

REVIEW ARTICLE | NOVEMBER 15 2023

## High-bandwidth density silicon photonic resonators for energy-efficient optical interconnects [F](#)

A. Novick  ; A. James  ; L. Y. Dai  ; Z. Wu  ; A. Rizzo  ; S. Wang; Y. Wang  ; M. Hattink  ; V. Gopal  ; K. Jang  ; R. Parsons  ; K. Bergman 



*Appl. Phys. Rev.* 10, 041306 (2023)

<https://doi.org/10.1063/5.0160441>

 CHORUS



View  
Online



Export  
Citation

### Articles You May Be Interested In

GaSb diode lasers tunable around  $2.6 \mu\text{m}$  using silicon photonics resonators or external diffractive gratings

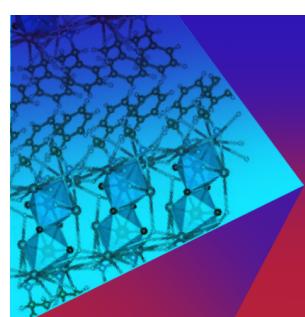
*Appl. Phys. Lett.* (February 2020)

Large-scale and energy-efficient tensorized optical neural networks on III–V-on-silicon MOSCAP platform

*APL Photonics* (December 2021)

Integrated magneto-photonic non-volatile multi-bit memory

*J. Appl. Phys.* (August 2024)



**Applied Physics Reviews**  
**Special Topics**  
**Open for Submissions**

**Submit Today**

 **AIP  
Publishing**

# High-bandwidth density silicon photonic resonators for energy-efficient optical interconnects

Cite as: Appl. Phys. Rev. **10**, 041306 (2023); doi: [10.1063/5.0160441](https://doi.org/10.1063/5.0160441)

Submitted: 1 June 2023 · Accepted: 23 October 2023 ·

Published Online: 15 November 2023



[View Online](#)



[Export Citation](#)



[CrossMark](#)

A. Novick,<sup>1,a)</sup> A. James,<sup>1</sup> L. Y. Dai,<sup>1</sup> Z. Wu,<sup>1</sup> A. Rizzo,<sup>2</sup> S. Wang,<sup>1</sup> Y. Wang,<sup>1</sup> M. Hattink,<sup>1</sup> V. Gopal,<sup>1</sup> K. Jang,<sup>1</sup> R. Parsons,<sup>1</sup> and K. Bergman<sup>1</sup>

## AFFILIATIONS

<sup>1</sup>Electrical Engineering at Columbia University in the City of New York, New York 10027, USA

<sup>2</sup>Air Force Research Laboratory Information Directorate, Rome, New York 13441, USA

<sup>a)</sup>Author to whom correspondence should be addressed: [asher.novick@columbia.edu](mailto:asher.novick@columbia.edu)

## ABSTRACT

The growth of artificial intelligence applications demands ever larger and more complex deep learning models, dominating today's—and tomorrow's—data center and high-performance computing systems. While traditional electronics are failing to keep pace with application demands, silicon photonic (SiPh) interconnects have emerged as a necessary technology to support these systems. SiPh-driven wavelength-division multiplexing (WDM) offers a particularly promising path toward supporting incredibly high-aggregate link bandwidth in a compact and efficient form factor. One of the basic building blocks of these integrated WDM interconnects is the SiPh resonator. Their inherent wavelength selectivity and compact footprint allow for efficient data transmission multiplexed across dozens of carrier wavelengths. Used as add-drop (AD) filters, SiPh resonators are critical to constructing integrated tunable wavelength-selective optical circuit switches as well as for demultiplexing the different carrier wavelengths toward independent wavelength-insensitive photodiodes in a dense wavelength-division multiplexing receiver. Resonators in the all-pass (AP) configuration are widespread as well, allowing for wavelength-selective modulation to drive aggregate link bandwidths far beyond the individual channel data rate. Unlike SiPh Mach-Zehnder modulators (MZM), resonant modulators can be driven using low, complementary metal-oxide-semiconductor drive voltages, allowing for tight co-integration between photonic integrated circuits, fabricated with larger process node technologies, and electronic integrated circuits, designed to exploit the advantages of the latest node. To push toward practical peta-scale interconnects, a comprehensive review of SiPh resonators is required, addressing bottlenecks and design constraints at both the architecture and device levels. We first describe the predominant integrated link architectures and identify their limits. We then discuss the device-level design considerations that can be made for both AD and AP configuration resonators to overcome the system level limits with novel resonator device designs. Analytical models and numerical simulation of resonators are validated by experimental measurement of devices fabricated in a commercial 300-mm foundry, showing a clear path toward volume manufacturing. The demonstrated resonant modulators and filters support the feasibility of increasing the aggregate bandwidth of resonator-driven SiPh interconnects into the peta-scale regime.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0160441>

29 April 2025 06:18:29

## TABLE OF CONTENTS

I. INTRODUCTION .....	1	B. All-pass modulators .....	11
II. RESONATOR-BASED INTERCONNECTS .....	2	IV. CONCLUSION .....	14
A. Dense wavelength-division multiplexing transceivers .....	3		
B. Wavelength-selective optical circuit switches .....	4	I. INTRODUCTION	
III. RESONANT DEVICES .....	5	Accompanied by massive increases in computation power over the past two decades, artificial intelligence (AI) applications have been widely deployed in many scientific and industrial fields such as	
A. Add-drop filters .....	5		

medical imaging, recommendation systems, and natural language processing. The ever-growing requirement for more accurate AI applications has driven the demand for larger and more complex deep learning (DL) models, which require an increasingly larger share of the computation power and energy consumption in today's data center (DC) and high-performance computing (HPC) systems.<sup>1</sup> In response to this rapid model and dataset growth, distributed deep learning (DDL) has been proposed. DDL helps mitigate the memory and computational limitations of individual devices by partitioning the large model (or the dataset) onto multiple computing nodes using various parallelization strategies. These computing nodes cooperate on a training task, each processing a sub-part of the model and synchronizing results with other nodes via an intermediary network. However, current hardware solutions can only provide high-bandwidth connections for a limited group of computing nodes, rendering the communication among the computing nodes a key bottleneck in the training process.<sup>2,3</sup> Each Nvidia DGX node connects 8–16 GPUs using high-speed NVSwitches and NVLinks for up to 900 GB/s aggregate bidirectional bandwidth<sup>4</sup> and relies on 400 Gb/s InfiniBand links to carry the inter-node communication traffic. This bandwidth discrepancy has been shown to severely limit the communication efficiency during the training process.<sup>2</sup> The newest Nvidia DGX system (H100) aims to bridge this gap by extending NVSwitches and NVLinks for inter-node communication but is currently only able to scale up to 256 GPUs in a DGX-Super-POD. In addition to the bandwidth bottleneck, the energy consumption of these DL training models has become prohibitively high and brings severe impact onto the environment.<sup>5</sup>

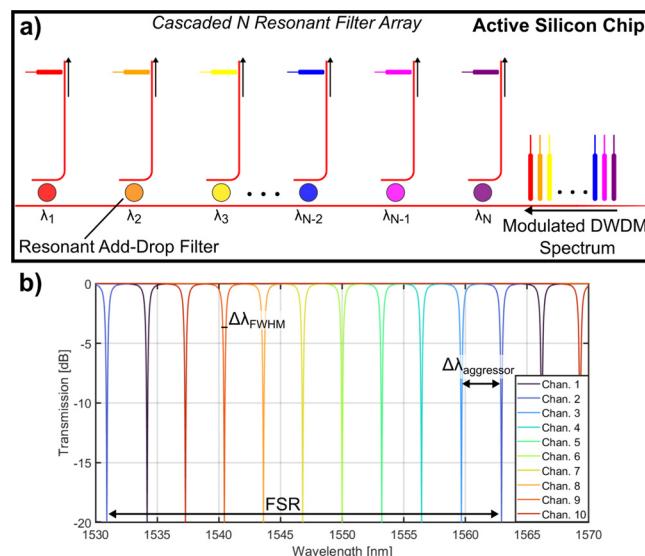
To tackle the bandwidth and energy problems posed by DL trainings, an emerging trend is to incorporate embedded silicon photonics (SiPh) technologies in the network as a means to achieve peta-scale bandwidth interconnects<sup>6–12</sup> with high energy efficiency and high-bandwidth densities.<sup>10,13</sup> These SiPh chips can leverage the mature complementary metal-oxide-semiconductor complementary metal-oxide-semiconductor (CMOS) foundries used to fabricate microelectronics chips, resulting in a clear path to low-cost volume scaling and compatibility with standard semiconductor packaging techniques.<sup>14</sup> SiPh combined with dense wavelength-division multiplexing (DWDM) is a particularly promising direction toward achieving these interconnects. With relatively modest per-carrier baud rates ( $\leq 25$  Gb/s), the energy per bit can remain low, while aggregate bandwidth density scales via massive parallelism; this is primarily achieved by increasing the number of parallel optical carriers. An example of a commercial SiPh interface is TeraPHY,<sup>9</sup> which supports up to 2 Tb/s bandwidth per chiplet. Recently, a promising Kerr frequency comb-driven SiPh transceiver architecture<sup>15</sup> has been reported, driven by an integrated frequency comb<sup>16</sup> DWDM light source and using (de-) interleavers to split and combine wavelength channels in order to achieve massive parallelism and scale up the data transmission bandwidth within a single fiber. In both of these examples, SiPh resonators are used to both perform wavelength-selective modulation and photo-detection on the transmitter and receiver ends of the photonic integrated circuit (PIC) architecture, respectively. The inherent wavelength selectivity of these resonators, along with their compact footprint and energy-efficient operation,<sup>7</sup> is critical toward realization of the aggressive bandwidth density and energy efficiency metrics (1 Pb/s/mm<sup>2</sup> at 100 fJ/b<sup>7,17</sup>) required for the next generation of DC and HPC interconnects.

Resonant devices provide a further advantage over non-resonant devices—for example, the Mach-Zehnder interferometer (MZI)—as the resonator field enhancement enables drastically more compact footprint, compatibility with leading-edge CMOS node driving voltages, and substantially lower energy consumption. Due to these attributes, a resonator-driven system's bandwidth density scalability is dramatically increased since the compact footprints and low driving voltages allow for close co-integration of PIC with electronics at unprecedented bandwidth densities.<sup>18,19</sup> Here, we aim to describe the design trade-offs at both the system and device levels for these scalable resonator-driven optical interconnects, which must be made to push the limits at both levels and achieve the next generation of integrated optical interconnects and the systems that rely upon them.

## II. RESONATOR-BASED INTERCONNECTS

For any DWDM system, the aggregate bandwidth-per-fiber ( $f_{\text{total}}$ ) is calculated as the per-channel modulation bandwidth ( $f_{\lambda}$ ) multiplied by the number of multiplexed wavelength channels ( $N_{\lambda}$ ). For the most part,  $f_{\lambda}$  is determined by the bandwidth limitations of individual devices, to be discussed in greater depth in Sec. III. Maximizing  $N_{\lambda}$  is strongly dependent on the high-level circuit architecture. Consider the most basic scalable resonator-based DWDM architecture, a single array of cascaded resonators, illustrated in Fig. 1(a). The limitation on  $N_{\lambda}$  can be calculated as

$$N_{\lambda} = \text{floor} \left[ \min \left( \frac{\text{FSR}}{\Delta \lambda_{\text{ag}}}, \frac{IL_{\text{limit}}}{IL_{\text{off}}} + 1 \right) \right], \quad (1)$$



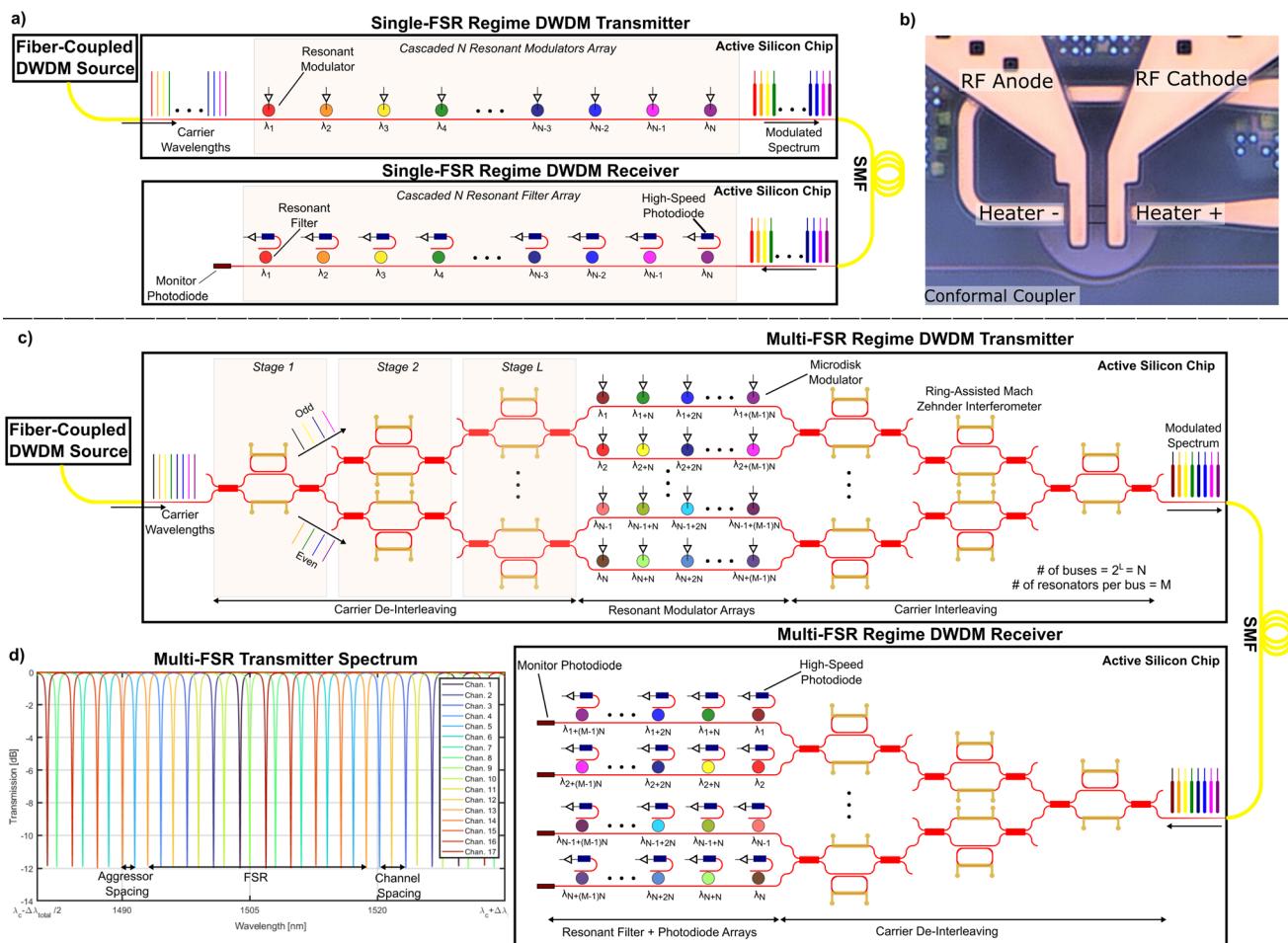
**FIG. 1.** (a) Basic schematic of a resonator-based  $1 \times N$  integrated wavelength demultiplexer, with a modulated multi-wavelength source input from the right. Demultiplexing devices in this case are illustrated as individual resonators in the add-drop configuration. (b) Annotated transmission spectrum at the leftmost end after the cascaded resonator array  $1 \times 10$  wavelength demultiplexer, with dropped spectral content of each resonator illustrated in a different color. Each resonators drops a Lorentzian-shaped section of spectral content with full-width half-maximum ( $\Delta \lambda_{\text{FWHM}}$ ), which repeats along the wavelength axis with the periodicity of the free spectral range (FSR). The number of wavelength channels that can be fit within a single FSR is determined by the spectral distance between adjacent channels ( $\Delta \lambda_{\text{aggressor}}$ ).

where free spectral range (FSR) is for the cascaded resonators,  $\Delta\lambda_{\text{ag}}$  is the distance between a channel carrier wavelength and an adjacent resonance corresponding to a different channel,  $IL_{\text{limit}}$  is the maximum tolerable difference in received power from the first to last channel in the array, and  $IL_{\text{off}}$  is the off-resonance IL of each resonator. FSR and  $IL_{\text{off}}$  are determined by the specific device design and will be discussed further in Sec. III A.

The two key interconnect architectures that must be kept in mind are transceivers and optical circuit switches (OCSes). For a resonator-based DWDM system, many of the physical layer design optimization choices are strongly correlated between these two systems, allowing for efficient co-design targeting extreme bandwidth density scaling within a fully reconfigurable interconnect fabric.

### A. Dense wavelength-division multiplexing transceivers

For transceivers, the transmitter and receiver architectures of resonator-based SiPh transceivers are typically near-symmetric, mirrored around the I/O interface, as illustrated in Fig. 2(a). At the receiver end, the PIC architecture strongly resembles an OCS, but with photo-detectors at the resonator drop ports rather than an off-chip fiber coupled interface. The transmitter side consists of a cascaded array of all-pass (AP) configuration resonators, such as the one shown in Fig. 2(b). The resonant modulators imprint information on respective carrier frequencies via modulation of the resonator effective index ( $n_{\text{eff}}$ ), shifting the resonance between blocking and non-blocking states.<sup>23–27</sup> As with wavelength-selective OCSes,  $f_{\text{total}} = f_{\lambda} \times N_{\lambda}$ , but both terms become further constrained by the additional architecture



**FIG. 2.** (a) Exemplary single-FSR resonator-based DWDM transceiver architecture, requiring all spectral channels to fit within a single resonator FSR. (b) Micrograph of an individual microdisk modulator, with RF contacts separated into quadrants and a rectangular integrated thermal tuner comprised of doped silicon.<sup>7,20</sup> (c) Proposed schematic of a multi-FSR resonator-based DWDM transceiver architecture, utilizing ring-assisted MZI based even-odd interleaving to increase the channel spacing at the cascaded resonator arrays.<sup>7,15,21</sup> (d) Spectral transmission of the multi-FSR regime transceiver from the input to output of the cascaded resonator array. Increased channel spacing leaves room for additional aliases to fall in the space between channels at the resonator arrays. In this case, the aggressor channels become a full FSR apart instead of the channels spectrally adjacent, as annotated, but otherwise the resonator design considerations remain the same as for a single-FSR regime link, with the exception of increased sensitivity to resonator FSR selection.<sup>22</sup>

and device-level design considerations for the resonant modulators. For example, at the transmitter side,  $\Delta\lambda_{\text{ag}}$  requires additional margin due to the reduction in worst case channel spacing introduced by the shift in resonance required to modulate.<sup>28</sup>

Integrated Kerr frequency combs are a promising solution for driving IMDD DWDM receivers at a large scale, since they offer several advantages over platforms based on DFB arrays and supermode laser diodes. However, it remains a challenge to achieve sufficient comb line power and conversion efficiency for line spacings of  $< 100$  GHz, which sets a practical lower limit on the achievable channel spacing. Furthermore, the maximum micro-resonator FSR is limited by practical design constraints and excess bend radiation losses. To this end, the resonator FSR must be carefully chosen to prevent resonance aliases from overlapping with non-target comb lines. The typical strategy is to design the resonator FSR such that it extends beyond the spectral bandwidth of the link, ensuring no alias resonances fall near a channel. Since the de-interleaved even and odd groups of comb lines nearly preserve the spectral bandwidth of the original comb, however, it is infeasible to design micro-resonators with FSRs greater than the comb bandwidth beyond a limit due to the prohibitively small physical dimension required that both raise manufacturability and bend loss concerns.<sup>23</sup> Thus, in this single-FSR regime, there is a practical limit on the number of usable channels assuming a constant channel spacing.

In contrast, a multi-FSR channel arrangement scheme aims to place resonance aliases in between the modulated channels while maintaining an adequate spacing between each channel and its nearest aggressor for crosstalk minimization. To mathematically formulate the problem of finding a valid multi-FSR channel arrangement, we define the following two auxiliary variables:<sup>7</sup>

$$\mathcal{S} = \frac{\Delta\lambda_{\text{ch}}}{\Delta\lambda_{\text{ag}}}, \quad (2a)$$

$$\mathcal{T} = \frac{\text{FSR}}{\Delta\lambda_{\text{ag}}}, \quad (2b)$$

where  $\Delta\lambda_{\text{ch}}$  is the effective channel spacing on each bus after de-interleaving,  $\Delta\lambda_{\text{ag}}$  is the reduced spacing between a channel and its nearest aggressor alias, and FSR is the resonator free spectral range. The quantities  $\mathcal{S}$  and  $\mathcal{T}$  represent the channel spacing and the resonator FSR normalized to the aggressor spacing. With these auxiliary variables defined, find a valid resonator FSR, one would need to find a proper pair of integers for  $\mathcal{S}$  and  $\mathcal{T}$  satisfying Eqs. (2a) and (2b) under the constraint of  $\Delta\lambda_{\text{ag}}$  being adequately large to minimize crosstalk. It has been shown that a valid multi-FSR solution occurs when the following two conditions are simultaneously met: (i)  $\mathcal{S}$  and  $\mathcal{T}$  are co-prime integers, i.e., having no common factors other than 1, and (ii)  $\mathcal{T}$  is greater than or equal to  $N_{\text{ch}}$ , the desired number of channels on each bus.<sup>7</sup> In this way, we can continue to scale up the number of channels per bus without being limited by the FSR of the resonators [Fig. 2(d)].

It is possible to further scale these SiPh-DWDM transceiver architectures by multiplexing over additional dimensions, beyond just the wavelength dimension. Multi-dimensional scaling allows for the multiplication of  $f_{\text{total}}$  by the number of orthogonal multiplexing axes utilized,  $K$ . The combined total multi-dimensional per-fiber-bandwidth becomes  $f_{\text{hybrid}} = K \times f_{\text{total}}$ . There has been particularly high

interest in exploring polarization division multiplexing (PDM) and mode division multiplexing (MDM) as paths toward achieving these multi-dimensional transceivers. While not compatible with one another, both PDM and MDM techniques are on their own fundamentally compatible with DWDM, only requiring the necessary margin in the link budget to accommodate additional penalties contributed by their respective non-ideal added (de)multiplexing devices. Physical realization of these hybrid links requires compact and low-loss, SiPh devices that are both broadband and demonstrate high channel isolation. There are many such integrated PDM devices that have been proposed and demonstrated.<sup>29–32</sup> It is now common for SiPh devices to demonstrate insertion loss (IL)  $< 1$  dB and polarization crosstalk suppression as high as 20 dB on-chip. Furthermore, PDM can be realized with standard single-mode fiber (SMF) interconnects between chips, assuming the link budget allows for increased crosstalk. In the case that the link budget does not provide such margin, PDM may still be feasible by substituting SMF for polarization-maintaining fiber (PMF), which utilizes a unique cross section that eliminates the cylindrical symmetry of SMF to create fast and slow propagation axes, maintaining polarization orthogonality along the fiber channel. Unlike PDM, MDM allows for  $K > 2$ , but can be limited by device performance and complications that arise from requiring few-mode fiber (FMF) or multi-mode fiber (MMF), rather than standard SMF, between PIC nodes. A number of recent studies have been shown simultaneous support for MDM and broadband DWDM supported by novel integrated devices, both for handling multiple propagating modes intra-PIC and inter-PIC.<sup>33–37</sup> While the inclusion of PDM or SDM with DWDM can introduce additional IL and crosstalk penalties to the link due to imperfect devices, significant work has been devoted over the years to reducing losses and crosstalk between orthogonal modes in devices such as waveguide bends,<sup>38,39</sup> mode (de-)multiplexers,<sup>33,35,40,41</sup> and polarization converters.<sup>42</sup>

## B. Wavelength-selective optical circuit switches

Aside from using resonator-based circuits to improve data transmission and reception, another area where they may be useful is in the interconnect fabric. DC and HPC traffic workloads have been shown to exhibit high degrees of spatial locality.<sup>43,44</sup> To cope with the traffic imbalance, researchers proposed using reconfigurable OCSes to dynamically allocate more bandwidth to traffic hotspots by simultaneously switching all the wavelength channels from any given input port to any output port (i.e., spatial switching).<sup>45</sup> However, these spatial OCSes lack the ability to selectively route individual wavelengths in the wavelength domain to achieve higher switching granularity, which is especially useful for DWDM architectures.

Wavelength-selective optical circuit switches enable one to establish all-to-all networks with low complexity, where computing resources are directly interconnected through leveraging the properties of DWDM. One of the most reliable and extensively researched realizations of wavelength-selective OCSes is the arrayed waveguide grating router (AWGR). Since AWGRs are passive optical components, reconfiguration within the routing fabric is not feasible. Consequently, to address different endpoints, rapid wavelength-tunable lasers must be employed at each node's transmitter, as AWGRs allocate spatial destinations based on an optical signal's carrier wavelength as well as spatial input. While the development of fast wavelength-tunable lasers continues to be a subject of ongoing research,<sup>46</sup> AWGRs are mature,

commercially accessible, with radix of  $370 \times 370$  well-demonstrated in the literature.<sup>47</sup> For a dynamic wavelength-selective switch fabric, one can look toward implementations which feature microring resonator (MRR) in its interconnects. The inherent wavelength selectivity of MRR switching cells lends itself toward the straightforward development of wavelength-selective switching topologies in silicon photonics. Currently, the MRR-based switches with the largest radix include the  $8 \times 8$  crossbar<sup>48</sup> and switch-and-select,<sup>49</sup> but have been experimentally emulated to include a  $16 \times 16$  Clos.<sup>50</sup> The metrics seen in Dai *et al.*<sup>50</sup> can be seen to correlate very closely with the scaling proposed in previous studies,<sup>49</sup> making a practical case for the  $128 \times 128$ .

Spatial-and-wavelength-selective switches are an extension on the classical wavelength-selective variants, as they allow for signal routing on both the wavelength and spatial domains. Particularly, they can steer any subset of channels in an input multi-wavelength signal to a given spatial destination; a capability that is currently lacking but desirable in both data center and high-performance computing applications for handling heterogeneous traffic patterns.<sup>51,52</sup> The inceptive designs of these novel switching architectures were straightforward: demultiplexing input DWDM signals on the wavelength domain, collecting and spatially reroute channels of the same wavelength, then re-multiplexing the result to form a DWDM output.<sup>53,54</sup> However, to do so in a non-blocking manner requires an unsustainable number of switch elements (SEs) at scale. More modern architectural iterations work to integrate spatial switches with wavelength-selective components—decreasing the hardware complexity of the interconnects considerably.<sup>55</sup>

### III. RESONANT DEVICES

#### A. Add-drop filters

As a result of their inherent wavelength selectivity, compact footprint, and efficient resonance tuning (thermally or otherwise), SiPh resonators are ubiquitous in integrated DWDM systems.<sup>15,33,35,56,57</sup> Furthermore, SiPh resonator-based DWDM systems boast unrivaled design flexibility at both device and system levels due to the development of parameterized models that accurately capture the physical device behavior.<sup>58,59</sup> These parameterized models can be used to develop resonators for both add-drop (de)multiplexing filters<sup>60,61</sup> as well as active high-speed modulators;<sup>23,24,62</sup> applications even exist utilizing their non-linear phase response near resonance, such as ring-assisted MZI (RAMZI) devices.<sup>63–65</sup>

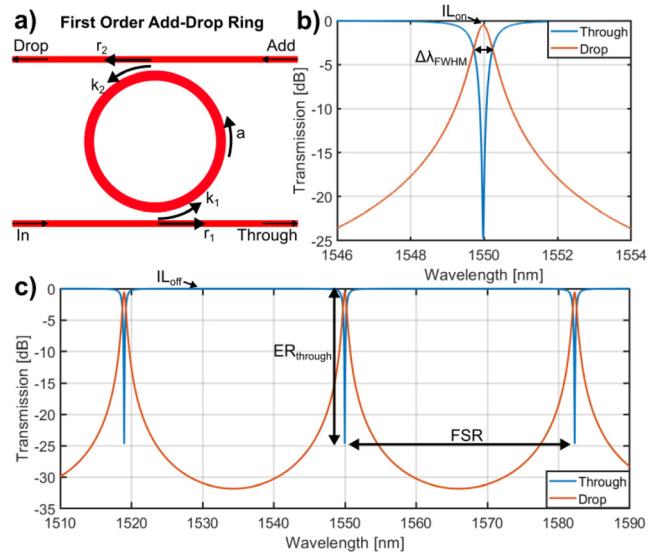
The most basic building blocks of the DWDM receiver or OCS are the resonant add-drop (AD) filter. Figure 3 illustrates a basic schematic of a first-order AD ring resonator, as well as typical transmission spectra, all annotated with the most relevant parameters. Assuming incident light at the “In” port of Fig. 3(a), spectral transmission for first-order resonators at the “Through” and “Drop” ports follows:<sup>59</sup>

$$\frac{E_{\text{through}}}{E_{\text{in}}} = e^{j(\pi+\phi)} \frac{r_2 a - r_1 e^{-j\phi}}{1 - r_1 r_2 a e^{j\phi}}, \quad (3a)$$

$$\frac{I_{\text{through}}}{I_{\text{in}}} \equiv T_{\text{through}} = \frac{r_2^2 a^2 - 2r_1 r_2 a \cos \phi + r_1^2}{1 - 2r_1 r_2 a \cos \phi + r_1^2 r_2^2 a^2}, \quad (3b)$$

$$\frac{E_{\text{drop}}}{E_{\text{in}}} = e^{j(\pi+\phi/2)} \frac{\sqrt{(1-r_1^2)(1-r_2^2)a}}{1 - r_1 r_2 a e^{j\phi}}, \quad (3c)$$

$$\frac{I_{\text{drop}}}{I_{\text{in}}} \equiv T_{\text{drop}} = \frac{(1-r_1^2)(1-r_2^2)a}{1 - 2r_1 r_2 a \cos \phi + r_1^2 r_2^2 a^2}, \quad (3d)$$



**FIG. 3.** (a) First-order radial ring filter with straight coupling waveguides in the AD configuration. Input-output waveguide ports are labeled, with self-coupling and cross-coupling coefficients annotated consistent with Eqs. (3) and (4). (b) Modeled spectral transmission near the resonance for a critically coupled first-order AD resonator with  $\lambda_{\text{res}} \approx 1550$  nm. (c) Spectral transmission of the same resonator model over a wider wavelength range, illustrating the resonance periodicity.

with  $r$  being the E-field self-coupling coefficient of the waveguide and  $r^2 = 1 - k^2$ , where  $k$  is the E-field cross-coupling coefficient between the waveguide and resonator,  $a$  being the E-field self-coupling coefficient of the resonator for each round trip around the cavity, relating to the power attenuation coefficient,  $\alpha_{\text{total}}$ , as  $a^2 = \exp(-\alpha_{\text{total}})$ , and  $\phi = \beta L$ , where  $\beta = 2\pi n_{\text{eff}}/\lambda$ ,  $n_{\text{eff}}$  is the effective index of the propagating optical mode in the resonator, and  $L$  is the round trip cavity propagation length. While not directly used as wavelength de-multiplexing filters, the AP resonator configuration is a sub-category of AD, lacking the add-drop coupled waveguide opposite the in-through coupled waveguide, and can be modeled with the same equations by setting  $r_2$  to 1. AP resonators and their applications in transceiver architectures as modulators will be further discussed in Sec. III B.

While seemingly independent, many of these parameters are fundamentally intertwined with one another along with the physical (sub-)device properties, thus becoming a complex design space with a slew of trade-offs that must be considered to achieve the target performance metrics. For example, while it is clear that  $\beta$  is strongly dependent on  $\lambda$ ,  $n_{\text{eff}}$  and the coupling coefficients are functions of  $\lambda$  as well. Before untangling these complex co-dependencies, it is helpful to define several additional relevant terms,<sup>59</sup>

$$\lambda_{\text{res}} = \frac{n_{\text{eff}} L}{m}, \quad m = 1, 2, 3, \dots, \quad (4a)$$

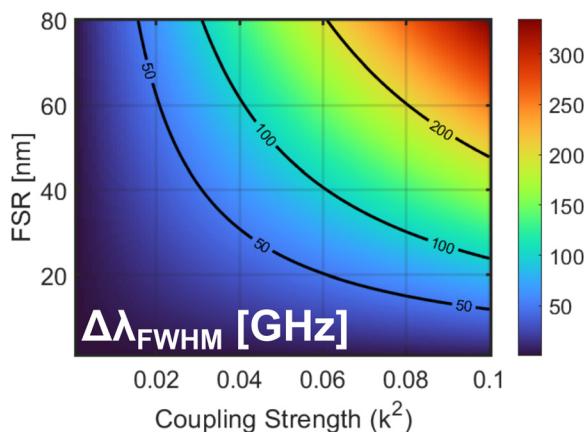
$$\text{FSR} = \frac{\lambda^2}{n_{\text{g}} L}, \quad (4b)$$

$$\Delta\lambda_{\text{FWHM}} = \frac{(1 - r_1 r_2 a) \lambda_{\text{res}}^2}{\pi n_{\text{g}} L \sqrt{r_1 r_2 a}}, \quad (4c)$$

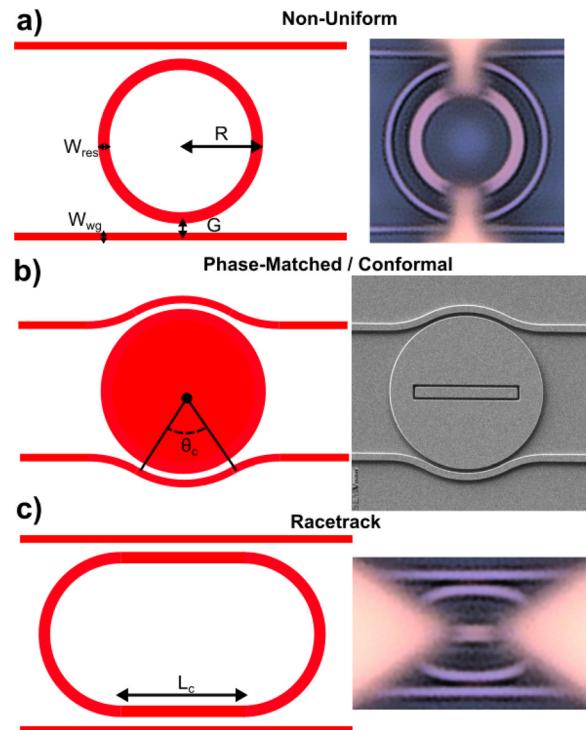
a resonance wavelength,  $\lambda_{\text{res}}$ , occurs whenever an integer multiple,  $m$ , of wavelengths fits within one optical round trip of the resonator cavity, defined as the geometric path length of the propagating mode,  $L$ , multiplied by the mode's effective index,  $n_{\text{eff}}$ . For a single device, repeated resonances will occur, with the periodicity equal to the average FSR over the wavelength span between a given set of  $\lambda_{\text{res},m}$  and  $\lambda_{\text{res},m+1}$ , which is a function of the mode's group index,  $n_g$ .  $\Delta\lambda_{\text{FWHM}}$  is the filter's Lorentzian full-width half-maximum (FWHM). Additionally, the resonator Q factor can be calculated as  $Q = \lambda_{\text{res}}/\Delta\lambda_{\text{FWHM}}$ .

We can use the two sets of equations, Eqs. (3) and (4), to make several initial observations regarding the design space. First, critical coupling will occur when  $r_1 = r_2a$ , meaning that the ER<sub>through</sub> will be maximized, while the on-resonance drop-port insertion loss,  $IL_{\text{on}}$ , will be minimized (a typically desired outcome). Second, in agreement with common intuition, minimizing  $\alpha$ —preferably such that  $a \approx 1$ —results in better performance (lower  $IL_{\text{on}}$ ) and a more simple set of equations to work with. Third, it is clear that for fixed coupling and loss coefficients, the FSR is proportional to  $\Delta\lambda_{\text{FWHM}}$ ; thus, to achieve the same filter bandwidth with a different FSR, the coupling strength must be designed accordingly, as illustrated in Fig. 4.

The terms  $n_{\text{eff}}$ ,  $n_g$ ,  $L$ , and  $\alpha$  present in Eqs. (3) and (4) are all determined by the resonator's physical waveguide geometry, such as width and bend radius, and  $n_g$  can be calculated from the wavelength-dependent  $n_{\text{eff}}$  as  $n_g = n_{\text{eff}} - \lambda_0 \frac{dn_{\text{eff}}}{d\lambda}$ . Referring to Fig. 3(a), an AD resonator is constructed out of several constituent sub-devices: the waveguides constructing the resonator itself (at least one pair of bent waveguides), and the coupling sections connecting the input–output waveguides to the resonator. Figure 5 illustrates several styles of AD resonator, showing examples of both single-mode rings and multi-mode disks, as well as a variety of coupling geometries. It should be noted that while the largest resonator FSR is typically achievable with the geometry shown in Fig. 5(b), such a geometry is also highly susceptible to increased  $IL_{\text{off}}$ , as described in Eq. (1), due to the tightly bent coupled waveguide geometry.



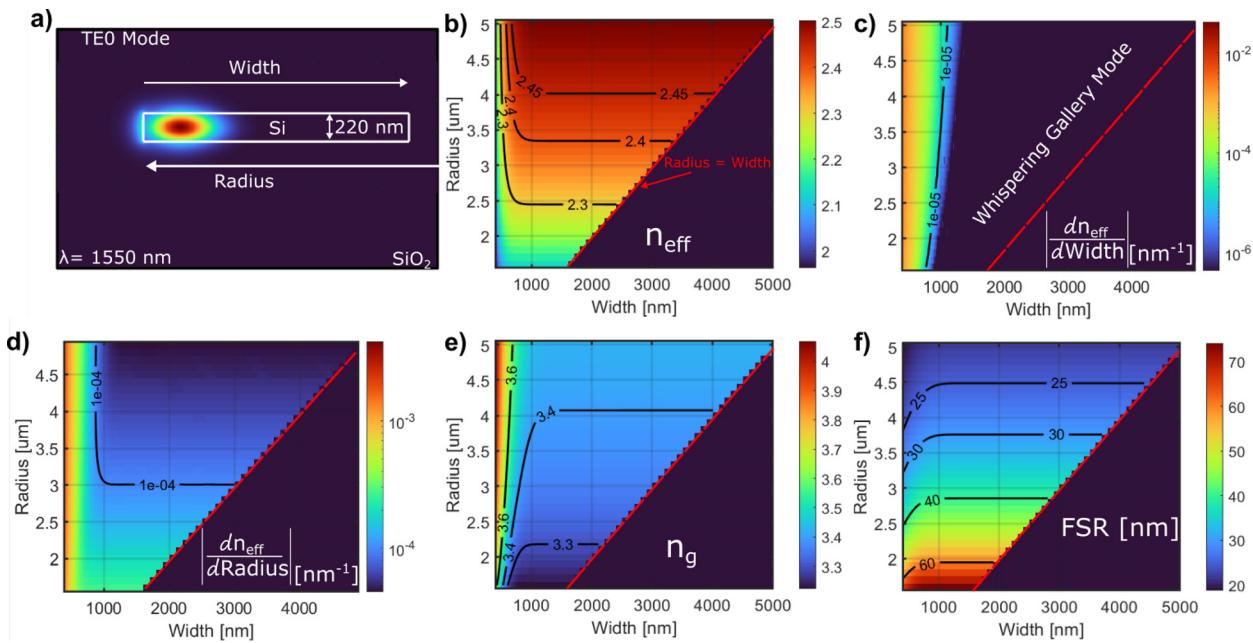
**FIG. 4.** Resonator FWHM as a function of resonator FSR and coupling strength—defined as the magnitude of optical power coupling coefficient ( $k^2$ )—for a lossless ( $a = 1$ ) and critically coupled ( $r_1 = r_2 \iff k_1^2 = k_2^2$ ) AD resonator. Note that the FWHM increases with FSR for a fixed  $k^2$ , meaning the filter Q factor is inversely proportional to FSR independent of coupling.



**FIG. 5.** Schematics and corresponding microscope images for (a) non-uniform coupling for ring resonator with radius  $R = 5 \mu\text{m}$  and  $G = 170 \text{ nm}$ , (b) phase-matched coupling for disk resonator  $R = 4.5 \mu\text{m}$ ,  $G = 200 \text{ nm}$ , and coupling angle  $\theta_c = 30^\circ$ , and (c) racetrack ring with  $R = 1.75 \mu\text{m}$ ,  $G = 150 \text{ nm}$ , and extended coupling length  $L_c = 0.5 \mu\text{m}$ .<sup>66</sup>

The bent waveguide sections create the all-important loopback path allowing for the resonance to occur, and their physical characteristics can dominate geometric design complexity. The Lumerical MODE solver is used to determine the fundamental transverse-electric (TE) mode parameters for a given bent 2D waveguide cross section, parameterized in Fig. 5(a). To allow for a unified geometric description consistent between single-mode rings and multi-mode disks, we describe the bent waveguide geometry with the radius defined as the origin curvature out to opposing waveguide sidewall, and the waveguide width measured back from the outer sidewall toward the inner sidewall. In these coordinates, the bent waveguide transitions from a ring to a disk when the radius is equal to the width, a threshold denoted by the dashed red line in Figs. 6(b)–6(f). The  $n_{\text{eff}}$  and  $n_g$  were calculated for the fundamental TE mode for the 2D cross section of bent waveguide at  $\lambda = 1550 \text{ nm}$  for a variety of waveguide widths and bend radii.

It can be seen in Figs. 6(b)–6(d) that as the width of the waveguide increases, the  $n_{\text{eff}}$  becomes less dependent on the inner portion of the waveguide, while as the radius decreases the  $n_{\text{eff}}$  becomes more sensitive to variations along the outer waveguide wall, effectively corresponding to local changes in the radius [as defined in Fig. 6(a)]. The impact of this relative sensitivity to variation is most apparent when measuring the standard deviation in  $\lambda_{\text{res}}$  for the same resonator across multiple copies of the same device. This measurement was done



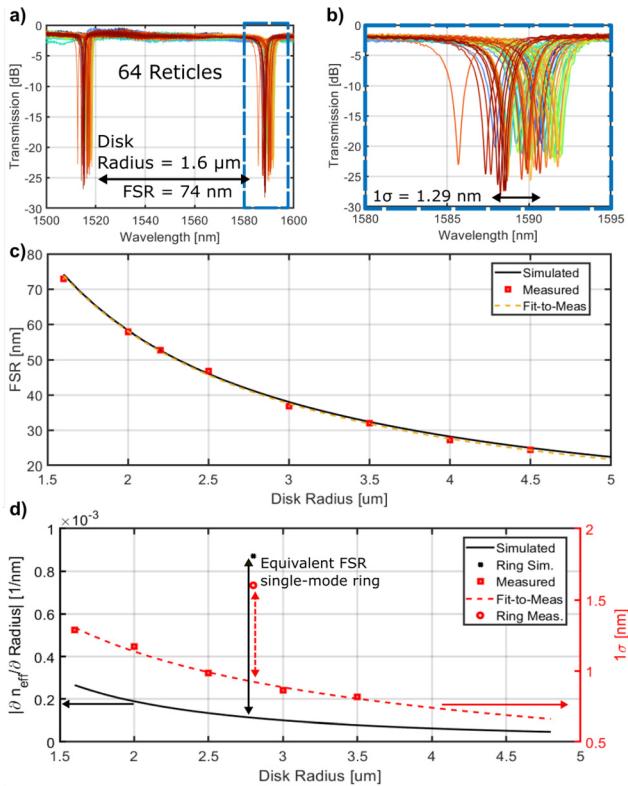
**FIG. 6.** (a) Numerically solved E-field profile at 1550 nm of the fundamental TE mode for a 2D cross section of a bent silicon-on-insulator (SOI) waveguide (silicon region outlined in white). The radius is measured from the origin of curvature (on the right) to the outer sidewall of the bent waveguide (on the left). The waveguide width is measured from the outer sidewall toward the inner sidewall. (b)  $n_{\text{eff}}$  of the TE0 mode for  $400 \text{ nm} \leq \text{width} \leq 5000 \text{ nm}$  and  $1.6 \mu\text{m} \leq \text{radius} \leq 5 \mu\text{m}$ . (c) The marginal change in  $n_{\text{eff}}$  with respect to the width to the right of the contour line ( $\frac{\partial n_{\text{eff}}}{\partial \text{Width}} = 1 \times 10^{-5}$ ) TE0 can be considered a whispering gallery mode, where the mode does not interact with the internal waveguide structure beyond that threshold. (d) The marginal change in  $n_{\text{eff}}$  with respect to small changes in outer radius, showing increased mode interaction with the outer sidewall at small radii. (e)  $n_g$  for the waveguide configurations, used in Eq. (4b) to calculate (f) the FSR of a resonator constructed out of a constant radially bent waveguide.

for both single-mode rings and multi-mode disks over a full 300 mm wafer fabricated by AIM Photonics, with the measurement results plotted in Fig. 7. In agreement with that intuition, Figs. 7(a) and 7(b) show the normalized through transmission for a disk resonator with a radius of  $1.6 \mu\text{m}$  with a conformal coupler, measured across 64 reticles. Figure 7(c) demonstrates close matching between the FSR for simulated and measured disk resonators over a range of radii. Figure 7(d) shows that for equivalent FSR, a single-mode ring resonator has substantially greater  $1\sigma$  variation in resonance location than a multi-mode disk resonator, in agreement with the expected increased fabrication sensitivity for single-mode bent waveguides from Fig. 6(d).

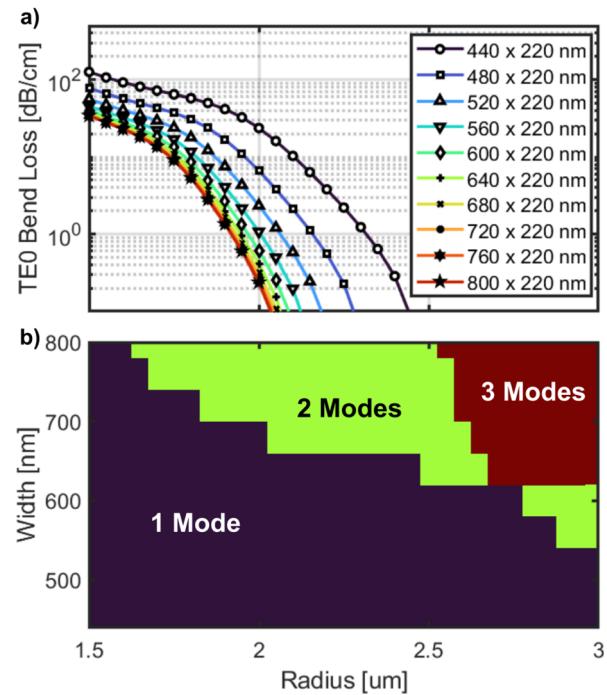
As stated before, minimizing the round trip resonator loss is attractive, and often a top priority for meeting design metrics. This loss can manifest from several sources, such as bend radiation loss, propagation loss (typically a combination of material loss and sidewall roughness), mode mismatch loss, and loss due to the couplers. The same Lumerical MODE solver was used to record the expected bend radiation loss of the TE0 mode in a bent waveguide. Illustrated in Fig. 8(a), radiative bend loss is inversely proportional to both the waveguide width and the waveguide radius. Furthermore, as shown in Fig. 6(d), as the waveguide width increases the mode becomes less sensitive to variations in outer sidewall ( $|\partial n_{\text{eff}}/\partial \text{radius}|$ ), also reducing losses due to sidewall roughness, especially for modes that have crossed into the whispering gallery territory and no longer interact with the inner wall of the bent waveguide. Thus, whispering gallery resonator modes typically have a much greater intrinsic Q than their non-whispering gallery counterparts.<sup>23</sup>

Loss is not necessarily always undesirable; particularly in the case of multi-mode resonators, it can be very helpful to induce mode selective loss. Referring back to Fig. 6(c), once TE0 can be considered a whispering gallery mode, the propagation characteristics (including loss) are no longer dependent on the internal bent waveguide interface. On the other hand, higher-order modes (such as TE1) are less spatially confined and so will still significantly interact with the internal waveguide beyond the whispering gallery cutoff radius of TE0, as shown in Fig. 9(a). Therefore, mode selective loss to attenuate higher-order mode resonances and preserve the disk FSR can be achieved through selective placement of metal vias and/or silicon dopants [due to free carrier absorption (FCA)], such as what might be used to form an integrated heater. By proper design of selective loss, a disk with radius =  $3.5 \mu\text{m}$ , shown in Fig. 9(b), can be treated as a single-mode resonator with an FSR and intrinsic Q factor greater than a single-mode radially bent waveguide would typically allow. Properly designing with mode selective loss in mind is particularly critical for multi-mode disk modulators, where there is a direct trade-off between distance from the metal contact and series resistance through doped silicon, which can limit modulation bandwidth.<sup>7,67</sup>

To couple into, and out of, a standard resonator, a four-port device is necessary. In most cases, this device is a variation of DC, though in the case of racetrack style rings, a  $2 \times 2$  multi-mode interferometer (MMI) is frequently employed.<sup>68</sup> Several coupling schemes are illustrated in Fig. 5, with schematic adjacent to microscopic or SEM image of a representative fabricated device. There are two key parameters to consider with any coupler, wavelength-dependent coupling

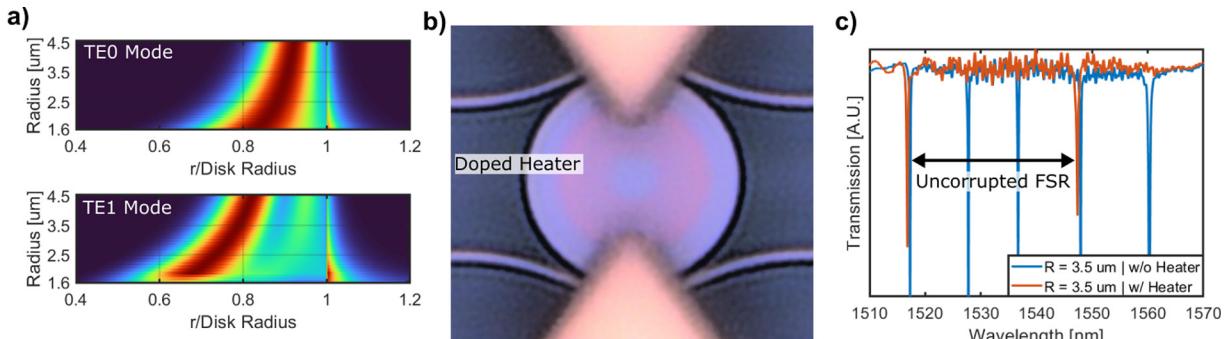


**FIG. 7.** (a) Measured and normalized through port spectral transmission of a disk filter with radius =  $1.6 \mu\text{m}$  and FSR = 74 nm, measured on 64 reticles sampled over a full 300 mm wafer. (b) Magnified plot of the L-band resonance of the disk filter on each reticle. (c) Disk FSR vs drawn radius showing simulation in close agreement with measured FSR. (d) Simulated disk  $|\partial n_{\text{eff}}/\partial \text{radius}|$  as a function of radius, plotted on the same horizontal axis as the measured  $1\sigma$  standard deviation in resonance location, showing a strong positive correlation. A single-mode ring's measured  $1\sigma$  resonance variation for an equivalent FSR is also shown for comparison.



**FIG. 8.** (a) Radially dependent fundamental TE0 mode radiative bend loss for a simulated bent SOI waveguide cross section in the Lumerical MODE solver. As the bend radius decreases, bend loss increases. Conversely, as waveguide width increases radiative bend loss decreases, allowing for a high Q resonator to be constructed out of more tightly bent, but wider waveguides. (b) The number of supported propagating modes found by the Lumerical MODE solver for the same radius and width. Higher-order modes fail to be supported as the bend radius reduces, reducing the risk of undesired coupling to those higher-order modes.<sup>65</sup>

strength,  $k^2(\lambda)$  and  $IL(\lambda)$ . Looking at Eq. (3), coupling and loss coefficients are strongly intertwined when it comes to device performance,  $a$  and  $r_2$ , in particular, are completely indistinguishable in the equation for  $T_{\text{through}}$ . In the case of a lossless resonator ( $a \approx 1$ ), the coupling



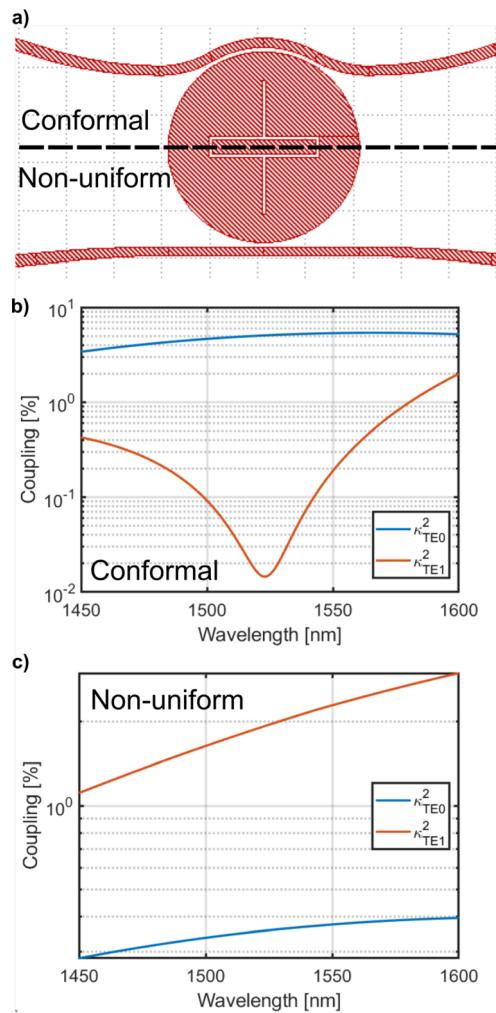
**FIG. 9.** (a) Stacked cross-sectional slices (at a height of 110 nm) for the mode E-field intensity of many 220 nm tall SOI microdisk resonators of different radii for the first two TE modes, simulated in the Lumerical MODE solver. The x axis is normalized relative to the disk radius so that the outer edge of the disk is always at  $x = 1$ . TE1 is more broadly distributed in the radial direction both inward and outward from the edge of the disk. (b) Microscope image of a fabricated AD microdisk filter, with the heavily doped integrated heater ring visible, designed to attenuate TE1+ resonances via FCA while minimally impacting TE0. (c) Transmission spectrum for the through port of the AD micro-disk filter with and without the integrated heater doping, showing the efficacy of creating mode selective loss. In this case, the heater attenuated the TE0 resonances more than intended, but more conservative placement of the doped heater will achieve the desired higher-order resonance suppression with minimal impact on TE0.

strength and path length will entirely determine the filter bandwidth, as described in Fig. 4. Only considering the case of DC, the maximum  $k^2(\lambda)$  is a function of the coupling length, coupling gap, and phase matching between arms. A fundamental convenience can be observed here, as the FSR is increased the required maximum  $k^2$  to support the same filter  $\Delta\lambda_{\text{FWHM}}$  goes down as well, allowing for a reduced coupling length, which is in geometric agreement with the resonator shrinkage that is tied to the increased FSR.

For the most uniform spectral resonator characteristics, the most uniform broadband coupling coefficients are desired. A symmetric DC tends to exhibit a strong wavelength dependence, but there are many alternative coupling schemes making use of asymmetric geometries that are demonstrated to be both more broadband and robust to fabrication variations.<sup>69–73</sup> Bent directional couplers (BDCs), which can be referred to as conformal couplers as shown in Fig. 5(b), are particularly promising as they allow for increasing the  $L_c$  and having a more mode selective coupling to TE0 for a multi-mode disk.<sup>7</sup> In the case of an AD resonator, increasing  $L_c$  and thus  $k^2$  primarily serve to increase the  $\Delta\lambda_{\text{FWHM}}$  or reduce the drop port's  $IL_{\text{on}}$  by increasing the ratio of  $a$  to  $r_2$  so that coupling out to the waveguide is much greater than coupling out due to loss. In terms of mode-selective coupling, Fig. 10 shows the coupling into the TE0 and TE1 modes of a microdisk resonator for a straight-coupled waveguide and conformal-coupled waveguide designed to couple into TE0. For the former, the difference between mode coupling is less pronounced and the more widely dispersed TE1 mode field's evanescent tail overlaps more with the coupling waveguide allowing  $k_{\text{TE1}}^2 > k_{\text{TE0}}^2$ . On the other hand, in the conformal coupler case, the angular phase matching between TE0 in the resonator and the coupling waveguide dominates the resultant coupling coefficients, despite the greater TE1 field overlap that is still present. By minimizing power coupled into higher-order resonator modes, suppressing the corresponding resonances becomes easier utilizing techniques such as induced mode selective loss as described before.

In most cases, the resonator cannot be considered completely lossless, especially in the coupling region. Excess loss of the coupler device has two primary effects, one being the degradation of the resonator Q factor as a result of  $a$  reducing,<sup>59</sup> and the other is the increase in  $IL_{\text{off}}$  that is accumulated in the through path.<sup>66</sup> While the unloaded Q of a disk resonator's TE0 mode is much larger than a single-mode rings unloaded Q for the same FSR,<sup>23</sup> the introduction of the waveguide coupling region disturbs that. The smaller the resonator, the more abrupt the transition into and out of the coupling region tends to be. In fact, the phase-matched coupler is especially lossy due to the mode mismatch at the S-bend inflection and phase matching requirement to narrow the waveguide resulting in substantially more bend loss, as described in Fig. 8(a). Recently, an adiabatically varied quasi-ellipsoid single-mode ring resonator was demonstrated [shown in Fig. 5(c)].<sup>66</sup> The straight waveguide coupling allows for reduced mode mismatch in the coupling region, resulting in substantial reduction of coupler loss. This permits off-resonance light to pass by the filter relatively unimpeded, as shown in Fig. 11, and an increase in the limit for  $N_\lambda$  in a DWDM system by allowing for the second term in Eq. (1) to grow.

To determine the minimum permissible  $\Delta\lambda_{\text{ag}}$  term from Eq. (1), we must calculate the channel crosstalk penalty as it relates to both the spacing between channels and the  $\Delta\lambda_{\text{FWHM}}$  of the filters themselves. For this purpose, we can calculate the drop-port transmission of an ideal lossless AD filter, with  $a \approx 1$  and  $r_1 = r_2 = r$ , as

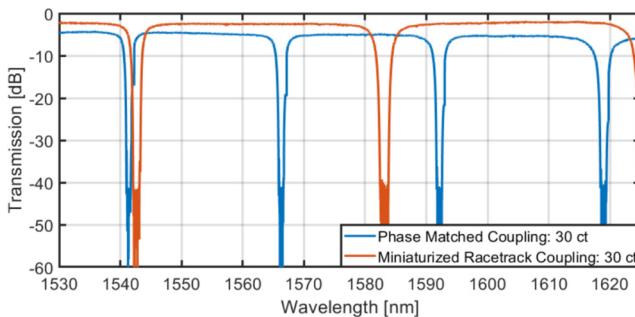


**FIG. 10.** (a) Layout of a microdisk resonator of  $R = 4 \mu\text{m}$  with a conformally coupled waveguide on the top side and a non-uniform straight waveguide coupling on the bottom side. Coupling strength from waveguide to microdisk simulated in Lumerical FDTD for (b) coupling via conformal coupler designed to selectively couple into the microdisk TE0 and (c) coupling via non-uniform coupler, much more strongly coupled into TE1 than TE0.

$$T_{\text{drop}_{\text{ideal}}} = \frac{(1 - r^2)^2}{1 + r^4 - 2r^2 \cos\left(\frac{2n_{\text{eff}}(1 - r^2)}{n_g r} \chi_0\right)}, \quad (5)$$

where  $\chi_0$  is a normalized detuning parameter,  $\chi_0 = \frac{\Delta\lambda_{\text{ag}}}{\Delta\lambda_{\text{FWHM}}}$ .

It is clear from Eq. (5) that the impact of channel spacing on crosstalk is directly proportional to the  $\Delta\lambda_{\text{FWHM}}$  of the filter. To reduce crosstalk between channels due to spectral overlap,  $\chi_0$  should be maximized, so  $\Delta\lambda_{\text{FWHM}}$  should be minimized. At the same time,  $\Delta\lambda_{\text{FWHM}}$  must be broad enough to capture the spectral content comprising the signal information for the targeted channel. To determine this minimum  $\Delta\lambda_{\text{FWHM}}$ , the experiment described in Fig. 12(a) was done, using a tunable bandpass filter (TBF) (II–VI Waveshaper) and variable



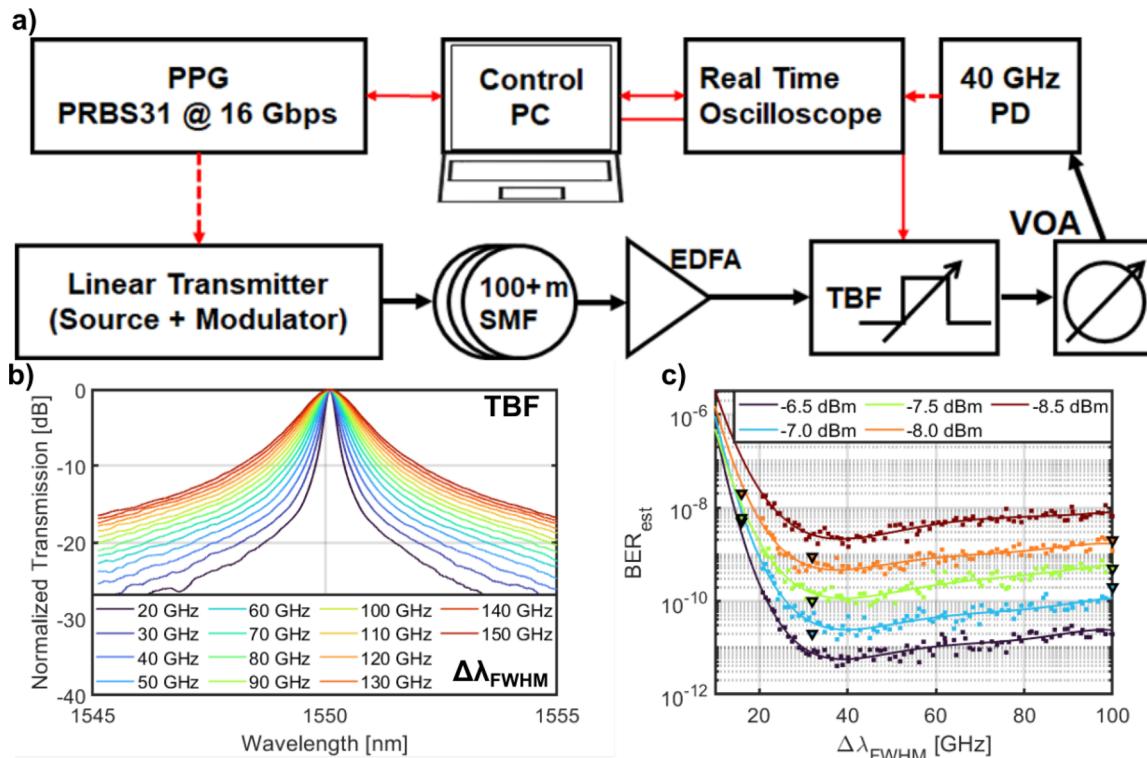
**FIG. 11.** Transmission of 30 cascaded identical disk resonators with conformal coupling and 30 cascaded miniaturized racetrack single-mode rings with straight waveguide coupling. Cutback measurements for each showed 0.1 and 0.02 dB  $\text{IL}_{\text{off}}$ , respectively, despite the single-mode ring having 60% larger FSR.<sup>66</sup>

optical attenuator (VOA) (Keysight N776A) to emulate a range of Lorentzian transmissions with constant peak transmission, shown in Fig. 12(b). A pulse pattern generator (PPG) (Anritsu MP1900A) generated a pseudo-random bit sequence (PRBS)  $2^{31} - 1$  bits long (PRBS31) at 16 Gb/s to drive a linear reference transmitter (Thorlabs MX35E) and output a modulated optical signal. An erbium doped fiber amplifier (EDFA) (Fiberlabs AMP-FL8021-CLB) was used to

overlay broadband spectral noise and the 80 GSa/s real-time oscilloscope (Keysight DSOZ632A) captures the eye diagrams of the filtered 16 Gb/s PRBS31 signal after being converted back to an electrical signal by a 40 GHz photoreceiver (Thorlabs RXM40AF). The eye diagram Q factor is converted to an estimated bit error rate (BER),  $\text{BER}_{\text{est}}$ , by Freude *et al.*<sup>74</sup>

$$\text{BER}_{\text{est}} = \frac{1}{2} \operatorname{erfc} \left( \frac{Q}{\sqrt{2}} \right). \quad (6)$$

The  $\text{BER}_{\text{est}}$  for a range of input powers and programmed  $\Delta\lambda_{\text{FWHM}}$  s is measured and plotted in Fig. 12 as dots, with a polynomial fit line. To validate  $\text{BER}_{\text{est}}$  the photoreceiver electrical output was connected to the receiver input of the PPG to measure the real BER for several configurations, which are plotted as triangular markers on Fig. 12, in good agreement with Eq. (6). Due to the filtering out of broadband EDFA noise, a 0.5 dB improvement in receiver sensitivity is observed to maintain a constant BER for the filter  $\Delta\lambda_{\text{FWHM}}$  being reduced from  $\geq 5 \times$  the data rate ( $\geq 80$  GHz) to  $2 \times$  the data rate (32 GHz). Once the  $\Delta\lambda_{\text{FWHM}}$  reduced below  $2 \times$  the data rate, significant signal degradation occurred. In the presence of high-frequency signal distortion, it has also been demonstrated that setting the filter  $\Delta\lambda_{\text{FWHM}}$  to  $2 \times$  the data rate can result in 3 dB gains in signal quality via equalization in the optical domain.<sup>75</sup> Setting the minimum



**FIG. 12.** (a) Experimental setup for measuring signal impact of optical filter. The VOA after the filter is set to automatically adjust attenuation to maintain constant output power to compensate for additional peak attenuation at narrower  $\Delta\lambda_{\text{FWHM}}$ . (b) Measured transmission spectrum of programmable optical filter set to Lorentzians of varying  $\Delta\lambda_{\text{FWHM}}$ . (c) BER at several constant received optical powers for a 16 Gbps NRZ optical signal. Triangular marker measurements are direct BER measurements simultaneously measured with the signal Q. Dots correspond to automated Lorentzian  $\Delta\lambda_{\text{FWHM}}$  sweeps that measured signal Q and converted to  $\text{BER}_{\text{est}}$  as defined by Eq. (6). Solid lines correspond to a polynomial fit.

$\Delta\lambda_{FWHM}$  to be twice the data rate, we can determine the minimum  $\Delta\lambda_{ag}$  to satisfy a specific threshold of crosstalk suppression, calculated by Eq. (5). Using  $-20$  dB as a common acceptable crosstalk suppression threshold,<sup>76</sup> it follows that the minimum  $\Delta\lambda_{ag}$  for a first-order resonator drop demultiplex would need to be at least  $5\times$  the  $\Delta\lambda_{FWHM}$  and  $10\times$  the data rate per wavelength.

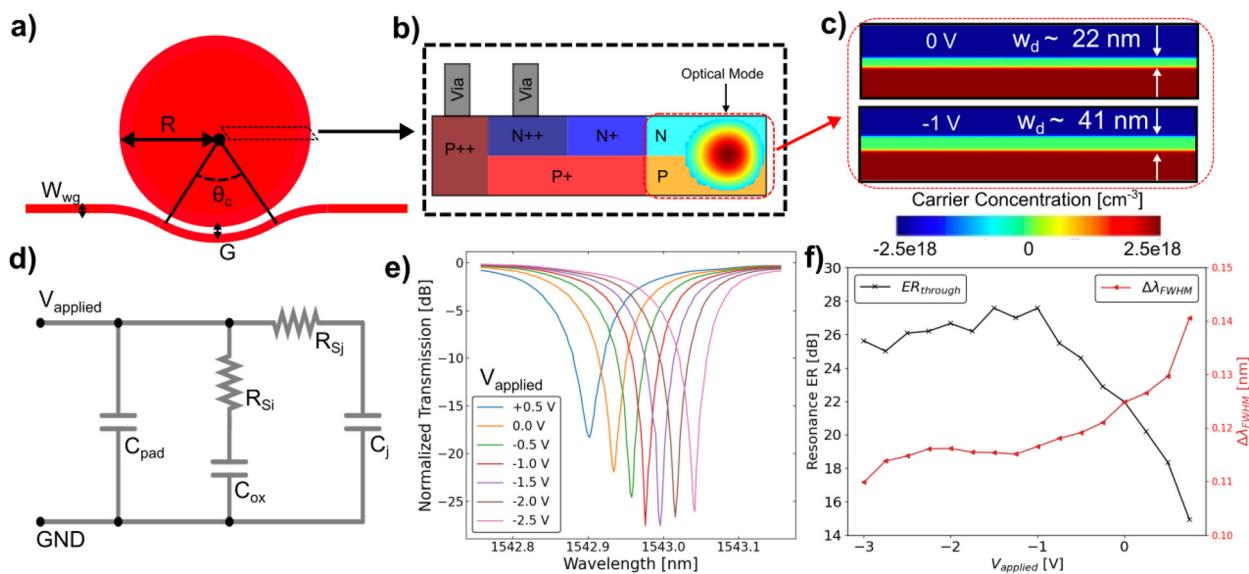
Taking the case of a single-FSR link configuration, as described in Fig. 2(a), and assuming an FSR covering the full C-band (4.8 THz), to maintain  $-20$  dB crosstalk at the receiver for a data rate of 25 Gbps, only 19 carrier frequencies could be utilized. In the case of a reduced crosstalk suppression requirement ( $>-20$  dB), reducing  $\Delta\lambda_{ag}$  relative to a given  $\Delta\lambda_{FWHM}$  is one option. Another alternative to increase the channel count would be to implement more complex architectures, such as the multi-FSR link configuration as described in Fig. 2(b) or implementing a multi-dimensional multiplexing scheme as described earlier. Recent work showed a promising path toward increased aggregate link bandwidth due to the combination of reduced channel data rate allowing for additional link margin, which in turn permitted increased crosstalk and signal truncation penalties and thus closer channel spacings.<sup>22</sup> As the crosstalk is more due to the Lorentzian filter shape than any significant spectral overlap of the signals propagation along the channel, utilizing an architecture like the multi-FSR link, particularly the flat-top even-odd interleavers, can efficiently double the channel spacing at each stage, multiplying the effective spectral utilization and bandwidth of the resonator-based optical interconnect.

## B. All-pass modulators

The relatively passive design principles for AD resonant filters equally apply to AP resonant modulators. Simply modifying Eq. (3) setting  $r_2 = 1$  converts the AD configuration to the AP configuration.

The same FSR, IL, and crosstalk constraints remain as with the AD filters, but additional optimizations and design choices are added to support high device electro-optic bandwidth. Most critically, these include optimization of the active electro-optic region and choosing operating bias targets, both electrical and optical. While there are three main categories of resonator modulation—index ( $n_{eff}$ ), loss ( $\alpha$ ), and coupling ( $r_1$ ) modulation—index modulation is by far the most common, especially for DWDM interconnects.<sup>25</sup> Index modulation is performed by creating a waveguide in which the  $n_{eff}$  of the propagating optical mode is dependent on an applied voltage gradient, most often by using the dopants to create a PN diode within the optical waveguide.<sup>7,23,77</sup> These p-type and n-type dopants are placed in close proximity to form the PN diode and can be forward or reverse biased by an external driving signal to modulate  $n_{eff}$ . While a greater change in  $n_{eff}$  can be achieved via forward biasing and injecting free carriers into the waveguide, the modulation bandwidth is limited and requires complex pre-emphasis techniques to exceed single-digit GHz bandwidths.<sup>77,78</sup> On the other hand, in reverse bias index modulation is performed by varying the width of the PN diode depletion region where it overlaps with the optical mode; this change in depletion width modulates the  $n_{eff}$  via the plasma dispersion effect. Depletion mode index modulators can support tens of GHz bandwidth per-carrier wavelength, and integrating them into resonant devices allows for error-free data transmission to be achievable at CMOS-compatible drive voltages.<sup>7,15,23,79</sup> Furthermore, modulation by charging and discharging the junction capacitance,  $C_j$ , does not consume DC power, and the energy per bit can be calculated as  $E_{bit} = 1/4C_j V_{pp}^2$ , allowing for reported  $\approx 1$  fJ/bit modulation efficiency up to data rates of 25 Gbps.<sup>23</sup>

Figure 13 illustrates the fundamental operating principles for a depletion mode resonant modulator. Figure 13(a) shows an aerial



**FIG. 13.** (a) Schematic of an AP disk modulator with a conformal coupler designed to selectively excite the TE<sub>0</sub> mode. (b) Cross section of a vertical junction doping profile for the disk modulator, showing junction, series, and contact dopings. (c) Simulated 2D free carrier profile of the cross section at 0 and 1 V reverse biases with corresponding depletion widths. (d) Small-signal circuit of a depletion mode resonant modulator. (e) A measured disk modulator resonance shifting due to various applied reverse biases. Deeper reverse bias corresponds with red shifting of the resonance and vice versa. (f) Loss due to FCA decreases as the depletion width grows with deeper reverse bias, showing a transition between undercoupled and overcoupled regimes at  $-1.5 \leq V_{applied} \leq -1$  V.

schematic view of a typical AP configuration disk modulator, including the mode selective conformal coupler, similar to the one shown in Fig. 5(b). Figure 13(b) is a possible cross section for the vertical PN junction in the plane of optical propagation, with the whispering gallery mode field plotted to show overlap with the lightest doped PN diode region. For an abrupt PN junction, the theoretical depletion width,  $w_d$ , is given by

$$w_d = \sqrt{\frac{2\epsilon(N_D + N_A)(\phi_{bi} - V)}{qN_AN_D}}, \quad (7)$$

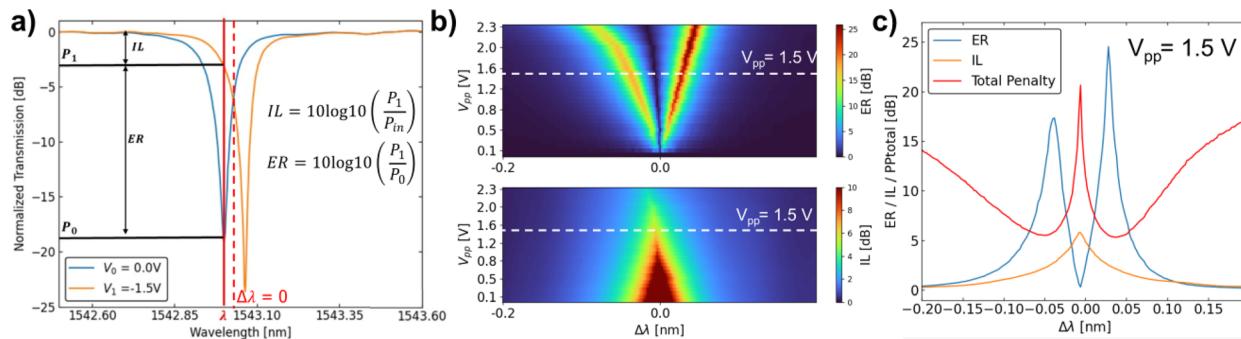
where  $\epsilon$  is the permittivity of silicon,  $N_D$  is the donor dopant concentration,  $N_A$  is the acceptor dopant concentration,  $\phi_{bi}$  is the built-in voltage,  $q$  is the elementary charge, and  $V$  is the applied bias. For silicon at 300 K with  $N_D = N_A = 2.5 \times 10^{18} \text{ cm}^3$ ,  $\phi_{bi} \approx 1 \text{ V}$ , Eq. (7) gives  $w_d \approx 30 \text{ nm}$ , similar to the numerically simulated vertical junction-free carrier profile results shown in Fig. 13(c).<sup>7</sup> For a disk modulator, contact and series dopings should be placed in accordance with the same principles as demonstrated in Fig. 9, in order to suppress higher-order mode resonances while adding minimal excess attenuation to TE0. We can use the whispering gallery cutoff threshold as described in Fig. 6(c) to determine the minimum distance between disk edge and higher concentration dopants, which can be defined as the minimum junction width ( $w_j$ ). Minimizing  $w_j$  or otherwise diminishing the electrical path length from the driver to the optical mode, reduces the series resistance of the junction ( $R_{Sj}$ ), thus increasing the modulator RC bandwidth. Another avenue of mode selective loss generation in the disk modulator configuration is to add a drop bus waveguide, converting the resonator to an asymmetric AD configuration that is much more strongly coupled to higher-order modes, such as the one shown in Fig. 10.

Figure 13(d) shows a small-signal RC approximation of the depletion mode modulator assuming a short trace length between electrical pads and the photonic device.<sup>67</sup> With increased reverse bias, the geometric path length of the resonator remains unchanged, but the change in  $n_{\text{eff}}$  directly affects the  $\phi$  term in Eq. (3), shifting the resonance and changing the transmission characteristics of a fixed carrier wavelength near the resonance. This resonance shift can be seen in Fig. 13(e) and is typically referred to as the depletion response,

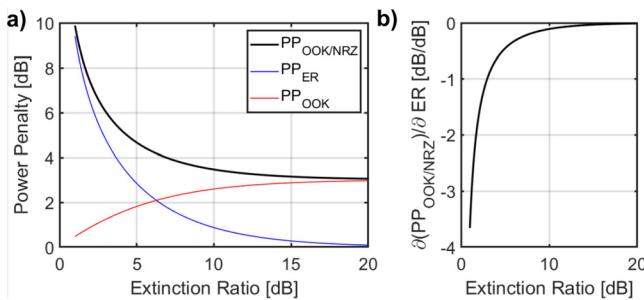
measured in GHz (or sometimes pm) per applied Volt. Depletion response measurements as high as 30 GHz/V have been reported for disk modulators utilizing a vertical PN junction due to the improved overlap between optical mode and depletion region relative to an equivalent lateral PN junction, which typically shows depletion responses far less than even 5 GHz/V.<sup>23,80</sup> Figure 13(f) plots the dependence of  $ER_{\text{through}}$  and  $\Delta\lambda_{\text{FWHM}}$  on  $V_{\text{applied}}$ , showing critical coupling is achieved at  $V_{\text{applied}} \approx 1.25 \text{ V}$ , with the resonator being undercoupled and overcoupled with shallower and deeper reverse biases, respectively. The DC depletion response measurement is used to characterize steady-state performance metrics of depletion mode modulators. For energy-efficient broadband DWDM, modulation is typically in the on-off keying (OOK) non-return to zero (NRZ) format, comprised of a sequence of high- and low-power optical signal strengths, representing a series of binary 1's and 0's.<sup>72,1</sup> A power penalty, defined as the difference in input power required to achieve the same BER at the receiver relative to an ideal signal, is added on top of the baseline IL of the device due to the non-infinite modulation extinction ratio (ER) and can be calculated as<sup>7</sup>

$$PP_{\text{OOK/NRZ}} = 10 \log_{10} \left( \frac{r+1}{r-1} \right) + 10 \log_{10} \left( \frac{2r}{r+1} \right), \quad (8)$$

where  $r$  is the linear ER and can be calculated as  $r = 10^{ER/10}$ . To modulate, a fixed detuning wavelength,  $\Delta\lambda_{\text{mod}}$ , is chosen relative to the mean resonance position where applying  $\pm V_{pp}/2$  will shift the resonance between blocking and non-blocking positions for the carrier wavelength,  $\lambda$ , as shown in Fig. 14(a). Figure 14(b) shows the ER and IL as a function of  $V_{pp}$  and  $\Delta\lambda$  based on the depletion response measurements in Fig. 13(e). Taking the ER and IL at a fixed  $V_{pp} = 1.5 \text{ V}$ , we can calculate the total penalty as  $PP_{\text{total}} = IL + PP_{\text{OOK/NRZ}}$  and determine an optimal operating point for  $V_{pp}$  and  $\Delta\lambda$ , minimizing  $PP_{\text{total}}$  as shown in Fig. 14(c). Looking more closely at  $PP_{\text{OOK/NRZ}}$ , we can separate the terms into the power penalty from having non-infinite ER [ $PP_{\text{ER}} = 10 \log_{10} \left( \frac{r+1}{r-1} \right)$ ] and the power penalty from reduced average optical power due to the expectation of equiprobabilistic 1's and 0's transmitted [ $PP_{\text{OOK}} = 10 \log_{10} \left( \frac{2r}{r+1} \right)$ ]. Figure 15(a) plots the dependence of  $PP_{\text{OOK/NRZ}}$  on signal ER and the contribution of these constituent parts, showing asymptotic approach to



**FIG. 14.** (a) Index modulated OOK NRZ is performed by shifting the modulator resonance between blocking and non-blocking states relative to a fixed carrier wavelength. The IL and ER are a function of both the peak-to-peak voltage swing ( $V_{pp}$ ) and detuning wavelength ( $\Delta\lambda$ ). Resonance extinction ratio,  $ER_{\text{through}}$ , as a function of coupling strength,  $k^2$ , into the resonator fundamental mode for a simulated microdisk modulator with dopant profile as shown in Fig. 13. Critical coupling occurs at  $k^2 = 0.029$ . (b) Derivative of  $ER_{\text{through}}$  with respect to  $k^2$ . For the same  $ER_{\text{through}} = 15 \text{ dB}$ , the sensitivity to variations in  $k^2$  is half as much on the overcoupled side ( $k^2 = 0.04$ ) as on the undercoupled side ( $k^2 = 0.02$ ).



**FIG. 15.** (a) The OOK NRZ modulation power penalty,  $PP_{\text{OOK/NRZ}}$ , and its constituent components plotted against the signal ER. As the power penalty associated with non-infinite ER,  $PP_{\text{ER}}$ , goes down with increased ER, the modulation insertion loss, represented by  $PP_{\text{OOK}}$ , increases toward 3 dB due to equally probable 0 and 1 bits in the transmitted data. Overall, as  $ER \rightarrow \infty$ ,  $PP_{\text{OOK/NRZ}} \rightarrow 3$  dB. (b) The derivative of  $PP_{\text{OOK/NRZ}}$  with respect to ER, showing rapidly diminishing returns for ER greater than  $\approx 12$  dB.

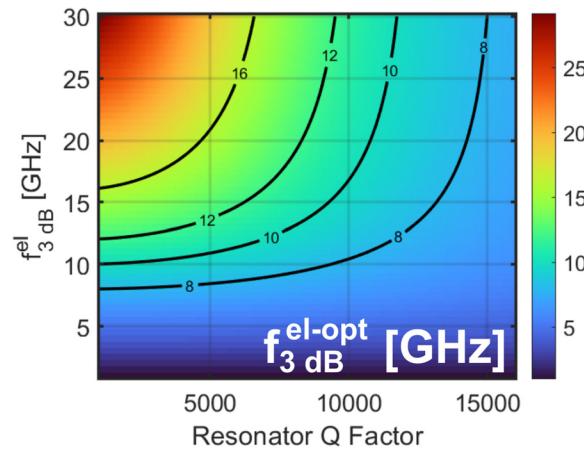
3 dB with  $ER \rightarrow \infty$ . Figure 15(b) plots the derivative of  $PP_{\text{OOK/NRZ}}$  with respect to signal ER, suggesting pursuing signal ER beyond  $\approx 12$  dB results in strongly diminishing returns. As shown in Fig. 14(c), the ER is fundamentally tied to IL and the optimal operating point is not necessarily where ER is maximized.

The DC depletion response measurements suggest expected modulation performance when the resonant modulator is operating well below the electro-optic bandwidth limits ( $f_{3\text{dB}}^{\text{el-opt}}$ ), which can be calculated as<sup>7</sup>

$$f_{3\text{dB}}^{\text{el-opt}} = \left[ \left( \frac{1}{f_{3\text{dB}}^{\text{el}}} \right)^2 + \left( \frac{1}{f_{3\text{dB}}^{\text{opt}}} \right)^2 \right]^{-\frac{1}{2}}, \quad (9)$$

where  $f_{3\text{dB}}^{\text{el}}$  is the 3 dB bandwidth of the modulator small-signal circuit model, illustrated in Fig. 13(d), and  $f_{3\text{dB}}^{\text{opt}}$  is determined by the modulator's photon lifetime limit and can be approximated as  $f_{3\text{dB}}^{\text{opt}} \approx \sqrt{\sqrt{2} - 1} \times \Delta\lambda_{\text{FWHM}}$ . While a higher modulator Q, and thus a lower  $\Delta\lambda_{\text{FWHM}}$ , would result in a more favorable modulation IL and ER trade-off in terms of detuning, there is a practical limit to the achievable modulation bandwidth, as shown in Fig. 16. The  $\Delta\lambda$  factor will also affect the cumulative electro-optic modulator bandwidth, as peaking occurs depending on the detuning frequency between the resonance and carrier which can extend the modulation bandwidth past the photon lifetime limit.<sup>25,26,67</sup>

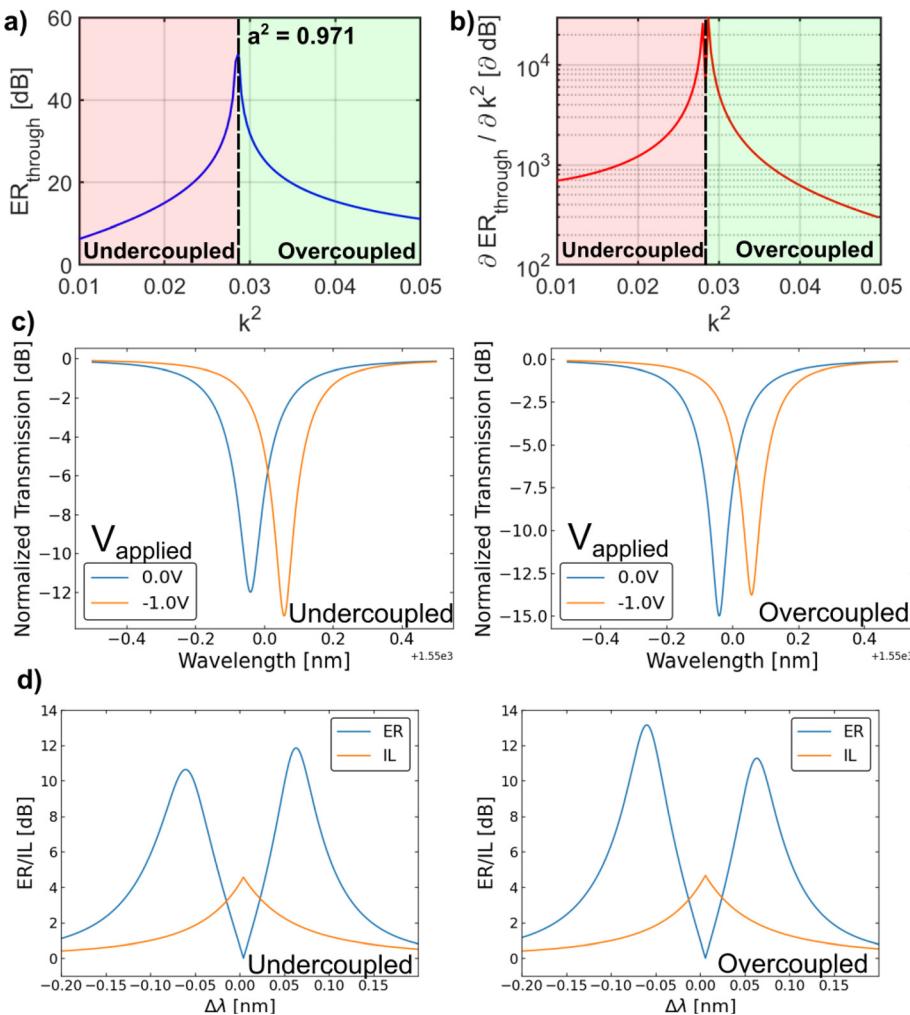
The resonator coupling regime also has a strong effect on modulation characteristics. While it is relatively straightforward to design symmetric coupling schemes to achieve broadband critical coupling in the AD filter configuration due to targeting  $a \approx 1$ , in the case of an AP modulator,  $a < 1$  is primarily due to the FCA loss from implanted dopants in the waveguide. As a result,  $a$  becomes relatively broadband, while  $r_1$  is strongly wavelength dependent, as shown in Fig. 10. The asymmetry between  $a$  and  $r_1$  with respect to optical wavelength means that an AP modulator designed for critical coupling at one wavelength will show substantial degradation of  $ER_{\text{through}}$  at adjacent FSRs, as well as when the resonance is thermally tuned. In the case of ultra-broadband DWDM links utilizing many individual wavelength channels, it is desirable to have relatively uniform spectral performance.<sup>21</sup> Figure 17(a) shows the resonance extinction ratio,  $ER_{\text{through}}$ , as a



**FIG. 16.** Electro-optic bandwidth based on a combination of the electrical bandwidth and optical photon lifetime limited bandwidth, which is determined by the resonator Q factor, as described in Eq. (9). The total modulation bandwidth for index modulation based resonant modulators is thus constrained by both the small-signal electrical circuit bandwidth and the modulator Q factor, setting a practical maximum Q for a resonant modulator designed to support a specific bandwidth.

function of coupling strength,  $k^2$ , into the resonator fundamental mode for a simulated microdisk modulator with 2.9% round trip loss. Critical coupling occurs at  $k^2 = 0.029$ . As shown by Fig. 17(b), the same  $ER_{\text{through}} = 15$  dB, the sensitivity to variations in  $k^2$  is half as much on the overcoupled side ( $k^2 = 0.04$ ) as on the undercoupled side ( $k^2 = 0.02$ ). At critical coupling the  $ER_{\text{through}}$  is orders of magnitude more sensitive to  $k^2$  and therefore also strongly varies with wavelength. In addition to reducing resonance wavelength dependence, it can also be seen in Figs. 17(c) and 17(d) that the characteristic trade-off between IL, ER, and  $\Delta\lambda$  is significantly more favorable for an overcoupled modulator. This is particularly the case for microdisk modulators, which are more susceptible to the self-heating effect due to their large intrinsic Q factor, forcing them to be required to modulate only in the  $\Delta\lambda < 0$  region at per line powers appropriate for telecommunication applications.<sup>81</sup>

An additional design consideration that is unique to the resonant modulator is requiring two sets of internal circuitry, one for the integrated thermal phase-shifter and the other to support high-speed RF modulation, both visible in Fig. 2(b).<sup>20</sup> Due to foundry constraints, such as minimum metal widths and gaps, this forces an asymmetry between the maximum FSR for thermally tunable AP modulators and AD filters, as the latter only requires the integrated heater. It is possible to place a doped silicon heater outside the resonator circumference instead, but a minimum distance from the resonator edge must be determined to avoid perturbing the optical mode. If not properly designed, this perturbation can degrade the resonator Q, change the coupling conditions, and create undesired back reflections in the cavity. Increasing this distance lowers the chance of detrimental optical coupling, but also effectively reduces the heater efficiency as more thermal power must be dissipated to result in an equivalent change in waveguide temperature. The larger the modulator FSR, the larger this distance must be, as it can be seen in Fig. 9(a) that the TE0 mode's evanescent tail extends further out of the resonator as radius decreases. The thermal conductivity of the material bridging the gap between the



**FIG. 17.** (a) Resonance extinction ratio,  $ER_{\text{through}}$ , as a function of coupling strength,  $k^2$ , into the resonator fundamental mode for a simulated microdisk modulator with dopant profile as shown in Fig. 13. (b) Derivative of  $ER_{\text{through}}$  with respect to  $k^2$ . (c) Characteristic depletion mode modulator resonance shifts for an undercoupled and an overcoupled resonant modulator, the undercoupled resonator gets closer to critical coupling with deeper reverse bias and vice versa. (d)  $ER$  and  $IL$  as a function of  $\Delta\lambda$  for fixed  $V_{pp} = 1.5$  V and an undercoupled and overcoupled resonant modulator.

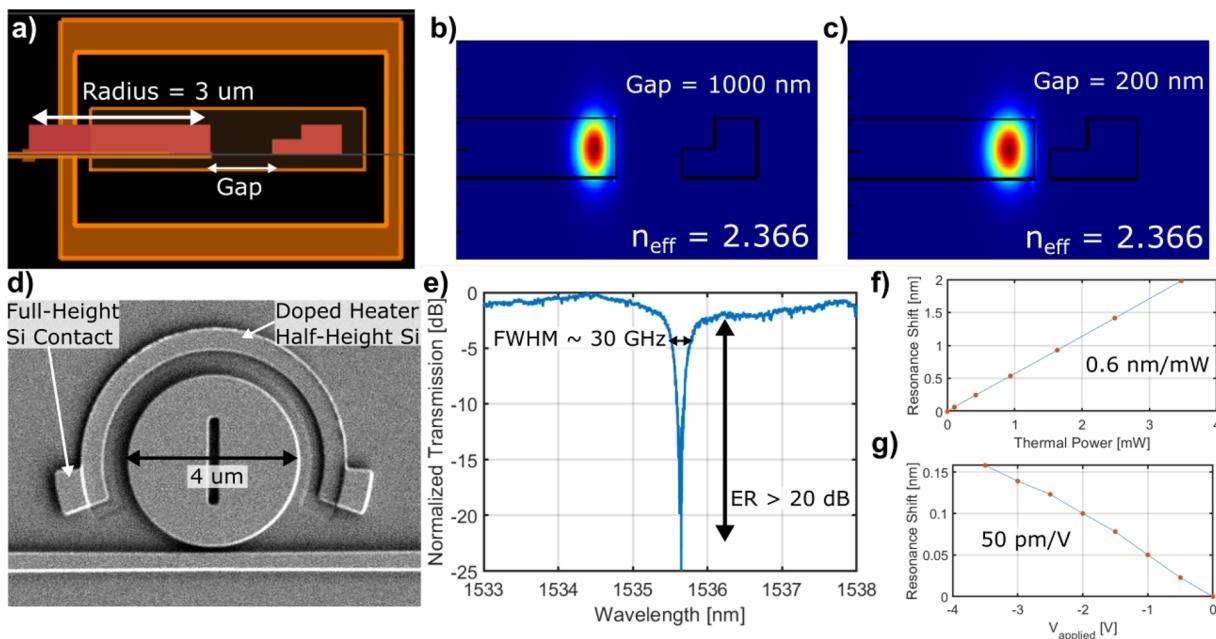
doped heater and resonator also plays a role in determining an external heater's efficiency, as silica is far less thermally conductive than silicon. Placing the integrated heater externally from the resonator, with the gap between heater and optical mode being entirely comprised of silica, further reduces thermal tuning efficiency and therefore link efficiency.<sup>82</sup> One possible solution to this asymmetry in modulator and filter FSR is to modify the transceiver architecture to accommodate the differential in transmitter and receiver resonator FSR, for example, halving the number of interleaver stages for the receiver relative to the transmitter in a multi-FSR implementation similar to the one shown in Fig. 2(c).<sup>21</sup>

Alternatively, a more aggressive external heater spacing can be pursued by targeting a heater silicon geometry that is phase mismatched as much as possible to the resonator's fundamental TE mode. Linear phase mismatched waveguides have been demonstrated in the context of tightly routed linear waveguides, allowing for drastically reduced crosstalk over hundreds of  $\mu\text{m}$  of near propagation.<sup>83,84</sup> Similarly, we can design the external doped silicon heater so that the angular  $n_{\text{eff}}$  is different enough as to be functionally transparent to the resonator TE0 mode, despite near physical proximity. An example of

one such device is shown in Fig. 18(a), where the microdisk modulator radius is  $2 \mu\text{m}$ , with an external heater that is only 500 nm away from the outer disk edge. By constructing the heater out of half-etched height silicon, a common foundry capability typically used for rib waveguides, an enormous angular  $n_{\text{eff}}$  mismatch is achieved, with the resulting modulator resonance showing near critical coupling and preserved Q in Fig. 18(b). The thermal tuning efficiency and depletion response are similar in magnitude to vertical junction micro-disk modulators manufactured in the same process with similar materials and dopant characteristics.<sup>7,80</sup> Numerical device simulations suggest even more aggressive external heater spacing may be possible, possibly as low 200 nm for microdisk modulators with radius  $\leq 2 \mu\text{m}$ , and FSR  $\geq 60$  nm. This and future demonstrations support the feasibility of fully symmetric resonator-based DWDM transceiver architectures, pushing the aggregate per-fiber-bandwidth limit to Tb/s and beyond.<sup>7,21</sup>

#### IV. CONCLUSION

In summary, we have explored the architecture and device-level design trade-offs for first-order resonator-driven integrated DWDM



**FIG. 18.** (a) Angular cross section of a disk resonator with radius = 3  $\mu\text{m}$  and an external doped half-etch silicon heater with an oxide gap separating the heater from the disk outer circumference. (b) Disk TE0 mode field profile solved for Gap = 1000 nm, finding an  $n_{\text{eff}} = 2.366$  with negligible coupling strength. (c) Disk TE0 mode field profile solved for Gap = 200 nm, with the same  $n_{\text{eff}} = 2.366$ , suggesting negligible optical coupling despite the close proximity. (d) SEM image of the fabricated microdisk modulator with a half-etched doped silicon external heater. (e) Measured resonance of the fabricated device, with  $\Delta\lambda_{\text{FWHM}} \approx 30$  GHz,  $ER_{\text{through}} > 20$  dB, and  $\text{FSR} \approx 58$  nm. (f) Thermal tuning efficiency of the half-etch heater disk, comparable to ring filter and disk modulator designs with an internal heater.<sup>36,80</sup> (g) Depletion response of the disk showing similar results to standard vertical junction microdisk modulators.<sup>7,80</sup>

29 April 2025 06:18:29

links that can support future peta-scale optical interconnects. A comprehensive analysis combining analytical models, numerical simulation, and experimental measurements is performed to demonstrate application-specific resonator design methodology for both AD filters and AP modulators. We show that, for a wavelength demux of cascaded first-order AD resonators, the minimum  $\Delta\lambda_{\text{ag}}$  should be at least 10 times the data rate to support OOK data communication. The case for designing resonant intensity modulators toward the overcoupled regime is also presented, particularly for microdisk modulators sensitive to the self-heating effect. Additionally, a novel microdisk modulator with a half-etched external doped silicon heater is demonstrated with larger than 20 dB resonance  $ER_{\text{through}}$ ,  $\Delta\lambda_{\text{FWHM}} \approx 30$  GHz,  $\text{FSR} \approx 58$  nm, with 0.6 nm/mW thermal tuning efficiency and 50 pm/V depletion response.

The combined results in this work provide a foundational framework for optimizing micro-resonator structures in DWDM systems. The wavelength-selective, energy-efficient, and footprint compact nature of these micro-resonator structures establish a basis for enabling flexible and energy-efficient scalability of communication systems in DC and HPC environments up to the peta-scale level. By leveraging these resonator structures, this massive bandwidth scaling can be directly integrated onto each compute node. This integration empowers the optimization of interconnect solutions to efficiently handle the substantial data movement in large-scale deep learning trainings, thereby mitigating energy consumption and alleviating the bandwidth discrepancy bottleneck, paving the way for the realization of increasingly advanced AI model capabilities.

## ACKNOWLEDGMENTS

This work was supported in part by the U.S. Advanced Research Projects Agency–Energy under ENLITENED Grant No. DE-AR000843 and in part by the U.S. Defense Advanced Research Projects Agency under PIPES Grant No. HR00111920014. We wish to acknowledge the support and manufacturing services of Lewis Carpenter, Chris Baiocco, Dan Coleman, and Jerry Leake at AIM Photonics in Albany, New York.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Asher Novick:** Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Software (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). **Kaylx Jang:** Conceptualization (supporting); Data curation (supporting). **Robert Parsons:** Data curation (supporting). **Keren Bergman:** Conceptualization (supporting); Funding acquisition (lead); Investigation (supporting); Project administration (lead); Supervision (lead); Writing – original draft (supporting); Writing – review & editing (supporting). **Aneek E. James:** Conceptualization (supporting); Data curation (supporting); Software (supporting); Writing – original draft (supporting). **Liang Yuan Dai:**

Conceptualization (supporting); Writing – original draft (supporting). **Zhenguo Wu:** Conceptualization (supporting); Writing – original draft (supporting). **Anthony J. Rizzo:** Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Writing – original draft (supporting). **Songli Wang:** Conceptualization (supporting); Data curation (supporting). **Yuyang Wang:** Conceptualization (supporting); Visualization (supporting). **Maarten Hattink:** Conceptualization (supporting); Writing – original draft (supporting). **Vignesh V. Gopal:** Conceptualization (supporting); Data curation (supporting).

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

- <sup>1</sup>M. Naumov, J. Kim, D. Mudigere, S. Sridharan, X. Wang, W. Zhao, S. Yilmaz, C. Kim, H. Yuen, M. Ozdal *et al.*, “Deep learning training in facebook data centers: Design of scale-up and scale-out systems,” [arXiv:2003.09518](https://arxiv.org/abs/2003.09518) (2020).
- <sup>2</sup>D. Narayanan, M. Shoeibi, J. Casper, P. LeGresley, M. Patwary, V. Korthikanti, D. Vainbrand, P. Kashinkunti, J. Bernauer, B. Catanzaro *et al.*, “Efficient large-scale language model training on GPU clusters using megatron-LM,” in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2021.
- <sup>3</sup>D. Mudigere, Y. Hao, J. Huang, Z. Jia, A. Tulloch, S. Sridharan, X. Liu, M. Ozdal, J. Nie, J. Park *et al.*, “Software-hardware co-design for fast and scalable training of deep learning recommendation models,” in *Proceedings of the 49th Annual International Symposium on Computer Architecture* (Association for Computing Machinery, 2022), pp. 993–1011.
- <sup>4</sup>See <https://resources.nvidia.com/en-us/tensor-core> for “NVIDIA h100 tensor core GPU architecture overview.”
- <sup>5</sup>D. Patterson, J. Gonzalez, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. So, M. Texier, and J. Dean, “Carbon emissions and large neural network training,” [arXiv:2104.10350](https://arxiv.org/abs/2104.10350) (2021).
- <sup>6</sup>K. Bergman, “Peta-scale embedded photonics for high performance computing,” in *European Conference on Optical Communication (ECOC)* (IEEE, 2022).
- <sup>7</sup>A. Rizzo, S. Daudlin, A. Novick, A. James, V. Gopal, V. Murthy, Q. Cheng, B. Y. Kim, X. Ji, Y. Okawachi, M. van Niekerk, V. Deenadayalan, G. Leake, M. Fanto, S. Preble, M. Lipson, A. Gaeta, and K. Bergman, “Petabit-scale silicon photonic interconnects with integrated Kerr frequency combs,” *IEEE J. Sel. Top. Quantum Electron.* **29**, 3700120 (2023).
- <sup>8</sup>Z. Wu, L. Y. Dai, A. Novick, M. Glick, Z. Zhu, S. Rumley, G. Michelogiannakis, J. Shalf, and K. Bergman, “Peta-scale embedded photonics architecture for distributed deep learning applications,” *J. Lightwave Technol.* **41**, 3737–3749 (2023).
- <sup>9</sup>M. Wade, E. Anderson, S. Ardalani, P. Bhargava, S. Buchbinder, M. L. Davenport, J. Fini, H. Lu, C. Li, R. Meade *et al.*, “TeraPHY: A chiplet technology for low-power, high-bandwidth in-package optical I/O,” *IEEE Micro* **40**, 63–71 (2020).
- <sup>10</sup>K. Hosseini, E. Kok, S. Y. Shumarayev, C.-P. Chiu, A. Sarkar, A. Toda, Y. Ke, A. Chan, D. Jeong, M. Zhang *et al.*, “8 Tb/s co-packed FPGA and silicon photonics optical IO,” in *Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, 2021).
- <sup>11</sup>M. Khani, M. Ghobadi, M. Alizadeh, Z. Zhu, M. Glick, K. Bergman, A. Vahdat, B. Klenk, and E. Ebrahimi, “SiP-ML: High-bandwidth optical network interconnects for machine learning training,” in *Proceedings of the ACM SIGCOMM 2021 Conference* (Association for Computing Machinery, 2021), pp. 657–675.
- <sup>12</sup>R. Mahajan, X. Li, J. Fryman, Z. Zhang, S. Nekkanty, P. Tadayon, J. Jaussi, S. Shumarayev, A. Agrawal, S. Jadhav *et al.*, “Co-packaged photonics for high performance computing: Status, challenges and opportunities,” *J. Lightwave Technol.* **40**, 379–392 (2021).
- <sup>13</sup>S. Daudlin, A. Rizzo, N. C. Abrams, S. Lee, D. Khilwani, V. Murthy, J. Robinson, T. Collier, A. Molnar, and K. Bergman, “3D-integrated multipchip module transceiver for terabit-scale DWDM interconnects,” in *Optical Fiber Communication Conference and Exhibition* (Optica Publishing Group, 2021).
- <sup>14</sup>S. Y. Siew, B. Li, F. Gao, H. Y. Zheng, W. Zhang, P. Guo, S. W. Xie, A. Song, B. Dong, L. W. Luo *et al.*, “Review of silicon photonics technology and platform development,” *J. Lightwave Technol.* **39**, 4374–4389 (2021).
- <sup>15</sup>A. Rizzo, A. Novick, V. Gopal, B. Y. Kim, X. Ji, S. Daudlin, Y. Okawachi, Q. Cheng, M. Lipson, A. L. Gaeta *et al.*, “Massively scalable Kerr comb-driven silicon photonic link,” *Nat. Photonics* **17**, 781–790 (2023).
- <sup>16</sup>B. Y. Kim, Y. Okawachi, J. K. Jang, M. Yu, X. Ji, Y. Zhao, C. Joshi, M. Lipson, and A. L. Gaeta, “Turn-key, high-efficiency Kerr comb source,” *Opt. Lett.* **44**, 4475–4478 (2019).
- <sup>17</sup>G. A. Keeler, “Optical microsystem technologies and applications,” in *Optical Fiber Communication Conference and Exhibition (OFC)* (Optica Publishing Group, 2021).
- <sup>18</sup>S. Daudlin, S. Lee, D. Kilwani, C. Ou, A. Rizzo, S. Wang, M. Cullen, A. Molnar, and K. Bergman, “Ultra-dense 3D integrated 5.3 Tb/s/mm<sup>2</sup> 80 micro-disk modulator transmitter,” in *Optical Fiber Communication Conference (OFC)* (Optica Publishing Group, 2023).
- <sup>19</sup>A. Samanta, P.-H. Chang, P. Yan, M. Fu, M. Berkay-On, A. Kumar, H. Kang, I.-M. Yi, D. Annabattuni, Y. Zhang, D. Scott, R. Patti, Y.-H. Fan, Y. Zhu, S. Palermo, and S. J. B. Yoo, “A direct bond interconnect 3D co-integrated silicon-photonics transceiver in 12 nm FinFET with -20.3dBm OMA sensitivity and 691fJ/bit,” in *Optical Fiber Communication Conference (OFC)* (Optica Publishing Group, 2023).
- <sup>20</sup>E. Timurdogan, C. M. Sorace-Agaskar, E. S. Hosseini, G. Leake, D. D. Coolbaugh, and M. R. Watts, “Vertical junction silicon microdisk modulator with integrated thermal tuner,” in *CLEO*, 2013.
- <sup>21</sup>Y. Wang, A. Novick, R. Parsons, S. Wang, K. Jang, A. James, M. Hattink, V. Gopal, A. Rizzo, C.-P. Chiu, K. Hosseini, T. T. Hoang, and K. Bergman, “Scalable architecture for sub-pJ/b multi-Tbps comb-driven DWDM silicon photonic transceiver,” *Proc. SPIE* **12429**, 124291F (2023).
- <sup>22</sup>A. James, A. Novick, A. Rizzo, R. Parsons, K. Jang, M. Hattink, and K. Bergman, “Scaling comb-driven resonator-based DWDM silicon photonic links to multi-Tb/s in the multi-FSR regime,” *Optica* **10**, 832–840 (2023).
- <sup>23</sup>E. Timurdogan, C. M. Sorace-Agaskar, J. Sun, E. Shah Hosseini, A. Biberman, and M. R. Watts, “An ultralow power athermal silicon modulator,” *Nat. Commun.* **5**, 4008 (2014).
- <sup>24</sup>A. Biberman, E. Timurdogan, W. A. Zortman, D. C. Trotter, and M. R. Watts, “Adiabatic microring modulators,” *Opt. Express* **20**, 29223–29236 (2012).
- <sup>25</sup>W. D. Sacher and J. K. S. Poon, “Dynamics of microring resonator modulators,” *Opt. Express* **16**, 15741–15753 (2008).
- <sup>26</sup>B. Pile and G. Taylor, “Small-signal analysis of microring resonator modulators,” *Opt. Express* **22**, 14913–14928 (2014).
- <sup>27</sup>R. Dubé-Demers, J. St-Yves, A. Bois, Q. Zhong, M. Caverley, Y. Wang, L. Chróstowski, S. LaRochelle, D. V. Plant, and W. Shi, “Analytical modeling of silicon microring and microdisk modulators with electrical and optical dynamics,” *J. Lightwave Technol.* **33**, 4240–4252 (2015).
- <sup>28</sup>K. Padmaraju, N. Ophir, A. Biberman, L. Chen, E. Swan, J. Chan, M. Lipson, and K. Bergman, “Intermodulation crosstalk from silicon microring modulators in wavelength-parallel photonic networks-on-chip,” in *23rd Annual Meeting of the IEEE Photonics Society* (IEEE, 2010), pp. 562–563.
- <sup>29</sup>D. Dai, L. Liu, S. Gao, D.-X. Xu, and S. He, “Polarization management for silicon photonic integrated circuits,” *Laser Photonics Rev.* **7**, 303–328 (2013).
- <sup>30</sup>D. Dai and J. E. Bowers, “Novel ultra-short and ultra-broadband polarization beam splitter based on a bent directional coupler,” *Opt. Express* **19**, 18614–18620 (2011).
- <sup>31</sup>D. Dai and H. Wu, “Realization of a compact polarization splitter-rotator on silicon,” *Opt. Lett.* **41**, 2346–2349 (2016).
- <sup>32</sup>W. D. Sacher, Y. Huang, L. Ding, T. Barwickz, J. C. Mikkelsen, B. J. F. Taylor, G.-Q. Lo, and J. K. S. Poon, “Polarization rotator-splitters and controllers in a Si<sub>3</sub>N<sub>4</sub>-on-SOI integrated photonics platform,” *Opt. Express* **22**, 11167–11174 (2014).
- <sup>33</sup>L.-W. Luo, N. Ophir, C. P. Chen, L. H. Gabrielli, C. B. Poitras, K. Bergman, and M. Lipson, “WDM-compatible mode-division multiplexing on a silicon chip,” *Nat. Commun.* **5**, 3069 (2014).
- <sup>34</sup>O. A. J. Gordillo, U. D. Dave, and M. Lipson, “Bridging between Si and few-mode fiber higher order modes,” in *Conference on Lasers and Electro-Optics* (Optica Publishing Group, 2020).

- <sup>35</sup>K. Y. Yang, C. Shirpurkar, A. D. White *et al.*, "Multi-dimensional data transmission using inverse-designed silicon photonics and microcombs," *Nat. Commun.* **13**, 7862 (2022).
- <sup>36</sup>O. A. J. Gordillo, A. Novick, O. L. Wang, A. J. Rizzo, U. D. Dave, K. Bergman, and M. Lipson, "Fiber-chip link via mode division multiplexing," *IEEE Photonics Technol. Lett.* **35**, 1071–1074 (2023).
- <sup>37</sup>C. Li, H. Zhang, G. Zhou, L. Lu, M. Jin, J. Dong, L. Zhou, and J. Chen, "Hybrid WDM-MDM transmitter with an integrated Si modulator array and a micro-resonator comb source," *Opt. Express* **29**, 39847–39858 (2021).
- <sup>38</sup>A. Novick, K. Jang, A. Rizzo, A. James, U. Dave, M. Lipson, and K. Bergman, "CMOS-foundry compatible, broadband, and compact routing of multimode SOI waveguides," in *Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, 2023).
- <sup>39</sup>L. H. Gabrielli, D. Liu, S. G. Johnson, and M. Lipson, "On-chip transformation optics for multimode waveguide bends," *Nat. Commun.* **3**, 1217 (2012).
- <sup>40</sup>C. Li, D. Liu, and D. Dai, "Multimode silicon photonics," *Nanophotonics* **8**, 227–247 (2018).
- <sup>41</sup>J. Wang, S. He, and D. Dai, "On-chip silicon 8-channel hybrid (de) multiplexer enabling simultaneous mode-and polarization-division-multiplexing," *Laser Photonics Rev.* **8**, L18–L22 (2014).
- <sup>42</sup>D. Dai and J. E. Bowers, "Novel concept for ultracompact polarization splitter-rotator based on silicon nanowires," *Opt. Express* **19**, 10940–10949 (2011).
- <sup>43</sup>T. Benson, A. Anand, A. Akella, and M. Zhang, "Understanding data center traffic characteristics," *ACM SIGCOMM Comput. Commun. Rev.* **40**, 92–99 (2010).
- <sup>44</sup>A. Roy, H. Zeng, J. Bagga, G. Porter, and A. C. Snoeren, "Inside the social network's (datacenter) network," in *Proceedings of the ACM Conference on Special Interest Group on Data Communication* (Association for Computing Machinery, 2015), pp. 123–137.
- <sup>45</sup>L. Poutievski, O. Mashayekhi, J. Ong, A. Singh, M. Tariq, R. Wang, J. Zhang, V. Beauregard, P. Conner, S. Gribble *et al.*, "Jupiter evolving: Transforming google's datacenter network via optical circuit switches and software-defined networking," in *Proceedings of the ACM SIGCOMM Conference* (Association for Computing Machinery, 2022), pp. 66–85.
- <sup>46</sup>S. Dhoore, G. Roelkens, and G. Morthier, "Fast wavelength-tunable lasers on silicon," *IEEE J. Sel. Top. Quantum Electron.* **25**, 1500908 (2019).
- <sup>47</sup>K.-I. Sato, "Realization and application of large-scale fast optical circuit switch for data center networking," *J. Lightwave Technol.* **36**, 1411–1419 (2018).
- <sup>48</sup>A. S. Khope, A. M. Netherton, T. Hirokawa, N. Volet, E. J. Stanton, C. Schow, R. Helkey, A. A. Saleh, J. E. Bowers, and R. C. Alferness, "Elastic WDM optoelectronic crossbar switch with on-chip wavelength control," in *Photonics in Switching* (Optica Publishing Group, 2017), pp. PTh1D-3.
- <sup>49</sup>Q. Cheng, M. Bahadori, Y.-H. Hung, Y. Huang, N. Abrams, and K. Bergman, "Scalable microring-based silicon clos switch fabric with switch-and-select stages," *IEEE J. Sel. Top. Quantum Electron.* **25**, 3600111 (2019).
- <sup>50</sup>L. Y. Dai, Y.-H. Hung, Q. Cheng, and K. Bergman, "Experimental demonstration of PAM-4 transmission through microring silicon photonic Clos switch fabric," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2020.
- <sup>51</sup>R. Proietti, G. Liu, X. Xiao, S. Werner, P. Fotouhi, and S. J. B. Yoo, "FlexLION: A reconfigurable all-to-all optical interconnect fabric with bandwidth steering," in *CLEO: Science and Innovations* (Optica Publishing Group, 2019).
- <sup>52</sup>Z. Wu, L. Y. Dai, Z. Zhu, A. Novick, M. Glick, and K. Bergman, "SiP architecture for accelerating collective communication in distributed deep learning," in *Optical Fiber Communication Conference (OFC)*, 2023.
- <sup>53</sup>T. J. Seok, J. Luo, Z. Huang, K. Kwon, J. Henriksson, J. Jacobs, L. Ochikubo, R. S. Muller, and M. C. Wu, "MEMS-actuated  $8 \times 8$  silicon photonic wavelength-selective switches with 8 wavelength channels," in *Conference on Lasers and Electro-Optics* (Optica Publishing Group, 2018).
- <sup>54</sup>Q. Cheng, M. Bahadori, M. Glick, and K. Bergman, "Scalable space-and-wavelength selective switch architecture using microring resonators," in *Conference on Lasers and Electro-Optics (CLEO)* (IEEE, 2019).
- <sup>55</sup>Y. Huang, Q. Cheng, A. Rizzo, and K. Bergman, "Push-pull microring-assisted space-and-wavelength selective switch," *Opt. Lett.* **45**, 2696–2699 (2020).
- <sup>56</sup>J. Xia, M. Chen, A. Rizzo, B. Sun, Z. Wang, M. Glick, Q. Cheng, K. Bergman, and R. Penty, "O-band microring resonator based switch-and-select silicon photonic switch fabric," in *Conference on Lasers and Electro-Optics* (Optica Publishing Group, 2022).
- <sup>57</sup>A. S. P. Khope, R. Zhang, R. Helkey, R. C. Alferness, A. A. M. Saleh, and J. E. Bowers, "Compact wavelength selective crossbar switch with cascaded first order micro-ring resonators," *Photonics* **9**, 60 (2022).
- <sup>58</sup>C. Manolatou, M. J. Khan, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, "Coupling of modes analysis of resonant channel add-drop filters," *IEEE J. Quantum Electron.* **35**, 1322–1331 (1999).
- <sup>59</sup>W. Bogaerts, P. De Heyn, T. Van Vaerenbergh, K. De Vos, S. Kumar Selvaraja, T. Claes, P. Dumon, P. Bienstman, D. Van Thourhout, and R. Baets, "Silicon microring resonators," *Laser Photonics Rev.* **6**, 47–73 (2012).
- <sup>60</sup>A. Biberman, M. J. Shaw, E. Timurdogan, J. B. Wright, and M. R. Watts, "Ultralow-loss silicon ring resonators," *Opt. Lett.* **37**, 4236–4238 (2012).
- <sup>61</sup>M. Bahadori, M. Nikdast, S. Rumley, L. Y. Dai, N. Janosik, T. V. Vaerenbergh, A. Gazman, Q. Cheng, R. Polster, and K. Bergman, "Design space exploration of microring resonators in silicon photonic interconnects: Impact of the ring curvature," *J. Lightwave Technol.* **36**, 2767–2782 (2018).
- <sup>62</sup>E. Timurdogan, C. M. Sorace-Agaskar, E. S. Hosseini, and M. R. Watts, "An interior-ridge silicon microring modulator," *J. Lightwave Technol.* **31**, 3907–3914 (2013).
- <sup>63</sup>L.-W. Luo, S. Ibrahim, A. Nitkowski, Z. Ding, C. B. Poitras, S. J. B. Yoo, and M. Lipson, "High bandwidth on-chip silicon photonic interleaver," *Opt. Express* **18**, 23079–23087 (2010).
- <sup>64</sup>A. Rizzo, Q. Cheng, S. Daudlin, and K. Bergman, "Ultra-broadband interleaver for extreme wavelength scaling in silicon photonic links," *IEEE Photonics Technol. Lett.* **33**, 55–58 (2021).
- <sup>65</sup>A. Novick, S. Wang, A. Rizzo, R. Parsons, K. Jang, and K. Bergman, "Tunable and compact SiP quasi-dichroic filter with  $\geq 10$  dB/nm roll-off across C- & L-bands," in *International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD)* (IEEE, 2022), pp. 171–172.
- <sup>66</sup>A. Novick, K. Jang, A. Rizzo, R. Parsons, and K. Bergman, "Low-loss wide-FSR miniaturized racetrack style microring filters for  $\geq 1$  Tbps DWDM," in *Optical Fiber Communication Conference and Exhibition (OFC)* (Optica Publishing Group, San Diego, California, 2023).
- <sup>67</sup>S. Karimelahi, W. Rahman, M. Parvizi, N. Ben-Hamida, and A. Sheikholeslami, "Optical and electrical trade-offs of rib-to-contact distance in depletion-type ring modulators," *Opt. Express* **25**, 20202–20215 (2017).
- <sup>68</sup>D.-X. Xu, A. Densmore, P. Waldron, J. Lapointe, E. Post, A. Delâge, S. Janz, P. Cheben, J. H. Schmid, and B. Lamontagne, "High bandwidth SOI photonic wire ring resonators using MMI couplers," *Opt. Express* **15**, 3149–3155 (2007).
- <sup>69</sup>Z. Lu, H. Yun, Y. Wang, Z. Chen, F. Zhang, N. A. F. Jaeger, and L. Chrostowski, "Broadband silicon photonic directional coupler using asymmetric-waveguide based phase control," *Opt. Express* **23**, 3795–3808 (2015).
- <sup>70</sup>S. Chen, Y. Shi, S. He, and D. Dai, "Low-loss and broadband  $2 \times 2$  silicon thermo-optic Mach-Zehnder switch with bent directional couplers," *Opt. Lett.* **41**, 836–839 (2016).
- <sup>71</sup>C. R. Doerr, M. Cappuzzo, E. Chen, A. Wong-Foy, L. Gomez, A. Griffin, and L. Buhl, "Bending of a planar lightwave circuit 2/spl times/2 coupler to desensitize it to wavelength, polarization, and fabrication changes," *IEEE Photonics Technol. Lett.* **17**, 1211–1213 (2005).
- <sup>72</sup>H. Morino, T. Maruyama, and K. Iiyama, "Reduction of wavelength dependence of coupling characteristics using Si optical waveguide curved directional coupler," *J. Lightwave Technol.* **32**, 2188–2192 (2014).
- <sup>73</sup>G. F. R. Chen, J. R. Ong, T. Y. L. Ang, S. T. Lim, C. E. Png, and D. T. H. Tan, "Broadband silicon-on-insulator directional couplers using a combination of straight and curved waveguide sections," *Sci. Rep.* **7**, 7246 (2017).
- <sup>74</sup>W. Freude, R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, and J. Leuthold, "Quality metrics for optical signals: Eye diagram, Q-factor, OSNR, EVM and BER," in *14th International Conference on Transparent Optical Networks (ICTON)*, 2012.
- <sup>75</sup>Y.-H. Hung, Q. Cheng, M. Glick, M. Bahadori, L. Y. Dai, and K. Bergman, "Silicon photonic switch-based optical equalization for mitigating pulsewidth distortion," *Opt. Express* **27**, 19426–19435 (2019).
- <sup>76</sup>P. Heyn, J. Coster, P. Verheyen, G. Lepage, M. Pantouvaki, P. Absil, W. Bogaerts, J. Van Campenhout, and D. Thourhout, "Fabrication-tolerant four-

- channel wavelength-division-multiplexing filter based on collectively tuned Si microrings,” *J. Lightwave Technol.* **31**, 2785–2792 (2013).
- <sup>77</sup>G. T. Reed and C. E. Jason Png, “Silicon optical modulators,” *Mater. Today* **8**, 40–50 (2005).
- <sup>78</sup>R. Wu, C.-H. Chen, T.-C. Huang, K.-T. Cheng, and R. Beausoleil, “20 Gb/s carrier-injection silicon microring modulator with SPICE-compatible dynamic model,” in International Conference on Photonics in Switching (PS), 2015.
- <sup>79</sup>V. Gopal, A. Rizzo, M. Hattink, A. Novick, J. Robinson, K. Hosseini, T. T. Hoang, and K. Bergman, “Simultaneous error-free data modulation with silicon microdisks in the multi-FSR regime for scalable DWDM links,” in *Optical Fiber Communication Conference (OFC)* (Optica Publishing Group, San Diego, California, 2023).
- <sup>80</sup>N. M. Fahrenkopf, C. McDonough, G. L. Leake, Z. Su, E. Timurdogan, and D. D. Coolbaugh, “The AIM photonics MPW: A highly accessible cutting edge technology for rapid prototyping of photonic integrated circuits,” *IEEE J. Sel. Top. Quantum Electron.* **25**, 8201406 (2019).
- <sup>81</sup>M. de Cea, A. H. Atabaki, and R. J. Ram, “Power handling of silicon microring modulators,” *Opt. Express* **27**, 24274–24285 (2019).
- <sup>82</sup>N. C. Harris, Y. Ma, J. Mower, T. Baehr-Jones, D. Englund, M. Hochberg, and C. Galland, “Efficient, compact and low loss thermo-optic phase shifter in silicon,” *Opt. Express* **22**, 10487–10493 (2014).
- <sup>83</sup>Z. Lu, K. Murray, H. Jayatilleka, and L. Chrostowski, “Michelson interferometer thermo-optic switch on SOI with a 50- $\mu$ w power consumption,” *IEEE Photonics Technol. Lett.* **27**, 2319–2322 (2015b).
- <sup>84</sup>S. Chung, M. Nakai, and H. Hashemi, “Low-power thermo-optic silicon modulator for large-scale photonic integrated systems,” *Opt. Express* **27**, 13430–13459 (2019).