

A Monolithically-Integrated Optical Receiver in Standard 45-nm SOI

Michael Georgas, *Student Member, IEEE*, Jason Orcutt, Rajeev J. Ram, *Member, IEEE*, and Vladimir Stojanović, *Member, IEEE*

Abstract—Integrated photonics has emerged as an I/O technology that can meet the throughput demands of future many-core processors. Taking advantage of the low capacitance environment provided by monolithic integration, we developed an integrating receiver front-end built directly into a clocked comparator, achieving high sensitivity and energy-efficiency. A simple model of the receiver provides intuition on the effects of wiring and photodiode capacitance, and leads to a photodiode-splitting technique enabling improved sensitivity at higher data rates. The receiver is characterized *in situ* and shown to operate with μA -sensitivity at 3.5 Gb/s with a power consumption of 180 μW (52 fJ/bit) and area of 108 μm^2 . This work demonstrates that photonics and electronics can be jointly integrated in a standard 45-nm SOI process.

Index Terms—Photonics, interconnect, monolithic integration, SOI, high-speed I/O, many-core, multi-core, sense-amplifiers, transimpedance amplifiers, integrating receivers, chip-to-chip links.

I. INTRODUCTION

IN ORDER TO harness the potential of emerging many-core processor systems, the communication fabric between cores and shared off-chip memory must provide high throughput at low power and footprint costs, overcoming chip power constraints and I/O pin limitations. Monolithically-integrated silicon photonics offers a dense wavelength-division-multiplexed (DWDM) fabric with orders of magnitude better energy-efficiency and bandwidth density than electrical interconnects [1]. However, existing circuit techniques are not geared towards leveraging the small photodiode (PD) capacitance and availability of a receive-side clock in a DWDM monolithically-integrated photonic link.

The optical receiver has traditionally been designed as a discrete component for optical fiber communication. To mitigate the gain-bandwidth limitation at the dominant pole of the input node without the availability of a receive-side clock, transimpedance amplifiers (TIA) were implemented to lower the input resistance, R_{in} , while preserving a large transimpedance gain, R_{TIA} [2].

Manuscript received November 17, 2011; revised January 19, 2012; accepted February 09, 2012. Date of publication May 22, 2012; date of current version June 21, 2012. This paper was approved by Guest Editor Patrick Reynaert. This work was supported in part by DARPA, NSF, FCRP IFC, Trusted Foundry, Intel, MIT CICS, APIC, Santec, and NSERC.

The authors are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: mgeorgas@mit.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSSC.2012.2191684

In [3] the photocurrent is integrated onto the capacitance at the receiver input, and the resulting voltage is evaluated by an energy-efficient clocked comparator. The input capacitance of 420 fF mitigates comparator kick-back and charge-sharing issues that would arise in a lower capacitance, more tightly integrated environment.

More recently, integrated photonics has addressed chip I/O bottlenecks through hybrid-packaged solutions [4], [5]. The decreased PD and parasitic capacitances have helped to improve the sensitivity and energy efficiency, but also caused designers to turn back to TIA-based designs that avoid the integrator problems described. A capacitance of 90 fF is reported in [4]. In [6] a 25 fF PD capacitance connected through a 20 fF microsoldier bump leads to a receiver sensitivity of 9 μA but energy-cost of 690 fJ/bit at 5 Gb/s. A TIA with a clocked comparator was implemented in [7] to achieve a lower energy-cost of 395 fJ/bit, but a sensitivity of only 22.1 μA . Further capacitance reduction will improve energy-cost and sensitivity, which maps directly to the system's laser power, resulting in designs more competitive with electrical solutions already at 1 pJ/bit [8].

In this paper we present a sense-amplifier-based optical data receiver with a monolithically-integrated photodetector. Building on [3], we make use of a source-forwarded clock and implement an integrating front-end and clocked comparator for high sensitivity and energy-efficiency. The low PD capacitance allows us to build the front-end directly into the comparator, avoiding kick-back and charge-sharing issues. We also develop the receiver sensitivity modeling and analysis, which provides intuition on the effects of wiring and PD capacitance, and leads to a PD-splitting technique that enables improved sensitivity at higher data-rates.

II. PHOTONIC LINKS

An example of a DWDM, monolithically-integrated photonic link is shown in Fig. 1. A continuous wave (CW), multi- λ laser is coupled from an optical fiber onto the chip through a vertical grating coupler. The light is then routed throughout the chip along waveguides fabricated using either gate poly-silicon or the SOI body.

Resonant drop rings form notch filters that can pull a particular wavelength-channel off of the optical bus. This can redirect the wavelength-channel to a different waveguide or be used to modulate a data signal. The modulator leverages the free-carrier-dispersion effect to modulate a P-N junction located around the ring in order to change the ring's refractive index, and therefore its resonant frequency. By tuning a particular ring's resonance to a wavelength channel, light is confined to the ring and

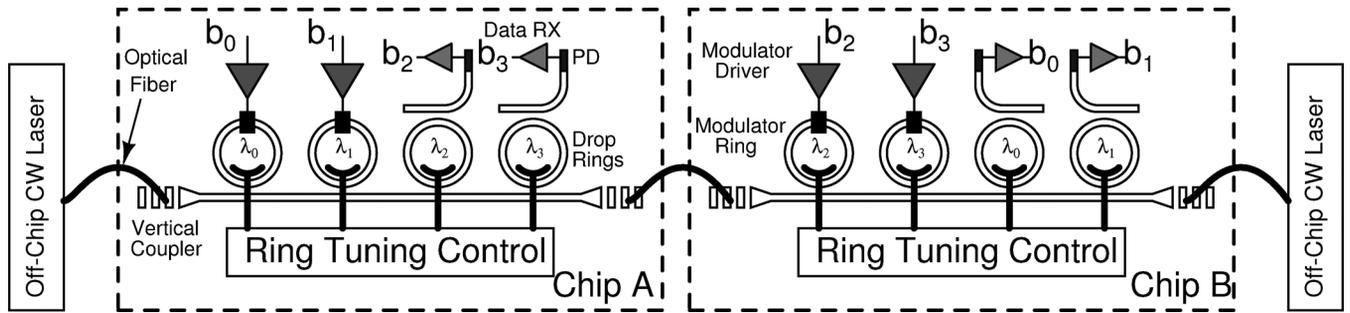


Fig. 1. An example optical link with chip-to-chip and intra-chip communication links shown. A CW laser source is coupled onto Chip A through a vertical grating. Two ring-resonant modulators imprint data onto two wavelength-channels, λ_0 and λ_1 , which propagate along the waveguide. The bus is routed over an optical fiber to Chip B. The drop rings on Chip B are each tuned to either λ_0 or λ_1 to select that channel from the bus and direct it to the correct data receiver. A second set of wavelengths, λ_2 and λ_3 carry data from Chip B to Chip A.

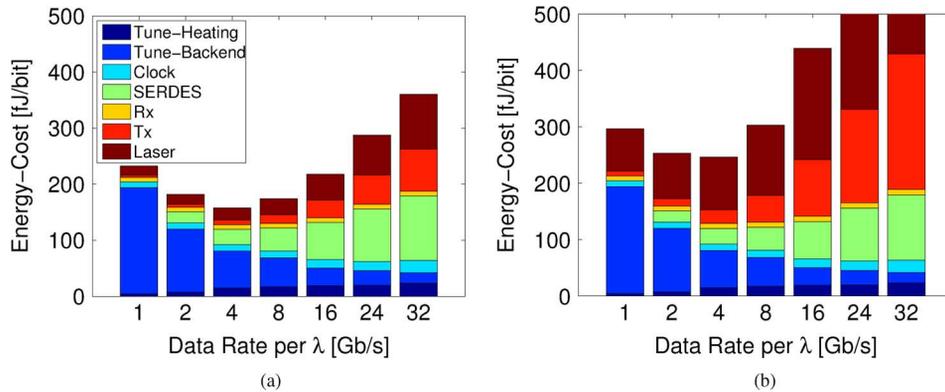


Fig. 2. WDM link power breakdown for a total throughput of 256 Gb/s. Two options are shown, representing monolithic and hybrid photonic integration. Modeling performed in equivalent 32-nm process. Legend entries follow the order of bar-graph sections from the bottom [9].

prevented from traveling down the waveguide, yielding an optical-0. De-tuning the ring shifts the notch filter away from the wavelength-channel and light propagates down the waveguide, yielding an optical-1.

The modulated light is routed to another location on the die (e.g., core-to-core) or to another die (e.g., socket-to-socket). At the destination, the ring-tuning control block selects the channel to be removed from the optical bus by setting the resonance of a drop ring filter to the particular wavelength-channel. An optical receiver, such as the one presented in this work, then converts the data back into the electrical domain by detecting the PD photocurrent. An optical clock signal can also be forwarded along with data, which is desirable as little interference exists in the optical domain.

One of the key characteristics of a monolithically-integrated photonic link is that all of the components are tightly integrated with each other on the same die. As such, power can be optimized at the system level, setting the specification of each component. The system-level design trade-offs are explored in [9] and summarized here.

Fig. 2 shows the WDM link power breakdown in an equivalent 32-nm CMOS process [9]. For a total throughput of 256 Gb/s, which is easily accommodated by a single optical fiber coupled to the chip, two integration scenarios are considered: 5 fF and 25 fF PD capacitances representing monolithic and hybrid integration, respectively. For each channel data rate, the system was optimized for energy-efficiency through a balance of component specifications.

Fig. 2(a) shows the energy-cost breakdown for $C_p = 5$ fF, where C_p is the sum of the PD capacitance, C_{PD} , and the wiring capacitance from the PD to the receiver, C_w . Focusing on any single channel rate, a high extinction ratio (ER) from the modulator will allow the receiver to operate with greater sensitivity, reducing the energy-cost at the laser. However, the cost for this ER is paid for in the modulator's energy budget. As the data rate per channel increases, the receivers become less sensitive, requiring an increase in laser power. Modulators try to compensate for this and become more expensive. Since the core frequency is kept fixed, serialization costs also grow significantly at higher rates. In contrast, as the data rate per channel increases, the total number of channels and rings decreases, requiring overall less tuning power. These two opposing energy-cost trends combine to form an optimal channel rate around 4 Gb/s to 8 Gb/s, motivating the design of receivers at moderate rates. Interestingly, this design point does not coincide with the traditional view in the optical community that channel rates should be pushed as fast as the circuit technology allows (often beyond 20 Gb/s).

As we move from a low total capacitance at the PD (Fig. 2(a)) to a higher one (Fig. 2(b)), the receiver's sensitivity degrades, requiring more laser power. The modulator tries to compensate for this by improving its insertion loss (IL) and ER, but in doing so consumes more power itself.

Fig. 2 shows that the use of moderate-rate DWDM enables energy-efficient link parallelism that leads to 200 fJ/bit link energy-costs, which is significantly less than other works in Table I.

TABLE I
 COMPARISON WITH PREVIOUS WORK

Work	Process	Rate (Gb/s)	RX Efficiency (pJ/bit)	Link Efficiency (pJ/bit)	Sensitivity RX (μ A)
[2]	80-nm CMOS	14 (20 GHz)	0.157	-	79.2
[3]	250-nm CMOS	1.6	1.88	-	11
[4]	90-nm CMOS	10	3.7	10	200
[6]	90-nm CMOS	5	0.690	1.01	9
[7]	40-nm CMOS	10	0.395	0.530	22.1
[16]	90-nm CMOS	16	3.62	8.1	52.4
[22]	130-nm CMOS	5	2.8	-	-
This Work	45-nm SOI	3.5	0.052	-	<10

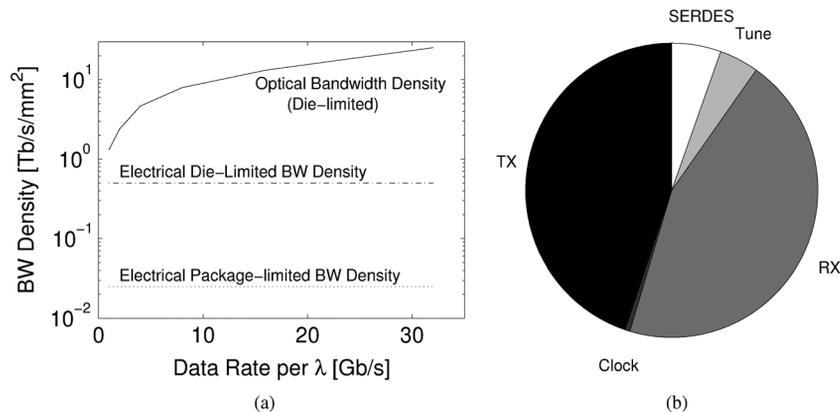


Fig. 3. WDM link area and bandwidth density for a total throughput of 256 Gb/s and a PD capacitance of 5 fF. (a) Link bandwidth density. (b) WDM link area breakdown for a channel rate of 4 Gb/s.

In addition to energy-cost, bandwidth density is an important metric for off-chip I/O. Consider the electrical I/O at the die level. An aggressive 100 μ m C4 bump pitch results in 100 bumps/mm², of which optimistically only half will be available for I/O (the other half being used for power and ground). 25 differential links operating at 20 Gb/s each yields a bandwidth density of 500 Gb/s/mm². At the package level, we assume a very ambitious pin count of up to 8000, resulting in 2000 differential I/Os again at 20 Gb/s. A 40 mm by 40 mm socket then achieves a bandwidth density of only 25 Gb/s/mm². On the other hand, at the package-level, photonic interconnects offer bandwidth densities of around 1 Tb/s per fiber, where fibers can be positioned at a roughly 100 μ m pitch to the chip. Fig. 3(a) shows the component-area-limited die-level bandwidth density achieved in the 256 Gb/s WDM link. While not as high as the 100 Tb/s/mm² promised by optical fiber pitch density, at the energy-optimal channel rate of 4 Gb/s, the die-limited bandwidth density is still more than two orders of magnitude better than the electrical limitation.

Fig. 3(b) shows the WDM link component area breakdown. It is clear that at these moderate channel rates where the number of parallel channels is relatively high, the TX and RX block areas grow to dominate the total. The design of the receiver must be mindful of this area limitation, avoiding the use of area-hungry components such as inductors.

A. Optical Data Receivers

In contrast to large PD parasitic capacitances that cause traditional optical receivers to utilize various power-hungry TIA

topologies, monolithic integration offers low PD parasitic capacitances. WDM provides a forwarded clock, enabling the use of a clocked comparator and avoiding the need for clock-and-data-recovery circuits after expensive limiting-amplifier stages [10], [11]. We build on the insight from [3] that an integrating receiver with a clocked comparator and reset will be more sensitive than a TIA.

Here, we explore this further by illustrating the relationship between sensitivity and power consumption across ranges of data rates and parasitic capacitances for various receiver topologies, including TIA and integrating receivers. The analysis presented in the remainder of this section closely follows [9], summarizing the key results.

The PD is modeled as a current source in parallel with a capacitance and a series resistance (Fig. 4(a)). The channel, consisting of the wiring between the PD and the frontend, is modeled as a series resistance and parallel capacitance. The value of the capacitance depends on the integration scenario: e.g., 5 fF for monolithic integration and 25 fF for hybrid integration with a Through-Silicon Via (TSV). Series resistances are assumed to be negligible.

A sense amplifier (SA) is used as a comparator to regenerate the full-swing digital signal. The main factors affecting SA sensitivity are mismatch, settling time, supply noise, and circuit noise. The minimum input signal that allows the latch's decision nodes to settle to the rails is $v_{\text{sense}} = V_{DD}e^{-T_{\text{bit}}/(2\tau)}$, where τ is the time constant of the exponential regeneration. Residue offset due to mismatch is compensated by a 5-bit DAC [12] resulting in $v_{\text{OS,res}} = 3v_{\text{OS}}/2^5$, with v_{OS} of 40 mV taken from [13]. Deterministic ($v_{\text{supply,det}}$) and

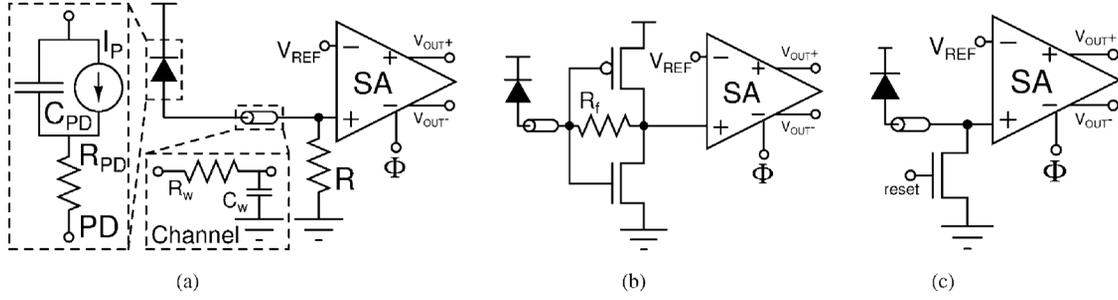


Fig. 4. Optical receiver front-end topologies. (a) Resistor. (b) TIA. (c) Integrator.

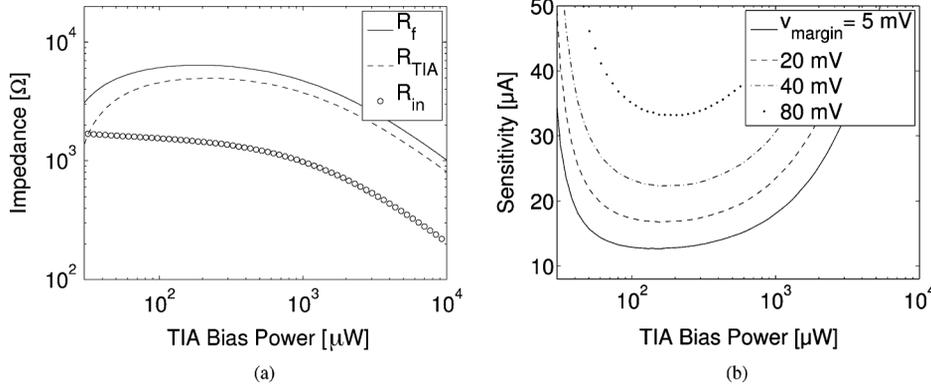


Fig. 5. TIA design example at $C_p = 25$ fF, DR = 5 Gb/s. Model computed for bandwidth >0.7 DR. (a) $v_{\text{margin}} = 20$ mV. (b) Effect of v_{margin} .

random ($v_{\text{supply,rand}}$) supply noise estimates were taken from [14], [15]. The PD's noise is shot-noise dominant [2], and is given by $\sigma_{i,\text{PD}} = \sqrt{2qI_{\text{PD}}\Delta f}$. v_{margin} accounts for any other un-modeled noise or non-idealities, and was set to 20 mV.

By input-referring the SA's input swing requirements across each of the front-ends considered through the front-end's transimpedance, R_{FE} , (1) describes the receiver's input current-swing requirement in terms of input photocurrent [9], [16].

$$\Delta I = \underbrace{\frac{i_{\text{sense}}}{v_{\text{sense}}/R_{\text{FE}}}}_{\text{SA}} + \underbrace{\frac{i_{\text{OS,res}}}{v_{\text{OS,res}}/R_{\text{FE}}}}_{\text{SA}} + \frac{v_{\text{supply,det}}}{\text{CMRR} \cdot R_{\text{FE}}} + \underbrace{\frac{i_{\text{margin}}}{v_{\text{margin}}/R_{\text{FE}}}}_{\text{SA}} + \sqrt{\text{SNR}}\sigma_n \quad (1)$$

where

$$\sigma_n = \sqrt{\sigma_{i,\text{circuit}}^2 + \frac{v_{\text{supply,rand}}^2}{\text{CMRR}^2 \cdot R_{\text{FE}}^2} + \sigma_{i,\text{PD}}^2} \quad (2)$$

The input sensitivity of the receiver can then be computed as $I_{\text{ON}} = \Delta I / (1 - 10^{-\text{ER}/10})$, where ER is the extinction ratio of the modulator and $\Delta I = I_{\text{ON}} - I_{\text{OFF}}$ is the difference in photocurrents required to meet a given BER requirement.

1) *Resistive Receiver*: The simplest receiver is the resistor, across which the PD's photocurrent is converted to a voltage that can be detected by a SA (Fig. 4(a)). The transimpedance gain and input resistance, R_{in} , are both equal to the resistance, R . With the dominant pole at the input node, the bandwidth is given by (3). Once the data-rate is determined, the largest resistance that satisfies the bandwidth constraint should be used

to maximize sensitivity. Note that the resistor is penalized for its size by including parasitic capacitances through the parameter k_R [17].

$$\text{BW} = \frac{1}{2\pi R_{\text{in}}(C_p + k_R R)} \quad (3)$$

2) *TIA*: Due to the large input parasitics associated with discrete optical receivers, a TIA (Fig. 4(b)) can be implemented to reduce the impedance at the dominant input-node-pole, while keeping the transimpedance gain high [2].

$$R_{\text{TIA}} = \frac{g_m - g_f}{g_f(g_m + g_{ds})} \quad (4)$$

$$R_{\text{in}} = \frac{g_{ds} + g_f}{g_f(g_m + g_{ds})} \quad (5)$$

Equations (4) and (5) compute the transimpedance gain and input resistance, respectively, where g_m and g_{ds} are the total transconductances of the frontend's NMOS and PMOS transistors and $g_f = 1/R_f$ is the transconductance of the feedback resistor. Equation (3) can be used to compute the bandwidth, with $R = R_f$. The impedances are plotted as a function of TIA bias power (Fig. 5(a)). Both the resistive receiver and TIA bandwidths are limited by their input resistance. While the resistive receiver's gain is equal to this input resistance, Fig. 5(a) shows that the TIA is able to achieve a transimpedance gain larger than its input resistance, and therefore has superior gain for the same bandwidth. The resulting sensitivity and the impact of the additional noise term, v_{margin} , is shown in Fig. 5(b).

A drawback to the TIA is the static-current biasing, which hurts the link energy-efficiency. The TIA is also fundamentally not as sensitive as a well-designed integrate/reset receiver.

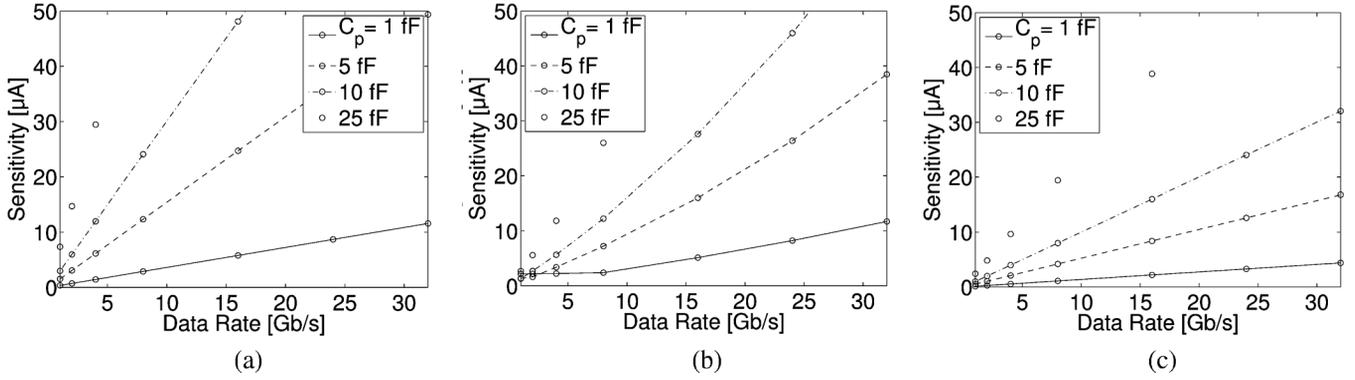


Fig. 6. Receiver sensitivity comparison. (a) Resistor. (b) TIA. (c) Integrator.

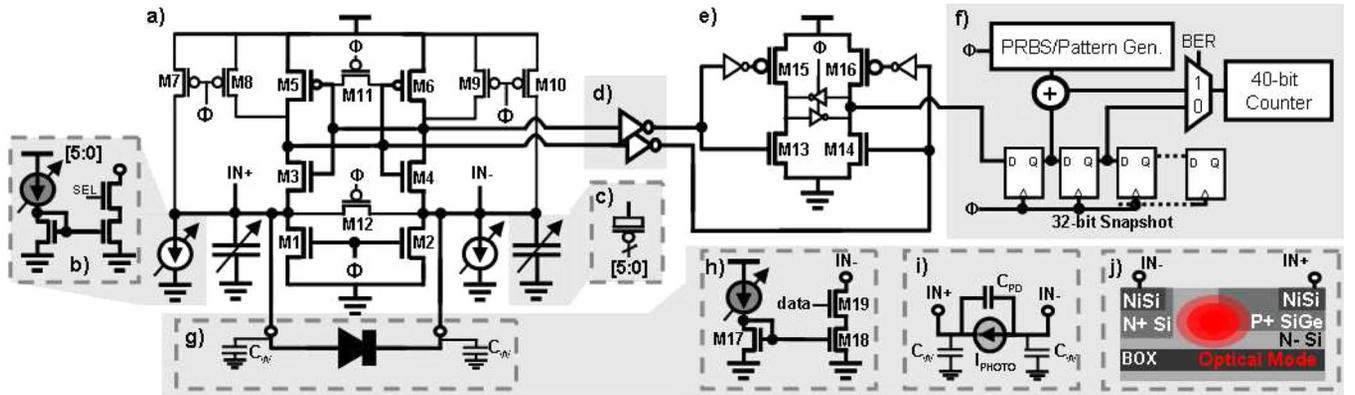


Fig. 7. Optical data receiver architecture. The LSA (a) is followed by an output buffer stage (d) and dynamic-to-static converter (e), before being fed into the digital backend infrastructure (f). The chip has receivers connected to integrated PDs (g) or electrical diode-emulation circuits (h). The simulation model is shown in (i). A cross-section of the implemented PD and optical mode is shown in (j).

3) *Integrating Receiver*: The third topology considered is an integrating receiver (Fig. 4(c)) that takes advantage of decreasing PD capacitances and the presence of an RX clock. The photocurrent is integrated onto the capacitance at the input node, $C_{INT} = C_{PD} + C_w + C_{in,circuit}$, over a fraction ($k_{INT} = 0.7$) of a bit time yielding the front-end gain given by (6).

$$R_{INT} = \frac{k_{INT} \cdot T_{bit}}{C_{INT}} \quad (6)$$

Fig. 6 shows a comparison of the receiver sensitivity performance. From the plots, it is clear that the integrating receiver is more sensitive in both C_p cases. Furthermore, its power consumption will be dominated by the SA and so will be considerably lower than that of the TIA while still scaling with frequency due to the digital design of the SA.

It should be noted that this simple model of the integrating receiver has several hidden challenges remaining. The voltage on C_{INT} must be reset or at least charge-shared [3], which is partially accounted for through k_{INT} . A small C_{INT} will also suffer from comparator kickback, while increasing C_{INT} degrades sensitivity.

III. RECEIVER ARCHITECTURE

Building on insight from the previous section, we develop a receiver architecture that employs an integrate/evaluate/reset scheme. We take advantage of the availability of a receiver-

side clock to implement an energy-efficient regenerative comparator. In order to avoid the charge-sharing and comparator-kickback issues in this low PD capacitance environment [3], we leverage the low parasitics offered by monolithic integration to build the integrating front-end directly into the regenerative latch.

The receiver architecture (Fig. 7) consists of a PD connected differentially across a latching sense-amplifier (LSA), followed by a dynamic-to-static (DS) converter and an on-chip high-speed digital testing backend. The receiver operates in two clock phases, receiving one bit per clock period.

In the PD (Fig. 7(g), (j)) we make use of P + SiGe, which is integrated in the SOI process for PMOS strain engineering and is suitable for optical absorption in the near-IR range [18]. The PD is extremely compact and has an estimated capacitance of 10 fF. Since the PD is not transit-time limited, increasing the reverse bias does not increase the speed of the device, but will increase the dark current.

The LSA (Fig. 7(a)) senses the differential photocurrent and makes a bit decision. During the reset phase ($\Phi = 0$), the LSA's nodes pre-charge high. During the decision phase ($\Phi = 1$), the two branches, $M_{1,3,5}$ and $M_{2,4,6}$, discharge. If an *optical-1* is received, photocurrent flows from node IN- to IN+, slowing the discharge of branch $M_{1,3,5}$ and causing it to latch *high*. Otherwise, imbalance programmed through offset compensation causes branch $M_{1,3,5}$ to latch *low*. Without the programmed

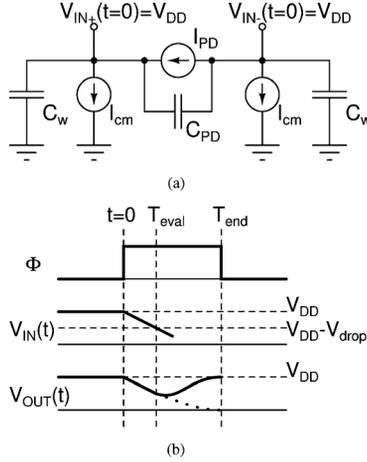


Fig. 8. LSA sensitivity model. (a) LSA sensitivity circuit model. (b) LSA sensitivity waveforms.

imbalance, an *optical-0* will cause the branches to discharge at the same rate, resulting in a random bit decision.

The LSA transistors are carefully sized according to [19] in order to adjust the sampling aperture (time resolution) of the receiver. In particular, transistors $M_{3,4}$ are sized large relative to $M_{5,6}$. This lowers the trip-point voltage of the cross-coupled inverters and ensures that they do not activate too early, which would increase the noise bandwidth of the LSA. Offset compensation is implemented as programmable current-steering (Fig. 7(b)) and capacitive (Fig. 7(c)) DACs [12], for coarse- and fine-compensation, respectively.

Fig. 7(h) shows a diode-emulation circuit that is used to characterize the receiver's performance when decoupled from the optical devices. When the input *data* is 1, the circuit pulls current from $IN-$, emulating the photocurrent sourced from that node. A 0-bit sources no current. The diode-emulation circuit is driven by a pattern generator on a separate, programmable clock phase from the rest of the receiver.

The output of the LSA is buffered (Fig. 7(d)) to isolate the LSA decision nodes from the data-dependent capacitance looking into the DS (Fig. 7(e)).

The bits stored in the DS are fed into the on-chip digital test backend for *in situ* processing (Fig. 7). The backend, consisting of synthesized PRBS and pattern generators, snapshots, and counters, gathers bit-error-rate and receiver decision threshold data and exports only the collected statistics off-chip.

A. Effect of Capacitance on Receiver Settling

To provide qualitative analysis of the impact of parasitic capacitances and operation frequency on the receiver decision settling time, an equivalent model of the LSA is shown in Fig. 8.

Fig. 8(a) shows the input nodes at the end of LSA reset ($t = 0$), pre-charged *high*. I_{cm} models $M_{1,2}$ pulling down on the input nodes until cross-coupled inverters M_{3-6} turn on. C_w represents the wiring capacitance from the PD to the receiver. The model divides the decision phase into two steps: integration, and evaluation (Fig. 8(b)). During the integration phase of duration T_{eval} (7), the photocurrent is integrated across $C_{int} = C_{PD} +$

$C_w/2$, resulting in a voltage difference, $V_{diff} = V_{IN+} - V_{IN-}$, at the onset of evaluation (9).

$$T_{eval} = \frac{C_w V_{drop}}{I_{cm}} \quad (7)$$

$$V_{diff} = \frac{I_{PD} T_{eval}}{C_{int}} \quad (8)$$

$$= \frac{I_{PD} C_w V_{drop}}{C_{int} I_{cm}} \quad (9)$$

$$V_{out} = AV_{diff} e^{\frac{T_{end} - T_{eval}}{\tau}} \quad (10)$$

$$I_{PD} = \frac{V_{out} C_{int} I_{cm}}{AC_w V_{drop}} e^{-\frac{T_{end}}{\tau}} e^{C_w \frac{V_{drop}}{I_{cm} \tau}} \quad (11)$$

The output voltage of the LSA is related to the input voltage difference, V_{diff} , through a proportionality constant, A . During the evaluation phase, V_{out} regenerates exponentially until T_{end} according to (10). Rearranging the formulas, the current-sensitivity of the receiver can be expressed by (11).

Fig. 9(a) shows through noiseless extracted simulation that for high data-rates where the exponential is not completely settling, wire capacitance, C_w , delays the onset of evaluation, shortening the evaluation time and therefore demanding exponentially more input photocurrent. Sensitivity is computed based on an output voltage settling constraint. The proposed topology may suffer in scenarios where a second die provides the optical transport layer, necessitating TSV or microsoldier bumps where C_w may increase above 20 fF [20]. Fig. 9(a) shows that for C_w in this range and data rates above 4 Gbps, the sensitivity becomes prohibitively poor. As our PD was implemented on the same die as the receiver, the low-metal-layer routing between the PD and the receiver results in a small C_w , ≈ 2.5 fF, exploiting the benefits of monolithic integration. Fig. 9(b) shows that PD capacitance, C_{PD} , reduces V_{diff} linearly, demanding only proportionally more photocurrent (Fig. 9(b)).

B. Receiver Sensitivity

In addition to an output voltage settling-time constraint, it is critical to evaluate the impact of noise and mismatch on the sensitivity of the receiver. We compute the minimum input current signal from a BER requirement as we did in Section II-A.

The minimum input signal required for the exponential to evaluate to the rails is given by $i_{sense} = V_{DD} G e^{-(T_{bit} - T_{eval})/\tau}$. The time constant of the exponential term, τ , and conductance, G , are measured in simulation. As in our equivalent model, as the end of the bit time starts to approach T_{eval} , the receiver's sensitivity degrades exponentially.

Mismatch in the differential latching receiver also leads to a threshold offset. In order to avoid large latch sizing, which results in increased power, we employ offset compensation circuitry in the form of a 6-bit capacitive DAC. The threshold offset is measured through Monte Carlo simulation. The residual threshold is then given by $i_{OS,res} = i_{OS}/2^6$, found to be negligably small.

The circuit noise was computed in a transient noise simulation by sweeping the receiver's input photocurrent threshold and recording the decision statistics. The resulting input-referred noise cumulative distribution function as a function of the LSA's

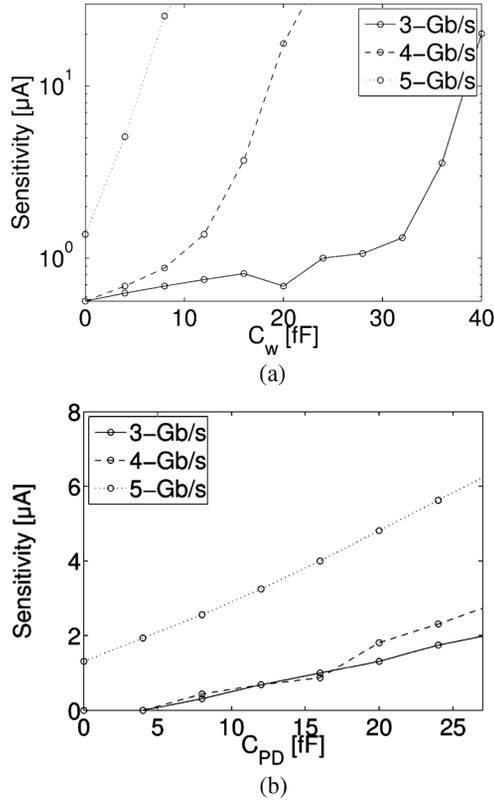


Fig. 9. LSA sensitivity model. (a) Effect of C_w on receiver sensitivity. $C_{PD} = 10$ fF. (b) Effect of C_{PD} on receiver sensitivity. $C_w = 2.5$ fF.

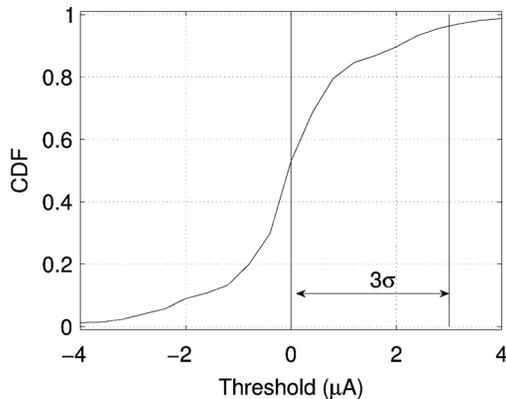


Fig. 10. Receiver circuit noise CDF.

decision threshold is shown in Fig. 10, depicting a standard deviation of $\approx 1 \mu\text{A}$.

PD shot noise was computed using the receiver's sampling bandwidth [21]. Supply noise sources were ignored and no additional margin was added.

Fig. 11 shows the receiver's predicted sensitivity as a function of data-rate for $C_{PD} = 10$ fF and $C_w = 5$ fF. We can see that as the data-rate increases, T_{end} in our model decreases, demanding exponentially more input photocurrent starting around 5 Gb/s.

C. Photodiode Splitting

Understanding the impact of the settling time and noise, it is possible to further optimize the receiver, leveraging the mono-

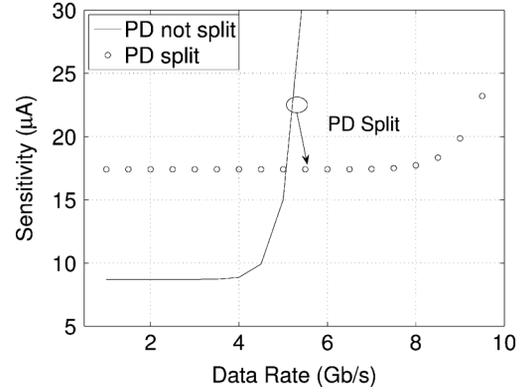


Fig. 11. Receiver sensitivity and PD split. $C_{PD} = 10$ fF, $C_w = 5$ fF.

lithic integration once again for a closer interaction between the PD and the receiver circuit.

The limiting sensitivity factor at higher data rates is the settling-time term (i_{sense} in (1)). By operating the receiver at half the rate, we can double the value of T_{end} , giving the exponential regeneration phase much more time to settle. In order to keep the data-rate on the channel the same, we need two receivers and a DEMUX. Since monolithic integration affords us a high degree of control over the design of the PD, we can simply interdigitate metal contacts to break it into two separate PDs, each connected to a separate receiver. While one receiver is integrating and evaluating the input signal, the other is resetting. As a result of the photodiode splitting, only half of the total photocurrent will go to each receiver, requiring 2X the laser power, but this is still better than a higher exponential factor.

Fig. 11 plots the sensitivity of the receiver for both PD split and unsplit cases. By splitting the photodiode and doubling T_{end} , the exponential term begins to dominate only at much higher data rates. A factor of 2 is applied to the total sensitivity computed in (1), reflecting that each receiver only gets half of the photocurrent. Careful partitioning of the PD fingers ensures that each receiver's PD gets a roughly equivalent share of the incident optical power.

This diode-splitting enables double-data-rate (DDR) receivers which are very useful in parallel source-forwarded links, where a data-pattern of 101010 on one of the transmitted wavelength-channels can be used as a receive clock and directly applied to all of the DDR receivers.

IV. MEASURED RESULTS

The monolithically-integrated data receiver was fabricated in a standard 45-nm SOI process, as a part of a flexible electronic-photonics test vehicle.

Fig. 12 shows different PD designs implemented on the test chip. In Fig. 12(a) the PD is implemented as an absorption-type detector where the optical power is absorbed along the length of the device. In order to maximize the photocurrent, the device must be relatively long. As a result, the PD capacitance is increased, reducing the receiver's sensitivity. Fig. 12(b) shows a second option where the detector is integrated into a resonant ring. When the ring is tuned to the particular wavelength channel of the incoming signal, the light becomes confined in the ring. During each round-trip, part of the light is absorbed, allowing

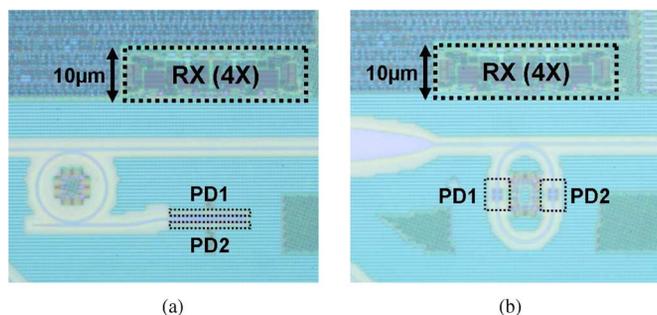


Fig. 12. Die photo showing photodetector layout options. (a) Ring with absorption detector. (b) Resonant-ring detector.

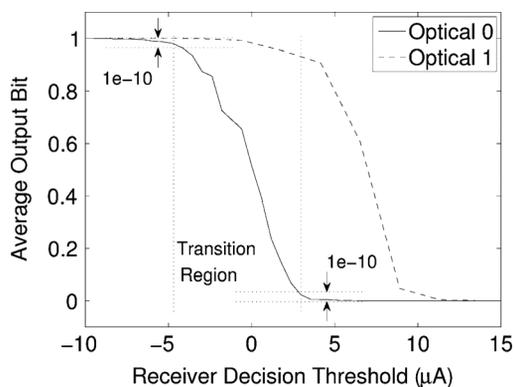


Fig. 13. The measurement shows the receiver's ability to distinguish between a DC *optical-1* and *optical-0*.

for a smaller PD length and therefore less capacitance and better sensitivity. The figure also shows how two PDs can be implemented in the same wavelength-channel, enabling the PD split described.

Fig. 13 shows two DC photocurrents generated by a 1310-nm wavelength laser, coupled into the chip through a vertical coupler and horizontal waveguide made with front-end body Si. The receiver's threshold is swept using the offset circuitry (Fig. 7(b), (c)) while recording the output decision statistics. Photocurrent values were de-embedded from DAC settings through simulation. Though the receiver was able to detect photocurrent from the PD, a foundry error in the SiGe mask definition limited the achievable PD bandwidth.

Fig. 14(a) shows the receiver's sensitivity vs. data rate for different supply voltages. Sensitivity is measured on a PD-connected receiver (Fig. 7(g),(j)) as the width of the transition region (Fig. 13) of an optical-0. As clock frequency increases, sensitivity degrades exponentially as predicted by our model due to the decrease in T_{end} . The sensitivity degrades with reduced supply voltage, as I_{CM} in (11) will decrease with reduced V_{DD} , increasing the length of the integration phase, but decreasing the length of the exponential evaluation phase.

Fig. 14(b) shows the energy-cost of the receiver. The linearity emphasizes the digital design, with power following $P_{\text{digital}} = fCV_{DD}^2$, keeping the receiver energy-cost ≈ 50 fJ/b across a range of frequencies. The power breakdown is shown in Fig. 15. While the latch power is dominant, experimental infrastructure such as capacitive DACs used for link analysis and eye-diagram measurement are shown to be equally expensive. This cost can

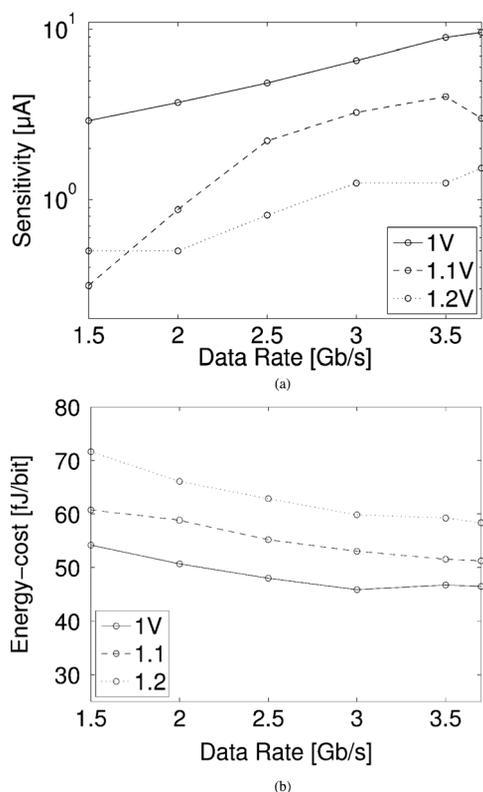


Fig. 14. Measured data receiver performance. (a) Sensitivity. (b) Energy cost.

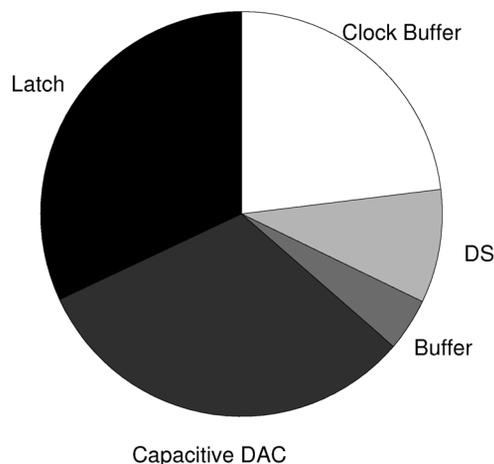


Fig. 15. Receiver power breakdown (de-embedded from simulation).

be reduced in the future. The receiver's current-DAC was not needed and was shut off, consuming no power.

Fig. 16 shows the bit-error-rate eye diagram of the receiver when configured with a PD-emulation circuit (Fig. 7(h)). Clock phase and receiver threshold were swept for a 31-bit PRBS data sequence at 3.5 Gb/s and a supply of 1.1 V, and error statistics were gathered *in situ* using the digital backend. Clock rates above 3.7 GHz caused timing violations in the digital testing backend.

A die photo is shown in Fig. 17. The chip contains 72 test cells that implement combinations of optical modulators and receivers. Each receiver has a circuit area of 108 μm² and PD area of 416 μm².

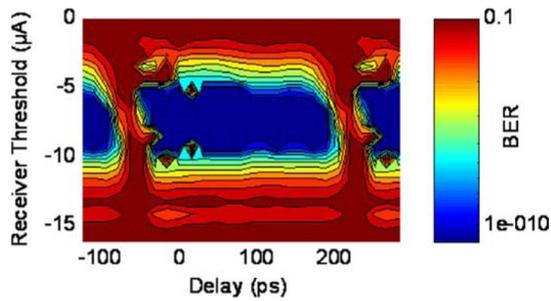
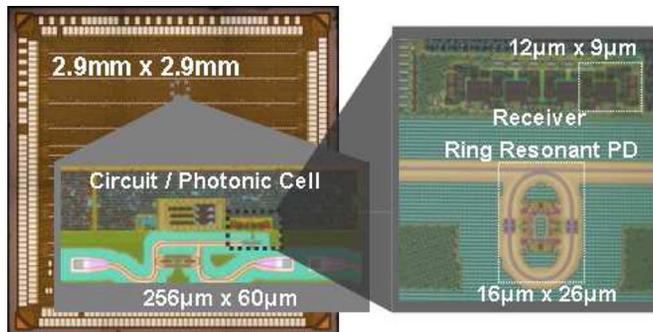

 Fig. 16. Electrical *in situ* eye-diagram at 3.5 Gb/s.


Fig. 17. Die, backside photonic cell and receiver photos.

V. CONCLUSION

Integrated photonics has emerged as an I/O technology that can meet the throughput demands of future many-core processors. In this work, the monolithic integration of the photodetector enables the design of a fully-digital, low-energy receiver with high input sensitivity. WDM provides a forwarded clock that enables the implementation of an energy-efficient clocked comparator. The integrating receiver front-end is built into the comparator, taking advantage of the low PD and wiring capacitances. The developed receiver model provides intuition for the impact of different PD integration scenarios on the receiver's sensitivity performance. This insight led to the development of a PD-splitting technique that enables operation at higher data rates.

The receiver is shown to operate with μA -sensitivity at 3.5 Gb/s with an energy-efficiency of 52 fJ/b. This work demonstrates the first monolithic electronic-photonics receiver integration in a sub-100-nm standard SOI process.

ACKNOWLEDGMENT

The authors would like to thank the Integrated Photonics teams at MIT and UC Boulder.

REFERENCES

- [1] C. Batten, A. Joshi, J. Orcutt, C. Holzwarth, M. Popovic, J. Hoyt, F. Kartner, R. Ram, V. Stojanovic, and K. Asanovic, "Building manycore processor-to-DRAM networks with monolithic CMOS silicon photonics," *IEEE Micro*, vol. 29, no. 4, pp. 8–21, Jul.-Aug. 2009.
- [2] C. Kromer, G. Sialm, T. Morf, M. Schmatz, F. Ellinger, D. Erni, and H. Jackel, "A low-power 20-GHz 52-dB-Ohm transimpedance amplifier in 80-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 39, no. 6, pp. 885–894, Jun. 2004.
- [3] A. Emami-Neyestanak, D. Liu, G. Keeler, N. Helman, and M. Horowitz, "A 1.6 Gb/s, 3 mW CMOS receiver for optical communication," in *Symp. VLSI Circuits Dig. Tech. Papers*, 2002, pp. 84–87.
- [4] I. Young, E. Mohammed, J. Liao, A. Kern, S. Palermo, B. Block, M. Reshotko, and P. Chang, "Optical I/O technology for tera-scale computing," *IEEE J. Solid-State Circuits*, vol. 45, no. 1, pp. 235–248, Jan. 2010.
- [5] A. Rylyakov, C. Schow, B. Lee, W. Green, J. Van Campenhout, M. Yang, F. Doany, S. Assefa, C. Jahnes, J. Kash, and Y. Vlasov, "A 3.9 ns 8.9 mw 4×4 silicon photonic switch hybrid integrated with CMOS driver," in *IEEE ISSCC Dig.*, 2011, pp. 222–224.
- [6] G. Li *et al.*, "Ultralow-power silicon photonic interconnect for high-performance computing systems," in *Proc. SPIE Optoelectronic Interconnects and Component Integration IX*, 2010, vol. 7607, pp. 703–760.
- [7] F. Liu, D. Patil, J. Lexau, P. Amberg, M. Dayringer, J. Gainsley, H. Moghadam, X. Zheng, J. Cunningham, A. Krishnamoorthy, E. Alon, and R. Ho, "10 Gbps, 530 fJ/b optical transceiver circuits in 40 nm CMOS," in *Symp. VLSI Circuits Dig.*, 2011, pp. 290–291.
- [8] K. Fukuda, H. Yamashita, G. Ono, R. Nemoto, E. Suzuki, N. Masuda, T. Takemoto, F. Yuki, and T. Saito, "A 12.3-mW 12.5-Gb/s complete transceiver in 65-nm CMOS process," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2838–2849, Dec. 2010.
- [9] M. Georgas, J. Leu, B. Moss, C. Sun, and V. Stojanovic, "Addressing link-level design tradeoffs for integrated photonic interconnects," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, 2011, pp. 1–8.
- [10] J. S. Yun, M. Seo, B. Choi, J. Han, Y. Eo, and S. M. Park, "A 4 Gb/s current-mode optical transceiver in 0.18 μm CMOS," in *IEEE ISSCC Dig.*, 2009, pp. 102–103.
- [11] A. Narasimha, B. Analui, Y. Liang, T. Sleboda, and C. Gunn, "A fully integrated 4×10 Gb/s DWDM optoelectronic transceiver in a standard 0.13 μm CMOS SOI," in *IEEE ISSCC Dig.*, 2007, pp. 42–586.
- [12] K.-L. Wong and C.-K. Yang, "Offset compensation in comparators with minimum input-referred supply noise," *IEEE J. Solid-State Circuits*, vol. 39, no. 5, pp. 837–840, May 2004.
- [13] S.-H. Woo, H. Kang, K. Park, and S.-O. Jung, "Offset voltage estimation model for latch-type sense amplifiers," *IET Circuits, Devices Syst.*, vol. 4, no. 6, pp. 503–513, 2010.
- [14] E. Alon, V. Stojanovic, and M. Horowitz, "Circuits and techniques for high-resolution measurement of on-chip power supply noise," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 820–828, Apr. 2005.
- [15] S. Naffziger, B. Stackhouse, T. Grutkowski, D. Josephson, J. Desai, E. Alon, and M. Horowitz, "The implementation of a 2-core, multi-threaded Itanium family processor," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 197–209, Jan. 2006.
- [16] S. Palermo, A. Emami-Neyestanak, and M. Horowitz, "A 90 nm CMOS 16 Gb/s transceiver for optical interconnects," *IEEE J. Solid-State Circuits*, vol. 43, no. 5, pp. 1235–1246, May 2008.
- [17] S. Assefa, F. Xia, W. Green, C. Schow, A. Rylyakov, and Y. Vlasov, "CMOS-integrated optical receivers for on-chip interconnects," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1376–1385, Sep.-Oct. 2010.
- [18] J. Polleux and C. Rumelhard, "Optical absorption coefficient determination and physical modelling of strained SiGe/Si photodetectors," in *8th IEEE Int. Symp. High Performance Electron Devices for Microwave and Optoelectronic Applications*, 2000, pp. 167–172.
- [19] S. Mukhopadhyay, R. Joshi, K. Kim, and C.-T. Chuang, "Variability analysis for sub-100 nm PD/SOI sense-amplifier," in *Proc. ISQED*, 2008, pp. 488–491.
- [20] J.-S. Kim, C. S. Oh, H. Lee, D. Lee, H.-R. Hwang, S. Hwang, B. Na, J. Moon, J.-G. Kim, H. Park, J.-W. Ryu, K. Park, S.-K. Kang, S.-Y. Kim, H. Kim, J.-M. Bang, H. Cho, M. Jang, C. Han, J.-B. Lee, K. Kyung, J.-S. Choi, and Y.-H. Jun, "A 1.2 V 12.8 GB/s 2 Gb mobile wide-I/O DRAM with 4×128 I/Os using TSV-based stacking," in *IEEE ISSCC Dig.*, 2011, pp. 496–498.
- [21] M. Jeeradit, J. Kim, B. Leibowitz, P. Nikaen, V. Wang, B. Garlepp, and C. Werner, "Characterizing sampling aperture of clocked comparators," in *Symp. VLSI Circuits Dig.*, 2008, pp. 68–69.
- [22] S. Goswami, J. Silver, T. Copani, W. Chen, H. Barnaby, B. Vermeire, and S. Kiaei, "A 14 mW 5 Gb/s CMOS TIA with gain-reuse regulated cascode compensation for parallel optical interconnects," in *IEEE ISSCC Dig.*, 2009, pp. 100–101, 101a.



Michael Georgas (S'07) is currently working towards the Ph.D. degree in electrical engineering and computer science at Massachusetts Institute of Technology (MIT), Cambridge. He received the M.S. degree in electrical engineering and computer science from MIT in 2009 and the B.A.Sc. from the University of Toronto, Canada, in 2007.

His research interests relate to the circuit and system level design of high-speed interconnects, and in particular, optical interconnects.



Jason Orcutt received the Ph.D. degree in electrical engineering and computer science from Massachusetts Institute of Technology (MIT), Cambridge, in 2012, the M.S. degree in electrical engineering and computer science from MIT in 2008, and the B.S. degree in electrical engineering from Columbia University, New York, in 2005.

He is a research scientist with the Research Laboratory of Electronics at MIT. His research interests focus on novel uses of CMOS processing technology for diverse applications and photonic device design

and testing.



Rajeev J. Ram (S'94–M'96) received the B.S. degree in applied physics from the California Institute of Technology, Pasadena, in 1991, and the Ph.D. degree in electrical engineering from the University of California, Santa Barbara, in 1997.

In 1997, he joined the faculty of the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, and is currently a Professor of electrical engineering. He is the Associate Director of the Research Laboratory of Electronics, MIT, and the Director of the

Center for Integrated Photonics Systems. His research focuses on physical optics and electronics, including the development of novel components and systems for communications and sensing, novel semiconductor lasers for advanced fiber optic communications, and studies of fundamental interactions between electronic materials and light.



Vladimir Stojanović (M'04) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2005, and the Dipl.Ing. degree from the University of Belgrade, Serbia, in 1998.

He is the Emanuel E. Landsman Associate Professor of Electrical Engineering and Computer Science at Massachusetts Institute of Technology (MIT), Cambridge. His research interests include design, modeling and optimization of integrated systems, from CMOS-based VLSI blocks and interfaces to system design with emerging devices like NEM

relays and silicon-photonics. He is also interested in design and implementation of energy-efficient electrical and optical networks, and digital communication techniques in high-speed interfaces and high-speed mixed-signal IC design. He was with Rambus, Inc., Los Altos, CA, from 2001 through 2004.

Dr. Stojanović received the 2006 IBM Faculty Partnership Award, and the 2009 NSF CAREER Award as well as the 2008 ICCAD William J. McCalla, 2008 IEEE Transactions on Advanced Packaging, and 2010 ISSCC Jack Raper best paper awards.