Multi-channel control for microring weight banks

Alexander N. Tait,[*] **Thomas Ferreira de Lima, Mitchell A. Nahmias,**
Bhavin J. Shastri, and Paul R. Prucnal

Princeton University, Princeton, NJ 08544, USA

[*]*atait@princeton.edu*

Abstract: We demonstrate 4-channel, 2GHz weighted addition in a silicon microring filter bank. Accurate analog weight control becomes more difficult with increasing number of channels, N, as feedback approaches become impractical and brute force feedforward approaches take $O(2^N)$ calibration measurements in the presence of inter-channel dependence. We introduce model-based calibration techniques for thermal cross-talk and cross-gain saturation, which result in a scalable $O(N)$ calibration routine and 3.8 bit feedforward weight accuracy on every channel. Practical calibration routines are indispensible for controlling large-scale microring systems. The effect of thermal model complexity on accuracy is discussed. Weighted addition based on silicon microrings can apply the strengths of photonic manufacturing, wideband information processing, and multiwavelength networks towards new paradigms of ultrafast analog distributed processing.

References and links

1. 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**(6), 751–779 (2010).
2. A. Mekis, S. Gloeckner, G. Masini, A. Narasimha, T. Pinguet, S. Sahni, and P. De Dobbelaere, "A grating-coupler-enabled CMOS photonics platform," IEEE J. Sel. Top. Quantum Electron. **17**(3), 597–608 (2011).
3. J. S. Orcutt, B. Moss, C. Sun, J. Leu, M. Georgas, J. Shainline, E. Zgraggen, H. Li, J. Sun, M. Weaver, S. Urošević, M. Popović, R. J. Ram, and V. Stojanović, "Open foundry platform for high-performance electronic-photonic integration," Opt. Express **20**(11), 12222–12232 (2012).
4. Y. Vlasov, "Silicon CMOS-integrated nano-photonics for computer and data communications beyond 100G," IEEE Commun. Mag. **50**(2), s67–s72 (2012).
5. A.-J. Lim, J. Song, Q. Fang, C. Li, X. Tu, N. Duan, K. K. Chen, R.-C. Tern, and T.-Y. Liow, "Review of silicon photonics foundry efforts," IEEE J. Sel. Top. Quantum Electron. **20**(4), 405–416 (2014).
6. P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, B. Brezzo, I. Vo, S. K. Esser, R. Appuswamy, B. Taba, A. Amir, M. D. Flickner, W. P. Risk, R. Manohar, and D. S. Modha, "A million spiking-neuron integrated circuit with a scalable communication network and interface," Science **345**(6197), 668–673 (2014).
7. G. Indiveri, B. Linares-Barranco, T. J. Hamilton, A. van Schaik, R. Etienne-Cummings, T. Delbruck, S.-C. Liu, P. Dudek, P. Hfliger, S. Renaud, J. Schemmel, G. Cauwenberghs, J. Arthur, K. Hynna, F. Folowosele, S. Saighi, T. Serrano-Gotarredona, J. Wijekoon, Y. Wang, and K. Boahen, "Neuromorphic silicon neuron circuits," Front. Neurosci. **5**(73), 1–23 (2011).
8. J. Hasler and H. B. Marr, "Finding a roadmap to achieve large neuromorphic hardware systems," Front. Neurosci. **7**(118), 00118 (2013).
9. P. Asthana, G. P. Nordin, J. Armand R. Tanguay, and B. K. Jenkins, "Analysis of weighted fan-out/fan-in volume holographic optical interconnections," Appl. Opt. **32**(8), 1441–1469 (1993).

10. M. Hill, E. E. E. Frietman, H. de Waardt, G.-D. Khoe, and H. Dorren, "All fiber-optic neural network using coupled SOA based ring lasers," IEEE Trans. Neural Networks **13**(6), 1504–1513 (2002).
11. 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. Lightwave Technol. **32**(21), 3427–3439 (2014).
12. R. Ramaswami, "Multiwavelength lightwave networks for computer communication," IEEE Commun. Mag. **31**(2), 78–88 (1993).
13. D. Marpaung, C. Roeloffzen, R. Heideman, A. Leinse, S. Sales, and J. Capmany, "Integrated microwave photonics," Laser Photonics Rev. **7**(4), 506–538 (2013).
14. M. Chang, M. Fok, A. Hofmaier, and P. Prucnal, "Optical analog self-interference cancellation using electro-absorption modulators," IEEE Microwave Wireless Compon. Lett. **23**(2), 99–101 (2013).
15. 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 J. Sel. Top. Quantum Electron. **19**(5), 1800212 (2013).
16. W. Coomans, L. Gelens, S. Beri, J. Danckaert, and G. Van der Sande, "Solitary and coupled semiconductor ring lasers as optical spiking neurons," Phys. Rev. E **84**(3), 036209 (2011).
17. F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, "Relative refractory period in an excitable semiconductor laser," Phys. Rev. Lett. **112**, 183902 (2014).
18. B. J. Shastri, M. A. Nahmias, A. N. Tait, A. W. Rodriguez, B. Wu, and P. R. Prucnal, "Spike processing with a graphene excitable laser," Sci. Rep. **5**, 19126 (2015).
19. A. N. Tait, M. A. Nahmias, Y. Tian, B. J. Shastri, and P. R. Prucnal, "Photonic Neuromorphic Signal Processing and Computing," in *Nanophotonic Information Physics*, M. Naruse, ed., pp. 183–222 (Springer, 2014).
20. E. Klein, D. Geuzebroek, H. Kelderman, G. Sengo, N. Baker, and A. Driessen, "Reconfigurable optical add-drop multiplexer using microring resonators," IEEE Photonics Technol. Lett. **17**(11), 2358–2360 (2005).
21. J. C. C. Mak, W. D. Sacher, T. Xue, J. C. Mikkelsen, Z. Yong, and J. K. S. Poon, "Automatic Resonance Alignment of High-Order Microring Filters," IEEE J. Quantum Electron. **51**(11), 1–11 (2015).
22. J. A. Cox, A. L. Lentine, D. C. Trotter, and A. L. Starbuck, "Control of integrated micro-resonator wavelength via balanced homodyne locking," Opt. Express **22**(9), 11279–11289 (2014).
23. J. Cardenas, M. A. Foster, N. Sherwood-Droz, C. B. Poitras, H. L. R. Lira, B. Zhang, A. L. Gaeta, J. B. Khurgin, P. Morton, and M. Lipson, "Wide-bandwidth continuously tunable optical delay line using silicon microring resonators," Opt. Express **18**(25), 26,525–26,534 (2010).
24. C. T. DeRose, M. R. Watts, D. C. Trotter, D. L. Luck, G. N. Nielson, and R. W. Young, "Silicon microring modulator with integrated heater and temperature sensor for thermal control," in *Conference on Lasers and Electro-Optics* (Optical Society of America, 2010), paper CThJ3.
25. H. Jayatilleka, K. Murray, M. Ángel Guillén-Torres, M. Caverley, R. Hu, N. A. F. Jaeger, L. Chrostowski, and S. Shekhar, "Wavelength tuning and stabilization of microring-based filters using silicon in-resonator photoconductive heaters," Opt. Express **23**(19), 25084–25097 (2015).
26. A. Tait, T. Ferreira de Lima, M. Nahmias, B. Shastri, and P. Prucnal, "Continuous calibration of microring weights for analog optical networks," IEEE Photonics Technol. Lett. **28**(8), 887–890 (2016).
27. A. Tait, M. Nahmias, B. Shastri, M. Chang, A. Wu, E. Zhou, E. Blow, T. Ferreira de Lima, B. Wu, and P. Prucnal, "Balanced WDM weight banks for analog optical processing and networking in silicon," in *Summer Topicals*, (IEEE/OSA, 2015), paper MC2.3.
28. M. A. Nahmias, A. N. Tait, B. J. Shastri, T. F. de Lima, and P. R. Prucnal, "Excitable laser processing network node in hybrid silicon: analysis and simulation," Opt. Express **23**(20), 26800–26813 (2015).
29. A. Tait, A. Wu, E. Zhou, T. Ferreira de Lima, M. Nahmias, B. Shastri, and P. Prucnal, "Multi-channel microring weight bank control for reconfigurable analog photonic networks," in *Optical Interconnects*, (IEEE, 2016), paper WC4.
30. S. Friedmann, N. Frémaux, J. Schemmel, W. Gerstner, and K. Meier, "Reward-based learning under hardware constraints - using a RISC processor embedded in a neuromorphic substrate," Front. Neurosci. **7**(160) (2013).
31. 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 Des. Integr. Circuits Syst. **34**(10), 1537–1557 (2015).
32. L. Chrostowski and M. Hochberg, *Silicon Photonics Design: From Devices to Systems* (Cambridge University Press, 2015).
33. R. J. Bojko, J. Li, L. He, T. Baehr-Jones, M. Hochberg, and Y. Aida, "Electron beam lithography writing strategies for low loss, high confinement silicon optical waveguides," J. Vac. Sci. Technol., B **29**(6), 06F309 (2011).
34. Y. Wang, X. Wang, J. Flueckiger, H. Yun, W. Shi, R. Bojko, N. A. Jaeger, and L. Chrostowski, "Focusing sub-wavelength grating couplers with low back reflections for rapid prototyping of silicon photonic circuits," Opt. Express **22**(17), 20652–20662 (2014).
35. 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**(10), 12758–12765 (2015).
36. M. S. Dahlem, C. W. Holzwarth, A. Khilo, F. X. Kärtner, H. I. Smith, and E. P. Ippen, "Reconfigurable multi-channel second-order silicon microring-resonator filterbanks for on-chip WDM systems," Opt. Express **19**(1),

306–316 (2011).

37. H. Park, Y. hao Kuo, A. W. Fang, R. Jones, O. Cohen, M. J. Paniccia, and J. E. Bowers, "A hybrid AlGaInAs-silicon evanescent preamplifier and photodetector," Opt. Express **15**(21), 13539–13546 (2007).

38. J. D. B. Bradley and E. S. Hosseini, "Monolithic erbium- and ytterbium-doped microring lasers on silicon chips," Opt. Express **22**(10), 12226–12237 (2014).

39. P. Dong, S. Liao, D. Feng, H. Liang, D. Zheng, R. Shafiiha, C. Kung, W. Qian, G. Li, X. Zheng, A. Krishnamoorthy, and M. Asghari, "Low V_{pp}, ultralow-energy, compact, high-speed silicon electro-optic modulator," Opt. Express **17**(25), 22484–22490 (2009).

40. H. Jayatilleka, K. Murray, M. Caverley, N. Jaeger, L. Chrostowski, and S. Shekhar, "Crosstalk in SOI microring resonator-based filters," J. Lightwave Technol. **PP**(99), 1–1 (2015).

41. O. Schwelb, "Generalized analysis for a class of linear interferometric networks. I. Analysis," IEEE Trans. Microwave Theory Tech. **46**(10), 1399–1408 (1998).

---

1. Introduction

Advances in photonic integrated circuit (PIC) technologies [1–5] will generate opportunities for large-scale, low-cost optical processing systems. At the same time, a revival is underway in unconventional (neuro-inspired) microelectronic computing architectures, aimed to address energy efficiency limitations inherent in von Neumann computers [6–8]. Neuro-inspired hardware addresses these issues, in part, by distributing processing among many nodes, and, as such, rely heavily on multi-access networking strategies in which connection strengths (i.e. "weights") are reconfigurable. It has long been recognized that optical physics are well-suited to the analog interconnect problem, yet solutions based on holograms [9] and fiber [10] circuits have not led to integrated systems. We have previously proposed a PIC-compatible multi-access analog network called "broadcast-and-weight" [11], which combines results in multiwavelength networks [12], analog photonic links [13, 14], and photonic neurons [15–18]. Broadcast-and-weight networks could open processing domains with unprecedented speed and complexity [19]. Figure 1(a) depicts the concept of a broadcast-and-weight network.

Broadcast-and-weight relies heavily on wavelength-division multiplexed (WDM) weighted addition. The analog network topology is "programmed" by controlling weight values. Microring resonator (MRR) implementations of weight banks, drawn in Fig. 1(b), have the advantages of compactness, WDM capability, and ease of tuning. On the other hand, MRR sensitivity to fabrication variations, thermal fluctuations, and thermal cross-talk presents a control problem. MRR control is an important topic for WDM demultiplexers [20], high-order filters [21], modulators [22], and delay lines [23]. MRR controllers are often based on online feedback control [24, 25], but the unique requirements of an analog weight bank (continuous range of weights, input signals of unknown amplitude and shape, etc.) call for a feedforward control approach with offline pre-calibration performed at least once per fabricated device [26].

Prior work demonstrated feedforward control of an add/drop MRR filter edge for effecting a continuous range of transmission values. This enabled a single photonic weight with a range of –1 to +1 [27] and precision of 3.1 bits [26] (i.e. a maximum error of ±0.117 over the range ±1, or a dynamic range of 9.33dB). Calibration consisted of recording weights over the filter edge tuning range and interpolating. The results in [26] are preliminary in the sense that interpolation-based techniques are impractical for simultaneously controlling more than one MRR weight, an essential requirement for photonic weight banks and weighted optical networks. When weight interdependency or cross-talk are present, the control problem can not be separated into N isolated channels. The dimensionality of the full tuning range increases with N, necessitating $O(2^N)$ calibration measurements in general. The major contributions of the present work are full, simultaneous control of MRR weight *banks* and the development of tractable, $O(N)$, calibration methods for banks with any number of channels.

Model-based calibration is required for the two predominant sources of weight interdependency: thermal cross-talk and cross-gain saturation. In this work, we expand on preliminary
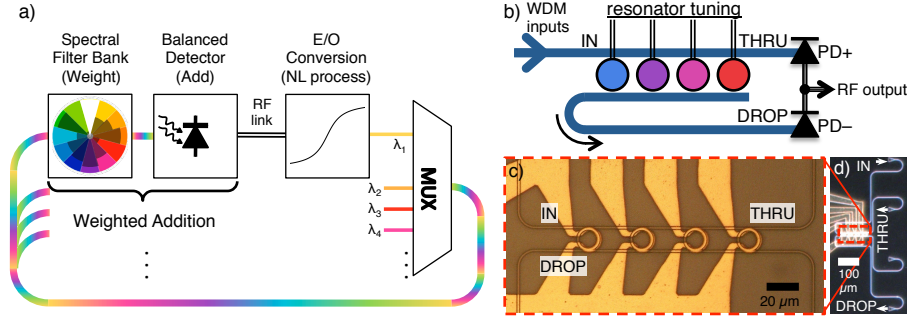
Fig. 1. a) Role of WDM weighted addition in a proposed on-chip analog photonic processing network [11]. Each E/O node – which could be any nonlinear modulator, direct-driven laser, or dynamical laser neuron – produces a signal modulated on a unique wavelength. Weighted addition banks produce electrical signals that drive the E/O converters [28]. b) Microring resonator (MRR) implementation of a WDM weight bank. Tuning MRRs between on- and off- resonance switches a continuous amount of optical power between drop and through ports. A balanced photodetector (PD) yields the sum and difference of weighted signals. c) Optical micrograph of the device under test, showing a bank of four thermally-tuned MRRs. d) Wide area micrograph, showing fiber-to-chip grating coupler ports.

results in [29], developing models whose parameters can be fit (i.e. calibrated) with a $O(N)$ routine of spectral and oscilloscope measurements. Whereas an interpolation-only approach with 20 points resolution would require $20^4 = 160,000$ calibration measurements, the presented calibration routine takes roughly $4 \times [10(\text{heater}) + 20(\text{filter}) + 4(\text{amplifier})] = 136$ total calibration measurements. We then assess factors affecting weight precision, including the complexity of the thermal cross-talk model. We demonstrate simultaneous 4-channel MRR weight control with an accuracy of 3.8 bits and precision of 4.0 bits (plus 1.0 sign bit) on each channel. While optimal weight resolution is still a topic of discussion in the neuromorphic electronics community [8], several state-of-the-art architectures with dedicated weight hardware have settled on 4-bit resolution [30, 31]. Practical, accurate, and scalable MRR control techniques are a critical step towards large scale analog processing networks based on MRR weight banks.

2. Methods

MRR weight bank samples were fabricated on silicon-on-insulator wafers at the Washington Nanofabrication Fabrication through the UBC SiEPIC rapid prototyping group [32]. Silicon thickness is 220 nm, and buried oxide thickness is 3 μm. 500 nm wide WGs were patterned by Ebeam lithography and fully etched to the buried oxide [33]. A hydrogen silsesquioxane resist (HSQ, Dow-Corning XP-1541-006) was spin-coated at 4000 rpm, then hotplate baked at 80 C for 4 minutes. Electron beam lithography was performed using a JEOL JBX-6300FS system operated at 100 keV energy, 8 nA beam current, 500 μm exposure field size, and exposure dose of 2800 μC/cm². The resist was developed by immersion in 25% tetramethylammonium hydroxide (TMAH) for 4 minutes. Silicon was etched from unexposed areas by inductively coupled plasma etching in an Oxford Plasmalab System, with a chlorine gas flow of 20 sccm, pressure of 12 mT, ICP power of 800 W, bias power of 40 W, and a plate temperature of 20 °C, resulting in a bias voltage of 185 V. Thermo-optic MRR resonance tuning is accomplished by ohmic heating of Ti/Pt/Au contacts (100 Å:Ti, 50 Å:Pt, 2000 Å:Au), which were deposited on top of a 3 μm oxide passivation layer using an Angstron Engineering electron beam evaporation
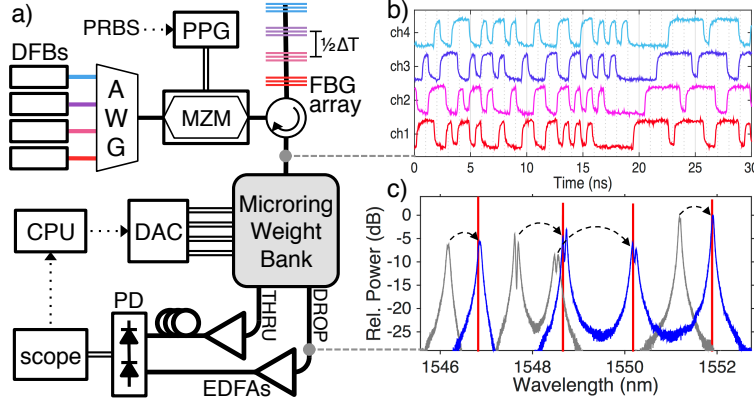
Fig. 2. a) Experimental setup. An input generator creates uncorrelated signals on different wavelengths by time delaying a single PRBS. DFB: distributed feedback laser; AWG: arrayed-waveguide grating; PPG: pulse pattern generator; MZM: Mach-Zehnder Modulator; FBG: fiber Bragg grating. The microring weight bank is thermally tuned by a current-mode DAC (digital-to-analog converter). Drop and thru outputs are amplified by erbium doped fiber amplifiers (EDFAs) and delay-matched before detection by a balanced photodetector (PD). A computer (CPU) executes the calibration routine. b) Time domain traces of reference input signals on different wavelength channels. c) Optical spectrum of WDM inputs (red) and transmission spectra of the drop port when tuning current is off (gray) and tuned onto resonance (blue), measured with a drop port spectrum analyzer (not shown).

system.

The weight bank device pictured in Fig. 1(c) consists of two bus waveguides and four MRRs in a parallel add/drop configuration , each of which controls a single wavelength channel by tuning on or off resonance. The radii of the MRRs are [6.37, 6.90, 7.43, and 7.96] µm, respectively. Coupling region gaps were 200 nm, and neighboring MRRs are separated by 20 µm. Q-factors are approximately 10,000. The free spectral range of the first MRR is measured to be 15 nm, indicating an effective TE refractive index of 4.2. . The sample is mounted on a temperature-controlled alignment stage and coupled to fiber with TE focusing subwavelength grating couplers [34].

The experimental setup shown in Fig. 2(a) consists of a multiwavelength reference input generator [35] that produces statistically independent signals by imparting channel-dependent delays on a 2Gbps pseudo-random bit sequence (PRBS). These reference signals are shown in Fig. 2(b). A 4-channel 13-bit digital-to-analog converter (DAC), NI PCI-6723, buffered to provide up to 80 mA per channel, tunes the electrical power dissipated in each MRR heater. The heaters share a common connection to reduce electrical I/O count. Since this common wire is not perfectly conducting, the effective common voltage can fluctuate with total current flow. Current-mode drivers are used to avoid this issue. The drop and through outputs of the MRR weight bank are amplified, their net delays matched, and detected by a balanced photodiode (PD). A transmission spectrum analyzer (not shown) is also connected to the device to simultaneously monitor the filter resonance peaks, tune them onto resonance with the WDM input signals , and assist in thermal model calibration. Figure 2(c) depicts tuning the bank from the initial state to the all channels on-resonance state.

Although input signals to the MRR weight bank are not necessarily known during an operation phase, the calibration phase can take advantage of known reference inputs in order to simultaneously measure the effective weight of each channel. In this case, references were de-

layed PRBS signals, each of which is stored as $x_i(t)$. If channel delays exceed one bit period, then the correlation $\langle x_i(t) \cdot x_j(t) \rangle_t$ approaches zero for a sufficiently long pattern (in this work, 2^7 bits). All weights μ_i can then be determined by decomposing a single measurement $m(t)$ in terms of stored references:

$$\mu_i = \frac{\langle x_i(t) \cdot m(t) \rangle_t}{\langle x_i(t) \cdot x_i(t) \rangle_t} \tag{1}$$

The calibration routine estimates a mapping of applied current to weight $\vec{i} \rightarrow \vec{\mu}$. The inverse of this mapping becomes the feedforward control rule for effecting a desired weight vector. We separate the map into physical stages for thermal tuning ($\vec{i} \rightarrow \vec{\Delta\lambda}$), MRR bank transmission ($\vec{\Delta\lambda} \rightarrow \vec{T}$), and actual detected weight ($\vec{T} \rightarrow \vec{\mu}$).

2.1. Thermal cross-talk model

The temperature of an MRR waveguide is affected predominantly by the heater directly above, but heat can also leak between nearby MRRs. The relationship between dissipated electrical power, $i^2 R$, and resonant wavelength shift, $\vec{\lambda} - \vec{\lambda}_0$ is linear and can be modeled by a matrix, $\boldsymbol{K}$ [36]. Assuming heater resistance is constant,

$$\vec{\lambda} - \vec{\lambda}_0 = \boldsymbol{K}\vec{i^2} \tag{2}$$

where λ_0 is the resonant wavelength at zero tuning current, and $\boldsymbol{K}$ is a nearly diagonal matrix that describes the thermo-optic effect, heat transfer coupling, and heater resistance. Off-diagonals of $\boldsymbol{K}$ describe unintended heat transfer from a given heater to filter of different channels, a.k.a. thermal cross-talk. Substituting $q_j \equiv i_j^2$ for notational clarity, this equation can be put in a differential form around WDM signal wavelengths, $\vec{\lambda_{sig}}$, and the tuning current needed to bias filters on-resonance with these signals, $\vec{q_{bias}}$,

$$\vec{\lambda}_{sig} - \vec{\lambda}_0 \;=\; \boldsymbol{K}\vec{q}_{bias} \tag{3}$$
$$\vec{\lambda} - \vec{\lambda}_{sig} \;=\; \boldsymbol{K}(\vec{q} - \vec{q}_{bias}) \tag{4}$$
$$\vec{\Delta\lambda} \;=\; \boldsymbol{K}\vec{\Delta q} \tag{5}$$

This linear model is simple to calibrate and invert, but it relies on an assumption of constant heater resistance. In general, heater resistance is also temperature dependent due to thermo-electric self-heating. For a single current-driven heater with ambient resistance of R_0 and thermo-electric coefficient α,

$$R(q) \;=\; R_0\left[1 + \alpha q R(q)\right] \tag{6}$$
$$\;=\; \frac{R_0}{1 - \alpha R_0 q} \tag{7}$$

which is certainly not constant, and even has a singularity at $q = (\alpha R_0)^{-1}$, signifying a thermal runaway. Instead of combining the multivariate and nonlinear equations above, we simply note that non-constant resistance means that second and higher-order derivatives of $\vec{\Delta\lambda}$ in terms of $\vec{\Delta q}$ are non-zero but small enough that a Taylor expansion can incorporate the nonlinearities.

$$\vec{\Delta\lambda} \approx \sum_{d=1}^{D} \boldsymbol{K_d}\vec{\Delta q}^d \tag{8}$$

where D is model order, the exponentiation of $\vec{\Delta q}$ is element-wise, and the model now contains D distinct $\boldsymbol{K}$ matrices.

The Taylor approximation's main advantage for calibration is a simple method to fit $\boldsymbol{K}$ matrices. The tuning current of each channel is swept over an operating range of interest ($\sim$4 filter linewidths), while the wavelength shift of every filter peak is measured with the spectrum analyzer. Each peak shift function is fit with a D-order polynomial to obtain one element of each $\boldsymbol{K}$ matrix. The process is repeated for every channel. $\boldsymbol{K}$ values found in this experiment are shown in Fig. 3. To prevent overfitting, there must be at least D spectrum measurements per channel. We make $5DN$ measurements for added robustness and found a $D = 2$ made thermal modelling error sufficiently small so as not to limit overall precision, which is revisited in Section 3.

The polynomial mapping $\vec{\Delta q} \to \vec{\Delta \lambda}$ must be inverted to provide a feedforward control rule. While it does not have a closed-form inverse, the following iterative solution converges quickly.

$$\vec{\Delta q}_{[0]} = \vec{0} \tag{9}$$

$$\cdots$$

$$\vec{\Delta q}_{[n+1]} = \boldsymbol{K_1}^{-1} \left[\vec{\Delta \lambda} - \sum_{d=2}^{D} \boldsymbol{K_d} \left(\vec{\Delta q}_{[n]} \right)^d \right] \tag{10}$$

The iteration takes advantage of the fact that thermo-electric effect on heater resistance, represented by the $\boldsymbol{K}_{d>1}$ matrices, are relatively small perturbations. As heaters are biased closer to thermal runaway singularities from Eq. (7), the thermo-electric effects become stronger. This means more steps are needed to converge, and higher orders of Taylor approximation must be used, necessitating more calibration measurements.

2.2. Optical transmission interpolation

The transmission effect of each MRR filter edge is treated as an independent function, f_i : $\Delta \lambda_i \to T_i$, and calibrated with an interpolation-based approach orgininally developed for a single channel [26]. 20 samples per filter are interpolated to get a continuous estimate of the forward function, $\hat{f}_i(\Delta \lambda_i)$, and inverse, $\hat{f}_i^{-1}(\hat{T}_i)$. This estimate is refined by taking a second set of samples that are nominally uniform in T_i. Calibrated edge transmission functions are shown in Fig. 3. The advantage of the interpolation approach is robustness to arbitrary and non-ideal filter edge shapes; however, it requires that the channel spacing be large enough that filters do not interact optically. In this case, a minimum channel spacing of about 150GHz, seen in Fig. 2(c), gives sufficient isolation, but future work to increase channel density must reexamine the edge calibration approach.

2.3. Cross-gain saturation model

EDFAs at the output of the weight bank in Fig. 2(a) are subject to slow timescale cross-gain saturation, which depends on the weight of each channel in addition to absolute power levels that can fluctuate with polarization, ambient temperature, and fiber strain. The present fiber experiment must model this cross-gain saturation to obtain unbiased weight bank results. While optical amplifiers are not yet widely available on silicon PICs, semiconductor and rare earth ion amplifiers in silicon have been investigated [37, 38], and could potentially use a similar model. We model the cross-gain saturation effect, assuming two homogeneously broadened EDFAs in

Bias	Heater	Filter	Amplifier

i | Δq | $\Delta\lambda$ | T | μ

Bias

| $\lambda_{sig,1}$: 1547nm |
| $q_{bias,1}$: (51.8mA)² |
| $\lambda_{sig,2}$: 1549nm |
| $q_{bias,2}$: (49.0mA)² |
| $\lambda_{sig,3}$: 1550nm |
| $q_{bias,3}$: (39.9mA)² |
| $\lambda_{sig,4}$: 1552nm |
| $q_{bias,4}$: (11.5mA)² |

Heater — Eq. 7

K_1 (×100)

7.38	0.02	0.06	0.12
0.18	10.5	0.04	0.41
0.14	0.63	7.52	0.05
0.07	0.04	0.22	5.91

K_2 (×1,000)

0.03	0.85	0.77	0.08
1.54	-3.13	-1.10	-0.23
0.24	1.39	-1.29	-0.39
0.07	-0.19	-0.32	-2.52

Filter — f_1, f_2, f_3, f_4 (0 to 0.2 nm)

Amplifier — Eq. 14

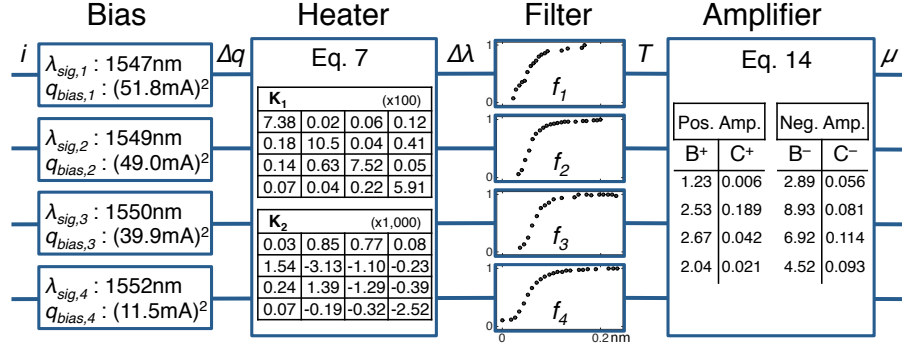| Pos. Amp.		Neg. Amp.	
B^+	C^+	B^-	C^-
1.23	0.006	2.89	0.056
2.53	0.189	8.93	0.081
2.67	0.042	6.92	0.114
2.04	0.021	4.52	0.093

Fig. 3. Diagram of modelling stages showing calibrated parameter values fit during this experiment. **Bias** stage puts variables in differential form around the state of all filters being on-resonance with signals, $\vec{\lambda}_{sig}$. **Heater** stage models thermo-electric, heat transfer, and thermo-optic effects with a predominantly diagonal, linear K_1 matrix and nonlinear corrections (order $D = 2$ shown). **Filter** stage consists of four independent interpolation-based estimates of the transmission along each MRR filter edge. **Amplifier** stage models absolute optical powers and fiber amplifier saturation characteristics preceding photodetection.

non-depleted pump regimes, as

$$\mu_i = P_{in,i}T_{c,i}\left[T_i\frac{g_{i,ss}^+}{1+\frac{P_{amp}^+}{P_s^+}} - \gamma_i(1-T_i)\frac{g_{i,ss}^-}{1+\frac{P_{amp}^-}{P_s^-}}\right] \tag{11}$$

$$P_{amp}^+ = \sum_j P_{in,j}T_{c,j}T_j \tag{12}$$

$$P_{amp}^- = \sum_j P_{in,j}T_{c,j}\gamma_j(1-T_j) \tag{13}$$

where i indicates channel number and T is the tunable microring through port transmission. P_{in} is input power, T_c is net coupling efficiency, γ is drop efficiency, g_{ss} is amplifier small-signal gain, and P_s is saturation power, which is not channel-dependent. P_{amp} signifies total power incident on an EDFA. (+,–) superscripts respectively indicate the amplifiers on through and drop output ports. Not all physical parameters are observable from weight measurements, but the following parameterization yields a fittable model:

$$\mu_i = \frac{B_i^+}{1+\sum_j C_j^+ T_j}T_i - \frac{B_i^-}{1+\sum_j C_j^-(1-T_j)}(1-T_i) \tag{14}$$

The parameter vectors $\vec{B}^{(+,-)}$ and $\vec{C}^{(+,-)}$ (totaling $4N$ parameters) can be fit (i.e. calibrated) with a series of $4N$ measurements at particular tuning states. We introduce a notation $\mu_i^{(xy)}$ to signify the measured weight of channel i when the transmission of channel $T_{j=i}$ is x and the transmission of other channels $T_{j\neq i}$ are y. For example, $\mu_2^{(10)}$ signifies the weight of channel 2 when it is transmitted to the through port ($T = 1$) and channels 1, 3, and 4 are coupled to the drop port ($T = 0$). The calibration procedure starts by measuring $\mu_i^{(11)}$ and $\mu_i^{(10)}$:

$$\mu_i^{(11)} = \frac{B_i^+}{1+\sum_j C_j^+} \qquad \mu_i^{(10)} = \frac{B_i^+}{1+C_i^+} \tag{15}$$
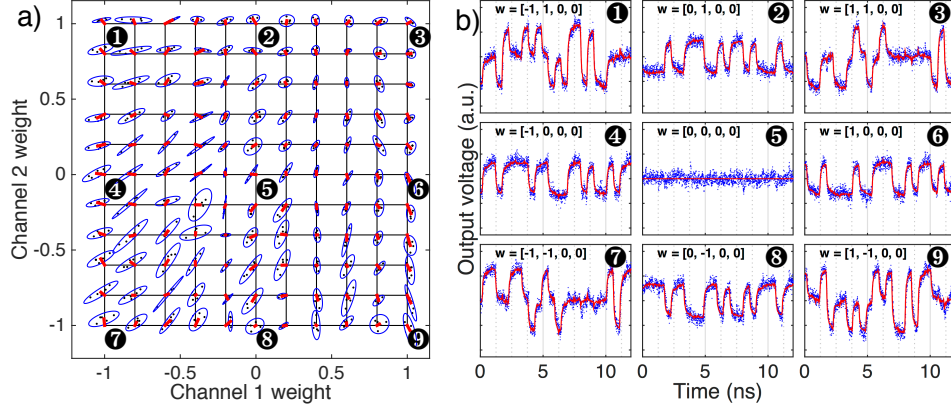
Fig. 4. a) Two-dimensional weight sweep showing controller accuracy and precision. After the calibration procedure, the target weight is swept 5 times over a grid of values from –1 to 1 (black grid). Black points are measured weight data. Red lines show the mean offset from each target grid point. Blue ellipses indicate one standard deviation around the mean. Mean error magnitude is less than 0.072 over the span. Standard deviation remains below 0.063, with a tendency to be larger for negative weights. From this plot, we estimate that the weight can be controlled with an accuracy of 3.8bits. b) Measured traces of 2Gbps signals. [①-⑨] Output signals corresponding to points labeled in (a). The expected signal is in red, while measured traces are in blue. All time and voltage axes have identical scales.

These equations containing $2N$ unknown parameters and $2N$ known measurements can be solved analytically as follows.

$$\frac{\mu_i^{(11)}}{\mu_i^{(10)}} = \frac{1+C_i^+}{1+\sum_j C_j^+} \tag{16}$$

$$C_i^+ = \frac{\mu_i^{(11)}}{\mu_i^{(10)}}\left(1+\sum_j C_j^+\right) - 1 \tag{17}$$

By summing this equation over all i and rearranging, the sum of C^+ can be stated entirely in terms of measured weights,

$$\sum_i C_i^+ = \frac{N - \sum_j \frac{\mu_j^{(11)}}{\mu_j^{(10)}}}{\sum_j \frac{\mu_j^{(11)}}{\mu_j^{(10)}} - 1} \tag{18}$$

at which point it can be substituted into Eq. (17) to recover individual C^+ parameters. The B^+ parameters then fall trivially from Eq. (15). Drop port amplifier parameters, C^- and B^-, follow an identical procedure upon measuring $\mu_i^{(00)}$ and $\mu_i^{(01)}$. The ability to decompose single measurements of $m(t)$ into all weights via Eq. (1) means that $\vec{\mu}^{(11)}$ and $\vec{\mu}^{(00)}$ only require one measurement each, while the dissimilar measurements call for distinct tuning states and therefore N measurements per amplifier. In this derivation, it was assumed that complete switching down to $T = 0$ is possible, which is not always the case in practice. A more algebraically complex calibration technique with nonzero T_{min} can be derived similarly, but is omitted here in the interest of space. Calibrated parameter values found for this experiment are shown in Fig. 3.

Once the forward model parameters have been calibrated, we must invert the mapping $\vec{T} \rightarrow \vec{\mu}$, Eq. (14), in order to work as a feedforward controller rule.

$$\mu_i + \frac{B_i^-}{1 + \sum_j C_j^- (1 - T_j)} = \frac{B_i^+}{1 + \sum_j C_j^+ T_j} T_i + \frac{B_i^-}{1 + \sum_j C_j^- (1 - T_j)} T_i \qquad (19)$$

$$T_i = \frac{\mu_i + \frac{B_i^-}{1 + \sum_j C_j^- (1 - T_j)}}{\frac{B_i^+}{1 + \sum_j C_j^+ T_j} + \frac{B_i^-}{1 + \sum_j C_j^- (1 - T_j)}} \qquad (20)$$

This is solved iteratively as follows:

$$T_{i[0]} = 1 \qquad (21)$$

$$\dots$$

$$T_{i[n+1]} = \frac{\mu_i + \frac{B_i^-}{1 + \sum_j C_j^- \left(1 - T_{j[n]}\right)}}{\frac{B_i^+}{1 + \sum_j C_j^+ T_{j[n]}} + \frac{B_i^-}{1 + \sum_j C_j^- \left(1 - T_{j[n]}\right)}} \qquad (22)$$

This iteration converges quickly when C parameters are small, as in Fig. 3, which is the case when signal powers are less than amplifier saturation power.

3. Results

After the above calibration procedure is performed, the four-dimensional command weight is swept in two-dimensions at a time while the actual weight is recorded. A sweep of command weight values over two channels is shown in Fig. 4. Sweeps over other pairs of channels (not shown) were seen to produce similar results. Figure 4(b) shows time traces compared to expectation at several weight values. Traces 2 and 6 represent the original inputs and traces 8 and 4 their respective inverses. The sweep in Fig. 4(a) is used to analyze accuracy, a.k.a. mean error or repeatable error (red lines), and precision, a.k.a. dynamic error or non-repeatable error (blue ellipses). Mean error is less than 0.072 over the range, corresponding to a weight accuracy of 3.8 bits (i.e. 11.4dB dynamic range), and dynamic error is less than 0.062 for a weight precision of 4.0 bits. Accuracy is 0.7 bits higher than in prior work with a single channel [26] due to procedure and setup changes that minimize polarization drift.

The effects of using simplified models for thermal physics are shown in Fig. 5. When thermal cross-talk and self heating are completely neglected (i.e. $D = 1$ and K_1 is diagonal), accuracy is reduced to 2.8 bits. A constant resistance model (i.e. $D = 1$) is used for Fig. 5(b), yielding a small improvement to 3.0 bits. In both cases, mean errors in Fig. 5 show no clear trend, besides being less accurate towards more negative weight values. Surprisingly, introducing a linear cross-talk model barely improves weight accuracy. This can be explained by the sharp sensitivity of filter transmission to resonant wavelength. The sensitive response of the MRR filter edge necessitates very accurate thermal modelling, in this case, $D = 2$ provided significant improvement. For the devices in this paper, we found $D = 3$ to yield negligible improvement over $D = 2$ since other factors limited precision; however, MRR weight banks with different biases, heater designs, materials, etc. may need increased Taylor orders for sufficient thermal model accuracy. An alternative to thermal tuning is depletion modulation [39], which could eliminate thermal cross-talk and the current-squared dependence, yet requires a more involved fabrication process with a partial etch of the top silicon layer and four dopant levels.
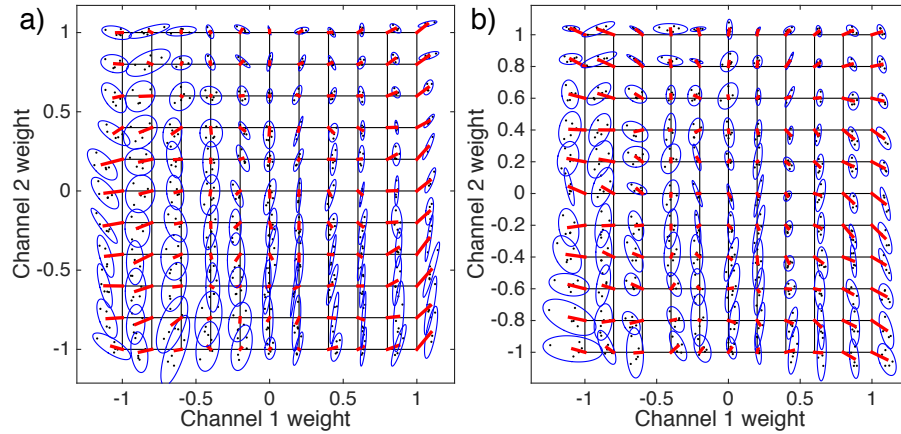
Fig. 5. Weight sweeps for simplified thermal cross-talk models over 5 iterations. The target grid (black), mean error vectors (red), and standard deviation ellipses (blue) are used as in Fig. 4(a). a) No model of thermal cross-talk is applied, and weight accuracy is 2.8 bits (8.4dB dynamic range). b) A first-order $D = 1$ (i.e. constant resistance) model of thermal cross-talk is applied, and weight accuracy is 3.0 bits (9.0dB dynamic range). In both sweeps, the amplifier cross-gain calibration model is applied in order to isolate the effect of thermal cross-talk modeling.

4. Discussion

We can make several conclusions about weight accuracy and precision by examining Fig. 4. Negative weight values are accurate on average, but have more variability because they correspond to the on-resonance condition where sensitivity to fluctuations is greater. The orientation of error ellipses indicate a positive correlation between dynamic errors, indicating intra-sweep power level fluctuations affecting all channels. Most likely, this is explained by drift in fiber temperature and strain affecting polarization and therefore fiber-to-chip coupling efficiency of all channels (T_c in Eq. (11)). Random non-repeatable errors with standard deviation 0.063 are likely attibutable to polarization drift; however, accuracy is still limited by systemic errors, which repeat over multiple sweeps and have a maximum of 0.072.

Several mechanisms could explain much of the discrepancy between DAC resolution (13 bits) and weight accuracy (3.8 bits), providing evidence that DAC resolution is limiting. Firstly, the transmission T dependence on $\Delta\lambda$ is sharply nonuniform over the tuning range of interest (Fig. 3). The ratio of maximum slope to mean slope is 5.3 in the worst case (ch. 2), representing 2.4 bits lost. This effect is somewhat intrinsic to filter edge tuning and difficult to improve without affecting performance. Secondly, the DAC driver is designed to have an 80mA range in order to cover large initial biases; however, the weight tuning range (in this case $\sim$1mA) is much less. This dynamic range mismatch reduces the usable resolution by a factor of 80, or 6.3 bits, accounting for the remaining discrepancy. Controller accuracy is therefore expected to improve by reducing the mismatch between tuning range of interest and driver range.

The range of interest for Δq_1 was $53^2 - 52^2 = 105\text{mA}^2$. Supposing MRRs could start on-resonance with zero bias, either by careful fabrication or by flexible WDM wavelengths, a maximum current of $\sqrt{105} = 10.2\text{mA}$ would be needed to get the same wavelength shift (provided identical device/heater design). If the driver's full range could then be set to 10.2mA, the dynamic range mismatch could be entirely eliminated, in theory. Reducing dynamic range mismatch is an important direction for making systemic controller errors negligible compared

with dynamic errors, and also reducing the DAC resolution needed to achieve a given accuracy.

A natural question to ask of MRR weight banks is minimum channel spacing. Just as the introduction of thermal cross-talk dictated a switch from interpolation-based calibration to model-based calibration, weight banks with dense channel spacing will be subject to optical cross-talk, requiring optical transmission modeling. The weight bank cannot be broken into N independent models or interpolated functions. Since all filters couple into the same output ports, a channel that partially couples through the "wrong" filter can still end up at the intended output, unlike in the case of a demultiplexer wherein unavoidable cross-talk dictates minimum channel spacing [40]. This suggests that model-based calibration of optical cross-talk in a MRR weight bank could be instrumental for increasing channel density and number. Further work in this direction will likely benefit from generalized models of waveguide circuits, such as described in [41].

5. Conclusion

We have demonstrated simultaneous feedforward control of a 4-channel microring weight bank, which could play a major role in large scale processing networks on silicon photonic platforms. The primary enablers of this result were scalable calibration models of thermal cross-talk and amplifier cross-gain saturation. A weight accuracy of 3.8 bits was demonstrated, on par with corresponding state-of-the-art digital electronic neuromorphic hardware [31]. Thermal models that neglect cross-talk and thermo-electric self-heating were found to be insufficient, reducing this accuracy by 1.0 and 0.8 bits, respectively. Parameterized calibration models for optical cross-talk could be developed for more advanced weight bank controllers. Further work could explore the limits of channel density in a single MRR weight bank, and the integration of multiple weight banks into a broadcast-and-weight network.