms of the received mixtures (left column) and corresponding ICs (right column) associated with 4 typical power ratios of the received power levels of ICs: (a) SIR = 0 dB, (b) SIR = 4.44 dB, (c) SIR = 7.96 dB, (d) SIR = 12.04 dB. Here, the square wave is the signal of interest (i.e., far source), while the sinusoidal wave is the signal of interference (i.e., near source). The decrease of the SIR (from top to bottom) makes it more difficult for FastICA (black curves, IC1) and PhotonicICA (red curves, IC1) to retrieve square waves. The recovery of sinusoidal wave (IC2) is largely successful thanks to the pure copy of it on the second mixture channel (RX2).

Fig. 5. Numerical investigations of the photonic ICA performance considering the channel

Algorithm 1: ICA algorithm

Input: weighted addition output **yF34 9.9626 Tf 121.62141**

5. Q. V. Le, W. Y. Zou, S. Y. Yeung, and A. Y. Ng, "Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis," in "*CVPR 2011*," (2011), pp. 3361–3368.
6. A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," Neural Netw. **13**(4-5), 411–430 (2000).
7. M.-Z. Poh, D. J. McDu□, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," Opt. Express **18**(10), 10762–10774 (2010).
8. G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Gri□anti, and S. M. Smith, "Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers,"

35. H. Jayatilleka, K. Murray, M. Á. Guillén-Torres, M. Caverley, R. Hu, N. A. Jaeger, L. Chrostowski, and S. Shekhar,