Scaling Optoelectronic-VLSI Circuits into the 21st Century: A Technology Roadmap

Ashok V. Krishnamoorthy, *Member, IEEE,* and David A. B. Miller, *Fellow, IEEE*

***Abstract*— Technologies now exist for implementing dense surface-normal optical interconnections for silicon CMOS VLSI using hybrid integration techniques. The critical factors in determining the performance of the resulting photonic chip are the yield on the transceiver device arrays, the sensitivity and power dissipation of the receiver and transmitter circuits, and the total optical power budget available. The use of GaAs–AlGaAs multiple-quantum-well p-i-n diodes for on-chip detection and modulation is one effective means of implementing the optoelectronic transceivers. We discuss a potential roadmap for the scaling of this hybrid optoelectronic VLSI technology as CMOS linewidths shrink and the characteristics of the hybrid optoelectronic tranceiver technology improve. An important general conclusion is that, unlike electrical interconnects, such dense optical interconnections directly to an electronic circuit will likely be able to scale in capacity to match the improved performance of future CMOS technology.**

NOMENCLATURE

$V_{\rm dd}$	Supply voltage
$V_{\rm ds}$	FET drain-to-source voltage
$g_{m,N\rm MOS}$	Transconductance of NMOSFET
$g_{m,P\rm MOS}$	Transconductance of PMOSFET
$g_{\rm ds,}N\rm MOS$	Output conductance of NMOSFET
$g_{\rm ds,}P\rm MOS$	Output conductance of NMOSFET
g_m	Transconductance of front-end amplifier
$C_{\rm gs}, C_{\rm gd}, C_{\rm gb}, C_{\rm ds}$	Gate-to-source, gate-to-drain, gate-to-bulk, and drain-to-source capacitance of FET
β	MOS transistor gain factor
$V_{\rm gs}$	FET gate-to-source voltage
V_t	FET threshold voltage
B	System clock-speed (bit rate per optical channel)
i_n	RMS noise current of CMOS receiver front end
f_T	Unity-gain frequency of MOS transistor
A	Open-loop gain of CMOS inverter stage
R_f	Impedance of transimpedance feedback element
I_d	Peak input photocurrent
$I_{\rm rec}$	Current in biased transimpedance front-end stage
$n_{\rm int}$	Total optical interconnect system efficiency (optical loss factor)
$n_{\rm optics}$	Optical system link efficiency
$n_{\rm mod}$	MQW modulator efficiency
C_T	Total capacitance of receiver front end
C_g	Input capacitance of MOSFET
W	FET gate width
L	FET gate length
$t_{\rm ox}$	Gate oxide thickness
λ	FET technology scaling factor
$I_{\rm ds}$	FET drain current
$C_{\rm diode}$	MQW diode capacitance
$C_{\rm bump+pad}$	Sum of solder-bump and flip-chip pad capacitances
C_D	Detector capacitance
A_c	MQW diode capacitance per unit area
d	Linear dimension of flip-chip pad
$t_{\rm bump}$	Sum of solder-bump and flip-chip pad thicknesses
t_d	Thickness of oxide between flip-chip pad and ground plane
ε	Permittivity of dielectric oxide
μ	Permiability of dielectric oxide
$\Delta T_{\rm diode}$	Diode temperature differential
$\Delta P_{\rm diode}$	Diode power dissipation differential
$\Delta I_{\rm diode}$	Diode absorbed current differential
ΔR	Diode reflectivity change
f_T'	Unity voltage gain bandwidth of identical cascaded inverters
V_1	Input voltage of transimpedance front end
V_2	Output voltage of transimpedance front end
$f_{-3\,\rm dB}$	3-dB frequency of transimpedance front end
Z_T	Low-frequency impedance of transimpedance front end
ΔL	Overlap of gate over the source and drain regions
B_0	Cut-off bit rate of classical receiver model
q	Electron charge constant
I_l	Sum of photodiode and FET leakage currents
I_2, I_3, I_f	Normalized Personick integrals
Γ	FET excess channel noise factor
f_c	$1/f$ noise corner frequency
$C_{\rm in}$	Input capacitance of receiver front end

Manuscript received March 15, 1996; revised August 13, 1996.

The authors are with Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733 USA.

Publisher Item Identifier S 1077-260X(96)08126-9.

λ'	FET channel length modulation parameter
η	Detector responsivity
C_{out}	Output capacitance of transimpedance front end
C_f	Feedback capacitance of transimpedance front end
C_{trans}	Total transmitter capacitance including diodes and driver
R_O	Output resistance of transimpedance front end
C_{Load}	Load capacitance of amplifier stage following front end
$I_{l,\text{Fet}}$	FET subthreshold leakage current
$I_{d,\text{sat}}$	FET drain-current at onset of saturation
S	Subthreshold conduction slope
ΔT	Duration of bit
ΔV	Input voltage swing at receiver front end
P_{avg}	Average optical power in received input
N_b	Number of transmitted bits
N_p	Number of light pulses
V_{mod}	MQW bias voltage
$P_{\text{Hi/In}}$	Normalized modulator power efficiency of high-reflectivity state
$P_{\text{Low/In}}$	Normalized modulator power efficiency of low-reflectivity state
$P_{\text{diss,rec}}$	Power dissipation of receiver circuit
$P_{\text{diss,trans}}$	Power dissipation of transmitter circuit
R_{diode}	Thermal resistance of NQW diode
E_{opt}	Optical energy-per-bit at receiver
T	Operating temperature

I. Introduction

THE tremendous progress in high-performance very-large-scale integrated circuit (VLSI) technology has made possible the incorporation of several million transistors onto a single silicon chip with on-chip clock rates of over 400 MHz. By 2001, the integration density for silicon complementary metal oxide semiconductor (CMOS) field-effect transistor (FET) logic is expected to be up to 13 million transistors and the projected on-chip clock rate to be 600 MHz [1]. Recent estimates made by the Semiconductor Industry Association indicate that the number of transistors available for logic chips and memory chips will respectively double and quadruple every three years. Two factors drive this trend: the shrinking feature size of silicon VLSI, resulting in a higher density of gates per unit area, and the improving yield of integrated circuits, resulting in more silicon real estate per chip. The enormous bandwidth that will be available for computation and switching on a silicon integrated circuit will create an increasing demand for high-bandwidth input and output (I/O) to a VLSI circuit. Novel interconnect technologies will be needed to meet this challenge.

One possible solution is the use of three-dimensional (3-D) optical interconnect technologies via surface-normal optical transmitters and receivers. Hybrid integration of several light-transmitter technologies has been investigated for parallel optical interconnections to silicon. These include electrooptic polarization modulators such as ferro-electric liquid crystal (FLC) and lead-lanthanum-zirconium-titanate (PLZT) modulators [2]–[5], indium-phosphide light emitting diodes [6], surface-emitting lasers [7]–[10], and gallium-arsenide (GaAs) multiple-quantum-well (MQW) electroabsorption modulators [11]–[16].

The integration of GaAs–AlGaAs p-i(MQW)-n diodes to CMOS circuits provides the ability to both transmit and receive data optically. This can be accomplished by flip-chip bonding photodetectors and light-modulators onto a prefabricated silicon integrated circuit containing receiver and transmitter (modulator) driver circuits. The attachment operation is followed by substrate removal of the GaAs chip, which allows the modulators to operate at a wavelength of approximately 850 nm [17]. The intimate connection between these optoelectronic MQW diodes and commodity CMOS devices represents an important step in the evolution of self-electrooptic-effect device (SEED) technology [18]. This method has been used to fabricate high-density optically interconnected submicron CMOS integrated circuits by bonding directly above active silicon gates [19], [20]. This technique effectively decouples the design of the silicon from the placement and bonding of the surface-normal optical I/O, making the technology more accessible to a system architect, and allowing high-performance silicon design tools to be used for rapid prototyping of optoelectronic circuits.

Arrays of four thousand operational optical devices can now be bonded to a single CMOS chip [21]. Individual transimpedance receiver/transmitter circuits, based on two-beam differential data encoding, have been built in 0.8-μm CMOS and are capable of $\geq$1 Gb/s transmission of digital data at a bit-error rate below 10^{-10} [22]. These results suggest that free-space optical interconnect technologies could soon provide over a terabit-per-second of optical input/output to a conventional silicon VLSI integrated circuit. Although care must be taken in comparing a research-level optical technology against a commodity electrical one, it is by no means clear that the electrical interconnect technology is practical for Tb/s or higher capacities, especially on and off a single chip.

The emergence of this integration technology and allied fiber and free-space optical technologies for steering and focusing light beams presents new opportunities and challenges to the system designer. One particular opportunity is that optical interconnections do not have the problem that their bit-rate capacity falls off rapidly with distance; electrical interconnections quite generally have a bit-rate capacity that falls as the square of the length for a given cross-sectional area [23],[1] making long, thin, high-speed electrical interconnects impractical. Optical interconnections will allow densities of information flow much larger than the few Gb/s/cm^2 typical of connectorized coaxial cable a few meters in length. In fact, the use of optics with the hybrid CMOS-MQW modulator

[1]This scaling limit arises both in high-speed (e.g., coaxial) cables and striplines limited by skin-effect losses and also in the resistive/capacitive lines found on chips.

technology described here allows large numbers of high-speed connection directly to and from a single chip to cabinets that are meters away, a concept that is not feasible electrically at high bit rates.

In this paper, we present a roadmap that will detail a possible evolution of this optoelectronic-VLSI (OE-VLSI) technology as silicon feature sizes shrink, and the dimensions of the GaAs–AlGaAs MQW diodes are reduced. The main objectives are: 1) to delineate an anticipated *design-space* for smart-pixel systems that are based on OE-VLSI circuits used in conjunction with free-space optical interconnects; and 2) to expose potential technology and circuit design challenges that will allow the boundaries of this design space to be met or even exceeded. The progress of silicon CMOS technology is expected to follow the projections of the Semiconductor Industry Association (SIA) roadmap [1]. Assumptions are made for the expected evolution of the accompanying optoelectronic MQW devices. Based on these data, and certain assumptions on the availability of sufficient laser power to drive the optoelectronic devices, the system loss budget, and ability to dissipate the power consumed on-chip, we project limits on the maximum number of optical I/O and the electrical circuit complexity per optical I/O for current and future generations of OE-VLSI chips. For projections of optical and optoelectronic device performance and yield, where there is no roadmap comparable to the SIA electronic roadmap, we have been very conservative in anticipating the capability of future optical and optoelectronic technologies. In many cases, the "predictions" of some future performance fall substantially below current research results; this is deliberate because we are interested in credible projections of performance for systems that could realistically be manufactured at reasonable cost. Section II reviews the device assumptions. Section III summarizes methodology used in this paper. Section IV discusses the receiver model assumed for this study and presents the derivation for the maximum number of optical receivers. Results of the study are presented in Section V. A summary and conclusions constitutes Section VI.

II. Technology Assumptions

The performance limits of the optoelectronic VLSI technology can be expressed in terms of: 1) the maximum number of optical I/O; 2) the bandwidth per I/O; and 3) the complexity (in terms of the number of transistors or gates) per optical I/O. We will attempt to quantify these expressions and expose tradeoffs that may occur between these quantities. In general, the performance-limiting factors typically include: 1) the available laser power; 2) the yield on the MQW diodes; 3) the system loss; 4) the receiver bandwidth; 5) the power consumed by the receivers; 6) the on-chip power-dissipation capability of the chip; 7) the chip size; and 8) the electrical clock speed. Refer to the nomenclature for the relevant symbols and expressions used in this paper.

A. Submicron CMOS VLSI Technologies

The technology parameters assumed for future generations of silicon CMOS (Table I) are based on those predicted for high-performance electronic integrated circuits (IC's) [1], [24]–[28]. Hot carrier effects, electromigration, and subthreshold leakage are all significant obstacles as CMOS feature sizes are scaled to 0.1 μm and below [28]–[31]. Among the critical parameters for transistor performance is the thickness of the gate oxide, $t_{\rm ox}$. The physical limitations of transistors that are scaled to 0.1-μm gate-lengths are not the central subject of this paper, but they have direct bearing on the analysis of the optoelectronic transceiver circuits. A popular model for CMOS scaling is based on the constant-field model [32] that assumes the transistor dimensions, supply voltage, gate-oxide thickness, gate capacitance, and substrate doping scale by a constant factor (λ), while the FET transconductance, the field across the gate oxide, and the electron and hole mobilities remain relatively constant with scaling. In practice, constant field scaling is typically not employed. The reason is due to the nonscaling of the threshold voltage of the FET's at room temperature; a scaled supply voltage does not provide adequate performance gain because the drain-source current in the FET, $I_{\rm ds}$, depends on the gate-source overdrive $V_{\rm gs} - V_t$. As a result, constant field scaling is only being pursued for the very low power technologies and may not be appropriate for the high-performance circuits being discussed in the paper. On the other hand, constant voltage scaling is also not possible due to the hot carrier effects that are evident in the deep submicron FET's that have very thin gate oxides. To make the analysis more relevant to technologies that are being developed, we use empirical data on submicron CMOS technologies that have been demonstrated in recent years. Data from a number of experimental technologies have been collected and trend lines have been fit to this data [33]–[58].

Figs. 1–5 show the empirical scaling of the gate oxide thickness, voltage, peak transconductance of the PMOS and NMOSFET's, nominal clock frequency of each technology, and the inverse gate-delay based on ring oscillator data. As mentioned above, the threshold voltage of the FET's is an important parameter that affects the power-delay product of the gates. A higher ratio of the threshold voltage V_t to the supply voltage $(V_{\rm dd})$ reduces the power dissipation of the circuits but also increases their delay. It will be assumed that the $V_t/V_{\rm dd}$ ratio is maintained at about 1/4 to ensure adequate performance improvement with each new generation [59]. The empirical data suggests that the observed scaling is closer fit to the quasi-constant voltage scaling model [60], [61], where the voltage scales as approximately $\lambda^{1/2}$ and the other parameters scale as λ. The resulting cutoff frequency, f_t, of the FET's

TABLE I
Assumed CMOS IC Technology Parameters

FET Gate-Length (μm)	# Gates (M)	Area (cm^2)	Voltage (Vdd)	Clock Speed (MHz)	Oxide Thickness (Å)	Max. Power Dissipation (W)
0.7	0.15	1	5	100	130	5
0.5	0.3	2	3.3	150	100	10
0.35	0.8	4	2.8	200	80	15
0.25	2	6	2.2	350	60	30
0.18	5	8	1.8	500	50	40
0.12	10	10	1.5	700	40	90
0.1	20	12	1.25	1000	35	180

Generations of CMOS are expected to be spaced three years apart [1].

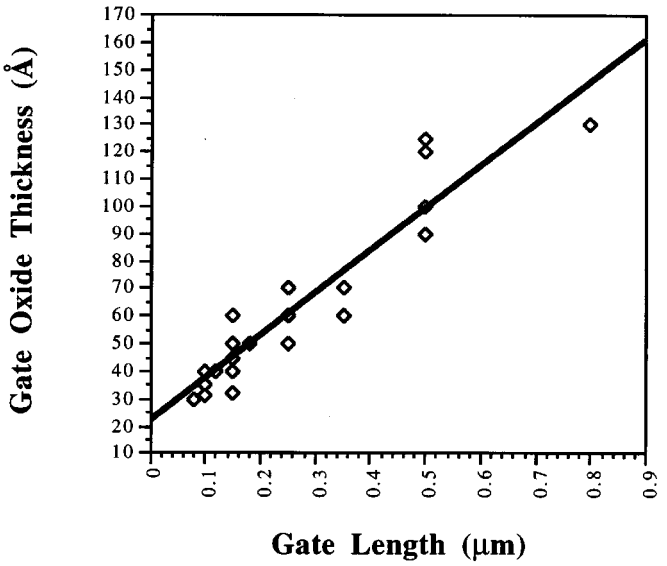


Fig. 1. Scaling trend for the gate-oxide thickness as a function of the gate length of the CMOS technology. (Data points correspond to demonstrated technologies).

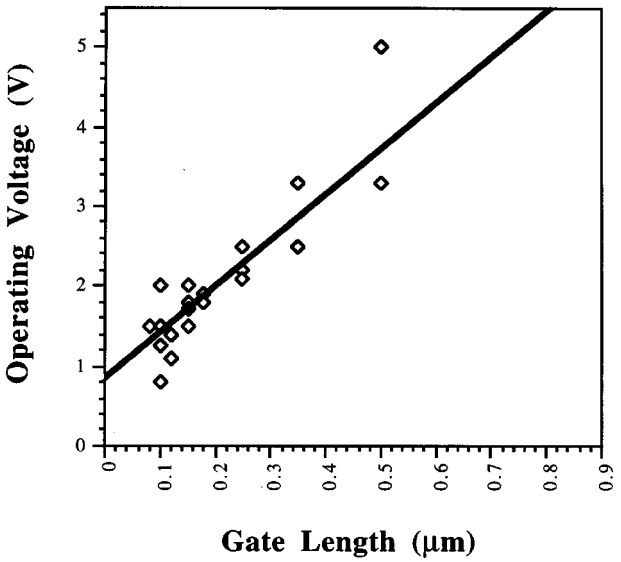
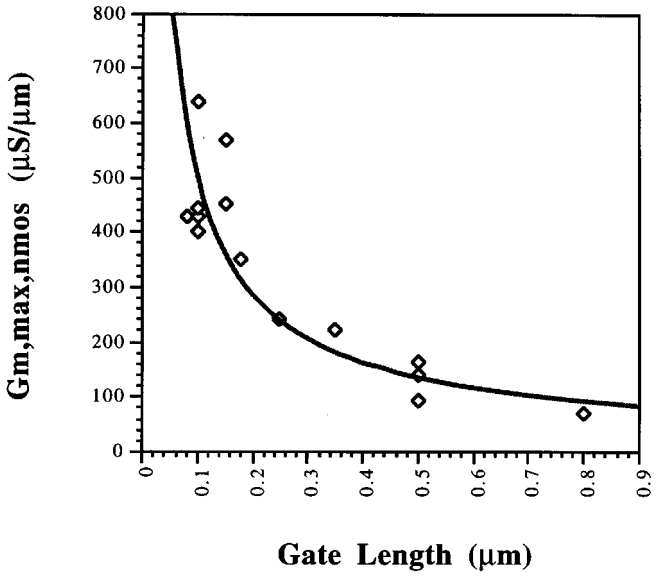
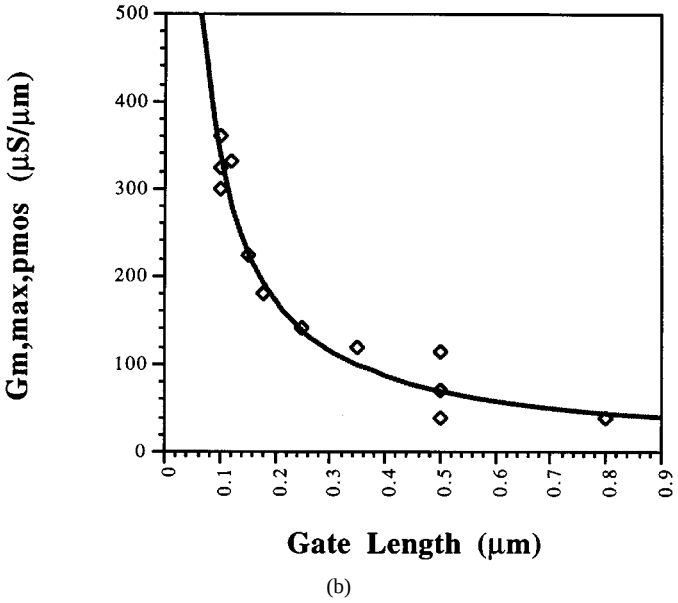


Fig. 2. Scaling trend for the operating voltage as a function of the gate length of the CMOS technology. (Data points correspond to demonstrated technologies).



(a)



(b)

Fig. 3. Scaling trend for the maximum transconductance-per-unit-length of (a) an NMOS transistor and (b) a PMOS transistor as a function of the gate length of the CMOS technology. (Data points correspond to demonstrated technologies).

increases with every generation due to the reduction in the gate capacitance.

B. Optoelectronic Technologies Based on MQW-Modulators

The basic structure of the hybrid CMOS-SEED technology is presented in Fig. 6. The details of the flip-chip bonding operation of the MQW diodes onto a fabricated CMOS IC followed by the removal of the GaAs substrate are presented elsewhere [17]. It is assumed that the placement and bonding of the MQW diodes occurs independently of the design of the CMOS IC, providing a 3-D structure with the sole constraint being the reservation of the topmost layer of metal on the CMOS circuit for flip-chip bonding pads and alignment marks [19], [20].

The anticipated technology parameters for succeeding generations of hybrid MQW modulator arrays are detailed in Table II. One of the most critical technology challenges is to increase the yield of the hybridized MQW-on-CMOS devices to a level where the large device arrays projected values in Table II will routinely be possible. We note that early work in this area produced bonded optoelectronic device arrays of size 8×8; the overall yield, obtained by bonding multiple smaller 8×8 arrays was 99.84 [12]. We note that much larger hybrid MQW arrays (128×128 pixels) have been attempted to date for application to spatial-light modulators (SLM). This effort produced a device yield of 99.87% (20 failed devices out of 16 K) [16]. Because the application of MQW diodes for interconnections within high-speed digital CMOS systems

TABLE II
ASSUMED OPTOELECTRONIC TECHNOLOGY PARAMETERS

Feature Size (microns)	Flip-Chip Pad Dimension (μm)	Bonded Diode Capacitance (fF)	Thermal Resistance (°C/W)	Laser Power Per Chip (mW)	# Optical Diodes	Optical Loss (dB)
0.7	16	150	3,500	50	1,000	-12
0.5	12	100	6,000	100	3,000	-11
0.35	10	70	7,500	200	6,000	-10
0.25	8	50	10,500	400	12,000	-9
0.18	6	36	17,000	800	24,000	-8
0.12	5	24	24,000	1600	40,000	-7
0.1	3-4	20	36,000	3200	50,000	-6

Bonded diode capacitance is quoted for a differential input pair. Note that the optoelectronic technology is assumed to be integrated with CMOS one generation old; 1995 optoelectronic technology (3200 diodes) being associated with 0.7-μm CMOS was commercially available in 1992. Thermal resistance is that of a single MQW diode.

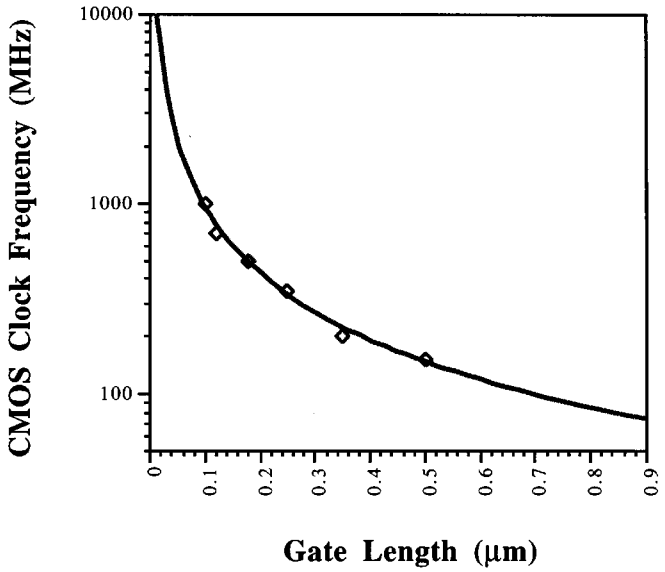


Fig. 4. Scaling of the qualified CMOS clock frequency as a function of gate length according to the SIA predictions.

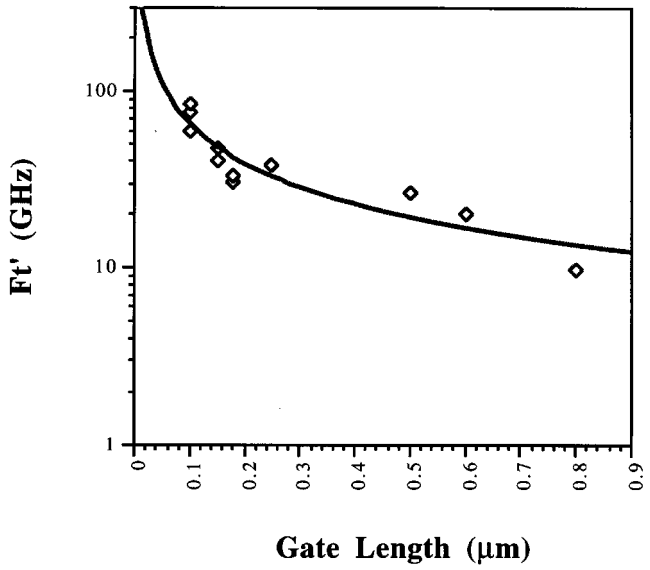


Fig. 5. Scaling trend for the inverse gate-delay as a function of the gate length of the CMOS technology. (Data points correspond to measured delays from ring-oscillator data).

has a low tolerance for failed devices, the projected number of devices in Table II is somewhat conservative relative to the SLM application. The current yield of the hybrid OE-VLSI device technology is approximately one failed device out of 4000 (or 99.97%) [17]. The crucial yield barrier to the hybrid CMOS-MQW process will most likely be due to defects that occur during the epitaxial growth of the GaAs–AlGaAs material. This leads to defect densities associated with the total area of the array that will likely limit the number of optoelectronic devices per chip to under 10^5.

The optoelectronic technology parameters of Table II represent an *estimate* as to the evolution of the photonic technology based on our current understanding of device fabrication and system issues. The available laser power is expected to increase by a factor of two every generation; the optical system loss is expected to reduce by 1 dB each generation and the maximum number of diodes is projected to increase rapidly in the first few generations and then grow gradually as the device yield, issues of bonding large arrays, and manipulating large numbers of light beams become increasingly significant. As the bonding technology is improved and flip-chip bond pad sizes are reduced, reductions in MQW diode capacitance and corresponding increases in diode thermal-resistance will also accompany each new generation.

It was our intent to be deliberately conservative in assessing the maturity of the optical technology available to the system designer. This is evidenced, for instance, in our assumptions for the expected magnitude of available laser power per stage. The full 3.2-W laser power suggested for the 0.1-μm generation (around 2007–2010) is in fact available today, though the system that generates it (e.g., ion-laser-pumped Ti:sapphire) is unreasonably large and expensive. A major technological evolution is taking place in high-power lasers with available powers continually increasing and cost ($/Watt) dropping by a factor of four every three years (i.e., every

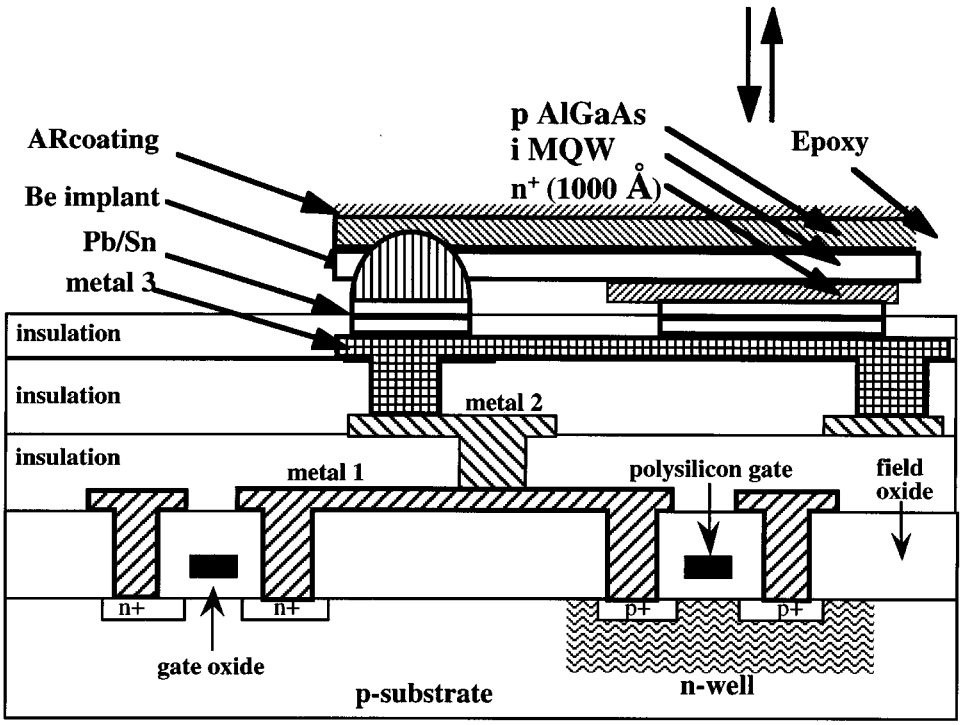


Fig. 6. Structure of the hybrid GaAs MQW/silicon CMOS circuit. Modulators may be bonded directly over active CMOS gates.

generation) [62]. Hence, the assumption of a factor of two improvement in laser power per stage is conservative and consistent with continual reduction in system cost over time. We have also assumed that the optical interconnection system will impose a limit to the minimum resolvable optical spot size; accordingly, the minimum linear dimension of the MQW diode is expected to be between 5 μm and 7 μm. This likely does not represent the minimum size at which MQW diodes could be fabricated, though at dimensions of the order of 2 μm or less there might be changes in the electroabsorption mechanism in the diodes because of fringing fields that could cause exciton broadening.

One of the more significant optoelectronic device parameters is the MQW diode capacitance. Each new generation of CMOS technology will be accompanied by a reduction in the dimensions of the flip-chip pads and of the active MQW diode area. The diode capacitance will reduce in proportion to its active area as long as the area-overlap capacitance of the MQW and of the flip-chip pads is the dominant factor. However, continued reductions will ultimately be limited by the diode capacitance arising from fringe components, the pad capacitance arising from fringe components, and stray capacitance due to the flip-chip bump bonds, etc. The scaling of the hybrid flip-chip bonded MQW detector capacitance, (C_D), has been studied in detail elsewhere [63]–[65]. Making the assumptions that the GaAs substrate has been etched away, and that the contribution of the remaining GaAs chip connecting the p-contact and n-contact pads is small, the diode capacitance and the flip-chip bump-plus-pad capacitance are the two primary components of the detector capacitance:

$$C_D \approx C_{\text{diode}} + C_{\text{bump+pad}}. \quad (1)$$

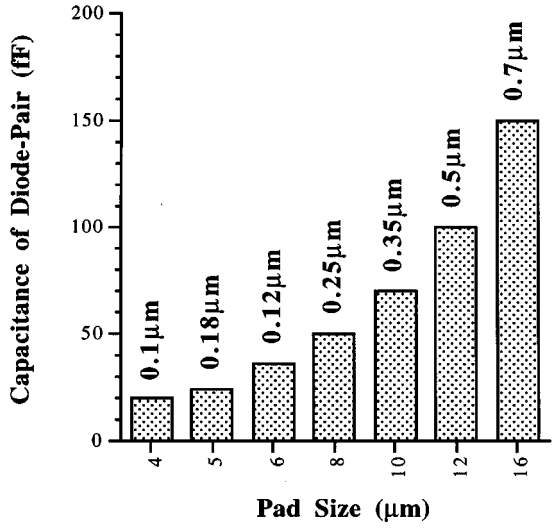
Based on an analysis for a parallel-plate over a ground plane [66], closed-form expressions for the MQW diode capacitance (C_{diode}) and the bump + pad capacitance $(C_{\text{bump+pad}})$ have been derived in [65] as

$$C_{\text{diode}} = A_c[1.15(d+2)(d+4) + 4.2d + 16] \quad (2)$$

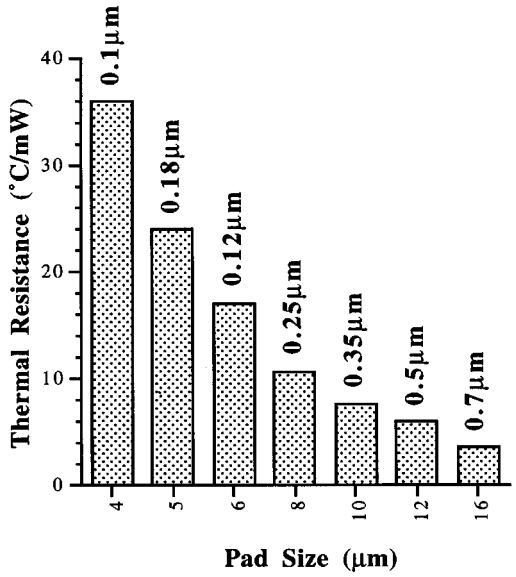
and

$$C_{\text{bump+pad}} = \varepsilon\big[1.15(d^2)/t_d + 5.6(t_{\text{bump}}/t_d)^{0.222} + 4.12t_d(t_{\text{bump}}/t_d)^{0.728}\big] \quad (3)$$

where A_c is the MQW diode capacitance per unit area (approximately 0.11 aF/μm^2), d is the linear dimension of the flip-chip pad, t_{bump} is the combined thickness of the metal pad and the thermocompression-bump, t_d is the height of the dielectric (oxide) separating the pad from the ground plane, and ε_r is the relative permittivity of the dielectric. The dominant term in (1) is typically C_{diode}. The first term in (2) corresponds to the parallel-plate overlap capacitance of the diode; the active area of the diode is typically several (2–4) microns larger than the associated flip-chip pad, using current design rules. The second and third terms correspond to the contributions from fringe-capacitance and corner-capacitance, respectively [65]. The precise value of $C_{\text{bump+pad}}$ in (3) depends on the underlying geometry beneath the pad. Fig. 7(a) and Table II display the expected reduction of the capacitance for a pair of detectors as the technology scales. It has been assumed that the pad size will gradually shrink from $d = 17$ μm (current generation), ultimately to $d = 3$–4 μm (with active diode windows of 5–7 μm). This provides approximately linear scaling of diode capacitance with feature size. As mentioned previously, the ultimate size of the pads will be governed by both the bonding technology and the optical interconnection



(a)



(b)

Fig. 7. Calculated scaling of (a) capacitance and (b) thermal resistance versus flip-chip pad size [65].

system. We note that MQW devices with optical windows of 5-μm$\times$ 5-μm have already been demonstrated in relatively large arrays (32 K devices [67]) for optical switching demonstration systems [68].

Power dissipation of the OE-VLSI chip is another important systems constraint. A temperature-induced shift of approximately 0.28 nm/°C in the exciton peak restricts the temperature swings of the MQW modulator during operation. There are two potential thermal effects that are of concern. The first is the ability of the package to extract sufficient heat in order to stabilize the overall temperature of the OE-VLSI chip. Note that the silicon substrate has a high thermal conductivity, and there is some freedom in choosing the mean operating temperature of the MQW modulators. Predictions by the SIA suggest that aggressive cooling technologies, capable of removing 5–16 W/cm^2 of power from a single chip, will be available for high-performance applications. As long these power dissipation limits are not exceeded, these external cooling techniques can be used to prevent large overall temperature swings on the OE-VLSI chip's surface during operation.

The second concern is the point-source heating of an individual modulator due to its finite thermal conductivity and the data-dependent photocurrent that flows through it during operation. Point-source heating effects are thus a potential source of local temperature swings, particularly for large input powers [69]. The scaling of the thermal resistance of a hybrid MQW diode versus flip-chip-bond-pad size has been investigated in [65]. Estimates of the thermal resistance of the MQW diode are included in Table II and Fig. 7(b); these figures assume that 3–5 μm of solder will be used in the bonding procedure. Assuming that the silicon chip is the heat sink, the change in temperature that a diode will experience, ΔT_{diode}, is proportional to the photocurrent and to the thermal resistance of MQW device (R_{diode}, which in turn is inversely related to the device size [63]. This can be expressed as

$$\Delta T_{\text{diode}} = \Delta P_{\text{diode}} \cdot R_{\text{diode}} = V_{\text{mod}} \Delta I_{\text{diode}} \cdot R_{\text{diode}} \quad (4)$$

where ΔP_{diode} is the difference in the electrical power dissipation in the diode between the low- and high-reflectivity states, and ΔI_{diode} is the corresponding difference in absorbed photocurrent. For dc-coupled NRZ data, these swings will typically be pattern-dependent; long strings of ones or zeros are capable of creating temperature excursions proportional to the peak photocurrent in the diode. At the largest values of photocurrent (lowest sensitivities), and at small device sizes, this corresponds to several milliWatts of power dissipation in the diode. Because the temperature swing experienced by the modulator must be kept within certain bounds ($\approx\pm 5$ °C), its thermal resistance presents a limit to the maximum photocurrent and hence the maximum optical energy that its corresponding receiver can demand. This presents an argument for operating at low optical energies (high sensitivity).

The operation and physical characteristics of the optical devices have been widely investigated in the literature. Saturation intensities of approximately 80 kW/cm^2 have been reported [13]. The arrayed MQW devices typically operate with a responsivity of approximately 0.5 A/W, a capacitance of approximately 0.11 aF/μm^2, and a contrast ratio of 3:1 when the input swing is 5 V. Reflectivity changes of about 35% (on-state) to 10% (off-state) have been measured across device arrays [70]. This represents a minimum reflectivity change, ΔR, of 25% that can be expected to be reproducible over large arrays. It should be noted that much better performance has been measured on individual devices, with ΔR of 50% and 70% with voltage swings of 5 and 10 V, respectively [71].

An important issue for continued compatibility of the optoelectronic devices to smaller line-width CMOS technologies is the voltage drive requirement. One of the technological challenges is to reduce the required voltage swing across the modulator diodes in order to maintain compatibility with mainstream IC technologies, without suffering a large penalty in contrast ratio. Reducing drive-voltage requirements is currently an active area of research. One possible approach to reduce voltage drive requirements is to employ asymmetric

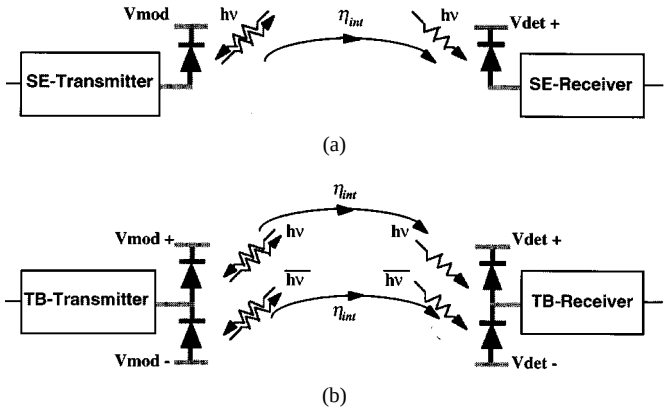


Fig. 8. Schematic of (a) single-ended (SE) optical interconnection link, (b) two-beam (TB) optical interconnection link; $V_{\rm det}$ and $V_{\rm mod}$ are the detector and transmitter bias voltages, respectively. $\eta_{\rm int}$ is the product of the optical system power efficiency and the modulator power efficiency.

Fabry–Perot cavity structures [72]–[75]. The penalty is the reduced tolerance to device nonuniformity [76], [77]. Another method is to use a stacked-diode modulator design that reduces drive voltage at the expense of increased diode capacitance [71]. We will assume that the MQW design innovations such as those described above will enable the technology to evolve according to Table II.

III. Limits on the Number of Optical I/O Circuits

In this paper, we will assume cascaded operation of hybrid OE-VLSI circuits with unity fanout. Information from one IC is transcribed to another, using a generic one-to-one free-space optical interconnection system that incorporates a certain amount of optical loss. In order to accomplish reliable data communication, we assume (without loss of generality) two-beam differential operation of receivers and transmitters (see Fig. 8), with each receiver and transmitter circuit requiring two MQW diodes, and the total optical power needed for all the modulators on the transmitting IC being defined in Table II. Experiments in free-space multistage photonic switching systems [68], [78] have demonstrated that two-beam differential optical signaling between switching stages is a reliable means of communication between arrays. If one were to use the transimpedance receivers described in this paper with single-ended input, a relatively low-power input light signal would be required in the "zero"-state (i.e., a high-input contrast ratio). This is, in fact, the mode of operation used in a recent switching systems demonstration [21], where the input was delivered directly from electronically modulated lasers to single-ended receivers on a CMOS chip.

Indeed, cascaded systems with high-contrast modulators may be able to run single-ended. However, two-beam operation is generally preferred for low-contrast devices and has advantages even with high-contrast devices; for instance, it provides a fair amount of rejection to common-mode noise in the input beams. Because the on-state and off-state currents are physical mirrors, a two-beam receiver can also provide a constant mark-to-space ratio (or pulsewidth) over a large dynamic range. In single-ended receivers, the mark-to-space ratio typically depends on the input power. It should be noted that both single-ended and two-beam receivers are susceptible to electrical noise in the supply lines, particularly when designed for high sensitivity; because noise-margins diminish when supply voltages are reduced, circuit techniques that provide supply noise rejection (not explicitly discussed in this paper) will eventually become a necessity.

For the purposes of this study, we will further assume that the hybrid OE-VLSI chips contain the maximum allowable number of receivers and transmitters based on the number of diodes available and power dissipation limits, and it is the combined total *bandwidth*, into and out of the optoelectronic IC, that is of interest to the system designer. One method of increasing the number of optical inputs and outputs to the chip is to allow a receiver and a corresponding transmitter circuit to share the use of the diode pair by time-multiplexing its operation as either an input detector or an output modulator [79]. Such a design would leave the bandwidth of the IC unaffected (except for a small penalty due to the multiplexing) but would potentially double the number of optical I/O to the chip compared to solutions that use separate devices for each receiver and transmitter. Because we examine the limits to the number of receivers and transmitters on a single IC, the results of this paper are also applicable to noncascaded single-stage systems. In systems where cascaded operation is not required, single-ended operation can be used to further double the number of optical I/O.

The total number of optical I/O that can be supported on an optoelectronic IC is constrained by several key factors: the maximum number of MQW diodes, the availability of laser power, and the power dissipation of the optical transceiver circuits, namely the receivers and transmitters. The results of this study indicate that the static current in the receivers is a dominant source of power dissipation for the optical interconnect. This has been corroborated experimentally for 0.8-μm CMOS [19]–[22], [80]. The scaling of the voltage supply ensures that the dynamic power dissipation of the receiver running at the system clock speed will constitute only a small fraction of the total receiver power. Section IV will examine the dissipation of the receiver and transmitter circuits in greater detail.

The methodology used in this paper is outlined in the influence diagram (Fig. 9) that shows how the assumed and derived quantities affect the maximum number of receivers on the optoelectronic IC. The fundamental FET parameters (i.e., linewidth, oxide thickness, supply voltage) determine the FET transconductance, the available gain per stage, the FET capacitance, and transistor gain bandwidth (f_T'). The MQW diode capacitance is added to the front-end FET capacitance to determine the total input capacitance of the receiver. This is then used to determine the receiver noise current, and the minimum required input current to guarantee a specific bit-error rate (assumed to be 10^{-15}). This current sets the minimum photocurrent (maximum sensitivity) for the receiver. For a given sensitivity, the required input photocurrent at the operating bit rate is then derived. This current results in an input voltage swing at the transimpedance receiver front end. The required number of additional gain stages is then calculated assuming that the receiver must restore the signal

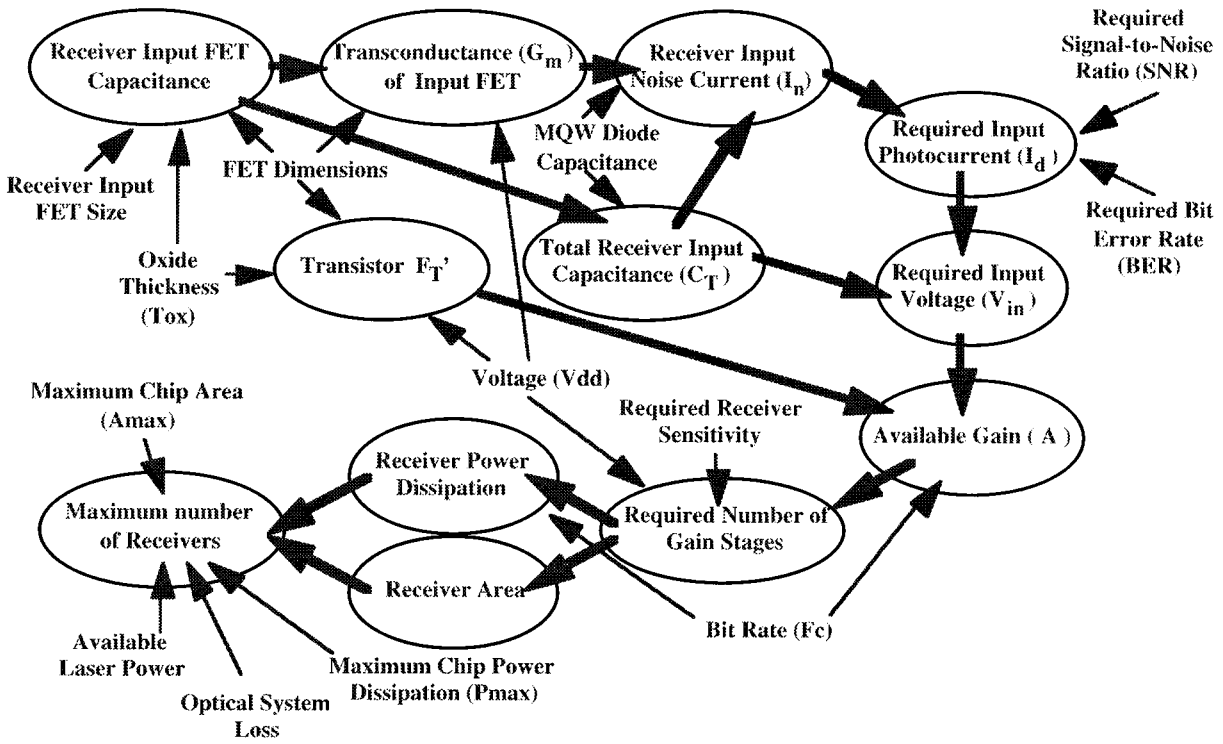


Fig. 9. Influence diagram for determining the maximum number of optical receivers. Assumed constants and derived expressions (in bubbles) are included. Arrows show relationships between terms.

to logic levels. The required number of gain stages then allows the receiver dissipation and area to be calculated. Limits on the maximum number of diodes, the available laser power, and the maximum receiver power-dissipation and receiver area are then used to determine the maximum number of optical I/O to the IC.

The maximum allowable area and power dissipation of an OE-VLSI chip follows the SIA roadmap predictions. The issue of power budgets for logic, clocking, and chip I/O has been examined in [81], where it is estimated that approximately half of the power budget is typically used for the chip-I/O in full-custom designed electronic chips built in 1-μm CMOS. Because on-chip voltage swings are decreasing faster than off-chip electrical signaling standards, this number can be expected to increase for electronic chips (up to 75%) as feature sizes shrink [81]. We note, as an example, that for switching and I/O intensive smart-pixel systems, with minimal logic for computing, the *nominal* I/O power budget fraction may be closer to 75% [21]. For convenience, we will make the assumption that half the power budget will be dedicated to the optoelectronic transceivers for all the OE-VLSI technologies.

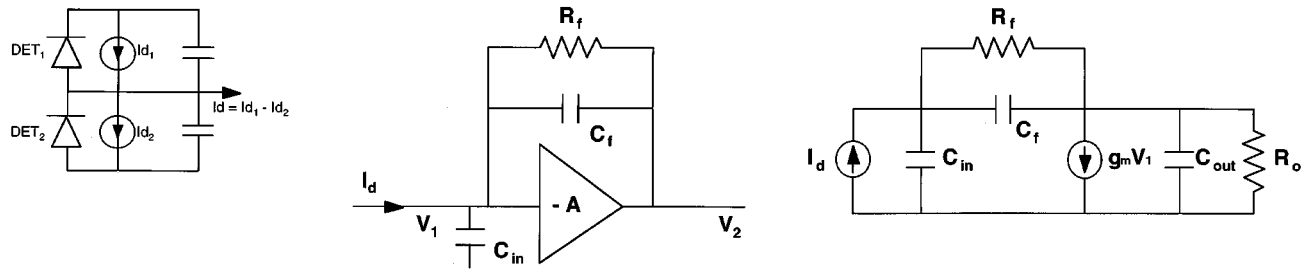
IV. Receiver and Transmitter Considerations for 2-D Array Optical Interconnects

In this section, we discuss the receiver model assumed for this study and presents the derivation for the maximum number of optical receivers on the OE-VLSI circuit. Section IV-A describes the transimpedance receiver design used for this study. Section IV-B reviews the general Smith–Personick noise model for receivers and an improved noise model that takes into account the transimpedance amplifier transfer function. Section IV-C then presents the scaling of receiver power dissipation and area.

A. Performance of CMOS Transimpedance Receiver Front End

For the purposes of this study, we will restrict the discussion to transimpedance-amplifier based receivers. Transimpedance receivers have been widely studied and implemented for optical data links [82]–[89]. The transimpedance receiver is chosen because it provides good sensitivity and dynamic range and because it allows the bandwidth to be increased, compared to an equivalent high-impedance (integrating) configuration, by the open-loop gain of the circuit. If we asked the simple question, "Given the gain-bandwidth product of the specific FET technology, how many simple gain stages would it take to raise the signal to a level compatible with CMOS logic at the speed we wish to operate," we would get answers on power dissipation and area that are essentially similar to those calculated here. The transimpedance circuit gives a real, workable circuit with a known transfer function.

The transimpedance receiver front end can be operated in single-ended mode where a single diode is used to generate positive photocurrent, as well as in a differential operation mode with an additional diode at the input. In the latter case, the logic states of the receiver output correspond to light shining on either one or the other diode. Small input swings in the front end are then amplified to logic levels at the receiver output using additional biased stages of inverters. In the following, we will assume that the transimpedance front-end stage consists of a pair of reverse-biased MQW diodes connected to an inverting amplifier with feedback [Fig. 10(a)–(b)]. The use of an active FET as the bias feedback



(a) (b) (c)

Fig. 10. (a) Differential input diode pair and (b) equivalent circuit for a typical transimpedance receiver. The feedback resistor R_f, is a PMOS inverter with a tunable gate voltage. Alternatively, the feedback element can be a parallel combination of PMOS and NMOS devices [19]. (c) Small-signal equivalent circuit for the transimpedance amplifier.

resistor, R_f, has the advantages of low area and capacitance as well as the ability to vary the resistance over a wide range [90]. The feedback resistor can be accomplished using a PMOS device. In practice, a parallel combination of a diode-connected NMOS device (gate attached to drain) together with the PMOS device provides a simple form of automatic gain control and hence a large dynamic range [19]. The small-signal equivalent circuit is provided in Fig. 10(c). Circuits based on a transimpedance front end followed with a single amplification stage have been demonstrated to operate at a bit rate of 375 Mb/s in 0.8-μm CMOS with a dc power consumption of approximately 3.5 mW, an area of 17-μm $\times$ 18-μm, a dynamic range of over 16 dB, and a sensitivity of approximately 60 fJ [19].

The operation of the transimpedance receiver is governed by the following equations, where V_1 and V_2 are, respectively, the voltages at the input and output of the front end, R is the transimpedance feedback, I_d is the photodetector current, and A is the open-loop gain of the amplifier

$$V_2 = -A \cdot V_1 \tag{5}$$

with

$$\frac{V_2 - V_1}{I_d} = R_f. \tag{6}$$

The resulting bandwidth (dominant input pole) of the transimpedance receiver is

$$f_{-3\,\mathrm{dB}} = \frac{A+1}{2\pi R_f C_T} \tag{7}$$

with

$$C_T = C_{\mathrm{in}} + (A+1)C_f \tag{8}$$

where C_{in} is the sum of the diode capacitance and the input capacitance of both the NMOS and PMOSFET's, C_f is the capacitance of the feedback element, and C_T is the total input capacitance of the receiver [91]. The low-frequency transimpedance of the receiver, Z_T, can be defined as

$$Z_T = \frac{-R_f}{1 + \frac{1}{A}}. \tag{9}$$

The receiver design considered in this paper uses a *single* gain stage within the closed loop transimpedance amplifier front end, followed by one or more additional inverter gain stages that will produce logic-level output. Because the gain in a single inverter stage is restricted to relatively small values (in the range of 10–30 with the assumed FET geometry), the stability of the system is then governed by the value of the feedback resistor. For a given gain, A, a higher value of this resistance increases the transimpedance gain and reduces noise, but also reduces the bandwidth and possibly induces ringing and/or instability. The value of the resistor, $R_f(\approx$10–80 k$\Omega)$, is bounded from above to meet the bit rate and stability requirement. It is also bounded from below to limit the noise in the receiver front end for a given sensitivity. No fanout is assumed so that the output capacitance of the front end is relatively low; in contrast, the input capacitance is relatively high due to the MQW diode loading.

There are several simple options for the gain stage as depicted in Fig. 11(a)–(c). These include an inverter amplifier with an active PMOS load, an inverter with a current source load (PMOS transistor), and a push–pull inverter amplifier. The push–pull inverter amplifier has the ability to both sink and source current and will provide the largest gain for a given drain current. Hence, this will be the preferred gain stage, although the discussions will also be valid for a current-source load. The small-signal equivalent circuit for the generic transimpedance circuit is shown in Fig. 11(c).

As (8) suggests, the total receiver input capacitance, C_T, plays an important role in determining the maximum bandwidth of the transimpedance receiver. Ignoring stray parasitics (that include local interconnect and fringing capacitance of the FET's), C_T consists of the detector capacitance, the gate capacitance of the NMOS and PMOSFET's, and the capacitance of the feedback element used in the receiver front-end amplifier. The total input capacitance of a MOSFET (C_g) is approximately given by the sum of the gate-to-bulk (channel) capacitance (C_{gb}) and the overlap capacitance as

$$C_g = C_{\mathrm{gb}} + C_{\mathrm{gs}} \tag{10}$$

where C_{gs} is the gate-source overlap capacitance. Ignoring fringing, we can estimate the values of these capacitances assuming the transistor is in saturation

$$C_g \approx \frac{\varepsilon}{t_{\mathrm{ox}}} \cdot \left[\frac{2}{3} W \cdot L + 2(W \cdot \Delta L)\right] \quad \text{(saturation)} \tag{11}$$

where ΔL is the overlap of the gate over the source and drain regions due to lateral diffusion under the gate. As previously mentioned, an important design objective is to minimize the power consumption of the receiver. This in turn implies that

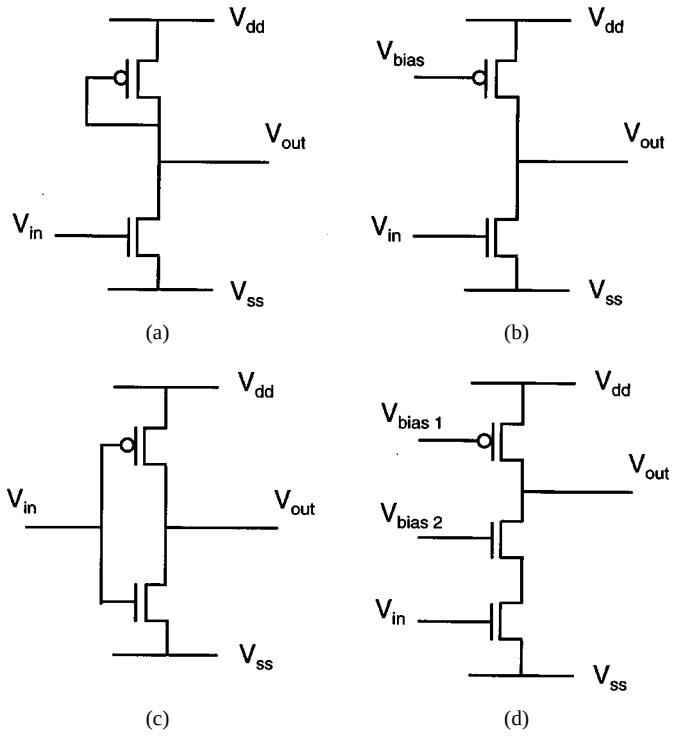


Fig. 11. (a) Inverter amplifier with an active PMOS load, (b) inverter amplifier with a PMOS current source load, (c) push–pull CMOS inverter amplifier, and (d) cascode inverter amplifier.

the static current in the input FET's of the transimpedance amplifier cannot be excessive. To examine the effect of the input FET sizing on receiver performance, it is instructive to examine the noise current flowing in this FET.

The effect of the dominant input pole of the receiver has been considered in (3). Because the transimpedance receiver front end has a single gain stage within the loop, a large input capacitance, and a relatively low output capacitance (with unity fanout), the stability of the front end is not expected to be an issue. The stability criterium sets a limit for the transimpedance-bandwidth product [from (7) and (9)] of a circuit for a given technology. For the closed-loop circuit to be stable, a minimum phase margin of 45° is required, with a 60° phase margin or larger preferred [92]. The latter is valid when the output pole is at least twice the frequency of the input pole. To verify stability, SPICE simulations of the front-end circuit geometry assumed in this paper were performed for 0.8-μm and 0.35-μm CMOS technologies.[2] Fig. 12(a)–(c) show the loopgain, the phase margin, and the transimpedance of the receiver front end, respectively. In the simulations, the input and output were respectively loaded with the appropriate detector capacitance and output load capacitance (another identical gain stage) for each technology. A minimum size PMOSFET of equal width and length was used as the feedback resistor. As shown by the simulations, a phase margin of 60° is obtained in both cases, with a transimpedance over 90 dB (32 kΩ) at the unity loop-gain frequency. The unity-gain frequency of the front end is more than twice the assumed clock frequency. These simulations provide evidence that the single gain-stage transimpedance front end will be stable with sufficient phase margin.

[2] SPICE simulations were performed using HP 0.8-μm and ATT 0.35-μm CMOS process parameters.

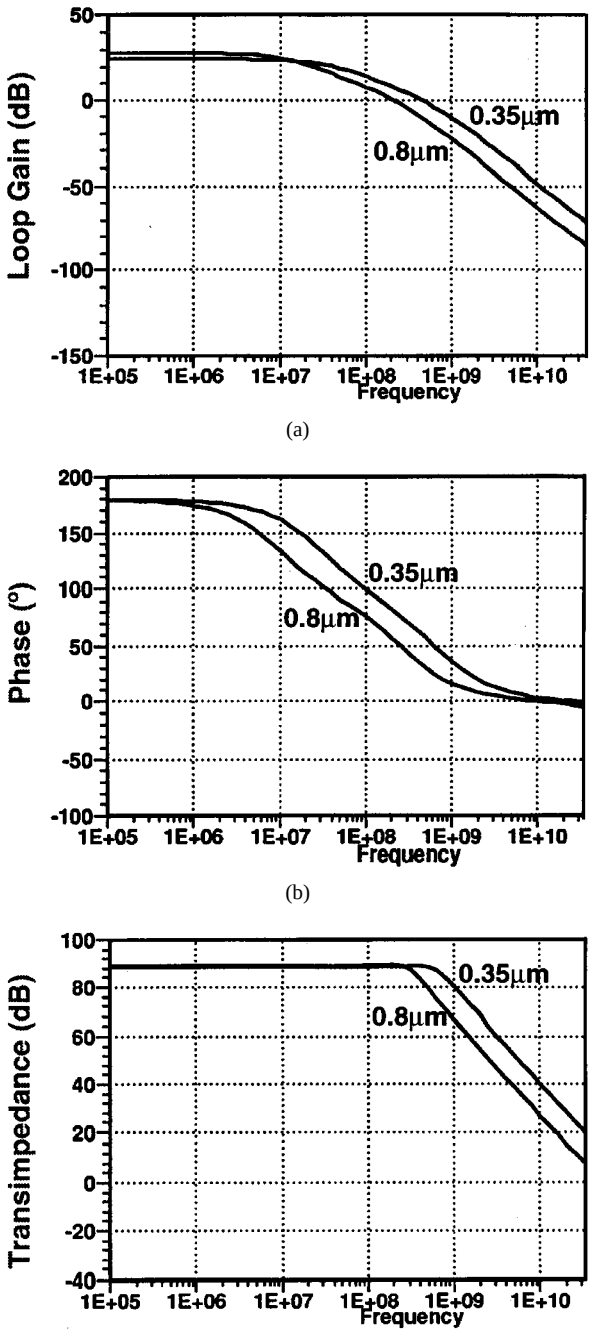


Fig. 12. Simulated performance of the input (front-end) stage of the transimpedance receiver using submicron CMOSFET's, showing (a) loop gain (dB), (b) phase margin (degrees), and (c) transimpedance (dB) of the front end for 0.8-μm and 0.35-μm technologies.

For a given operating bit rate, the excess phase margin can be traded in for reduced noise by increasing the transimpedance of the tunable feedback FET. The use of a PMOS transistor having a tunable bias is recommended for use in the feedback loop so that some optimization of the gain bandwidth of the front end can be performed during operation. To a limited extent, the gain of the overall amplifier can be increased by tuning R_f. To further increase the overall

gain of the receiver, additional inverter stages can be added after the transimpedance amplifier stage to increase the gain. The use of differential encoding together with appropriate transistor sizing allows these further gain stages to be biased at the midpoint $(V_{\rm dd}/2)$. However, these additional gain stages increase the power dissipation of the receiver; this tradeoff will be explored in the following sections. We have assumed that these stages are effectively biased by the transimpedance amplifier front end and calculate the static power dissipation of these additional stages based on this mid-point biasing. In practice, additional diode-connected FET's can be used between each pair of additional gain stages as resistor elements to bias these stages; these additional resistors also broaden the overall gain of each added stage and reduce the sensitivity of the receiver to input noise and process variations.

B. Receiver Front-End Noise Analysis

A generic receiver circuit is shown in Fig. 10(b). The noise behavior of this circuit has been studied in [93], where it is shown that the total mean-square equivalent input noise due to the input devices of the receiver (the photodiode, input FET, and bias resistor) front end has contributions arising from the Johnson noise of the bias resistor, the photodetector dark-current, the FET gate-leakage currents, the $1/f$ or "flicker" noise in the FET's, and the channel noise of the input FET [89], [91], [93]:

$$\langle i_n^2\rangle \cong \frac{4kT}{R_f}I_2B_0 + 2qI_lI_2B_0 + 4kT\Gamma\frac{(2\pi C_T)^2}{g_m}f_cI_fB_0^2 + 4kT\Gamma\frac{(2\pi C_T)^2}{g_m}I_3B_0^3 \tag{12}$$

where B_0 is the cutoff bit rate, I_l is the sum of the photodiode and FET leakage currents, I_2, I_3, and I_f are the normalized Personick noise bandwidth integrals, respectively, q is the electron charge constant, f_c is the $1/f$ noise corner frequency, and Γ is the excess channel noise factor term associated with short-channel transistors. The basic formula for noise in (12) applies to any receiver design; it can be verified from the fundamental noise mechanisms in the devices by ignoring the constants (I_2, I_3, and Γ), assuming a very simple model of biased gain stages without feedback, and replacing the cutoff bit rate, B_0, by the bandwidth.

The $1/f$ or flicker noise in CMOSFET's is a result of fluctuations in the number of trapped carriers in the gate oxide near the MOS interface. Typical flicker noise corner frequencies for MOSFET's are on the order of a few megahertz. There is evidence that $1/f$ noise will increase in short-channel FET's, as hot carriers degrade the channel [94], [95]. Although $1/f$ noise may become potentially disastrous in terms of high-sensitivity (dc-coupled) receiver performance, the use of encoding techniques to remove the low-frequency components in the data stream can be an effective means to ameliorate this potential problem. It should be noted that ac-coupled data encoding will also reduce point-source heating effects in the modulators and may ultimately prove quite valuable for OE-VLSI systems integration.

The precise expression given in (12) and the calculation of the Personick integrals, I_2, I_3, and I_f, are strictly valid only for a family of rectangular input pulses and a raised cosine output pulses in a linear, equalized channel that is filtered prior to the decision circuit. According to the Smith–Personick analysis, the values of I_2 and I_3 are relatively insensitive to the fraction of the time slot occupied by the pulse, or to the exact shape of the raised cosine pulse, i.e., $I_2 \approx 0.6, I_3 \approx 0.09, I_f \approx 0.2$ over a wide range of parameters. But in order to estimate the noise, we would be required to assume that the receiver has a cutoff frequency, B_0, exactly equal to the operating bit rate, B. Furthermore, we would also be required to assume that the output waveform at this bit rate has the shape of a raised cosine. Most practical receivers, such as the one considered in this paper, instead employ NRZ input and output pulses without equalizer stages or a well-defined cutoff bit rate.

The interpretation of the Smith–Personick analysis has been addressed more carefully in [96] where the noise performance of an unequalized transimpedance receiver has been considered without making assumptions on the input/output pulse shapes. Indeed, their results verify that the use of the Smith–Personick analysis, without regard to the points mentioned above, can lead to a significant underestimation of the rms noise current. Hence, it is important to consider the effect of the transfer function of the specific receiver front end when performing the noise calculation. While these refinements will be valid only for transimpedance front ends, they provide a more accurate estimate of the expected sensitivity of the receivers.

From [96], the mean square input noise current for a transimpedance receiver [Fig. 10(b)] can be written as

$$\langle i_n^2\rangle \cong \frac{4kT}{R_i}\Phi_1 + 2qI_l\Phi_1 + 4kT\Gamma\frac{(2\pi C_T)^2}{g_m}\Phi_2 \tag{13}$$

where

$$\Phi_1 = \frac{1 + g_mR_O}{4[R_O(C_{\rm in}+C_{\rm out}) + R_f(C_f+C_{\rm in}) + g_mR_oR_fC_f]} \tag{14}$$

and where Φ_2 is defined as (15), shown at the bottom of the next page. C_T is the total input capacitance of the receiver from (8), and R_O is the output resistance of the front-end inverter amplifier. The input $(C_{\rm in})$, output $(C_{\rm out})$, and feedback (C_f) capacitance in (14) and (15) for a two-beam inverter-amplifier-based receiver are given by

$$C_{\rm in} \approx C_{g,N\rm MOS} + C_{g,P\rm MOS} + 2\cdot C_D \tag{16}$$

$$C_{\rm out} \approx C_{{\rm ds},N\rm MOS} + C_{{\rm ds},P\rm MOS} + C_{\rm Load} \tag{17}$$

$$C_f \approx C_{{\rm ds},P\rm MOS} + \{C_{{\rm gd},N\rm MOS} + C_{{\rm gd},P\rm MOS}\} \tag{18}$$

where $C_{\rm Load}$ is the load capacitance of the following stage (consisting of an identical inverter amplifier). The feedback capacitance in (18) includes the drain-source capacitance of the feedback FET and the gate-drain capacitances (in brackets) of the NMOS and PMOSFET's in the inverter amplifier. From (8), we see that these last two terms in (18) are multiplied by

the gain of the amplifier, due to the Miller effect. One method of eliminating this Miller capacitance is to use a cascode amplifier configuration [Fig. 11(d)] [92]. The second term in (12) and (13) represents shot noise contributions of the diode dark current and FET leakage for nonreturn-to-zero coded data. Photodiode dark currents for the hybrid MQW modulators are expected to be on the order of a few nanoamperes or less. The FET leakage is directly related to the threshold voltage of the FET's. Low threshold voltages directly result in high subthreshold conduction and leakage currents in the FET's. Due to the exponential nature of the subthreshold current, the threshold voltage V_t cannot typically be scaled down linearly with feature size. Higher leakage currents will have detrimental effects on the performance of the analog circuits and will also increase the overall static power dissipation of the chip [24]–[31], [61]. The expression for subthreshold conduction of a FET, $I_{l,\mathrm{Fet}}$, can be written as

$$I_{l,\mathrm{Fet}} = I_{d,\mathrm{sat}} 10^{\frac{-V_t}{S}} \tag{19}$$

where S is the subthreshold conduction slope of the submicron MOSFET (defined as the change in gate voltage required to induce a $10\times$ change in drain current), and $I_{d,\mathrm{sat}}$ is the drain current of the FET at the onset of saturation. $I_{d,\mathrm{sat}}$ is on the order of 100 nA multiplied by the width-to-length (W/L) ratio of the FET. We will assume that the subthreshold slope will be 80 mV/decade or below, which is consistent with the expected performance of deep submicron MOSFET's at room temperature. As mentioned in Section II-A, a V_{dd}/V_t ratio of approximately four will be assumed in this paper. While FET leakage currents of 1–10 pA/μm are expected for the longer channel devices, an evaluation of (19) suggests that very short channel transistors (0.1 μm) may exhibit leakage currents on the order 100 pA–1 nA. Even at these levels, the overall effect of the leakage current on the receiver noise is small.

The significant sources of input noise are the first and last terms in (12) and (13). The first term is due to the thermal noise in the feedback resistor used in the transimpedance amplifier; for the high-resistance (e.g., $R_f \approx 40$ kΩ) FET-based feedback resistor, this term is typically small; it dominates only when $R_f \ll 10$ kΩ. As a result, the dominant term for wide-bandwidth receivers is the last term, corresponding to the mean-square input-noise current due to the channel (Johnson) noise in the input FET. Notice that this noise current is inversely proportional to the transconductance of the input-stage FET and directly proportional to the square of the total input capacitance of the receiver front end. A smaller FET in the front end also has lower capacitance, but the transconductance of the FET is proportionately reduced when its dimensions are shrunk. The conventional wisdom is to increase the transconductance of the FET (by optimizing its width-to-length ratio) to the point where the capacitance of the input FET is equal to the detector capacitance (plus the feedback capacitance). It can be shown that the optimum noise current is obtained when this occurs [89], [91], [93]. This maximizes the gain bandwidth of the front-end amplifier and minimizes the well-known figure-of-merit of the receiver, $g_m/(C_T)^2$, where g_m is the front-end-inverter transconductance and C_T the total input capacitance. This is typically done in telecommunications receivers [97], where the primary objective is to maximize the sensitivity of the receiver and hence to reduce the noise current to a minimum. However, the resulting input FET (for a current-source load) is typically large, resulting in large static current, and hence this sensitivity performance can result in unacceptably high power consumption per receiver. It was recently pointed out in [98] that the input FET capacitance could be reduced to approximately 20% of the sum of the photodetector capacitance plus the stray capacitance due to the feedback FET, without a significant change in the overall noise of the receiver. In this paper, we suggest further reductions in the input capacitance (compared to the photodetector capacitance) for arrays of receivers; this leads to a significant reduction in power dissipation in the receiver array, with only a small added noise penalty.

For large arrays of (smart-pixel) receivers, a general conclusion follows that the FET technology cannot be exercised so as to obtain true noise-limited receiver sensitivity. The smart pixel receivers discussed here will typically be gain-limited at the given bit rates due to the need for low power consumption. The addition of gain stages can improve the sensitivity up to the noise limit. To see this, one can write the low-frequency gain of an inverter as

$$A = g_m \cdot R_O = \frac{(g_{m,N\mathrm{MOS}} + g_{m,P\mathrm{MOS}})}{(g_{\mathrm{ds},N\mathrm{MOS}} + g_{\mathrm{ds},P\mathrm{MOS}})}. \tag{20}$$

From (20), one can see that the small-signal gain of the CMOS inverter amplifier is limited by its own output conductance. Typical open-loop inverter-gains are in the range of 10–30. Hence, at the highest quoted sensitivities, a single transimpedance gain stage will not be sufficient to produce the voltage swings necessary for a receiver decision stage to reliably restore to logic levels. To reduce the required swing at the input (and hence improve the sensitivity), additional stages of inverter amplifiers or gain-broadened inverters can be used in the front end. This allows smaller input photocurrents without reducing the bandwidth of the receiver. In this paper, the assumed geometry for the transimpedance front-end amplifier with push–pull inverter amplifiers, will be an NMOS transistor three times the minimum size $(3\times)$ in series with a PMOS transistor approximately nine times the minimum geometry $(9\times)$. This would provide approximately similar transconductances for the PMOS and NMOSFET's (and hence similar noise performance). Receivers that use this front-end geometry have been fabricated and tested in 0.8-

$$\Phi_2 = \frac{(1+g_m R_O)^2}{16\pi^2[R_O(C_{\mathrm{in}}+C_{\mathrm{out}}) + R_f(C_f+C_{\mathrm{in}}) + g_m R_O R_f C_f][R_O R_f\{(C_{\mathrm{in}}+C_{\mathrm{out}})C_f + C_{\mathrm{in}}C_{\mathrm{out}}\}]} \tag{15}$$

μm CMOS and show reasonable bandwidth and sensitivity [19]–[21].

As the feature size and gate length of the FET technology is scaled down, the input capacitance of the FET's that constitute the receiver front end will proportionately decrease. Because we have assumed a constant geometry (in terms of the gate length) for the front-end amplifier stage, the total receiver input capacitance will always be dominated by the capacitance of the MQW diode pair (quoted in Table II). Because we also assume that the MQW diode capacitance is scaled down with the CMOS feature size, the overall effect is that the contribution of the input FET's is between 5–15% of the total receiver input capacitance throughout the scaling.

The excess channel noise factor, Γ, is typically quoted as a constant for receiver noise calculations in a given technology (e.g., $\Gamma \approx 1.2$ for 1-μm gate-length FET's [91]). The scaling of this noise factor in submicron MOSFET's has been studied in greater detail in [99]–[101]. The exact value of Γ depends on the precise biasing and transconductance value of the FET. The noise factor is higher when the FET enters saturation; this is likely caused by hot electron effects in the pinched-off (high-field) part of the channel. As the FET linewidth is reduced below 1 μm, the studies indicate that the noise factor in submicron MOSFET's can be significantly higher (by a factor of two to four) than that ideal value of $\Gamma = 0.7$ predicted by long-channel theory [100]. For the receiver front end considered in this paper, we include a linear scaling of Γ with gate length, ranging from $\Gamma = 1.8$ to $\Gamma = 3$, to account for this increased channel noise as the device size is reduced. These values are consistent with empirically observed behavior in submicron FET's down to 0.25-μm gate lengths [99].

Because the excess channel noise factor of the FET is a function of technology scaling (noise factor increases as the linewidth is reduced), we can rewrite the noise figure-of-merit from (5) and (6) to $g_m/\Gamma \cdot (C_T)^2$, where Γ is the excess channel noise factor. If we make the further assumptions that the capacitance of the input FET's is a known fraction of the detector capacitance, C_D, and that the feedback capacitance is not significant, then we can further refine this figure-of-merit by substituting the expression for the unity-gain frequency (f_T) of the FET's [102] as

$$f_T = \frac{g_m}{2\pi(C_{\rm gs} + C_{\rm gd})}. \tag{21}$$

Equation (21) permits the figure-of-merit for the sensitivity of the receiver to be simplified to $(f_T/\Gamma \cdot C_D)$, which represents the performance of the specific OE-VLSI technology in terms of the detector capacitance, the unity-gain bandwidth of the FET's, and the channel noise factor of the FET's. This figure-of-merit provides an indication of the improvement in receiver performance that we may expect as the OE-VLSI technology is scaled. The extent to which the actual receiver sensitivity will track this technology figure-of-merit (assuming the technology scales are as expected) depends upon limits placed on the specific receiver circuit, including power dissipation constraints, and area.

The noise current can be calculated from (13)–(15), where the transconductance g_m is given by [103]

$$\begin{aligned} g_m &= \beta(V_{\rm gs} - V_t)(\text{saturation}) \\ &= \frac{\mu\varepsilon}{t_{\rm ox}}\left(\frac{W}{L}\right)(V_{\rm gs} - V_t)(1 + \lambda' V_{\rm ds}) \end{aligned} \tag{22}$$

where λ' is the channel-length modulation parameter; second-order effects due to gate electric field and source-drain electric fields have been neglected. Empirical data and trend lines for the saturation transconductance of both NMOS's and PMOSFET's are shown in Fig. 3(a)–(b). Equations (12)–(14) determine the minimum input current required to achieve the requisite bit-error rate (BER). In calculating the noise limited receiver input current, it has been assumed that the $1/f$ noise source associated with CMOS circuits could be neglected (or otherwise circumvented, for instance, by appropriate data-encoding). As previously mentioned, the receivers in an array are typically operated well above the noise-limited regime, so that the input photocurrent I_d obeys the condition $I_d \geq 8i_n$ for a BER of 10^{-15}. Differential signaling is assumed so that the photocurrent swing is doubled to a factor of 16 larger than the rms noise current. For a given input current at the operating bit rate (or switching energy), the effective input voltage swing (ΔV) can be written as

$$\Delta V = \frac{I_d \Delta T}{C_T} = \frac{I_d}{C_T B} \tag{23}$$

where ΔT is the bit duration. The output voltage of the receiver front end (V_2) can then be calculated from (5) as a function of the gain of the amplifier.

The sensitivity of the receiver can be expressed as either an optical power required at the given bit rate (in dBm) or as an optical energy/bit $(E_{\rm opt})$ required at the receiver for the given bit rate. For receivers with electronic gain, $E_{\rm opt}$ will typically be dependent on the bit rate. The optical energy required per bit can be expressed as a function of the average optical power in the beam, the fraction of time that the light is present during the received data stream, and the bit period as [104]

$$E_{\rm opt} = P_{\rm avg} \cdot \Delta T \cdot \frac{N_b}{N_p} \tag{24}$$

where $P_{\rm avg}$ is the average optical power in the input beam that will guarantee a specific bit-error rate, N_b is the number of transmitted bits, N_p is the number of light pulses, and ΔT is the bit period. $P_{\rm avg}$ in (24) can be replaced by $I_{\rm avg}/\eta$, where $I_{\rm avg}$ is the average photocurrent in the receiver, and η is the responsivity of the detector. For random NRZ data with single-ended receivers, the ratio N_b/N_p is equal to two, and ΔT is the inverse of the operating bit rate B. For NRZ data with two-beam, differential receivers, the ratio N_b/N_p per detector is equal to two, and an effective $E_{\rm opt}$ per beam can be defined. It will be the convention in this paper to quote the energy *per* beam when calculating sensitivities.

C. Scalability of OE-VLSI Transceiver Circuits

Given the requirements for the input photocurrent in the receivers, we can calculate the number of cascaded amplifier (inverter) stages at the given bandwidth needed to

produce the full logic swing at the output of the receiver. This will determine the overall power dissipation and area of the receiver circuit. The transistor f_T calculated in (21) is a small-signal parameter corresponding to unity current gain bandwidth (whereas we are using a voltage amplifier). The unity voltage-gain frequency of a series of identical cascaded inverter stages, denoted by f_T', represents the highest frequency at which gain can be extracted from the inverter [105]; this is typically layout-dependent and difficult to express in closed form. However, it can be approximated by the gate delay of the technology as determined by ring-oscillator frequency measurements (Fig. 5); this empirical data also captures the effect of parasitic interconnect capacitances. From this data, the gain available from one stage of the amplifier can be calculated. The logarithm of the ratio of the voltage swing required at the receiver output $(V_{\rm dd})$ to the effective input voltage swing from (23), taken to the base of the gain available from one such amplifier stage, represents the number of such stages that will be needed in the receiver front end. The total power dissipation of the receiver follows from the dissipation per stage and the number of required stages, the latter growing logarithmically with reducing receiver input energy. A similar scaling applies for the receiver area.

We assume that the receiver front end is biased in its saturation region at approximately half the supply voltage. The static power dissipation of the two-beam receiver with a single biased input stage can be calculated from

$$P_{\rm diss,rec} = I_{\rm rec} \cdot V_{\rm dd} = \frac{1}{2}\beta \cdot (V_{\rm gs} - V_t) \cdot V_{\rm dd} \qquad (25)$$

where $I_{\rm rec}$ is the current flowing through the biased transimpedance input stage. Equation (25) is valid for current-source load or for a push–pull inverter biased to allow the drain-source currents in the PMOS and NMOS devices to be equal. Making the simplifying assumption that the current-gain factors (β) and the absolute thresholds $(|V_t|)$ are equal for both NMOS and PMOS devices, this allows the inverter to be biased at $V_{\rm dd}/2$. Note that this expression applies only when the circuit is in its quiescent state. Once the receiver switches to a stable state corresponding to a received bit, this static power dissipation component reduces due to a small excursion from its quiescent bias point; the corresponding dynamic power dissipation of the receiver circuit increases according to the bit rate. At the operating bit rates assumed in this paper, the dominant source of the receiver power dissipation is the large static current; the worst-case power dissipation is thus dictated by (25). Fig. 13 shows the receiver power dissipation versus optical energy required at the receiver. For a specific linewidth technology, there is a tradeoff between the optical energy-per-bit (or sensitivity) required by a receiver and its electrical power dissipation. Reducing the optical energy (increasing the sensitivity) results in a lower input voltage swing (ΔV). This can be compensated by increasing the gain in the receiver, which leads to a larger number of gain stages and hence a higher power dissipation. The calculated data points represent an integer number of gain stages in the receiver for the given front-end FET geometries; the front-end FET's can be changed to allow different values of gain. The calculated data points as well as a logarithmic fit to the data (representing different values of gain per stage) have been plotted.

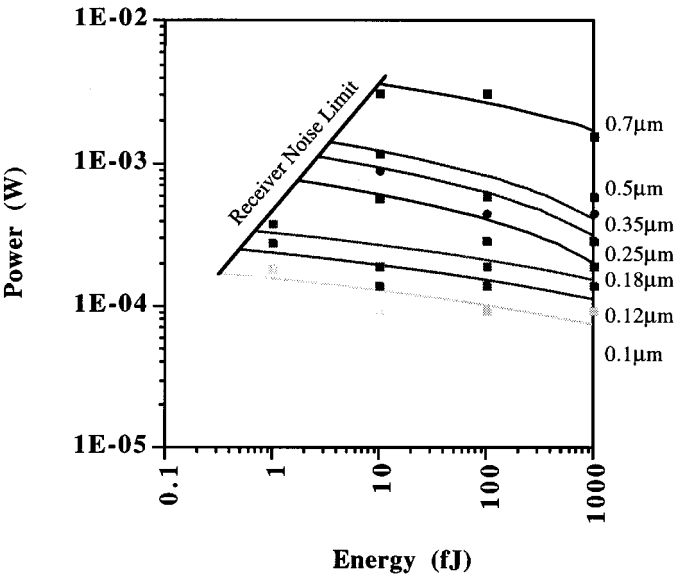


Fig. 13. Receiver power dissipation versus optical energy-per-bit required at receiver. Receiver noise limits represent an operating BER of 10^{-15}.

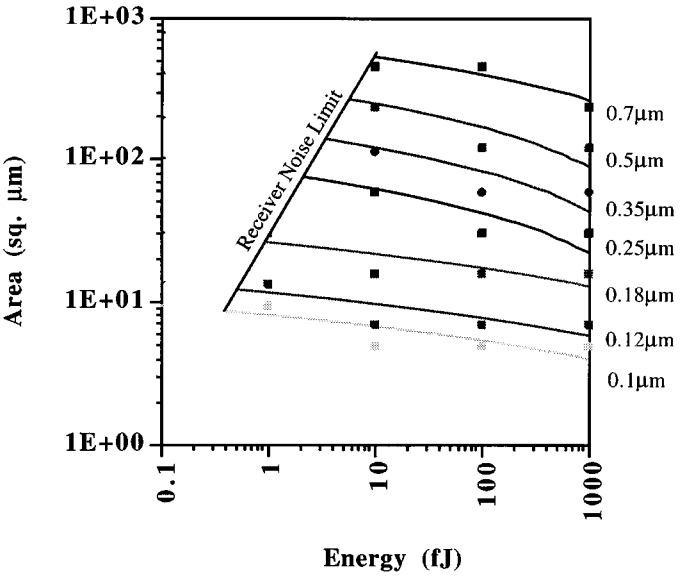


Fig. 14. Receiver area versus optical energy-per-bit required at receiver.

Fig. 14 similarly plots the receiver area versus optical energy required at receiver. Because of the very simple receiver and transmitter circuits, the number of transistors used for the optoelectronic transceivers is only a negligible fraction of the overall number of transistors available. Thus, the total number of transistorson the CMOS chip is not expected to limit the optoelectronics. From Fig. 14, it is clear that the area of the receivers is typically a minute fraction of the total chip area and therefore is not a limiting factor for the optical I/O bandwidth. However, the issue of optical array size may be a potential concern. Large-area chips can present difficulties in terms of the optical system packaging. Traditional smart-pixel designs have emphasized uniform and possibly dilute arrays of receivers and transmitters with circuits placed in

close proximity to their receivers and transmitters. These and other design choices, optimized for the optical interconnection system, are also possible since there are few restrictions in the placement of the optical diodes with respect to the underlying CMOS circuits.

We note that the use of optical interconnects can promote a specific technology to operate at bit rates well beyond the qualified speed of the technology given in Table I. Experiments in 0.8-μm CMOS technology have demonstrated the ability to operate receiver and transmitter circuits beyond 500 MHz (1 Gb/s NRZ data [22]), which is over five times the qualified clock frequency of the CMOS technology. This frequency is approximately $f_{T'}/20$, where $f_{T'}$ is the reciprocal gate delay. When operating at such high bit rates, the dynamic switching in the tranceivers will lead to significant power consumption. However, the opportunity exists to reduce the number of tranceivers on the chip in favor of a higher bit rate per link, while keeping the overall optical I/O bandwidth to the chip constant. Dynamic power dissipation data versus bit rate for an experimental 0.8-μm CMOS receiver and transmitter are provided in [80].

The transmitter (modulator) driver circuits will also consume a fraction of the total power budget. Driver circuits for the hybrid MQW modulators are typically comprised of a simple inverter buffer sized to drive the modulator capacitance and sink the photocurrent absorbed in the low-reflectivity state of the modulator [106], [107]. The modulator driver circuit power dissipation (assuming a two-beam transmitter) can be written as

$$P_{\text{diss,trans}} = C_{\text{trans}} \cdot (V_{\text{dd}})^2 \cdot f + \frac{I_d |V_{\text{mod}}|}{n_{\text{int}}} \tag{26}$$

where

$$n_{\text{int}} = n_{\text{optics}} \cdot n_{\text{mod}} \tag{27}$$

where C_{trans} is the total transmitter capacitance to be switched (including driver capacitance), I_d is the peak photocurrent required in the following detector, and $|V_{\text{mod}}|$ is the voltage difference (bias) across the modulator in its highly absorbing state. n_{int} is the product of the optical system link efficiency (n_{optics}), and the modulator efficiency (n_{mod}), given by the reflectivity difference between its two operating states:

$$n_{\text{mod}} = P_{\text{Hi/In}} - P_{\text{Low/In}} \tag{28}$$

where $P_{\text{Hi/In}}$ and $P_{\text{Low/In}}$ are, respectively, the fractions of the optical input power that are available in the corresponding high and low reflectivity output states of the modulator. Note that the modulator efficiency includes the effect of the insertion loss and the finite contrast ratio of the modulator device. It is assumed that the voltage swing across the modulators is the supply voltage, V_{dd}, and that the bias voltage V_{mod}, across the absorbing diode in the transmitter circuit is twice the supply voltage. The second term in (26) reflects the fact that half the combined input power to both diodes of a differential transmitter is always being absorbed, regardless of the bit pattern. For single-ended transmitters, (26) is valid when I_d is replaced with the average photocurrent. A calculation of the modulator-driver dissipation indicates that the second term of (26) dominates when the receiver to which it is optically connected has a relatively large switching energy (≈1 pJ). This is because the modulator-driver circuit is forced to sink a large current during the low-reflectivity state of the modulator. When this happens, the transmitter power dissipation (Fig. 15) can equal or even exceed the receiver power dissipation. But as the receiver sensitivity is increased, the corresponding transmitter power dissipation becomes a smaller fraction of the receiver power dissipation. It should be noted that MQW absorption saturation at high optical intensity levels will not typically be a concern for the scaling parameters assumed in Table II. Even at the lower sensitivities (100 fJ/bit–1 pJ/bit) and the smallest geometries considered (4-μm × 4-μm active area for 0.1-μm CMOS), the required light intensity at the MQW modulators is below the saturation limit (80 kW/cm^2). At sufficiently large optical energies, one might then imagine it feasible to operate in a receiver-less mode, with the incoming optical energy being sufficient to swing the input capacitance of the receiver by a threshold voltage or higher. Such a mode would eliminate the high receiver power dissipation. However, for receiver switching energies above 1 pJ, the modulator-driver power dissipation due to the absorbed photocurrent, and modulator inefficiencies related to absorption saturation, can make the overall transmitter circuit dissipation unacceptably high. Saturation would also be an issue for interconnect systems that use a large optical fanout per channel [107]. As will be discussed below, of even greater concern are heating effects due to the thermal resistance of the hybrid devices. For cascaded systems, these arguments present a case against operating the optical interconnect without a receiver front end, in a mode that would require high optical powers. Even if sufficient laser power was available, the benefit of the reduction in receiver power dissipation is negated by the increase in modulator driver dissipation.

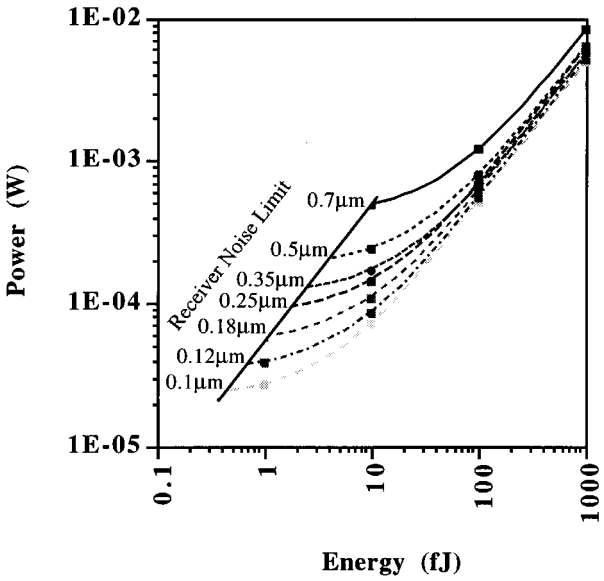


Fig. 15. Transmitter (modulator) power dissipation versus optical energy-per-bit required at receiver.

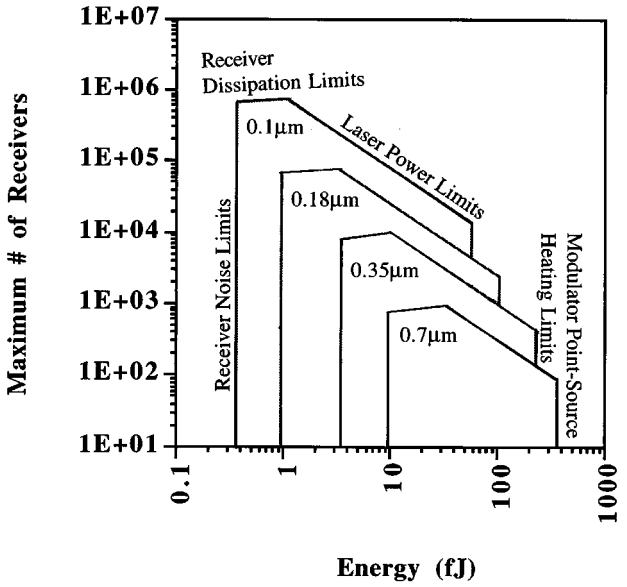


Fig. 16. Maximum number of receivers that can be supported on an IC versus optical energy required at individual receiver. Receiver power dissipation and optical power limit the total number of receivers. Circuit noise limits and diode thermal limits constrain the sensitivity of each receiver. Curves are shown only for 0.7-μm, 0.35-μm, 0.18-μm, and 0.1-μm technologies.

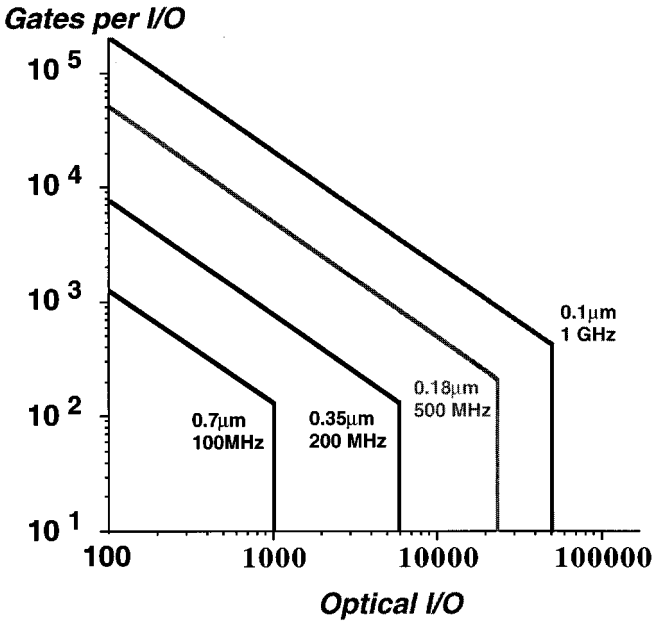


Fig. 17. Limits on the number of optical I/O devices and gates-per-I/O device that may be supported on the OE-VLSI chip. Sloping lines represent VLSI limits, while vertical lines correspond to OE limits. Curves are shown only for 0.7-μm, 0.35-μm, 0.18-μm, and 0.1-μm technologies.

V. Grain-Size Complexity and I/O Bandwidth Limits for Optoelectronic-VLSI Circuits

Assuming two IC's populated with modulator-drivers and receivers, respectively, (or IC's containing both modulator-drivers and receivers and sharing the use of the diode array), the aggregate I/O bandwidth to the chip will primarily be limited by: 1) area and power consumption of the receivers; 2) the availability of sufficient laser power to interrogate the modulators on the transmitting IC and transmit the information to their corresponding receivers on the receiving IC; and 3) the total number of optical diodes available. We will focus first on the maximum number of receivers that can be supported on an optoelectronic IC.

The limit on the number of receivers that can be accommodated on a single IC can be determined from (25) and (26), assuming that the total power dissipation budget is given (Table I) and that 50% of the budget may be used for the optical transceivers.

Fig. 16 shows the tradeoff between the maximum number of receivers on an optoelectronic IC and the optical energy required to switch each receiver in a cascaded system, assuming up to 50% of the available power budget (from Table I) can be spent on I/O to the chip. Limits due to receiver power dissipation and due to optical power limits (diagonal lines) are shown for 0.7 μm, 0.35 μm, 0.18 μm, and 0.1 μm technologies, respectively. All receivers on an OE-VLSI chip are assumed to be identical. As the figure indicates, the power dissipation of the receiver versus its optical input switching energy can be optimized in order to maximize the number of optical I/O to the IC. A useful conclusion is that a careful optimization of receiver power dissipation versus sensitivity is an important part of the system design. This is especially true in the earlier CMOS generations as well as in situations where advanced cooling techniques cannot be used, and the acceptable power dissipation (and therefore the number of receivers) may differ from the values assumed in Table I.

The design space for the hybrid OE-VLSI technology can be represented in terms of the number of optical I/O per IC and the complexity per optical I/O (or "grain size") measured in terms of the number of logic gates or transistors per I/O. The issue of optimum complexity per optical I/O has been the subject of several studies [108]–[111] because of its special relevance to systems design; typically the choice of grain size will be driven by the application. However, bounds on this complexity can be derived based on the scaling of the respective optoelectronic and VLSI technologies. As shown in Fig. 17, these bounds lead to a user design space for OE-VLSI technologies. For clarity, only curves for alternate linewidth technologies (i.e., 0.7 μm, 0.35 μm, 0.18 μm, and 0.1 μm) are shown. The expected limits on the number of optical I/O and gates-per-I/O that may be supported for each generation of MQW-modulator-based OE-VLSI technology are delineated on the graph. Sloping lines represent VLSI limits while vertical lines correspond to OE limits.

Assuming that the receiver-power-to-laser-power tradeoff was optimized according to Fig. 16, the limits to the number of optical I/O (vertical line) arise primarily from the maximum number of diodes, the receiver power dissipation, and the available laser power. The range of acceptable receiver dissipation and sensitivity grows as a function of the difference between the diode-yield limit to the number of optical I/O and the corresponding limit due to receiver/transmitter circuit power dissipation. As shown in Fig. 17, the diode limit in succeeding generations of the OE-VLSI technology causes the maximum grain size at the maximum number of optical I/O to monotonically increase. This is because the CMOS

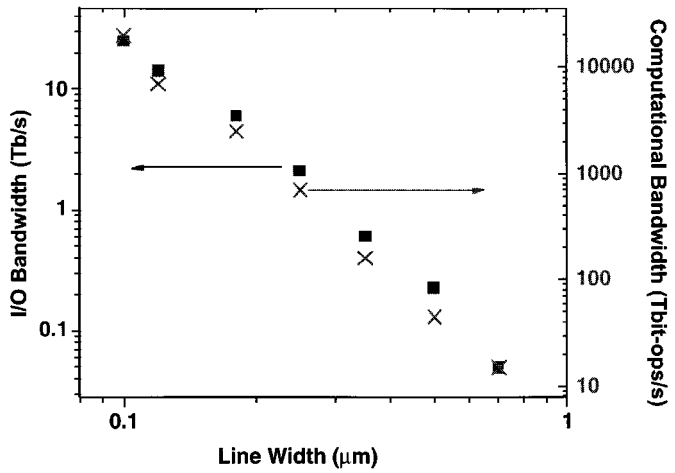


Fig. 18. The I/O bandwidth and computational bandwidth of an OE-VLSI chip versus CMOS linewidth, assuming differential optical signalling. The growth of the I/O bandwidth of the OE-VLSI chip matches the growth of the computational bandwidth of the silicon technology.

technology yield is expected to improve slightly faster than the corresponding hybrid GaAs MQW technology. Nevertheless, it appears that hybrid OE-VLSI technologies not only have substantial room for growth in I/O capabilities but will also be able to track the progress in the computation power of the silicon technology itself. This is shown is Fig. 18, where the I/O bandwidth and computational bandwidth of an OE-VLSI chip is plotted versus CMOS linewidth. Computational bandwidth is defined as the product of the number of gates and the clock frequency of the CMOS technology; the I/O bandwidth is defined as the product of the maximum number of optical I/O and the CMOS clock frequency. Assuming a fixed transceiver power dissipation budget of 50%, Figs. 17 and 18 suggest that optical interconnects can be expected to provide an aggregate data I/O bandwidth of 0.5 Tb/s to 0.7-μm CMOS VLSI circuits, and continue to match the exponential increase in computational bandwidth providing as much as 25 Tb/s data I/O to future VLSI circuits in 0.1-μm technology.

VI. Discussion and Conclusion

We have detailed a possible evolution of an OE-VLSI technology that is based on the hybrid attachment of GaAs–AlGaAs MQW diodes over submicron silicon CMOS circuits. This paper concentrated on the circuits and device technology that will be critical to this evolution. Technologies that can deliver a large number of optical beams to such OE IC's were not explicitly discussed here, and we acknowledge that progress in optical interconnect technologies such as diffractive optics and microoptics will also play a key role in future systems. The assumptions made in this study pertain to cascaded system operation, which is generally more demanding on optical and optoelectronic technology performance than single-stage systems. The discussion of laser power limits and particularly the discussion of the tradeoff between laser power and transceiver circuit dissipation are, in general, relevant only to cascaded systems. However, the treatment of the scaling of the power dissipation of a large array of high-performance receivers is also germane to single-stage systems and to systems based on other transmitter technologies.

The main barriers to continued scaling of OE-VLSI circuits based on modulator technology can be expected to come from the yield of the modulators, the finite laser source power, and the on-chip power-dissipation of the I/O circuits. Operation in a receiver-less mode with energies much greater than 1 pJ/bit will not be feasible due to large transmitter power dissipation and related thermal effects in the modulators, as well as limited laser power available for illumination. When operating at small optical energies-per-bit ($\ll$1 pJ/bit), the receivers are a greater concern in terms of electrical power dissipation than the modulator driver circuits. The key reason for the receiver power dissipation is that the circuits must be biased as small-signal amplifiers, resulting in a steady dissipation of power (in contrast to the CMOS logic circuts that generally dissipate only during switching). Indeed, a general conclusion is that the receiver circuits will be a significant source of electrical power dissipation for any OE-VLSI technology that attempts to provide a large number of low-energy, high-speed surface-normal optical links.

As expected, we found that reductions in silicon feature sizes and optoelectronic device dimensions would serve to increase the aggregate optical interconnect bandwidth to a VLSI chip by improving the sensitivity and bit rate of the receivers and by reducing the power dissipation of the transceivers. For a given receiver front end, a technology figure-of-merit, based on the gain bandwidth of the FET, the channel noise factor of the FET, and the detector capacitance, can be defined; this figure-of-merit is readily computed for different OE-VLSI technologies and provides a measure of the maximum receiver sensitivity.

For the simple receiver designs considered in this paper, the area of the receiver and transmitter circuits are not expected to be limiting factors. The results indicate certain operating ranges in terms of the sensitivity or optical energy-per-transmitted-bit for the optical interconnections; too low an energy per bit would result in an unacceptably high electrical power dissipation in the receivers and ultimately in errors due to insufficient signal-to-noise ratio; too large an energy per bit would result in excessive power dissipation in the corresponding transmitters and ultimately to link failure due to point-source heating in the optical modulators. By normalizing to the operating bit rate, this result can instead be interpreted as an operating range for input optical power, or detector photocurrent. Within this operating range for each FET technology, we found that the sensitivity of the receiver can be optimized to balance circuit power dissipation versus the required input optical power, thereby maximizing the number of I/O circuits on a chip. The $1/f$ or flicker noise in FET's was not explicitly considered in this paper. If it appears as a practical problem, it may be possible to circumvent the deleterious effects of this noise source (and also to reduce point-source heating effects in the modulators) with appropriate data-encoding techniques.

The assumption was made that 50% of the available power consumption budget would be reserved for the optoelectronic transceivers. Although this assumption is relatively conserva-

tive when compared to custom submicron electronic chips, individual OE-VLSI circuit applications would be expected to have differing requirements; this would directly affect the optimum receiver sensitivity. Assuming that the receiver sensitivity is optimized, the limit to the number of optical I/O is primarily due to the number of optical diodes available, though receiver power dissipation is close to the allowable limits also. Nevertheless, it appears that the growth of the I/O bandwidth of the OE-VLSI chip can be expected to match the growth of the computational bandwidth of the silicon technology. Indeed, an aggregate I/O bandwidth (assuming two-beam operation) on the order of 0.5 Tb/s with 0.7-μm CMOS and up to 25 Tb/s with future 0.1-μm CMOS technologies appears to be possible, despite some relatively conservative assumptions on optoelectronic device performance.

The receiver design considered in this paper was based on an asynchronous transimpedance feedback amplifier that has proven characteristics of high bandwidth, dynamic range, and tolerance to device mismatch. This choice of receiver front end allowed a reasonably accurate estimate to be made for the RMS noise currents, and hence the receiver sensitivity. However, like all circuits that require biasing to achieve high sensitivity, the transimpedance amplifier draws static current which leads to static power dissipation. It is possible to trade in the static power dissipation of the receiver for bandwidth, sensitivity, or system complexity. Designs for alternate receiver circuits are explored in greater depth elsewhere [112]. One method to reduce this static power dissipation is to use timing information that may be available in the system. This would allow the power supplies to be clocked so as to reduce the active biasing period. This is the rationale behind the use of synchronous sense-amplifier receivers circuits as discussed in [113]. Naturally, this improvement in power dissipation would come at the price of more exacting system timing constraints and possibly reduced receiver bandwidth. While this tradeoff may be particularly advantageous for synchronous system operation, it may be unsuitable in instances where precise timing information is unavailable. To preserve the generality of the paper, we chose to model an asynchronous circuit such as the transimpedance receiver. The details of the derivation and results presented in this paper would allow the reader to readily interpret the effects of a constant factor in power dissipation afforded by specific optimizations of the transceiver circuits.

In terms of the fundamental technology, the challenges are in reducing the drive voltages of the modulators to stay compatible with mainstream CMOS, to provide sufficient laser power, and to continue to improve the yield in the manufacturing and hybridizing of the MQW diodes. An important general conclusion that does not depend critically on the detailed assumptions is that it appears this hybrid optical I/O technology has substantial room for continued scaling to larger numbers of higher speed interconnects as silicon technology itself advances.

ACKNOWLEDGMENT

The authors express their gratitude to T. Woodward, A. Lentine, and B. Mansoorian for many illuminating discussions on receivers; to K. Devenport for his assistance in the analysis and simulations of transimpedance receiver stability; to K. Goossen for profitable discussions on the operating voltage and yield of the hybrid MQW modulators, and especially to R. Novotny for valuable input on point source heating effects and the scaling of the MQW diode capacitance and thermal resistance. The authors also thank two anonymous reviewers for their comments on the manuscript.

REFERENCES

[1] Semiconductor Industry Association, *The National Technology Roadmap for Semiconductors*, 1993–1994, p. B2.
[2] K. M. Johnson, M. A. Handschy, and L. A. Pagano-Stauffer, "Optical computing and image processing with ferroelectric liquid crystal," *Opt. Eng.*, vol. 26, no. 5, pp. 385–391, 1987.
[3] T. J. Drabik and M. A. Handschy, "Silicon VLSI/ferroelectric liquid crystal technology fro micropower optoelectronic computing devices," *Appl. Opt.*, vol. 29, no. 35, pp. 5220–5223, Dec. 1990.
[4] M. J. Goodwin, C. J. K. Kirkby, A. D. Parsons, I. Bennion, and W. J. Stewart, "8×8 element PLZT/silicon spatial light modulator," *Electron. Lett.*, vol. 25, p. 1260, 1989.
[5] B. Mansoorian, V. Ozguz, C. Fan, and S. Esener, "Design and implementation of flip-chip bonded Si/PLZT smart pixels," in *Proc. IEEE LEOS Summer Top. Meet. on Smart Pixels*, Santa Barbara, CA, Aug. 1992, Paper MC2.
[6] N. M. Jokerst, C. Camperi-Ginestet, B. Buchanan, S. Wilkenson, and M. A. Brooke, "Communication through stacked silicon circuitry using integrated thin film InP-based emitters and detectors," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 1028–1030, Sept. 1995.
[7] J. L. Jewell, K. F. Huang, K. Tai, Y. H. Lee, R. J. Fischer, S. L. McCall, and A. Y. Cho, "Vertical cavity single quantum well laser," *Appl. Phys. Lett.*, vol. 55, pp. 424–426, 1989.
[8] R. Geels, S. Corzine, and L. Coldren, "InGaAs vertical-cavity surface emitting lasers," *IEEE J. Quantum Electron.*, vol. 27, pp. 1359–1367, June 1991.
[9] C. Edge, R. M. Ash, G. G. Jones, and M. J. Goodwin, "Flip-chip solder bond mounting of laser diodes," *Electron. Lett.*, vol. 27, pp. 499–500, 1991.
[10] S. Daryanani, H. Fathollahnejad, D. L. Mathine, R. Droopad, A. Kubes, and G. N. Maracas, "Integration of a single vertical-cavity surface emitting laser onto a CMOS inverter chip," *Electron. Lett.*, vol. 31, no. 10, pp. 833–834, 1995.
[11] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus, "Novel hybrid optically bistable switch: The quantum well self-electro-optic effect device," *Appl. Phys. Lett.*, vol. 45, no. 1, pp. 13–15, 1984.
[12] M. J. Goodwin, A. J. Moseley, M. Q. Kearley, R. C. Morris, C. J. G. Kirkby, J. Thomson, R. C. Goodfellow, and I. Bennion, "Optoelectronic component arrays for optical interconnection of circuits and subsystems," *J. Lightwave Technol.*, vol. 9, pp. 1639–1645, Dec. 1991.
[13] K. W. Goossen, J. E. Cunningham, and W. Y. Jan, "GaAs 850 nm modulators solder-bonded to silicon," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 776–778, 1993.
[14] C. Kyriakakis, Z. Karim, A. R. Tanguay, Jr., R. F. Cartland, A. Madhukar, S. Piazolla, B. K. Jenkins, C. B. Kuznia, A. A. Sawchuk, and C. Von der Malsburg, "Photonic implementations of neural networks," *Tech. Dig. OSA Top. Meet. Optical Computing*, Salt Lake City, UT, Mar. 1995, pp. 128–130.
[15] M. Hansen, D. Shih, C. Fan, and S. Esener, "16×16 SLM with silicon CMOS drivers and III-V modulators," in *Proc. OSA Top. Meet. Spatial Light Modulators and Applications*, Salt Lake City, UT, Mar. 1995, pp. 5–8.
[16] T. L. Worchesky, K. J. Ritter, R. Martin, and B. Lane, "Large arrays of spatial light modulators hybridized to silicon integrated circuits," *Appl. Opt.*, vol. 35, no. 8, pp. 1180–1186, Mar. 1996.
[17] K. W. Goossen, J. A. Walker, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. D. Bacon, D. Dahringer, L. M. F. Chirovsky, A. L. Lentine, and D. A. B. Miller, "GaAs MQW modulators integrated with silicon CMOS," *IEEE Photon. Technol. Lett.*, vol. 7, no. 4, pp. 360–362, Apr. 1995.
[18] A. L. Lentine and D. A. B. Miller, "Evolution of the SEED technology: Bistable logic gates to optoelectronic smart pixels," *IEEE J. Quantum Electronics*, vol. 29, pp. 655–669, Feb. 1993.
[19] A. V. Krishnamoorthy, A. L. Lentine, K. W. Goossen, J. A. Walker, T. K. Woodward, J. E. Ford, G. F. Aplin, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L. M. F. Chirovsky,

and D. A. B. Miller, "3-D integration of MQW modulators over active sub-micron CMOS circuits: 375 Mb/s transimpedance receiver-transmitter circuit," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 1288–1290, Nov. 1995.

[20] A. V. Krishnamoorthy, J. E. Ford, K. W. Goossen, J. A. Walker, A. L. Lentine, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L. M. F. Chirovsky, F. E. Kiamilev, G. F. Aplin, R. G. Rozier, and D. A. B. Miller, "Photonic page buffer based on GaAs MQW modulators bonded directly over active silicon CMOS circuits," *Appl. Opt.*, vol. 35, no. 14, pp. 2439–2448, May 1996.

[21] A. L. Lentine, K. W. Goossen, J. A. Walker, L. M. F. Chirovsky, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, J. E. Cunningham, W. Y. Jan, J. M. Kuo, D. Dahringer, D. Kossives, D. Bacon, G. Livescue, R. L. Morrison, R. A. Novotny, and D. B. Buchholz, "Optoelectronic VLSI switching chip with greater than 4000 optical I/O based on flip-chip bonding of GaAs/AlGaAs MQW modulators and detectors to silicon CMOS," *Conf. Lasers and Electro-Opt. (CLEO-'96)*, May 1996, vol. 9, pp. 517–518; see also this issue, pp. 77–84.

[22] T. K. Woodward, A. V. Krishnamoorthy, A. L. Lentine, K. W. Goossen, J. A. Walker, J. E. Cunningham, W. Y. Jan, L. A. D'Asaro, L. M. F. Chirovsky, S. P. Hui, B. Tseng, D. Kossives, D. Dahringer, and R. Leibenguth, "1 Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW modulators bonded to 0.8 μm silicon CMOS," *IEEE Photon. Technol. Lett.*, vol. 8, pp. 422–444, Mar. 1996.

[23] D. A. B. Miller and H. Ozaktas, "Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture," *J. Parallel and Distributed Computing*, Special Issue on Optical Interconnects, to be published.

[24] E. J. Novak, "Ultimate CMOS ULSI performance," *IEDM Tech. Dig.*, pp. 115–118, 1993.

[25] B. Davari, R. H. Dennard, and G. G. Shahidi, "CMOS scaling for high performance and low power—The next ten years," *Proc. IEEE*, vol. 83, pp. 595–606, Apr. 1995.

[26] J. D. Meindl, "Low power microelectronics, retrospect and prospect," *Proc. IEEE*, vol. 83, pp. 619–635, Apr. 1995.

[27] M. Bohr, "MOS transistors: Scaling and performance trends," *Semicond. Int.*, pp. 75–78, June 1995.

[28] Y. Taur, Y.-J. Mii, D. J. Frank, H.-S. Wong, D. A. Buchanan, S. J. Wind, S. A. Rishton, G. A. Sai-Halasz, and E. J. Nowak, "CMOS scaling into the 21st century: 0.1 μm and beyond," *IBM J. Res. Develop.*, vol. 39, no. 1/2, Jan./Mar. 1995.

[29] G. Groeseneken, R. Bellens, G. Van den Bosch, and H. E. Maes, "Hot carrier degradation in submicrometer MOSFETs: From uniform injection toward the real operating conditions," *Semicond. Sci. Technol.*, vol. 10, pp. 1208–1220, 1995.

[30] C. A. Warwick, R. H. Yan, Y. O. Kim, and A. Ourmazd, "Toward giga-scale silicon integrated circuits," *AT&T Tech. J.*, pp. 50–59, Sept. 1993.

[31] R. H. Yan, D. Monroe, J. Weis, A. Mujtaba, and E. Westerwick, "Reducing operating voltage from 3, 2, to 1 volt and below: Challenges and guidelines for possible solutions," *IEDM Tech. Dig.*, pp. 55–58, 1995.

[32] N. H. E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design: A Systems Perspective*. Reading, MA: Addison-Wesley, 1993.

[33] A. Hori, H. Nakaoka, H. Umimoto, K. Yamashita, M. Takase, N. Shimizu, B. Mizuno, and S. Odanaka, "A 0.05 μm CMOS with ultra-shallow source/drain junctions fabricated by 5 KeV ion implantation and rapid thermal annealing," *IEDM Tech. Dig.*, pp. 485–488, 1994.

[34] M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, and H. Iwai, "Sub-50 nm gate-length N-MOSFET's with 10 nm phosphorus source and drain junctions," *IEDM Tech. Dig.*, pp. 119–122, 1993.

[35] Y. Taur, S. Wind, Y. J. Mii, Y. Lii, D. Moy, K. A. Jenkins, C. L. Chen, P. J. Coane, D. Klaus, J. Bucchignano, M. Rosenfeld, M. G. R. Thomson, and M. Polcari, "High-performance 0.1 μm devices with 1.5 V power supply," *IEDM Tech. Dig.*, pp. 127–130, 1993.

[36] G. E. Rittenhouse, W. M. Mansfield, A. Kornblit, R. A. Cirelli, D. Tomes, and G. K. Celler, "Sub-0.1 μm NMOS transistors fabricated using laser plasma point-source X-Ray lithography," *IEEE Electron Device Lett.*, vol. 16, pp. 322–324, July 1995.

[37] J. Lyu, B.-G. Park, K. Chun, and J. D. Lee, "A novel 0.1 μm MOSFET structure with inverted sidewall and recessed channel," *IEEE Electron Device Lett.*, vol. 17, pp. 157–159, Apr. 1996.

[38] M. Saito, T. Yoshitomi, M. Ono, Y. Akasaka, H. Nii, S. Matsuda, H. S. Momose, Y. Katsumata, Y. Ushiku, and H. Iwai, "An SPDD P-MOSFET structure suitable for 0.1 and sub 0.1 micron channel length and its electrical characteristics," *IEDM Tech. Dig.*, pp. 897–900, 1992.

[39] K. F. Lee, R. H. Yan, D. Y. Jeon, G. M. Chin, Y. O. Kim, D. M. Tennant, B. Razavi, H. D. Lin, Y. G. Wey, E. H. Westerwick, M. D. Morris, R. W. Johnson, T. M. Liu, M. Tarsia, M. Cerullo, R. G. Swartz, and A. Ourmazd, "Room temperature 0.1 μm CMOS technology with 11.8 ps delay," *IEDM Tech. Dig.*, pp. 131–134, 1993.

[40] R. H. Yan, K. F. Lee, D. Y. Jeon, Y. O. Kim, B. G. Park, M. R. Pinto, C. S. Rafferty, D. M. Tennant, E. H. Westerwick, G. M. Chin, M. D. Morris, K. Early, P. Mulgrew, W. M. Mansfield, R. K. Watts, A. M. Voshchenkov, J. Bokor, R. G. Swartz, and A. Ourmazd, "89 GHz f_T room-temperature silicon MOSFET's," *IEEE Electron Device Lett.*, vol. 13, pp. 256–258, May 1992.

[41] K. F. Lee, R. H. Yan, D. Y. Jeon, Y. O. Kim, D. M. Tennant, E. H. Westerwick, K. Early, G. M. Chin, M. D. Morris, R. W. Johnson, T. M. Liu, R. C. Kistler, A. M. Voshchenkov, R. G. Swartz, and A. Ourmazd, "0.1 μm p-channel MOSFET's with 51 GHz f_T," *IEDM Tech. Dig.*, pp. 1012–1014, 1992.

[42] G. G. Shahidi, J. Warnock, S. Fischer, P. A. McFarland, A. Acovic, S. Subbanna, E. Ganin, E. Crabbé, J. Comfort, J. Y.-C. Sun, T. H. Ning, and B. Davari, "High-performance devices for a 0.15 μm CMOS technology," *IEEE Electron Device Lett.*, vol. 14, pp. 466–468, Oct. 1993.

[43] K. Takeuchi, T. Yamamoto, A. Tanabe, T. Matsuki, T. Kunio, M. Fukuma, K. Nakajima, H. Aizaki, H. Miyamoto, and E. Ikawa, " 0.15 μm CMOS with high reliability and performance," *IEDM Tech. Dig.*, pp. 883–918, 1993.

[44] S. Subbanna, E. Ganin, E. Crabbé, J. Comfort, S. Wu, P. Agnello, B. Martin, M. McCord, H. Ng, T. Newman, P. McFarland, J. Sun, J. Snare, A. Acovic, A. Ray, R. Gehres, R. Schulz, S. Greco, K. Beyer, L. Liebmann, R. DellaGuardia, and A. Lamberti, "200 nm process integration for a 0.15 μm channel-length CMOS technology using mixed X-ray/optical lithography," *IEDM Tech. Dig.*, pp. 695–698, 1994.

[45] S. Shizumi, T. Kuroi, M. Kobayashi, T. Yamaguchi, T. Fujino, H. Maeda, T. Tsutsumi, Y. Hirose, S. Kusunoki, M. Inuishi, and N. Tsubouchi, "0.15 μm CMOS process for high performance and high reliability," *IEDM Tech. Dig.*, pp. 67–70, 1994.

[46] M. Rodder, S. Aur, and I.-C. Chen, "A scaled 1.8 V, 0.18 μm gate length CMOS technology: Device design and reliability considerations," *IEDM Tech. Dig.*, pp. 415–418, 1995.

[47] T. Mogami, H. Wakabayashi, Y. Saito, T. Matsuki, T. Tatsumi, and T. Kunio, "A novel salicide process for sub-quarter micron CMOS devices," *IEDM Tech. Dig.*, pp. 687–690, 1994.

[48] T. Izawa, M. Katsube, Y. Yokoyama, K. Hashimoto, E. Kawamura, A. Shimizu, H. Takagi, F. Inoue, H. Shimizu, K. Furumochi, H. Goto, S. Kawamura, K. Watanabe, and K. Aoyama, "A novel embedded SRAM technology with 10 μm^2 full CMOS cells for 0.25 μm logic devices," *IEDM Tech.Dig.*, pp. 941–943, 1994.

[49] M. Rodder, A. Amerasekera, S. Aur, and I. C. Chen, "A study of design/process dependence of 0.25 μm gate length CMOS for improved performance and reliability," *IEDM Tech. Dig.*, pp. 71–74, 1994.

[50] M. S. C. Luo, P. V. G. Tsui, W.-M. Chen, P. V. Gilbert, B. Maiti, A. R. Sitaram, and S.-W. Sun, "A 0.25 μm CMOS technology with 45 Å NO-nitrided oxide," *IEDM Tech. Dig.*, pp. 691–694, 1995.

[51] M. Bohr, S. U. Ahmed, L. Brigham, R. Chau, R. Gasser, R. Green, W. Hargrove, E. Lee, R. Natter, S. Thomson, K. Weldon, and S. Yang, "A high-performance 0.35 μm logic technology for 3.3 V and 2.5 V operation," *IEDM Tech. Dig.*, pp. 273–276, 1994.

[52] R. A. Chapman, C. C. Wei, D. A. Bell, S. Aur, G. A. Brown, and R. A. Haken, "0.5 micron CMOS for high performance," *IEDM Tech. Dig.*, pp. 52–55, 1988.

[53] P. G. Y. Tsui, P. V. Gilbert, and S. W. Sun, "Integration of power LDMOS into a low-voltage 0.5 μm BiCMOS technology," *IEDM Tech. Dig.*, pp. 27–30, 1992.

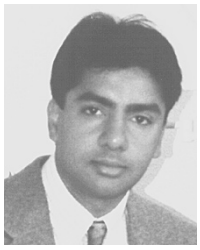
[54] T. M. Liu, G. M. Chin, D. Y. Jeon, M. D. Morris, V. D. Archer, H. H. Kim, M. Cerullo, K. F. Lee, J. M. Sung, K. Lau, T. Y. Chiu, A. M. Voshchenkov, and R. G. Swartz, "A half-micron super self-aligned BiCMOS technology for high speed applications," *IEDM Tech. Dig.*, pp. 23–26, 1992.

[55] Y. Kinoshita, K. Imai, H. Yoshida, H. Suzuki, T. Tatsumi, and T. Yamazaki, "Process integration technology for sub-30 ps ECL BiCMOS using heavily Boron doped epitaxial contact (HYDEC)," *IEDM Tech. Dig.*, pp. 441–444, 1994.

[56] D. L. Harame, E. F. Crabbé, J. D. Cressler, J. H. Comfort, J. Y.-C. Sun, S. R. Stiffler, E. Kobeda, J. N. Burghartz, M. M. Gilbert, J. C. Malinowski, A. J. Dally, S. Ratanaphanyarat, M. J. Saccamango, W. Rausch, J. Cotte, C. Chu, and J. M. C. Stork, "A high performance epitaxial SiGe-base ECL BiCMOS technology," *IEDM Tech. Dig.*, pp. 19–22, 1992.

[57] S. Subbanna, D. Harame, B. Chappell, J. Comfort, B. Davari, R. Franch, D. Danner, A. Acovic, S. Brodsky, J. Gilbreth, D. Robertson, J. Malinowski, T. Lii, and G. Shahidi, "A novel borderless contact/interconnect technology using aluminum oxide etch stop for high performance logic," *IEDM Tech. Dig.*, pp. 441–444, 1993.
[58] K. Kasai, Y. Akasaka, K. Nakajima, S. Suehiro, S. Suguro, H. Oyamatsu, M. Kinugawa, and M. Kakumu, "W/WNx/Poly-Si gate technology for future high speed deep submircron CMOS LSI's," *IEDM Tech. Dig.*, pp. 497–500, 1994.
[59] J. R. Pfiester, J. D. Shott, and J. D. Meindl, "Performance limits of CMOS ULSI," *IEEE J. Solid-State Circiuts*, vol. SC-20, p. 253, 1985.
[60] P. K. Chatterjee, W. R. Hunter, T. C. Holloway, and Y. T. Lin, "The impact of scaling laws on the choice of n-channel or p-channel for MOS VLSI," *IEEE Electron Device Lett.*, vol. EDL-1, pp. 220–223, Oct. 1980.
[61] S. Wong and C. A. T. Salama, "Impact of scaling on MOS analog performance," *IEEE J. Solid-State Circuits*, vol. SC-18, pp. 106–114, Feb. 1983.
[62] D. Scifres, *Spectra Diode Labs.*, private communication.
[63] R. A. Novotny, A. L. Lentine, D. B. Buchholz, and A. V. Krishnamoorthy, "Analysis of parasitic front-end capacitance and thermal resistance in hybrid flip-chip bonded GaAs SEED/Si CMOS receivers," in *Tech. Dig. OSA Top. Meet. on Optical Computing*, Salt Lake City, UT, Mar. 1995, pp. 207–209.
[64] A. V. Krishnamoorthy, T. K. Woodward, R. A. Novotny, K. W. Goossen, J. A. Walker, A. L. Lentine, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L. M. F. Chirovsky, G. F. Aplin, R. G. Rozier, F. E. Kiamilev, and D. A. B. Miller, "Ring oscillators with optical and electrical readout based on hybrid GaAs MQW modulators bonded to 0.8 micron CMOS," *Electron. Lett.*, vol. 31, no. 22, pp. 1917–1918, Oct. 1995.
[65] R. A. Novotny, "Analysis of smart-pixel digital logic and optical interconnection," Ph.D. Dissertation, Heriot Watt University, U.K., 1996.
[66] T. Sakurai and K. Tamaru, "Simple formulas for two- and three-dimensional capacitances," *IEEE Trans. Electron Devices*, vol. ED-30, pp. 183–185, Feb. 1983.
[67] L. M. F. Chirovsky, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, G. J. Przybylek, L. E. Smith, L. A. D. Asaro, A. L. Lentine, R. A. Novotny, and D. B. Huchholz, "Large arrays of symmetric self-electro-optic-effect devices," in *Tech. Dig. OSA Top. Meet. on Optical Computing*, Salt Lake City, UT, Mar. 1991, pp. 150–153.
[68] H. S. Hinton, T. J. Cloonan, F. B. McCormick, A. L. Lentin, and F. A. P. Tooley, "Free-space digital optical systems," *Proc. IEEE*, Nov. 1994, vol. 82, pp. 1632–1649.
[69] T. Sizer, II, R. E. La Marche, and T. K. Woodward, "Point-source heating effects in multiple quantum well modulators," *Appl. Phys. Lett.*, vol. 61, no. 4, pp. 420–422, July 1992.
[70] A. L. Lentine, K. W. Goossen, J. A. Walker, L. M. F. Chirovsky, L. A. D'Asaro, S. P. Hui, B. T. Tseng, R. E. Leibenguth, D. P. Kossives. D. W. Dahringer, D. D. Bacon, T. K. Woodward, and D. A. B. Miller, "Arrays of optoelectronic switching nodes comprised of flip-chip bonded MQW modulators and detectors on silicon CMOS circuitry," *IEEE Photon. Technol. Lett.*, vol. 8, pp. 221–223, Feb. 1996.
[71] K. W. Goossen, J. E. Cunningham, and W. Y. Jan, "Stacked-diode electroabsorption modulator," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 936–938, Aug. 1994.
[72] M. Whitehead, A. W. Rivers, G. Parry, and J. S. Roberts, "Very low voltage, normally-off asymmetric Fabry-Perot reflection modulator," *Electron. Lett.*, vol. 26, pp. 1588–1590, 1990.
[73] R. H. Yan, R. J. Simes, and L. A. Coldren, "Surface normal electroabsorption reflection modulators using asymmetric Fabry-Perot structures," *IEEE J. Quantum Electron.*, vol. 27, pp. 1922–1931, 1991.
[74] B. Pezeshki, D. Thomas, and J. S. Harris, "Optimization of modulation ratio and insertion loss in reflective electroabsorptive modulators," *Appl. Phys. Lett.*, vol. 57, pp. 1491–1492, 1990.
[75] P. Zouganeli, M. Whitehead, P. J. Stevens, A. W. Rivers, G. Parry, and J. S. Roberts, "High tolerances for a low voltage, high contrast, low insertion loss asymmetric Fabry-Perot modulator," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 733–735, Aug. 1991.
[76] P. Zouganeli and G. Parry, "Design tradeoffs and evalation of the performance attainable by GaAs–$Al_{0.3}Ga_{0.7}$As asymmetric Fabry-Perot modulators," *IEEE J. Quantum Electron.*, vol. 31, pp. 927–943, May 1995.
[77] K. W. Goossen, J. E. Cunningham, W. Y. Jan, and J. Centanni, "High-yield, low cost Fabry–Perot modulators utilizing correctable partial anti-reflection coatings," *Appl. Phys. Lett.*, vol. 66, no. 9, pp. 1041–1043, Feb. 1995.
[78] F. B. McCormick, T. J. Cloonan, A. L. Lentine, J. M. Saisan, R. L. Morrison, M. G. Beckman, S. L. Walker, M. J. Wojcik, S. J. Hinterlong, R. J. Crisci, R. A. Novotny, and H. S. Hinton, "Five-stage free-space optical switching network with field-effect transistor self-electro-optic-effect-device smart-pixel arrays," *Appl. Opt.*, vol. 33, pp. 1601–1618, 1994.
[79] A. V. Krishnamoorthy, T. K. Woodward, K. W. Goossen, J. A. Walker, B. Tseng, S. P. Hui, J. E. Cunningham, W. Y. Jan, F. E. Kiamilev, and D. A. B. Miller, "Operation of a dual-function receiver/transmitter hybrid CMOS SEED circuit," *Appl. Opt.*, to be published in Sept. 1996.
[80] A. V. Krishnamoorthy, T. K. Woodward, K. W. Goossen, J. A. Walker, A. L. Lentine, L. M. F. Chirovsky, S. P. Hui, B. Tseng, R. Leibenguth, J. E. Cunningham, and W. Y. Jan, "Operation of a single-ended 550 Mbit/s, 41 fJ, Hybrid CMOS/MQW receiver," *Electron. Lett.*, vol. 32, no. 8, pp. 764–765, Apr. 1996.
[81] D. Liu and C. Svensson, "Power consumption estimation in CMOS VLSI chips," *IEEE J. Solid-State Circiuts*, vol. 29, pp. 663–670, June 1994.
[82] A. A. Abidi, "Gigahertz transresistance amplifiers in fine line NMOS," *IEEE J. Solid-State Circiuts*, vol. SC-19, pp. 986–994, Dec. 1984.
[83] R. A. Novotny, M. J. Wojcik, A. L. Lentine, L. M. F. Chirovsky, L. A. D'Asaro, M. W. Focht, G. Guth, K. G. Glogovsky, R. Leibenguth, M. T. Asom, and J. M. Freund, "FET-SEED differential transimpedance amplifiers for two-dimensional optical data links," *J. Lightwave Technol.*, vol. 13, pp. 606–614, Apr. 1995.
[84] M. Ingels, G. Van der Plas, J. Crols, and M. Steyaert, "A CMOS 18 THzΩ 240 Mb/s transimpedance amplifier and 155 Mb/s LED-driver for low cost optical fiber links," *IEEE J. Solid-State Circiuts*, vol. 29, pp. 1552–1559, Dec. 1994.
[85] A. A. Ketterson, J.-W. Seo, M. H. Tong, K. L. Nummila, J. J. Morikuni, K.-Y. Cheng, S.-M. Kang, and I. Adesida, "A MODFET-based optoelectronic integrated circuit receiver for optical interconnects," *IEEE Trans. Electron Devices*, vol. 40, pp. 1406–1416, Aug. 1993.
[86] M. Govindarajan and S. R. Forrest, "Design considerations for wideband p-i-n/HBT monolithic transimpedance optical receivers," *J. Lightwave Technol.*, vol. 11, pp. 367–378, Feb. 1993.
[87] T. Vanisri and C. Toumazou, "Integrated high frequence low-noise current-mode optical transimpedance preamplifiers: Theory and practice," *IEEE J. Solid-State Circiuts*, vol. 30, pp. 677–785, June 1995.
[88] N. Scheinberg, R. J. Bayruns, and T. M. Laverick, "Monolithic GaAs transimpedance amplifiers for fiber-optic receivers," *IEEE J. Solid-State Circiuts*, vol. 26, pp. 1834–1839, Dec. 1991.
[89] B. I. Kasper, "Receiver design," in *Optical Fiber Telecommunications*, S. E. Miller and I. P. Kaminow Eds. New York: Springer Verlag, 1982, pp. 689–722.
[90] G. F. Williams and H. P. LeBlanc, "Active feedback lightwave receivers," *J. Lightwave Technol.*, vol. LT-4, pp. 1502–1508, Oct. 1986.
[91] G. F. Williams, "Lightwave receivers," in *Topics in Lightwave Transmission Systems*, T. Li, Ed. San Diego, CA: Academic, 1991.
[92] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design.* New York: Oxford Univ. Press, 1987.
[93] R. G. Smith and S. D. Personick, "Receiver design for optical fiber communication systems," in *Topics in Appl. Phys.*, H. Kressel, Ed. New York: Springer-Verlag, 1982, vol. 39.
[94] Z. H. Fang, S. Cristoloveanu, and A. Chovet, "Analysis of hot-carrier-induced aging from $1/f$ noise in short-channel MOSFET's," *IEEE Electron Device Lett.*, vol. EDL-7, pp. 371–373, June 1986.
[95] M. Aoki and M. Kato, "Hole-induced $1/f$ noise increase in MOS transistors," *IEEE Electron Device Lett.*, vol. 17, pp. 118–120, 1996.
[96] J. J. Morikuni, A. Dharchoudhury, Y. Leblebici, and S. M. Kang, "Improvements to the standard theory for photoreceiver noise," *J. Lightwave Technol.*, vol. 12, pp. 1174–1183, July 1994.
[97] L. D. Tzeng, R. E. Frahm, and W. Asous, "A high-performance optical receiver for 622 Mb/s direct-detection systems," *IEEE Photon. Technol. Lett.*, vol. 2, pp. 759–761, Oct. 1990.
[98] A. A. Abidi, "On the choice of optimum FET size in wide-band transimpedance amplifiers," *J. Lightwave Technol.*, vol. 6, pp. 64–66, Jan. 1988.
[99] R. P. Jindal, "High-frequency noise in fine-line NMOS field effect transistors," *IEDM Tech. Dig.*, pp. 68–71, 1985.
[100] A. A. Abidi, "High-frequence noise measurements on FET's with small dimensions," *IEEE Trans. Electron Devices*, vol. ED-33, pp. 1801–1805, Nov. 1986.
[101] K. Ogawa, B. Owen, and H. J. Boll, "A long wavelength optical receiver using a short channel Si-MOSFET," *Bell System Tech. J.*, vol. 62, no. 5, pp. 1181–1188, May-June 1983.
[102] S. M. Sze, *High-Speed Semiconductor Devices.* New York: Wiley, 1990.
[103] R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits.* New York: McGraw Hill, 1990, ch. 3.

[104] T. K. Woodward, A. L. Lentine, L. M. F. Chirovsky, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, L. E. Smith, L. A. D'Asaro, E. J. Laskowski, and S. S. Pei, "GaAs/AlGaAs FET-SEED receivers/transmitters," in *OSA Proc. Photon. Switching*, J. W. Goodman and R. C. Alferness, Eds. Washington, DC: Opt Soc. Amer., 1993, vol. 16, pp. 89–93.
[105] W. Fischer, "Equivalent circuit and gain of MOS field effect transistors," *Solid State Electron.*, vol. 9, pp. 71–81, 1966.
[106] T. Nakahara, S. Matsuo, S. Fukushima, and T. Kurokawa, "Performance comparison between multiple quantum-well modulator-based and vertical-cavity-surface-emitting-laser-based smart pixels," *Appl. Opt.*, vol. 35, no. 5, pp. 860–871, Feb. 1996.
[107] C. Fan, B. Mansoorian, D. A. Van Blerkom, M. Hansen, V. H. Ozguz, S. C. Esener, and G. C. Marsden, "Digital free-space optical interconnections: A comparison of transmitter technologies," *Appl. Opt.*, vol. 34, no. 17, pp. 3103–3115, June 1995.
[108] A. V. Krishnamoorthy, P. J. Marchand, F. E. Kiamilev, and S. C. Esener, "Grain-size considerations for optoelectronic multistage-interconnection networks," *Appl. Opt.*, vol. 31, no. 26, pp. 5480–5507, Sept. 1992.
[109] H. M. Ozaktas and J. W. Goodman, "Elements of a hybrid interconnection theory," *Appl. Opt.*, vol. 33, no. 14, pp. 2968–2987, 1994.
[110] T. J. Drabik, "Optoelectronic integrated systems based on free-space interconnects with an arbitrary degree of space variance," *Proc. IEEE*, vol. 82, pp. 1595–622, Nov. 1994.
[111] C. W. Stirk, "Cost-performance tradeoffs in optical interconnects," in *Tech. Dig. OSA Top. Meet. on Optical Computing*, Salt Lake City, UT, Mar. 1995, pp. 105–107.
[112] T. K. Woodward, A. V. Krishnamoorthy, A. L. Lentine, and L. M. F. Chirovsky, "Optical receivers for optoelectronic VLSI," this issue, pp. 106–116.
[113] T. K. Woodward, A. V. Krishnamoorthy, K. W. Goossen, J. A. Walker, J. E. Cunningham, W. Y. Jan, L. M. F. Chirovsky, S. P. Hui, B. Tseng, D. Kossives, D. Dahringer, D. Bacon, and R. E. Leibenguth, "Clocked-sense-amplifier-based smart pixel optical recievers," *IEEE Photon. Technol. Lett.*, vol. 8, Aug. 1996.



Ashok V. Krishnamoorthy (S'90–M'93) received the B.S. degree with honors from the California Institute of Technology, Pasadena, in 1986, the M.S.E.E. degree from the University of Southern California, Los Angeles, in 1988, and the Ph.D. degree in electrical engineering from the University of California, San Diego (UCSD), in 1993.

After spending a year as a Research Scientist in the Electrical & Computer Engineering Department at UCSD, he joined the Advanced Photonics Research Deparment at Bell Laboratories as a Member of Technical Staff in 1994. His research interests are in the area of optical interconnects, optoelectronic-VLSI smart pixels and their application to parallel computing, memory-access, and switching. He has published over 20 journal papers and one book chapter, presented over 10 conference invited talks, and filed more than 10 patents. He has served on several conference program committees for the IEEE Laser and Electro-Optics Society (LEOS), IEEE Computer Society (CS), and the Optical Society of America (OSA), and is the program chair for the 1997 IEEE LEOS Workshop on High-Speed Interconnections within Digital Systems.

Dr. Krishnamoorthy is a member of Tau Beta Pi, Sigma Xi, OSA, and SPIE.

David A. B. Miller (M'84–SM'89–F'95) received the B.Sc. degree in physics from St. Andrews University and the Ph.D. degree from Heriot-Watt University, U.K., in 1979.

He was a Carnegie Research Scholar at Heriot-Watt, where he continued to work, after receiving the Ph.D. degree, as a Lecturer in the Department of Physics. He moved to AT&T Bell Laboratories in 1981 as a Member of Technical Staff, and from 1987 to 1996 was Department Head, latterly of the Advanced Photonics Research Department. Since August 1996, he has been a Professor of Electrical Engineering at Stanford University, Stanford, CA. His research interests include optical switching and processing, nonlinear optics in semiconductors, and the physics of quantum-confined structures. He has published over 170 technical papers and four book chapters, delivered over 60 conference invited talks and over 25 short courses, and holds more than 30 patents. He has been a member or chair of over 25 technical conference committees, and was General Co-Chair for the Conference on Lasers and Electro-Optics in 1996. He has held posts or served on committees for the IEEE Lasers and Electro-Optics Society (LEOS), for the Optical Society of America (OSA), for the Optoelectronics Industry Development Association, for the Defense Advanced Research Projects Agency, and on several scientific journal editorial boards. He was President of the IEEE LEOS in 1995.

Dr. Miller is a fellow of the Royal Society of London, the Optical Society of America, and the American Physical Society. For his work on semiconductor nonlinear optics, quantum well optical properties, and novel devices, he was awarded the 1986 Adolph Lomb Medal of the OSA, was co-recipient of the 1988 R. W. Wood Medal, and received the 1991 Prize of the International Commission for Optics. He was also an IEEE LEOS Traveling Lecturer in 1986–87.