Petabit-Scale Silicon Photonic Interconnects with Integrated Kerr Frequency Combs

Anthony Rizzo, Stuart Daudlin, Asher Novick, Aneek James, Vignesh Gopal, Vaishnavi Murthy, Qixiang Cheng, Bok Young Kim, Xingchen Ji, Yoshitomo Okawachi, Matthew van Niekerk, Venkatesh Deenadayalan, Gerald Leake, Michael Fanto, Stefan Preble, Michal Lipson, Alexander Gaeta, and Keren Bergman

(Invited Paper)

Abstract—Silicon photonics holds significant promise in revolutionizing optical interconnects in data centers and high performance computers to enable scaling into the Pb/s package escape bandwidth regime while consuming orders of magnitude less energy per bit than current solutions. In this work, we review recent progress in silicon photonic interconnects leveraging chip-scale Kerr frequency comb sources and provide a comprehensive overview of massively scalable silicon photonic systems capable of capitalizing on the large number of wavelengths provided by such combs. We first consider the high-level architectural constraints and then proceed to detail the corresponding fundamental device designs supported by both simulated and experimental results. Furthermore, the majority of experimentally measured devices were fabricated in a commercial 300 mm foundry, showing a clear path to volume manufacturing. Finally, we present various system-level experiments which illustrate successful proof-of-principle operation, including flip-chip integration with a co-designed CMOS application-specific integrated circuit (ASIC) to realize a complete Kerr comb-driven electronic-photonic engine. These results provide a viable and appealing path towards future co-packaged silicon photonic interconnects with aggregate per-fiber bandwidth above 1 Tb/s, energy consumption below 1 pJ/bit, and areal bandwidth density greater than 5 Tb/s/mm².

Index Terms—Silicon photonics, optical frequency combs, wavelength division multiplexing.

I. INTRODUCTION

THE past decade has seen unprecedented growth of the digital world, spearheaded by the rapid rise to ubiquity of data-intensive workloads such as artificial intelligence and deep learning. Simultaneously, as Moore’s Law draws to a close and computing systems will no longer be able to rely on predictable advances in transistor density, the interconnect bandwidth in these systems will still continue to lag behind compute performance. Without significant intervention, the concert of these roadblocks will severely constrain future computing systems and lead to stagnation in performance which will be insufficient to keep pace with rapidly accelerating workload demands [1]. Silicon photonics is poised to alleviate this bandwidth bottleneck through providing highly scalable optical interconnects directly co-packaged with compute electronics [2], [3], effectively removing current constraints on the highly distance-dependent bandwidth density-energy efficiency product (Fig. 1). However, commercial silicon photonics solutions in the current co-packaged optics roadmap fundamentally lack avenues for extreme scaling to realize petabit-scale package escape bandwidths, which will be necessary in the coming decades to keep pace with exponentially growing workloads.

Optical frequency combs, consisting of equidistant tones in the frequency domain with intrinsically precise spacing, have been widely studied in a variety of systems beginning with mode-locked lasers [4], [5] and were the subject of the 2005 Nobel Prize in Physics [6]. The miniaturization of these frequency combs to the chip scale has sparked a second revolution in the field, promising to provide unprecedented performance in a mass-producible form factor which can be realized using the same complementary metal-oxide-semiconductor (CMOS) infrastructure currently used to fabricate commercial microelectronics chips [7]–[10]. Over the span of the past decade, these integrated Kerr frequency combs have found widespread use across a broad application space including optical metrology [11], optical frequency synthesis [12], quantum optics [13], spectroscopy [14], [15], long-haul telecommunications [16]–[21], and most recently short-reach data communications [22]–[26]. In the context of both long- and short-reach data communications, integrated optical frequency combs have drawn significant interest as wavelength-division multiplexing (WDM) sources, providing the alluring prospect of a single multi-wavelength source with tens to hundreds of equally-spaced carrier frequencies at standard ITU channel spacing. Furthermore, these sources have the potential to reside on a single chip [27], [28] which can either be directly integrated with the transmitter chip [29] or reside outside of the transceiver package as a remote light source [30].
Traditional optical interconnect approaches using pluggable transceivers require high-speed signals to propagate through centimeters of lossy metallic waveguides before conversion to the optical domain and typically provide a maximum of only four channels at high baud rates. These high-speed signals require complex serialization/deserialization (SerDes), digital signal processing (DSP), and clocking circuits to interface with the compute ASIC (CPU, GPU, etc.) which consumes large amounts of energy and adds significant latency. The combination of these requirements and the long electrical signal paths between pluggable transceivers and compute ASICs presents a severe bottleneck to further scaling the link bandwidth and energy efficiency. However, through leveraging the tens to hundreds of wavelengths provided by Kerr frequency combs, each WDM channel can be run at modest data rates to drastically reduce the gearboxing requirements at the compute ASIC interface while maintaining ultra-high aggregate bandwidth due to the massive wavelength parallelism. Thus, through realizing Kerr comb-driven co-packaged silicon photonic interconnects with millimeter-scale electrical signal paths between the electronics and photonics, fundamentally new computing architectures can be explored in data centers and high performance computing systems.

By removing the strong constraint of data locality imposed by lossy, distance-dependent electrical interconnects through providing ultra-high bandwidth, distance-agnostic WDM links with ultra-low energy consumption and latency, disaggregated architectures become feasible in which resources can be pooled and optically connected over long distances without relying on each server to be configured with a standard von Neumann architecture [31], [32]. Furthermore, with all data emerging from the compute package already in the optical domain, novel optical circuit switching topologies can be employed to partition and route the multi-wavelength data streams such that bandwidth is allocated to best match the system-level traffic patterns without costly O-E-O conversions [33], [34]. Therefore, the proposed Kerr comb-driven silicon photonic interconnects have the potential not only to alleviate the current data center bandwidth bottleneck, but to enable a fundamentally new class of distributed computing systems.

Here, we outline a comprehensive co-designed electronic-photonic interconnect leveraging the massive parallelism of integrated Kerr frequency combs capable of realizing 1 Tbit/s/fiber data transmission while consuming as low as 100 fJ/b with an areal bandwidth density greater than 5 Tbit/s/mm². While the strategies employed to reach these metrics represent a substantial departure from current state-of-the-art silicon photonic solutions, all of the described avenues are fundamentally compatible with high-volume manufacturing and thus represent a realistic path for continued scaling of optical interconnects into the Pb/s regime.

II. LINK ARCHITECTURE

The burden of (de-)multiplexing hundreds of comb lines into independent buses, where each bus supports only a single channel with a single modulator/receiver, is prohibitive in terms of footprint and energy consumption requirements. By taking advantage of the wavelength selectivity of resonant modulators and filters, the number of buses can be substantially reduced and thus relax these requirements. However, restrictions stemming from the resonator free spectral range (FSR) and off-resonance insertion loss typically limit the use of traditional single bus architectures in cases where the comb bandwidth exceeds the FSR. In these cases, we propose two distinct regimes of comb pre-division into sub-groups prior to traversing cascaded resonators: band interleaving and even-odd interleaving. This section details the benefits and drawbacks of all three permissible link architecture regimes (single bus, band interleaving, and even-odd interleaving), taking a holistic view including footprint, energy consumption, loss, and ease of implementation. The link-level trade-offs and considerations discussed in this section motivate the application-specific device designs and innovations in the subsequent sections.

A. Single Bus

Through cascading microresonators of varying radii on a single bus waveguide, individual resonators can address independent target wavelength channels while appearing transparent to other channels on the bus (Fig. 2a) [35]–[38]. While non-target channels will suffer from the off-resonance insertion loss associated with other resonators, this loss can be made exceptionally low with careful device design (∼ 0.1 dB). For this architecture, the predominant roadblock to scaling in channel count is not insertion loss but rather the FSR restriction associated with the spectral periodicity of resonant cavities. Since the spectrum of each resonator exhibits dips at each integer multiple of the FSR, these unwanted resonances prevent packing of more channels since they would result in high crosstalk due to overlap with non-target channels. While the FSR restricts the number of channels for a particular channel spacing, more channels can be added for a fixed FSR if the channel spacing is reduced. However, it has been shown that spacings below ∼ 100 GHz result in severe intermodulation crosstalk [39] and thus places a lower bound on how much the channel spacing can be reduced. Given these restrictions, for a state-of-the-art resonant modulator array at 100 GHz channel spacing with 25.6 nm FSR [40], this architecture can support a maximum of 32 WDM channels. While smaller resonators with a larger FSR are being explored...
as a means of further increasing the channel capacity of the single bus architecture [41]–[43], such devices have drawbacks including increased design/implant complexity and lack of space for integrated heaters due to their inherently small footprint. Despite high efficiency carrier depletion resonant modulators being capable of compensating ~ 10 K temperature swings without thermal tuning [43], [44], it is unlikely that this mechanism will be sufficient in real co-packaged thermal environments (\( \Delta T > 80 \) K) and thus necessitates the use of integrated heaters. Furthermore, despite the low off-resonance insertion loss, for high device counts cascaded on a single bus this can accumulate to an appreciable power penalty (i.e., \( \sim 0.1 \) dB = 6.3 dB). These restrictions on the scalability of the single bus architecture motivate the use of multi-bus architectures to support ultra-high channel counts.

B. Band Interleaving

Band interleaving is the natural extension of the single bus architecture in which the comb is first split into ‘bands’ of adjacent lines before traversing separate buses of cascaded resonators (Fig. 2b). As described for the single bus architecture, if the comb bandwidth is greater than the resonant modulator/filter FSR, pre-filtering is necessary to accommodate channels beyond the capacity of a single bus. Through sub-dividing the comb in this manner, the original comb channel spacing is preserved and the total bandwidth on each bus (and number of channels per bus) is reduced by a factor of \( 2^{S-1} \), where \( S \) is the number of interleaver stages. This architecture is particularly appealing since the number of interleaver stages can be chosen such that the comb bandwidth on each bus is smaller than the resonator FSR, thus mitigating any undesirable interactions between target and alias resonances. However, the main drawback lies at the device level rather than the system level; that is, the toolbox of compact, tunable silicon photonic devices which can perform band interleaving in this manner is scarce and demonstrated devices lack the necessary performance in terms of pass-band roll-off, crosstalk suppression, and tunability. Mach Zehnder-based wavelength filters have a pass-band roll-off proportional to their FSR, and thus the large FSR required for band interleaving (at least half of the total comb bandwidth) leads to unfeasible roll-off. Even for modified MZI designs targeted at flattening the pass-bands and sharpening the roll-off (such as ring-assisted MZIs [45], [46] and cascaded MZI lattices [47]), the demonstrated roll-off for large FSRs is still insufficient and would lead to severe crosstalk near the band transition. On-chip dichroic filters have been demonstrated as a candidate for similar filtering applications [48], but similarly have poor roll-off (2.82 dB/nm), insufficient crosstalk suppression (< 10 dB), and lack tunability to correct for fabrication imperfections and temperature swings.

However, contra-directional couplers have shown significant promise as band interleaving devices and typically occupy much smaller footprints than MZI-based filters. Demonstrations have shown 4 channel flat-top CWDM demultiplexers with ~ 1 mm² footprint [49], [50], device tunability through carrier injection in p-i-n junctions [51], and large stopband filters with bandwidths > 33 nm [52]. While such wavelength filters based on contra-directional couplers hold potential for band interleaving solutions, their implementation in such systems is currently restricted by their gradual passband roll-off which would result in lost comb lines near the band edges in demonstrated devices. Furthermore, fabricating sharp sub-wavelength structures in foundry processes using DUV lithography necessitates sophisticated optical proximity correction (OPC) to achieve the desired device geometry and requires the use of computational lithography to accurately simulate the device (accounting for lithography effects such as corner rounding) [53]. If these lithography effects are accurately calibrated for a given foundry process and device advancements lead to sharper passband roll-off, contra-directional coupler-based wavelength filters will provide a viable and appealing path to enabling massively wavelength-parallel silicon photonic links using band interleaving architectures.

C. Even-Odd Interleaving

Even-odd interleaving is an alternate path to sub-dividing the comb into separate buses in which the comb is split into ‘even’ and ‘odd’ groups, with each stage doubling the channel spacing of the output groups and preserving the total comb bandwidth on each bus (Fig. 2c). In this architecture, the pass-band roll-off is a more nuanced issue than in the band interleaving case. For band interleaving, it was shown that poor roll-off directly results in unusable comb lines and/or reduced performance of comb lines that reside near the pass-band edge.

![Fig. 2.](Image)
bandwidth, allowing all channels maximal protection from both crosstalk and dispersion walk-off. The multi-FSR regime, in contrast, uses microresonators with FSRs that are smaller than the total optical bandwidth on the bus. Alias resonances will fall in between each channel at regular intervals, effectively reducing the spacing between each channel and its nearest aggressor. The suitability for each regime for a given link depends on the optical bandwidth of the comb source and the physical constraints on the resonator FSR. While a single-FSR design is a simple solution to the alias resonance problem for small optical bandwidths, designing resonators for larger optical bandwidths runs the risk of requiring physical dimensions that may be infeasible due to both fabrication constraints and excess bend radiation losses [44]. Although the relaxed requirement on the resonator FSR for the multi-FSR regime is provided at the cost of additional crosstalk, multi-FSR designs are an attractive alternative that allows for aggressive scalability in total link bandwidth without requiring potentially infeasible resonator designs.

While any FSR larger than the link bandwidth will work in the single-FSR regime, only particular combinations of FSR and channel aggressor values can result in a valid multi-FSR design [24]. To properly analyze this trade-off between increased channel crosstalk and resonator FSR, we must define the requisite conditions which must be satisfied for a valid arrangement. We first define two auxiliary variables for designing the FSR:

\[ S = \frac{\lambda_{ch}}{\lambda_{ag}} \]  
\[ F = \frac{FSR}{\lambda_{ag}} \]  

For even-odd interleaving, this requirement is relaxed since the periodicity of the interleaver spectrum results in maximum transmission/rejection of desired comb lines regardless of rolloff, assuming perfect alignment and that the interleaver free spectral range (FSR) is both constant and exactly twice the channel spacing of the comb. For this reason, the library of suitable devices is much more expansive than that for band interleaving, making this architecture much more favorable for system implementation. Furthermore, each successive stage doubles the channel spacing and halves the number of channels per bus, relaxing constraints on crosstalk and insertion loss. However, since this architecture preserves the total comb bandwidth on each bus, if the comb bandwidth is greater than the resonator FSR then great care must be taken in constructing the channel arrangement such that device ‘alias’ resonances do not overlap with other ‘target’ resonances [22], [24].

1) Valid Multi-FSR Channel Arrangement: Since the channel bandwidth on each bus is approximately the same as the total link bandwidth for even-odd interleavers, consideration must be given to preventing ‘alias’ resonances from inducing severe crosstalk penalties for the other ‘target’ channels on a given bus. When choosing the design of an even-odd interleaver link, there exist two distinct regimes of operation: (i) single-FSR regime and (ii) multi-FSR regime. In the single-FSR regime, the microresonator FSR is designed to be larger than the total optical bandwidth on the bus. This design guarantees that any alias resonances fall outside of the optical bandwidth, allowing all channels maximal protection from both crosstalk and dispersion walk-off. The multi-FSR regime, in contrast, uses microresonators with FSRs that are smaller than the total optical bandwidth on the bus. Alias resonances will fall in between each channel at regular intervals, effectively reducing the spacing between each channel and its nearest aggressor. The suitability for each regime for a given link depends on the optical bandwidth of the comb source and the physical constraints on the resonator FSR. While a single-FSR design is a simple solution to the alias resonance problem for small optical bandwidths, designing resonators for larger optical bandwidths runs the risk of requiring physical dimensions that may be infeasible due to both fabrication constraints and excess bend radiation losses [44]. Although the relaxed requirement on the resonator FSR for the multi-FSR regime is provided at the cost of additional crosstalk, multi-FSR designs are an attractive alternative that allows for aggressive scalability in total link bandwidth without requiring potentially infeasible resonator designs.

While any FSR larger than the link bandwidth will work in the single-FSR regime, only particular combinations of FSR and channel aggressor values can result in a valid multi-FSR design [24]. To properly analyze this trade-off between increased channel crosstalk and resonator FSR, we must define the requisite conditions which must be satisfied for a valid arrangement. We first define two auxiliary variables for designing the FSR:

\[ S = \frac{\lambda_{ch}}{\lambda_{ag}} \]  
\[ F = \frac{FSR}{\lambda_{ag}} \]  

where \( \lambda_{ch} \) is the channel spacing on each bus, \( \lambda_{ag} \) is the reduced spacing between each channel and its nearest aggressor alias, \( N_{ch} \) is the total number of channels on each bus, and \( FSR \) is the resonator free spectral range.

The quantities \( F \) and \( S \) represent the resonator FSR and channel spacing normalized to the aggressor spacing. Together, they describe all possible multi-FSR channel arrangements. The normalized nature of both variables allows us to explicitly define the two conditions that \( F \) and \( S \) must follow such that they yield a valid channel arrangement: (i) \( F \) and \( S \) must be co-prime integers (i.e. they have no common factors other than 1), and (ii) \( F \) must be greater than or equal to \( N_{ch} \). By ensuring that \( F \) and \( S \) follow these conditions, it is guaranteed that no resonance aliases will land on any target wavelength within the bandwidth of the link (Fig. 3). Choosing \( F = N_{ch} \) corresponds to the fairest channel reduction case in which each channel experiences the same increase in crosstalk. Choosing \( F \) to be greater than \( N_{ch} \) means that one or more aggressor aliases won’t appear on the bus. The special case where \( S = 1 \) corresponds to the minimum FSR required for the single-FSR regime.

2) Multi-FSR Design Space Exploration: Once all constraints are defined, a design space boundary can be drawn such that solutions inside the boundary meet the design constraints. From this, a final optimization that maximizes the key priorities for the link (e.g. aggressor spacing) can be
performed. An example of this is provided in Fig 4a. Each co-prime $S\cdot F$ pair that falls within or on the boundary are valid link configurations that can be considered as a design candidate. Additionally, by using equations (1a) and (1b), each $S\cdot F$ pair can be converted into a corresponding FSR-$\lambda_{ag}$ pair, as shown in Fig. 4(b).

A key goal at this stage is to choose an optimal FSR and aggressor spacing which both meet the design requirements for the link and maximize the minimum aggressor spacing. For the same design constraints, additional solutions can be added by increasing the value for the channel bus spacing through either additional interleaver stages or increasing the original comb line spacing (Fig. 4(c-d)). By the definition of (1a), a lower value for $S$ corresponds to a larger aggressor spacing and therefore lower level of crosstalk. The quantity $S$ is normalized to the channel spacing, which is determined by $\lambda_{ch} = 2L \cdot \lambda_{comb}$, where $L$ is the number of interleaver stages and $\lambda_{comb}$ is the spacing of the comb source. In common integrated comb sources (e.g. Kerr combs and quantum dot mode locked lasers, discussed further in Section III), the channel spacing is directly related to the cavity path length which can’t be tuned post-fabrication to any appreciable fraction of the FSR. Therefore, choosing a link design which accounts for worst-case fabrication-induced walk-off in the source channel spacing is critical to ensuring yield.

Although the link design formalism outlined in this section omits dispersion for simplicity, typical silicon photonic resonators have non-zero dispersion which must be considered. This dispersion results in a wavelength-dependent FSR and will thus disrupt the placement of alias resonances over broad bandwidths. If the resonator FSR is designed with respect to a ‘nominal’ wavelength, the severity of this dispersion will increase the farther away an alias is from the nominal wavelength. The impact of this alias placement disruption is consequently a function of the total optical bandwidth on a given bus, which cannot be reduced when using even-odd interleavers as noted earlier. Though there is some variation in the dispersion between microdisk designs, the difference is too small for any choice in $S\cdot F$ pair to meaningfully compensate for using a large optical bandwidth. Considering both sources of non-idealities (imperfect comb spacing and dispersion), each $S\cdot F$ pair must be explicitly simulated to ensure that the resulting $\lambda_{ag}$ provides enough of a margin.

III. INTEGRATED KERR FREQUENCY COMB SOURCE

At the heart of such massively wavelength-multiplexed interconnects is the optical frequency comb source. In this section, we focus our analysis on Kerr frequency combs in the Si$_3$N$_4$ platform due to their highly appealing properties, but additionally provide comparisons to other platforms for comb generation. We then present the design and fabrication of dispersion-engineered coupled-resonator Si$_3$N$_4$ Kerr combs in the normal group velocity dispersion (GVD) regime, and conclude with experimental characterizations of a representative fabricated device.

A. Comparison of Frequency Comb Sources

Integrated optical frequency comb sources can be categorized under two broad classifications: semiconductor mode-locked lasers (SMLLs) and nonlinearly generated combs [56]. As this work is focused on short-reach data communications, we restrict our discussion to comb sources which can be integrated on-chip and fabricated in standard high-volume processes. Electrically-pumped quantum dot SMLLs have been widely demonstrated both as external comb sources [57] and integrated directly on-chip [58], [59] with sufficient performance for communications applications. While such sources can provide tens of high-quality wavelength channels, their total bandwidth is fundamentally restricted by the gain bandwidth of the active region material and thus scaling to the regime of hundreds of channels is infeasible. Additionally, due to restrictions on the cavity optical path length in such devices, typical channel spacings are much less than typical dense wavelength-division multiplexing (DWDM) standards.
(\sim 20 \text{ GHz versus } 100 \text{ GHz}) which can lead to significantly increased crosstalk. Furthermore, the native output power-per-line is typically much lower than the link budget and thus necessitates power-hungry pre-amplification prior to the transmitter.

Integrated comb sources based on parametric nonlinear generation have been demonstrated in various material platforms including thin film lithium niobate (using both $\chi^2$ [60] and $\chi^3$ [29] processes), AlGaAs-on-insulator [61], and silicon nitride ($\text{Si}_3\text{N}_4$) [8], [27], [28], [55]. Such sources are highly appealing since only a single continuous-wave (CW) pump laser is required to generate hundreds of new frequencies with inherently precise spacing. Electro-optic combs in the thin film lithium niobate platform have been demonstrated with excellent performance in terms of conversion efficiency and number of lines [60], [62], but with the drawbacks of low power-per-line, inherent spectral roll-off, and small channel spacing which depends on the speed of the microwave drive signal. However, combs generated using the $\chi^3$ nonlinearity (‘Kerr combs’) in $\text{Si}_3\text{N}_4$ microresonators represent an ideal platform due to their CMOS compatibility and since they do not require a high-speed microwave drive signal for comb generation. Additionally, since this physical process for comb generation is parametric, such sources are not restricted by an active material gain bandwidth and thus have been shown to be capable of spanning over an octave in the near IR, encompassing all telecommunication bands [63].

However, integrated Kerr frequency combs in $\text{Si}_3\text{N}_4$ have traditionally suffered from three major inherent drawbacks which were thought to preclude their use as short-reach DWDM sources: (i) low optical power per line, (ii) low pump-to-comb conversion efficiency, and (iii) non-uniformity in line power. Many of these properties are particular to soliton Kerr combs in the anomalous GVD regime, which have fundamental physical restrictions on conversion efficiency (< 5% [64]) and flatness (due to their inherent sech$^2$ envelope). However, nonsolitonic Kerr combs in the normal GVD regime have recently emerged as a promising alternative, as they have been experimentally shown to have pump-to-comb conversion efficiencies as high as 41%, high power-per-line ($> 0.5 \text{ mW}$), and tailorable spectral envelopes. Furthermore, we have recently shown that such nonsolitonic Kerr combs can be synchronized and coherently combined [65], [66] to achieve otherwise unattainable power-per-line and equalize the spectrum for improved flatness. Due to the combination of these highly favorable properties, such nonsolitonic Kerr combs provide a unique path to realizing chip-scale multi-wavelength sources capable of efficiently producing hundreds of evenly-spaced high power lines.

B. Coupled-Microresonator Design and Fabrication

Comb generation using localized anomalous GVD provided by an avoided mode crossing in coupled microresonator systems has been shown to be a reliable method for creating highly efficient multi-wavelength sources with high coherence [54], [55]. Furthermore, such devices can leverage independent thermal tuning of both resonators to alter the bus-ring and ring-ring coupling conditions post-fabrication to control the spectral position and splitting strength of the avoided crossing, allowing for a fixed-wavelength pump laser. Furthermore, the low-noise comb state can persist over a wide heater power range (in contrast to highly sensitive soliton states), providing stable operation in the presence of environmental temperature perturbations. For our representative device, we design the waveguide cross-section to be $780 \times 1,000 \text{ nm}$ such that the fundamental TE$_0$ mode experiences normal GVD near the pump wavelength. Furthermore, by designing the main ring FSR to be slightly offset from the auxiliary ring FSR (200 GHz and 206 GHz, respectively), we leverage the Vernier effect to control the spectral periodicity of the mode interactions to limit low-power comb lines in the bandwidth of interest.

To fabricate the $\text{Si}_3\text{N}_4$ microresonators, we first start from a 4-inch silicon wafer and thermally grow a 4-$\mu$m thick oxide layer to form the bottom cladding. Next, we deposit $\text{Si}_3\text{N}_4$ using low-pressure chemical vapor deposition (LPCVD) in two steps and then anneal at 1,200 $\text{°C}$ in an argon atmosphere for 3 hours. After the $\text{Si}_3\text{N}_4$ deposition, we then deposit a SiO$_2$ hard mask using plasma-enhanced chemical vapor deposition (PECVD). Electron beam lithography is then used to pattern the $\text{Si}_3\text{N}_4$ devices. We write the pattern using Ma-N 2403 resist and etch the nitride film in an inductively coupled plasma reactive ion etcher (ICP RIE) using a combination of CHF$_3$, N$_2$, and O$_2$ gases. After stripping the oxide mask, the devices are again annealed to remove residual N-H bonds in the $\text{Si}_3\text{N}_4$ layer. The devices are then clad with 500 nm of high-temperature silicon dioxide (HTO) deposited at 800 $\text{°C}$ followed by 2 $\mu$m of SiO$_2$ using PECVD. A deep etched facet and an inverse taper are designed and used to maximize the edge coupling loss. Further information and extended discussion regarding the device fabrication can be found in [67], [68].

C. Experimental Characterization

Due to the thermally tunable avoided mode crossing, we demonstrate deterministic computer-controlled mode-locked comb generation using a fixed frequency pump laser source ($\lambda = 1559.79$) [55]. Furthermore, to demonstrate fully general turn-key operation across different operating conditions, we show deterministic low-noise comb generation for multiple combinations of ring-ring coupling gap and pump wavelength. We experimentally characterized the pump-to-comb conversion efficiency of the fabricated normal GVD coupled-resonator Kerr combs to be as high as 41% when the comb is in the low noise mode-locked state and observed high coherence (> 94%) for all comb lines across the full spectrum, suitable for both intensity-modulated and coherent communications. Further details regarding the device characterization can be found in [55].

IV. Even-Odd (De-)Interleavers

In the proposed link architecture, the (de-)interleaver is the first device traversed by the comb for multi-bus configurations. Due to the difficulties associated with on-chip implementation of band interleaving, we focus our discussion on
even-odd interleavers. In particular, we examine ring-assisted MZI (RAMZI) interleavers due to their compact footprint, tunability, and flat-top pass-bands compared to the sinusoidal response of standard asymmetric MZIs. Despite the maturity of such devices, there exist a number of hurdles associated with standard designs which must be overcome for their consideration in ultra-low energy consumption silicon photonic links. In this section, we outline the exact challenges associated with their adoption and demonstrate solutions which reduce their energy consumption by orders-of-magnitude compared to standard designs.

A. Ultra-Broadband Performance

The broad bandwidth spanned by the comb requires all devices in the circuit to have predictable and uniform behavior over the full comb bandwidth. Often, this bandwidth can be greater than 100 nm and thus places considerable strain on the devices, which are often highly wavelength-dependent for standard designs. To accommodate combs with > 100 channels at 100 GHz spacing, we demonstrate an ultra-broadband silicon photonic interleaver capable of interleaving and de-interleaving frequency comb lines over a 125 nm bandwidth from 1525–1650 nm, spanning the full C- and L-bands and representing the broadest operational bandwidth to date for an on-chip interleaver [45]. The device consists of a ring-assisted Mach Zehnder interferometer (RMZI) with broadband coupling to the ring and broadband multi-mode interference (MMI) splitter/combiners, enabling consistent performance over the full bandwidth of interest. Previous demonstrations of silicon photonic interleavers have been restricted to usable spectral bandwidths less than 70 nm due to the use of standard directional couplers which are highly wavelength dependent [46], [69]. A schematic overview of the device is shown in Fig. 6a, highlighting the tunable broadband coupler to the ring needed for uniformity in the wavelength-dependent phase shift over the full comb bandwidth. The broadband 50-50 couplers in the balanced MZI are designed using phase-control sections as described in [70]. The device was taped out as part of a multi-project wafer (MPW) run in Advanced Micro Foundry’s (AMF) 200 mm silicon photonics process. An optical microscope image of the fabricated device is shown in Fig. 6b. The experimentally measured transmission spectrum of the device is shown in Fig. 6c-d, with flat-top pass- and stop-bands maintained over the full 125 nm bandwidth and a typical crosstalk suppression around 15 dB. This demonstration enables the previously proposed link designs using an even-odd interleaving architecture to operate with comb bandwidths spanning 125 nm and therefore support over 150 wavelength channels at 100 GHz spacing.

B. Ultra-Efficient Thermal Phase Shifters

Thermal phase shifters are widely used to correct for fabrication-induced phase errors and temperature swings in silicon photonic devices due to the large thermo-optic coefficient of silicon ($\frac{dn}{dT} = 1.8 \times 10^{-4}$ K$^{-1}$). This is particularly true for ring-assisted MZI interleavers, which are highly phase-sensitive in both ring resonance location and path length imbalance, necessitating at least two phase shifters per device. However, such phase shifters typically consume on the order of tens of milliwatts per element and are susceptible to large intra-chip parasitic crosstalk. For these reasons, elementary thermal devices do not meet the performance requirements inherent in large-scale, low-energy photonic interconnects. However, the efficiency of thermal devices can be dramatically increased by selectively removing the silicon substrate under the heater, effectively eliminating the path of lowest thermal...
Next, a vapor phase xenon difluoride (XeF$_2$) etch was used to selectively remove the substrate. This approach can be extended to entirely wafer-scale processing, solving thousands of phase shifters with minimal crosstalk and enabling future high-throughput fabrication of circuits containing thousands of phase shifters with minimal crosstalk and energy consumption.

Fig. 7. a, Cross-sectional view of a standard silicon waveguide with doped silicon heater lines running parallel. b, Cross-sectional view of the waveguide and heaters after substrate undercut, effectively air cladding the system for thermal isolation. c, Steady-state finite element simulation for the standard geometry, showing a $\Delta T \sim 4 \, ^\circ$K change in the waveguide temperature for 5 mW dissipated power. d, Identical simulation for the undercut geometry also with 5 mW dissipated power, demonstrating a dramatic increase in waveguide temperature ($\Delta T \sim 315 \, ^\circ$K). e, Microscope image of the unmodified foundry chip with air trenches down to BOX. f, Microscope image of the chip after postprocessing using inductively coupled plasma reactive ion etching (ICP RIE) to expose the substrate and a vapor phase xenon difluoride (XeF$_2$) etch to selectively remove the substrate. g, Measured frequency response of the standard and undercut devices. h, Measured thermal efficiency of the undercut device, demonstrating $P_{th} = 1.2 \, \text{mW}$.

to the thermal isolation, the full time of the thermal response is greatly increased, corresponding to a reduction in the device $S_{21}$ bandwidth compared to non-undercut designs (Fig. 7h). However, the thermal isolation increases both the response time for the integrated heater and the response time for ambient temperature fluctuations, so this reduction in $S_{21}$ does not impact the efficacy of standard thermal control loops used to stabilize devices. Finally, we have closely collaborated with AIM Photonics to enable the full undercut design flow entirely in-line at the wafer scale, eliminating the post-processing requirements and permitting large-scale silicon photonic circuits with thousands to tens of thousands of highly efficient thermal phase shifters on a single chip.

C. Robustness to Fabrication Variations

Despite the efficiency gains realized through undercut heaters as described in the previous section, the required thermal tuning energy to correct for fabrication-induced phase errors still represents a large fraction of the link energy budget and prevents sub-pJ/bit operation. To address this issue, we propose and demonstrate a fabrication-robust platform in standard sub-micron silicon-on-insulator (SOI) processes, yielding devices with dramatically reduced sensitivity to fabrication variations [74]. Furthermore, these devices are fully compatible with substrate undercuts, illuminating a promising path to sub-100 $\mu$W per device worst-case thermal power consumption. Our approach takes advantage of the well-known benefit of wide waveguides in terms of robustness to width variations, which are typically only used in straight sections of devices [75] due to wide waveguides supporting higher-order spatial modes which are easily excited in bends. In our...
platform, we enable the use of wide waveguides for entire devices through carefully designing all bends for adiabaticity, employing Euler spirals which have a radius of curvature which varies linearly along the path \([76]--[79]\). The condition of adiabaticity minimizes mode mismatch interfaces through the bend and thus ensures that the \(TE_0 \rightarrow TE_0\) transmission is maximized while \(TE_0 \rightarrow TE_{1,2,...,N}\) transmission is minimized, where \(N\) is the total number of supported modes. Using this methodology with 1.2 \(\mu m\) wide waveguides, we demonstrate compact RMZI interleavers with state-of-the-art performance and dramatically reduced sensitivity to fabrication variations \([74]\). We validate device performance for silicon devices fabricated using electron beam lithography in a university cleanroom setting as well as devices fabricated using deep ultra-violet (DUV) lithography in a 300 mm commercial foundry with excellent agreement between the two cases. Furthermore, we have recently taped out comprehensive DoE test structures on a dedicated 300 mm wafer run through AIM Photonics to collect comprehensive die-to-die, reticle-to-reticle, and wafer-to-wafer statistics for waveguide widths ranging from 400 nm to 3 \(\mu m\). The concert of ultra-broadband performance, highly efficient undercut thermal phase shifters, and fabrication-robust design, all demonstrated on foundry-fabricated wafers, will enable silicon photonic interleavers which meet the stringent performance and energy requirements for Tb/s/fiber and sub-pJ/bit operation.

V. Resonant Modulators

While silicon photonic resonant modulators have been widely studied by both academia and industry due to their small footprint, high bandwidth, and energy efficiency, the stringent requirements of the interconnects proposed in this work severely restrict the available design space. In particular, the devices are required to achieve state-of-the-art performance in terms of extinction ratio and insertion loss while being driven by an EIC fabricated in a leading-edge CMOS process, which is typically supply voltage-limited to \(V_{dd} < 1\) V. Furthermore, the devices must have a large FSR (\(> 20\) nm) and must be capable of being fabricated in high volume at a commercial CMOS foundry. Vertical junction microdisk modulators have been demonstrated with performance capable of satisfying this list of requirements \([44], [81]\), largely due to their low radiation loss at extremely small radii (\(< 5\) \(\mu m\)) and large overlap between the vertical junction and whispering gallery mode for standard 220 nm SOI processes. Furthermore, microdisk modulators possess a number of other favorable qualities compared to standard single-mode ring cavities, including reduced sensitivity to width variations. In this section, we explore the intricacies of the design and fabrication of microdisk modulators targeted for such systems, demonstrating a comprehensive design and simulation flow which closely matches experimental measurements from devices fabricated in a commercial 300 nm CMOS foundry.

A. Application-Specific Design Space

Resonant modulators have a notoriously large number of competing variables and design trade-offs which must be balanced in an application-specific manner, including (but not limited to) resonance enhancement versus photon lifetime, carrier concentration versus round-trip loss, and mode overlap versus junction capacitance. For example, a high quality factor (Q) resonator will have a more narrow resonance dip (increased resonance enhancement) resulting in a large dynamic extinction ratio (ER) when the resonance is shifted, with the caveat that the time-domain behavior of light circulating in the cavity will restrict the achievable modulation bandwidth \([82]\). The relation between optical and electrical bandwidth is given by \([44]\):

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

where \(f_{3dB}^{el-opt}\) is the total electro-optic bandwidth, \(f_{3dB}^{opt}\) is the optical bandwidth (photon lifetime-limited bandwidth), and \(f_{3dB}^{el}\) is the electrical RC-limited bandwidth. From this relation, it is clear that a small optical bandwidth cannot be compensated by a large electrical bandwidth; for a particular target bandwidth, a corresponding reduction in Q is required regardless of improvements in the RC time constant.

Since the proposed link implementations focus on high parallelism with natively error-free performance at modest per-channel data rates (chosen for optimal energy efficiency and latency, discussed further in Section VI), the modulators do not need to support data rates higher than \(\sim 20\) Gb/s/λ. Due to this eased performance requirement, the photon lifetime limit on bandwidth is no longer as restrictive and thus higher Q modulators (Q \(\sim 10,000\)) with improved resonance enhancement can be considered. While high-Q modulators are permissible for these applications requiring modest data rates (5–20 Gb/s), additional care must be taken in the physical-level modeling to account for optical power handing considerations such as self-heating \([83]\). However, in the context of Kerr frequency combs for which the optical power per line can vary significantly, the low optical power of particular comb lines allows for selective modulator design in which each modulator can be individually tailored in an \textit{a posteriori} manner to match the characteristics of the target comb line.

B. Carrier Profile

The vertical junction geometry is shown in Fig. 8a(ii), with n-doping in the top half of the waveguide and p-doping in the bottom half. Moderately doped p+ and n+ implants are used to reduce the series resistance and heavily doped p++ and n++ implants are used to form ohmic contacts between the silicon and metal vias \([44]\). Electrical carrier simulations were run using Ansys Lumerical CHARGE \([84]\) to optimize the target carrier concentrations and maximize the change in depletion width with applied voltage (Fig. 8a(iii-iv)) while keeping the optical loss experienced by the mode in a moderate range (\(\sim 30\) dB/cm) such that the intrinsic cavity Q remains high.

For an abrupt p-n junction, the theoretical depletion width \(w_d\) is given by:

\[
w_d = \sqrt{\frac{2e(N_D + N_A)(\phi_{bd} - V)}{qN_A N_D}}
\]
where $\varepsilon$ 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 $N_D = N_A = 2.5 \times 10^{18}$ cm$^{-3}$, $\phi_{bi} \approx 1$ V. From this, we can calculate the depletion width at zero bias for our vertical junction, yielding $w_d \approx 30$ nm, which agrees well with the TCAD simulated profile (Fig. 8a(iii)). The optimized 2D carrier concentration cross-sections were exported for electro-optic finite difference eigenmode (FDE) disk simulations in cylindrical coordinates using the well-known Soref and Bennett relations [85] (Ansys Lumerical MODE). Similarly, full 3D carrier simulations of the entire disk were run and exported for final electro-optic verification in 3D finite difference time domain (FDTD) simulations (Ansys Lumerical FDTD).

C. Mode-Selective Coupler

Microdisks are highly multimode systems and support a plethora of higher-order whispering gallery modes which can result in undesired spectral features leading to excess crosstalk and loss for non-target channels. To avoid exciting these higher-order modes and achieve an uncorrupted FSR, the coupler must be carefully designed to selectively excite the fundamental TE$_{0}$ mode close to critical coupling while transferring orders-of-magnitude less light from the bus into other modes. To achieve this, a radial pulley coupler is used to match the angular propagation constants of the symmetric and anti-symmetric supermodes of the disk-waveguide system for selective excitation of the fundamental TE$_{0}$ whispering gallery mode (Fig. 8a) [80]. First, an FDE solver is used to solve for the angular propagation constants (APCs) of the disk and curved bus waveguide in isolation to narrow the design space. However, coupling the bus and disk perturbs these APCs and gives rise to symmetric and anti-symmetric supermodes of the waveguide-disk system. FDE sweeps of the coupled system for concentric configurations are then simulated to extract the APC difference between the symmetric and anti-symmetric supermodes, yielding an anti-crossing region for which there is good phase-matching and thus selective excitation of the fundamental TE$_{0}$ mode (Fig. 8b). Next, sampled disk-coupler configurations from this anti-crossing region were simulated using 3D FDTD for a range of coupler angles to identify designs where coupling into TE$_{0}$ is close to critical ($\sim 2.5\%$) while coupling into TE$_{1}$ is minimized (Fig. 8c). The critical coupling condition of 2.5% power coupling was extracted from the free carrier-induced loss for the target doping profile as described in the previous section.

D. Electro-Optic Device Simulation

The full electro-optic performance of the device was simulated using two methods: (i) S-parameter-based circuit simulations (Ansys Lumerical INTERCONNECT) with exported carrier and mode results, and (ii) direct 3D FDTD simulations with exported 3D carrier profiles. S-parameter-based simulations were used to rapidly explore the electro-optic design space and create compact models for larger circuit-level simulations (eye diagram evaluation, bit error rate testing, etc.). The compact models consist of the exported 2D mode profile as a function of applied voltage (capturing the real and imaginary parts of $\Delta n_{eff}(V)$) and the exported coupler...
S-parameters from FDTD, which are used to form a transfer-matrix of the disk cavity to efficiently capture time-domain behavior [86]. Typically, high-Q resonant designs are difficult to simulate in FDTD since the simulation is computed in the time domain and convergence depends on the cavity photon lifetime. However, fully 3D electro-optic FDTD simulations are highly desirable to use as a final validation step for such designs, despite their high computational cost. Although the undoped silicon microdisks have extremely low material and radiation loss, the carrier-induced loss from the doping profile reduces the Q such that these simulations are viable to run on modern workstations for verification. The results for the 3D FDTD electro-optic simulations are shown in Fig. 8d-e, validating the design in terms of FSR (∼ 24.5 nm), critical coupling condition, modulation efficiency (∼ 67 pm/V), and higher-order mode suppression.

E. Integrated Heater

Doped silicon heaters, which are fabricated in the same silicon layer as the active and passive optical devices, require careful design as to not induce excessive optical loss while still maintaining close proximity to the optical mode for optimal thermo-optic efficiency. This design problem is additionally exacerbated in compact resonant modulators by the requirement of full electrical isolation between the heater and p-n junction used for carrier-depletion-based modulation. Previous designs have relied on engineering the intrinsic silicon region and thus result in highly undesirable electrical coupling between the heater and modulating diode.

We propose and demonstrate the use of a narrow silicon etch around an internal doped silicon heater in a compact microdisk modulator to provide full electrical isolation between the heater and the neighboring vertical junction p-n diode (Fig. 9) [88]. By introducing a 100 nm SiO₂ barrier between the heater and diode, we observe complete electrical isolation in both simulation and experiment with minimal degradation in thermo-optic efficiency. Furthermore, this degradation in efficiency can be made nearly negligible through the use of a selective substrate undercut similar to that described in Section IV [89]. This isolated heater enables reliable and fabrication-robust thermally-tunable microdisk modulators while maintaining a compact footprint and large FSR.

F. Experimental Characterization

The six new vertical junction implants, including series link and contact implants, were co-designed in collaboration with a commercial 300 mm foundry (AIM Photonics) to optimize doping conditions (implantation energy, dose, species, etc.). Doping quadrants were run across multiple wafers to optimize implant conditions and fully explore the electro-optic performance dependence on carrier concentration above and below the target design point. The design of experiment (DoE) devices were taped out in a dedicated wafer run and experimentally validated against simulation models, showing...
good agreement in key metrics including FSR, higher-order mode suppression, critical coupling condition, and thermal tuning efficiency (Fig. 10a-b). Scanning electron microscope (SEM) images of the devices prior to back-end-of-line (BEOL) processing were used to verify the integrity of the full silicon etch around the integrated heater and gap width between the bus and disk for both single bus and add-drop designs (Fig. 10c). The experimentally measured depletion response (60 pm/V, Fig. 10d) closely matches the electro-optic 3D FDTD simulations (67 pm/V, Fig. 8e), validating the design and simulation workflow. Finally, due to the high-Q design for moderate data rates, the devices can achieve large extinction ratio and low insertion loss operation for CMOS-compatible drive voltages (Fig. 10e).

VI. ENERGY EFFICIENCY CONSIDERATIONS

While a comprehensive analysis of the link energy efficiency trade-offs is beyond the scope of this work (see [90] and [91] for extended discussions), the device and system design choices outlined here are heavily influenced by general considerations regarding the energy consumption of high-speed electronic-photonics data communication links. In this section, we briefly discuss these considerations for all sources of energy consumption in our architecture and evaluate the requisite system characteristics necessary to achieve sub-pJ/bit operation.

A. Modulation Format, Data Rate, and Bit-Error Rate

While coherent communications (encoding information in the phase of light) can provide significantly more channel bandwidth than intensity modulated direct-detection (IM-DD) schemes, this increase comes at a cost in terms of latency and energy consumption. In particular, modest data rate IM-DD signaling with a native bit error rate (BER) below $10^{-12}$ provides ideal latency and energy efficiency characteristics since it removes the need for power-hungry digital signal processing (DSP) and forward error correction (FEC). Such signaling is typically infeasible for standard single or quad channel links due to the low aggregate bandwidth; however, in the proposed massively parallel WDM links, Tb/s/fiber aggregate bandwidth is achievable with modest data rates on the order of 10-20 Gb/s/λ. This wide-parallel approach keeps the signal format close to that of intra-die compute node signaling, reducing the need for DSP and (de-)serialization which in turn reduces latency and energy consumption.

B. Comb Source

Since Kerr combs rely on a CW pump to generate the comb, the total efficiency is determined by the wall-plug efficiency (WPE) of the pump and the pump-to-comb conversion efficiency of the microresonators. For this reason, it is critically important for both efficiencies to be large such that their product represents reasonable conversion from electricity to usable optical carriers. Pump conversion into unusable comb lines below the link budget can significantly degrade the overall system wall-plug efficiency and thus conversion should be spectrally confined to the bandwidth of interest with sharp roll-off beyond the band edges [92]. In this direction, we are pursuing synchronized Kerr combs [65], [93] possessing different spectral envelopes which are then coherently combined to boost and equalize the line powers. Furthermore, normal GVD Kerr combs exhibit inherently high pump-to-comb efficiencies (> 40%) which can be confined to a target spectral window and high power distributed feedback (DFB) lasers have been demonstrated with 35% WPE and output powers high enough for comb generation (∼ 200 mW) [94].

C. Modulators

Carrier-depletion-based resonant modulators (assumed to be a lumped-element, i.e. device size $< \lambda_{RF}$) act as a capacitive load which must be charged and discharged by the analog electronic driver. In contrast to carrier-injection modulators, no current flows under standard device operation and thus the only energy consumption is the switching energy associated with $0 \rightarrow 1$ transitions in the NRZ bit sequence ($E_b = \frac{1}{2}CV^2$ where $E_b$ is the energy per bit). Highly efficient vertical junction microdisk modulators with low peak-to-peak drive voltage have been shown to consume as low as $E_b = 1$ fJ [44] and thus do not contribute significantly to the total link energy.

D. Thermal Tuning

Thermal tuning of devices to correct for fabrication-induced phase errors and environmental temperature swings has been shown to be a dominant source of energy consumption in resonator-based silicon photonic systems subjected to co-packaged thermal environments [37]. The undercut thermal phase shifters and fabrication-robust design methodology discussed in Section IV can significantly decrease this source of energy consumption. Furthermore, other promising strategies for reducing thermal tuning power have been demonstrated in the silicon photonics platform, including post-fabrication trimming to near-perfectly correct for fabrication variations [95] and negative thermo-optic coefficient cladding designed to neutralize the strong thermo-optic effect in silicon [96]. However, both of these approaches require significant changes to the fabrication process, and trimming inherently reduces throughput since it requires post-processing of individual devices across the entire wafer.

E. Clock Distribution

Significant energy can be saved through using one of the comb WDM channels for clock forwarding since electronic clocking with digital phase-locked loops (PLLs) has been shown to be a dominant source of energy consumption in silicon photonic links [30]. As normal GVD Kerr combs have been shown to be capable of providing on the order of $> 100$ wavelength channels, the opportunity cost of the bandwidth gain associated with using one extra wavelength for data communications is far outweighed by the energy efficiency of using the same wavelength for clock forwarding [90].
F. Packaging Parasitics
Parasitic resistance, capacitance, and inductance associated with interfaces between electronic and photonic devices (wire-bonds, C4 solder bumps, copper pillars, etc.) can significantly degrade performance and must be considered in the overall system analysis. For this reason, we employ flip-chip bonding with copper pillar microbumps to form the connections between analog electronic drivers and photonic devices, creating a short electrical signal path with low parasitics. While monolithic integration provides the lowest parasitic connections between electronics and photonics (intra-chip BEOL metal traces), this gain is offset by the inability to leverage leading-edge process node transistors and thus has a significant impact on the energy consumption of the analog electronic circuits.

G. Aggregate Energy Efficiency Assumptions
In all reported energy efficiency numbers throughout this work, our analysis includes all sources beginning with an on-chip pseudo-random bit sequence (PRBS) generator in the flip-chipped CMOS ASIC proxy core and terminating at an on-chip BER checker on the same CMOS die. Therefore, the power consumption of all other electronic and photonic devices/circuits in the total signal path between the PRBS generation and checking are considered, including the copper pillar interconnects, electronic control circuits, and analog driving electronics. All electronic circuits are designed in the TSMC 28 nm process node.

VII. SYSTEM-LEVEL DEMONSTRATIONS
In this section, we detail comprehensive experimental demonstrations showing proof-of-principle operation for massively parallel silicon photonic links driven by Kerr frequency comb sources. Furthermore, we show dense heterogeneous integration between foundry-fabricated CMOS electronic and photonic chips, providing a platform for full system-level demonstrations of the proposed interconnect architecture.

A. Integrated Transmitter
We have successfully demonstrated error-free operation with a normal-GVD Kerr comb source and silicon photonic transmitter up to 16 Gb/s/λ for both single bus [23] and even-odd interleaved architectures [22], [24], with channel counts of 20 and 32, respectively. Furthermore, we have recently shown simultaneous modulation of adjacent comb lines at 200 GHz spacing on the same bus, achieving error-free operation for both channels with negligible crosstalk [97]. These initial proof-of-principle data transmission experiments were limited to 32 channels for a variety of pragmatic testing considerations including device insertion losses, packaging limitations, and the number of high-SNR comb lines generated in available fabricated devices. Through the previously detailed improvements of the photonic devices, Kerr combs, and packaging, future demonstrations will not have this restriction on channel count (see Appendix I). These experimental demonstrations validate the feasibility of the proposed architecture and approach, representing the first work of its kind showing WDM data transmission using active silicon photonic circuits with Kerr frequency comb sources.

B. Integrated Receiver
We demonstrated an experimental comb-driven receiver, shown in Fig. 11 and reported in [98], for evaluating the performance of receiving multiple comb lines with a PIC. Three comb lines at 1556.7, 1558.3, and 1559.9 nm are filtered from a Kerr comb source and modulated with a Mach Zehnder modulator at 16 Gb/s. Each data channel is offset by 0.5 unit intervals using 1.25 km of SMF-28 fiber and subsequently coupled into the PIC receiver chip. The PIC receiver drops the signals from a common bus with aligned ring filters to vertical p-i-n germanium photodiodes with 1.1 A/W responsivity, nearing the maximum quantum efficiency of 1.25 A/W responsivity in the C band [99]. A microscope image of the filter and photodiode is shown in Fig. 11b and the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
aligned filter spectrum is shown in Fig. 11c. The photodiodes were electrically probed with a 50 Ω S-G RF probe and reverse biased 2.5 V using a bias-T to minimize carrier transit time. A coaxial cable connects the probe to a real-time oscilloscope with 50 Ω termination, giving a 50 Ω transimpedance gain and yielding the resulting eye diagrams in Fig. 11d. The resulting received data was error-free at -2 dBm input power as measured with the DC current of the bias-T. The receiver demonstration was limited to three wavelengths due to the saturation output power of the EDFAs used in the experiment; the gain was saturated by the higher power lines and would not permit more lines at lower power into the system. A higher transimpedance gain, removal of the 1.25 km fiber spool, and closer integration with CMOS TIAs are expected to lower the required power for error-free reception.

C. 3D Heterogeneous Integration with CMOS Electronics

Silicon photonic devices can be arranged within a micron of each other due to the high index contrast between silicon and SiO₂. This vast potential for bandwidth density is constrained in applications by the packaging of the system. The foremost issue is integration of photonics and electronics. Photonic and CMOS devices may be monolithically integrated on the same die and connected with the wiring density of CMOS BEOL [30], [100]. Alternatively, 3D integration with flip-chip bonding places the EIC circuits on its own die, saving space and allowing for leading edge CMOS nodes with higher transistor density and performance [101]. In a 3D integrated transceiver, the density is limited by the connection pitch between the photonic and electronic die.

As a trial of dense 3D integration with low parasitics, we connected a CMOS EIC (TSMC 28 nm node) and photonic chip with solder-tipped copper pillar bumps at a 25 μm pitch (Fig. 12). The electronic and photonic chips are fabricated with arrays of aluminum terminated pads. In post-processing, these pads are plated with under-bump metallization, and copper pillar microbumps are fabricated on the EIC. A microscope image of the bumps is shown in Fig. 12c. The EIC is then flipped onto the PIC, with power and control signals delivered to the EIC by traces on the PIC wirebonded to a PCB. The combined die are shown in Fig. 12b. In this packaging platform, a link is designed with two stages of even-odd interleavers and floating edge couplers with sub-dB fiber-to-fiber alignment. The footprint of the 64 channel system fits into a total footprint of 0.4 mm², this platform could conceivably support 3.125 Pb/s package escape bandwidth. However, while densely bumped flip-chip packaging would scale trivially to accommodate the larger die size, optical I/O presents a bottleneck. Assuming the previously detailed 1 Tb/s/fiber bandwidth, the package would require 1,000 optical connections for Pb/s operation. Again assuming a reticle-limited die edge length of 25 mm with edge couplers at 127 μm pitch, each side would be limited to 196 connections and thus could only provide 784 total optical ports. However, using widely available reduced cladding fibers with 80 μm diameter [102], each side could support 312 connections and thus conceivably permit Pb/s-scale operation on a single chip.

D. Future Demonstrations

We have recently taped out a dedicated 300 mm wafer through AIM Photonics with various device- and system-level iterations improving on those described in this work (Fig. 13). Most notably, dedicated wafer fabrication (compared to singulated MPW die) will enable wafer-scale bumping and flip-chip bonding for more reliable and higher-yield packaging with CMOS electronics. Furthermore, the system-level circuit iterations taped out on this run include a combination of the device-level advances described in this work, including highly efficient microdisk modulators, undercut thermal phase shifters (for all thermally tuned devices), fabrication-robust interleavers, and floating edge couplers with sub-dB fiber-to-
chip loss for standard SMF-28 fiber. The MCM circuits on this wafer leverage all of the previously mentioned device advances and target an optical amplifier-free 64 channel system capable of achieving 0.5 pJ/bit energy efficiency with 1.024 Tbps/fiber aggregate bandwidth and 5 Tbps/mm\(^2\) areal bandwidth density. Through building a fully developed application-specific device library in a single foundry platform, unprecedented design complexity and performance can be achieved to meet the stringent metrics required for future Pb/s and sub-PJ/bit interconnect operation.

VIII. SUMMARY AND OUTLOOK

In this work, we have provided a comprehensive overview of massively parallel WDM optical interconnects based on integrated Kerr frequency combs and silicon photonic circuits. We detail the full link design hierarchy, including fundamental device design, circuit-level design considerations, and co-design of heterogeneously-integrated electronic-photonic systems. To support these design conclusions, we present a wide variety of representative results from simulation and experiment at each step of the hierarchy including numerous novel device and system demonstrations. While the work described here focused exclusively on extreme scaling in the wavelength domain, we have recently demonstrated a mode-division multiplexing (MDM) interface between silicon photonic chips and few-mode fiber [103], [104], leveraging the spatial mode domain as an orthogonal axis for multiplicative WDM and MDM scaling. Various WDM-compatible MDM architectures have been demonstrated in the silicon photonics platform [25], [105], which can be trivially extended to accommodate the WDM architectures described in this work. Achieving Pb/s package escape bandwidths will be crucial for continued scaling of computing systems to keep pace with exponentially growing demand over the next quarter century and beyond.

APPENDIX I: OPTICAL LINK BUDGET

In this Appendix, we present an analytical model for the optical link budget to demonstrate the scalability of the proposed massively parallel architecture. As an example, we consider an implementation driven by a 64 channel comb with 150 GHz channel spacing. To achieve Tb/s aggregate bandwidth per fiber while maintaining low single-channel NRZ data rates, we assume a single channel data rate of 16 Gb/s/\(\lambda\). Using two stages of even-odd interleaving, as illustrated in Fig. 2c, the initial 150 GHz-spaced comb is subdivided into four parallel buses, each with 16 channels at 600 GHz spacing. There are several sources of optical loss to be considered in such an architecture, which can be split among two main categories: bandwidth-dependent device IL and signal power penalties. The total optical link penalty is the sum of these two parts.

Bandwidth-Dependent Device IL

For coupling light on to and off of each chip, we use edge couplers rather than grating couplers due to their broadband low-loss performance. Previous state of the art demonstrations for silicon photonic edge couplers have shown broadband coupling efficiency \(\leq 1\) dB [106]. Limiting our scope to edge couplers that are compatible with the same fabrication processes as the rest of our devices, we estimate the per facet IL to be 1 dB. The link architecture uses a total of six RAMZIs performing even-odd interleaving to multiplex and de-multiplex the comb lines. To model the bandwidth dependent loss of the devices, we measured the transmission spectra on one of our test chips and fit a polynomial curve to the peaks. The 64 channel 150 GHz comb covers approximately 73 nm of optical bandwidth, corresponding to \(\sim 0.7\) dB IL per RAMZI.

For the 16 channel resonator arrays, two sources of device IL are considered: on- and off-resonance IL. For the modulators and filters on-resonance IL as low as 1 dB and 0.2 dB have been demonstrated respectively [40], [44]. For off-resonance IL, we have measured as low as 0.05 dB per device. Finally, the long on-chip routing results in non-negligible waveguide propagation loss, which we assume to be \(\sim 1\) dB/cm from reported foundry measurements [40]. We assume roughly 3.5 mm of waveguide routing between the transmitter and receiver, dominated by device electrical pad pitch.

Signal Power Penalties

The wide-parallel approach, using the lowest-order IM-DD format of non-return-to-zero on-off keying (NRZ-OOK), removes the need for power-consuming and latency-inducing DSP and FEC, but also introduces an inherent power penalty dependent on the signal extinction ratio (ER). The penalty can be calculated as

\[
PP_{ER/OOK} = PP_{ER} + PP_{OOK} = -10 \log_{10} \left( \frac{r - 1}{r + 1} \right) - 10 \log_{10} \left( \frac{r + 1}{2r} \right)
\]

Where \(r = 10^{ER/10}\) and \(ER\) is the logarithmic extinction ratio between the optical power of a ‘1’ and a ‘0’, in units of dB. Fig. 14a plots the individual components and total
TABLE I
SUM OF TOTAL INSERTION LOSSES AND POWER PENALTIES IN THE PROPOSED 1.024 Tbps LINK.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>2.80</td>
<td>1.75</td>
<td>3.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Receiver</th>
<th>Coupler IL [dB]</th>
<th>Interleaver IL [dB]</th>
<th>Ring Filter Array IL [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1.4</td>
<td>0.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.35</td>
<td>11.3</td>
<td>3.35</td>
<td><strong>15</strong></td>
</tr>
</tbody>
</table>

$P_{ER/OOK}$ as a function of ER, showing that for a perfect, infinite ER, the penalty converges to 3 dB (as a result of average power of the signal being half the input laser power for a data pattern with an equal distribution of ‘1’s and ‘0’s). Fig. 14b shows the derivative of modulation power penalty with respect to ER, emphasizing the diminishing returns in pursuing an $ER \geq 10$ dB.

![Fig. 14](image_url)

As illustrated in Fig. 10e, the device IL and ER of resonant modulators are inherently coupled, along with the electrical driving signal $V_{pp}$. For the purposes of this analysis, we use the IL and ER values demonstrated in [44], where $IL \sim 1$ dB and $ER \sim 8$ dB at data rates near 16 Gb/s with a CMOS compatible $V_{pp}$.

Inter-channel cross-talk limits can contribute substantial power penalties in WDM links, particularly if the channel spacing is made small. It has been demonstrated that for cascaded resonator-based WDM links, these penalties can be considered negligible ($\leq 0.1$ dB) when $\lambda_{ag} \geq 150$ GHz [39]. While a link operating across a single resonator FSR would be fundamentally limited, we can instead operate in the multi- FSR regime, and refer to Fig. 4c for a selection of resonator FSR for a 16 channel 600 GHz-spaced WDM bus. It is clear that a resonator FSR between 25 and 30 nm, corresponding to a disk radius between 3.6 µm and 4.5 µm, satisfies the $\lambda_{ag}$ constraint.

Finally, the bandwidth of the resonant filter at the receiver can introduce signal distortion, but has also been demonstrated as optically equalizing and mitigating pulse width distortion provided the Q-factor of the filter is designed so that the Lorentzian full-width half-max is approximately twice the channel data rate [107]. Pending further work on this topic, we do not attribute any positive or negative (an improvement in signal quality) penalty to the resonant filter spectral distortion effects.

Total Optical Link Budget

Present in the transmitter are 3 edge couplers, 4 RAMZI filter stages, 1 on-resonance and 15 off-resonance modulators per channel, and the modulation signal penalty. On the receiver side there is 1 edge coupler, 2 RAMZI filter stages, and similarly 1 on-resonance and 15 off-resonance filters per channel. The total link penalty is summarized in Table I.

Using state-of-the-art receiver designs we estimate a receiver sensitivity threshold of $\leq -19$ dBm at a channel data rate of 16 Gb/s [108]. Adding the estimated total optical link penalty, this results in a minimum optical power per line of -4 dBm to satisfy native error-free link performance.

IX. ACKNOWLEDGEMENTS

This work was supported in part by the U.S. Advanced Research Projects Agency–Energy under ENLITENED Grant DE-AR000843 and in part by the U.S. Defense Advanced Research Projects Agency under PIPES Grant HR00111920014. This work was performed in part at the Cornell NanoScale Facility, a member of the National Nanotechnology Coordinated Infrastructure (NNCI), which is supported by the National Science Foundation (Grant NNCI-2025233). The authors thank AIM Photonics and Advanced Micro Foundry for chip fabrication and Lewis Carpenter for SEM imaging.
Aneek James received his B.S. in Electrical and Electronics Engineering from the University of Georgia, Athens, GA in 2017 and his M.S., and M.Phil., in Electrical Engineering from Columbia University, New York, NY in 2019 and 2021, respectively. He is working as a Ph.D. candidate in Electrical Engineering in the Lightwave Research Laboratory under Professor Keren Bergman. He is also currently a part-time student researcher at Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA. His research interests include modeling fabrication variations in silicon photonic devices, as well as the testing and automated control of silicon photonic systems for high-throughput optical interconnects.

Vignesh Gopal received his B.S. in Electrical Engineering from the University of Illinois at Urbana Champaign in 2019 and his M.S. in Electrical Engineering from Columbia University in 2021. At Illinois, he graduated with high honors and was part of the James Scholar program. He is currently working in the Lightwave Research Laboratory at Columbia University under the mentorship of Professor Keren Bergman. His current research focuses on photonic transceiver design and high speed optical testing.

Vaihnavi Murthy received her B.S. degree in Electrical Engineering from University of Maryland, College Park, MD in 2019 and her M.S. degree in Electrical Engineering from Columbia University, New York, NY. She is currently pursuing a Ph.D. at Columbia University.

Qixiang Cheng received the B.S. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2010, and the Ph.D. degree from the University of Cambridge, Cambridge, U.K., in 2014. He joined the Shannon Laboratory, Huawei, China, where he researched future optical computing systems. From September 2016 to November 2019, he was a postdoctoral researcher and then a research scientist with the Lightwave Research Laboratory, Columbia University, New York, NY, USA. He has been appointed as a University Lecturer in photonic devices and systems at the University of Cambridge, since January 2020. His current research interests focus on system-wide photonic integrated circuits for optical communication and optical computing applications, including a range of optical functional circuits, such as packet-, circuit-, and wavelength-level optical switch fabrics, massively parallel transceivers, optical neural networks, and optical network-on-chip.

Bok Young Kim received his B.S. degree in Engineering Physics and Computer Science from Cornell University in 2014 and his Ph.D. in Applied Physics from Columbia University in 2021 under Professor Alexander Gaeta.

Xingchen Ji received his B.S. degree in Optics from Beijing Jiaotong University in 2013 and his Ph.D. in Electrical and Computer Engineering from Cornell University in 2018 under Professor Michal Lipson.

Yoshitomo Okawachi received his B.S. degree in Engineering Physics in 2002 and his Ph.D. in Applied Physics in 2008, both from Cornell University. He is currently a Research Scientist in the Quantum and Nonlinear Photonics Group at Columbia University with current research interests in optical frequency comb generation in silicon-based waveguides and microresonators, coherent computing based on degenerate optical parametric oscillation in microresonators, parametric nonlinear interactions in photonic devices, slow light, and all-optical signal processing using time-space duality techniques.

Matthew van Nickerk received the B.S degree in physics and mathematics from Roberts Wesleyan College, Rochester, NY, in 2017. He is currently pursuing a Ph.D. degree in Microsystems Engineering at the Rochester Institute of Technology under Professor Stefan Preble. He received a triplet award from the Quantum Information Science and Engineering Network in 2021. His doctoral research focuses on designing and implementing novel architectures for neuromorphic photonics and electronic-photonic co-packaging.

Venkatesh Deenadayalan received his Bachelor of Technology in Electronics and Communications Engineering from SRM University, Chennai, Tamilnadu, India in 2015. He got his Master’s degree in Microelectronics Engineering from Rochester Institute of Technology, Rochester, New York in 2019 and is currently a Ph.D. student in the same school working under Professor Stefan Preble. His doctoral research is focused on the design and fabrication of MEMS-based phase shifters for photonic circuits and development of packaging solutions for PIC chips.

Gerald Leake received the B.S. degree in material engineering from Penn State University, PA, USA, in 1999. He started his career at Atmel Corporation as a Process Engineer in Photolithography. He was with IBM Corporation as an Integration Engineer for 11 years working in the Semiconductor Research and Development Center in East Fishkill, NY, USA. His work spanned through CMOS Technology Nodes 130, 90, 65, 32, 20 nm. He is currently a Senior Integration Engineer of Emerging Technologies, SUNY Polytechnic Institute. His current function is on the AIM Photonics which is an institute of the National Network for Manufacturing Innovation.

Michael Fanto received his B.S. degree in Physics from Utica University in 2002 and his Ph.D. in Microsystems Engineering from Rochester Institute of Technology in 2020 with a focus in Quantum Integrated Photonics. He is currently a senior research physicist at the Air Force Research Laboratory Information Directorate in Rome, NY.

Michal Lipson (SM’07–F’13) is the Eugene Higgins Professor of Electrical Engineering and Professor of Applied Physics at Columbia University. She received her Ph.D. in Physics from the Technion in 1998. Her research focus is on nanophotonics and includes the investigation of novel phenomena, as well as the development of novel devices and applications. She pioneered critical building blocks in the field of silicon photonics, which today is recognized as one of the most promising directions for solving the major bottlenecks in microelectronics. She is the inventor of over 45 issued patents and has co-authored more than 250 scientific publications. She is a member of the National Academy of Sciences, a MacArthur Fellow, and a Fellow of the IEEE and OSA.

Alexander Gaeta is the David M. Ricky Professor of Applied Physics and Materials Science at Columbia University. He received his Ph.D. in Optics from the University of Rochester in 1991. He has made extensive contributions to quantum and nonlinear optics including nonlinear propagation of ultrashort pulses, nonlinear optics in photonic crystal fibers, few-photon nonlinear interactions, generation of optical frequency combs, and nonlinear nanophotonics. He is a Fellow of APS, IEEE and OSA.
Keren Bergman (S'87–M'93–SM'07–F'09) received the B.S. degree from Bucknell University, Lewisburg, PA, in 1988, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1991 and 1994, respectively, all in electrical engineering. Dr. Bergman is currently a Charles Batchelor Professor at Columbia University, New York, NY, where she also directs the Lightwave Research Laboratory. She leads multiple research programs on optical interconnection networks for advanced computing systems, data centers, optical packet switched routers, and chip multiprocessor nanophotonic networks-on-chip. Dr. Bergman is a Fellow of the IEEE and OSA.