Analysis, Synthesis, and Fabrication of VLSI Si Detector Arrays for Optoelectronic Interconnections

by

Edward Joseph Ouellette III

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of

Bachelor of Science

and

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

Author .....
Department of Electrical Engineering and Computer Science
May 6, 1994

Certified by .....
Thomas F. Knight
Thesis Supervisor (Academic)

Certified by .....
Alan Huang
Company Supervisor (AT&T Bell Laboratories)

Accepted by .....
Frederic R. Morgenthaler
Chair, Department Committee on Graduate Students

Analysis, Synthesis, and Fabrication of VLSI Si Detector Arrays for Optoelectronic Interconnections

by

Edward Joseph Ouellette III

Submitted to the Department of Electrical Engineering and Computer Science
on May 6, 1994, in partial fulfillment of the
requirements for the degrees of
Bachelor of Science
and
Master of Science

Abstract

Optics possesses many benefits over electronics for high-speed interconnections. Furthermore, the reliability and affordability of current Si CMOS VLSI technologies over other heterogeneous technologies, such as GaAs, suggests that Si optoelectronic microsystems can be a viable, low-cost solution to high-speed interconnect problems. This thesis focuses on the receiver end of optoelectronic interconnects, and more specifically addresses the issues of optimal Si photodetector design within the constraints of an unmodified CMOS VLSI technology.

Various photosensors are compatible with a standard Si CMOS VLSI process. The simple graded-junction N+N/P diode was shown to have the best optical and electrical performance of all the types investigated, hence, it was used as a template for a photodetector analysis, synthesis and parasitic extraction CAD tool. A fabricated Si photodetector array in a standard AT&T CMOS process had 0.56 A/W responsivity at 850 nm wavelengths with only 2.5V of reverse bias; the device capacitance was 70 pF/mm^2, and the dark current was 61 nA/mm^2, in agreement with simulations.

A new CAD tool was written to address the need for computer automation in the area of photodetector design and analysis. The tool provides analysis and optimization of photodiode parameters, synthesis of photodetector array mask geometries (including array wiring) for several photodiode styles, and parasitic model extraction for SPICE simulations. It is both process-technology and hardware-platform independent allowing portability. Moreover, the output format is compatible with the MOSIS shuttle service, aiding in quick, inexpensive turn-arounds of optoelectronic designs.

Thesis Supervisor (Academic): Thomas F. Knight
Title: Principle Research Scientist: Electrical Engineering and Computer Science

Thesis Supervisor (AT&T Bell Laboratories): Alan Huang

Acknowledgments

Alan Huang provided much inspiration, support, and personal encouragement with this thesis and has been a constant friend. Tom Knight also helped me think of fundamental concepts in a new light (excuse the pun). Thank God for creative thinkers!

Jay O'Neill provided extensive MULGA support, and has been a humorous (R.P. supporter) friend through difficult times. Ted Woodward assisted in the optical setup experiments; Jim Walker helped with fabrication techniques and taught me how to use a wire-bonding machine — a skill to add to my resume (not!)

Sam Reynolds at MOSIS provided the MOSIS Scalable CMOS Design Rule files presented in Appendix C.

To my freshman advisor and long-time friend, Hugh Secker-Walker, thanks for all the wonderful discussions. Also, many thanks to Fr. Dave, Fr. Campbell, and Tom Hereford — three wonderful spiritual advisors who have had a great influence on my life.

To the many wonderful friends in LAG, especially Derek and Lisa, who have provided me with extra spiritual impetus to realize my dreams — my heart goes out to you.

Thanks also to Brian, Kent, Ochida, and Steve who have always been special friends. To Bob and Susan, for years of constant support and encouragement, senseless humor, and for being great "parents away from home" — I love ya!

Extra thanks goes out to J.C. for helping me persist with this thesis even through extenuating circumstances. First, half-way into my research completely new research had to be started after my original advisor left AT&T to found a company. Second, departmental funding constraints limited fabricating only a few, simple chips through the AT&T shuttle. Finally, upon my return to MIT, the entire AT&T tool suite had to be ported to my home computer in order to finish the CAD tool design.

To Michelle and Grant — thanks for the "home away from home," for the great meals and excellent company. Extra special thanks go out to my family for their love, support and encouragement. Thanks, Dad, for helping to bring out the scientist in

me and for all the mechanics training! Mom, I'll never forget your comment to me that you once felt like the Pope speaking to Michelangelo when he was painting the Sistine Chapel: "it's been long enough — just get the thing finished!"

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

The Digital Optics Group at AT&T Bell Laboratories in Holmdel, New Jersey has been investigating low-cost, massively parallel monolithic optoelectronic interconnect microsystems, specifically developing Optical Data Link (ODL) ICs, using an in-house Si CMOS VLSI process technology [AD88, DPC+89]. Previous systems achieved very low communication rates (under 10 Mb/s per channel) — an order of magnitude slower than expected through simulation results —, and the cause for the low throughput has yet to be explored. We believe that a system redesign will yield much more interesting results. CAD tool deficiencies in the area of modeling OEIC designs were flagged as the downfall in the first design.

The standard AT&T VLSI design tool package, MULGA, [ACW87] uses symbolic mask editing for custom layouts, and cannot easily adapt to handle photodetector devices nor parasitic extraction of such geometries, nor the creation of accurate models for simulation. Other commercial CAD tools suffer the same deficiency: there is clearly a need for new CAD tools in the area of custom optoelectronic devices that interface cleanly with existing VLSI CAD packages, such as MULGA [McB90, Rub90b]. Moreover, in such a CAD package, many issues such as mechanical alignment, quantum efficiencies, transit response times, geometries of photodetectors, and receiver noise and sensitivity have yet to be addressed accurately.

The goals of this thesis are to further investigate several possible photodetector styles and geometries compatible with a standard AT&T Si VLSI CMOS process and to develop new CAD tools to aid in the design, analysis and synthesis of such devices for easy integration into receiver amplifier designs. A new photosensor CAD tool integrated into the MULGA package should allow designers to input required photodetector parameters such as area, response time, geometric configuration, light factors, and have complete mask specifications and accurate simulation models generated without having to rely on back-of-the-envelope calculations and hand layout of intricate detector mask geometries. Moreover, the tool developed will be designed for process-independence so that the standard MOSIS shuttle service may be used to lower costs.

1.2 The Need for Optical Interconnect Systems

Many researchers are investigating different aspects of optoelectronic interconnection systems for use in high-speed computing. [LHH+90, Hut86, Kim91, Wil91, CO90, KHK90, Ols93]. The ultimate speed of such systems is becoming limited not so much by processor speed, but rather by the speed of its inter-processor interconnects [CO90]. The reason is simple: electronic integrated circuit line widths continue to decrease, and gate densities are increasing with low power techniques being exploited to achieve more computations with a given amount of power. Interconnect technology development needs to improve at the same or better rate; however, it has become exceedingly difficult for electronics alone to meet the needs, and some sort of technology breakthrough is needed. [CO90]. The need for higher speeds, lower power dissipation, and greater packaging densities is the impetus for considering optics in an area where electronics currently dominates.

Optoelectronics is able to to improve upon electronics by combining the benefits of VLSI technology with the communications effectiveness of optical interconnects [KFC+90]. Current systems are, therefore, investigating the use of optics in the interconnects where passive or active electronic designs are currently used. If optics

is to survive in this arena, it must attain an integrability and cost comparable to or better than today's advanced electronics. [Mil91].

1.3 Thesis Outline

Chapter 2 provides background on optical interconnect systems. It begins with a discussion on the advantages of optics over electronics in interconnections, then concludes with an example of a typical free-space optical communication link.

Chapter 3 describes various designs of monolithic Si photodetectors which are compatible with standard CMOS processes. Mathematical expressions governing important device parameters will be used to analyze possible tradeoffs to optimize device performances under various constraints.

Chapter 4 describes the design and implementation of a process-independent photodetector design and analysis CAD tool which uses the design and analysis parameters from Chapter 3.

Chapter 5 includes results of testing a fabricated Si photodetector array.

Chapter 6 presents conclusions and topics for future study.

Appendix A is a derivation of the fundamental equations governing photodetector device characteristics for PN and PIN style photodetectors.

Appendix B is a table listing of all the important photodetector design parameters discussed in Chapter 3.

Appendix C is a listing of the MOSIS SCMOS design rules used in the photodetector CAD tool.

Chapter 2

The Advantages of Optics

2.1 Overview

Both optical and electrical information propagate in the form of electromagnetic waves. For power and cost considerations, applications exist where pure optics wins out over electronics (typically in long-haul telecommunications), and applications exist where electronics is the clear choice (especially in short distance on-chip signaling). However, in the area of interconnections, where the communication distances are much less than long-haul, but sometimes not much more than on-chip distances, one must make careful selection of the medium of communication by being fully aware of the advantages and disadvantages of using each technique [Cau90].

Electrical layering, thin film wiring, and bump bonding techniques have improved electrical interconnects, but often off-chip drivers must be slowed down in order to accommodate slow off-chip links. Lately, in order to achieve the ever-increasing requirements on bandwidth for off-chip communications, designs are being pushed into the optical domain (especially for low-cost optical solutions) for the benefits described herein this chapter.

2.2 Wiring Density

Because of differences in the fundamental particles used to carry information in optics and electronics, some interesting phenomena result. Electronics rely upon electrons which are fermions; they carry charge, and thus, interact with each other strongly. Optics, on the other hand, rely upon photons which are bosons; they carry no charge, and thus, react with each other weakly.[1] The direct consequence of the nature of interactions is that optical beams can cross the same point in space without interference, whereas separate electrical wires cannot without shorting. Moreover, high-speed electrical transmission lines must be routed above a ground plane or with surrounding ground lines (both of which consume considerable area especially for dense circuits) in order to lower crosstalk noise and control impedances; optical transmission is not constrained by the need for nearby ground potentials, and therefore, can be propagated anywhere with free-space optics. For these two reasons, orders of magnitude more signals can be carried in optical form in a given volume than can be carried electrically [Hua91].

2.3 Available Bandwidth

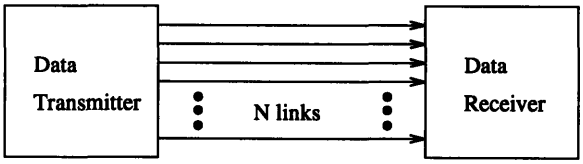
Optics has a much higher available bandwidth than electronics primarily because optical wavelengths are much shorter than the free-space wavelengths associated with typical electrical circuits. Since the carrier frequencies of optical systems tend to be in the petahertz (10^{15} Hz),[2] as opposed to the gigahertz range for electronics, optical systems have more bandwidth available for modulation by an information-bearing signal. Electrical information, on the other hand, propagates with free-space wavelengths on the same order as the size of the circuit or chip itself, and hence, wires must be used to confine the energy and prevent interference with adjacent signals. For this reason, even free-space microwave links between chips would be impractical.

[1] Photons may react with each other through a photon-electron-photon process; thus, there are no "purely optical" devices.

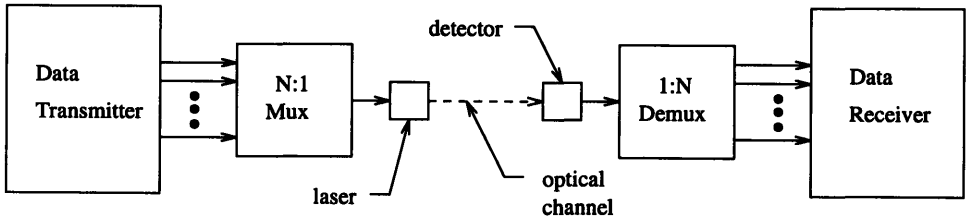
[2] For example, a light beam with wavelength of 800nm has a frequency of $\frac{c}{\lambda} = 3.8$PHz (petahertz)

Optical signals, however, may be precisely guided through a waveguide or via free-space means between chips without mutual interference. Thus, electrical interconnect wiring densities are not limited by current lithography, but are rather constrained by noise considerations [Rub90a]. Current electrical interconnect systems operate at rates up to several gigabits per second, whereas optical systems can easily operate in the terahertz range [CO90]. Moreover, the fastest electrical pulses are produced with light [Cau90].

Typically, the bandwidth available on an optical link is not limited by the link itself, but rather by the speed of the electronics. One approach to exploiting the available bandwidth on an optical link is to time-multiplex electrical signals. Another is to use wavelength-division multiplexing. Figure 2-1 shows how time-division multiplexing can be used in optical interconnects [Hut86]. This method works well as long as the multiplexors can be built from low-cost components which also handle the bandwidth.



Electrical Multiline Interconnect



Optical Multiplexed Interconnect

Figure 2-1: Time Division Multiplexing of Optical Interconnects

One major boon to monolithic optoelectronic interconnect systems is that they

can exploit hardware parallelism, creating a multiplicity[3] of optical links on a die using simple VLSI replication techniques. Figure 2-2 shows such a configuration for 40 optical channels. Jurgen successfully demonstrated 1024 such optical channels using planar microlens arrays [Jah90].

In Figure 2-2, a separate laser array is used to transmit optical signals to an optoelectronic receiver chip. The microlens arrays are used to correct beam divergence for this free-space configuration. Alternatively, the detectors on the receiver chip can be pitch-matched to a fiber ribbon array carrying the optical signals. The fiber array is then be butted up against the receiver chip's surface, and then the whole assembly is sealed into a package. The fiber cable could then be separated at an inline optical fiber connector in case disassembly was required. Alternatively, the fiber could be mounted up against a transparent optical window in the chip's package.

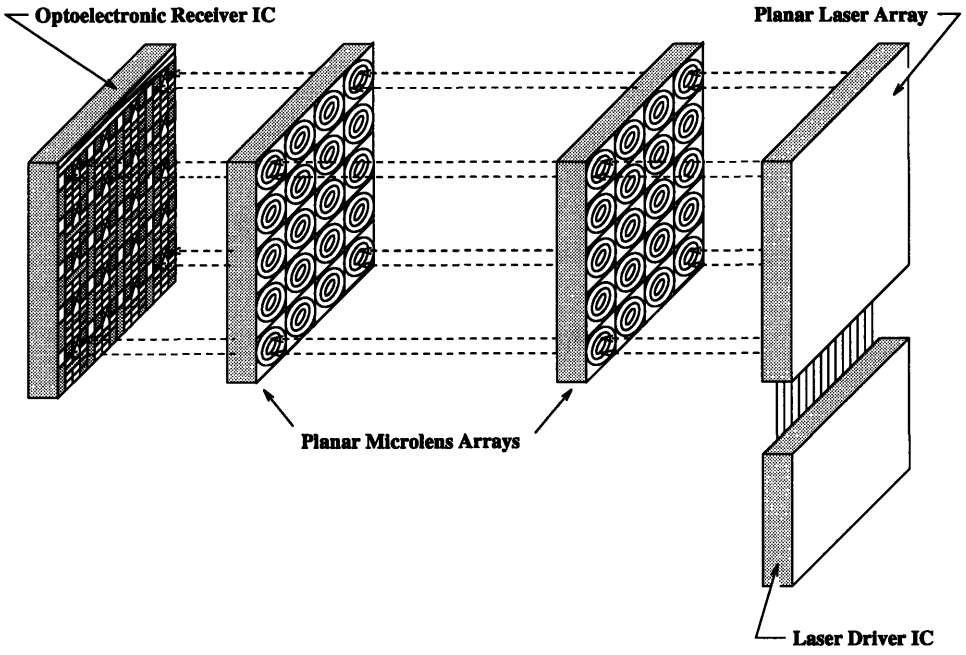


Figure 2-2: Parallel Optoelectronic Interconnect Links

The parallel optical link shown in Figure 2-2 uses separate transmitter and receiver array chips rather than bidirectional chips. This hybrid approach is required

[3]hundreds or even thousands of interconnects depending upon die size, and the area of overhead electronics required

for a technology such as Si because integrated photodiode detectors have a peak response to light with a wavelength around 850 nm.[4] However, successful microlaser arrays which emit light near 850 nm are typically made from GaAs-AlAs or GaInAs materials [LJ+89, JL+89, JML+89]. For each side of a bidirectional link, the receiver electronics for incoming signals and the laser driver circuitry for the return direction can be integrated into one IC. The laser array chip is then mounted near the electronics chip. Next, the two chips are interconnected through wire-bonding or bump-bonding technologies.[5] Obviously, this electrical interconnection degrades overall link performance. If lasers with wavelengths near 850 nm could be reliably and cheaply integrated with Si circuitry, however, a bidirectional Si optoelectronic VLSI chip could, theoretically, be produced.

Not only does parallelism in optoelectronic links lower costs, especially for technologies like Si, but it also provides a convenient medium for fault-tolerant, reliable interconnections through redundant spares and error detection and correction techniques which are readily implemented in VLSI electronics. Furthermore, in Si monolithic optoelectronic VLSI designs, where the available bandwidth is lower than in other, heterogeneous technologies, like GaAs, several slower, parallel Si optical links can replace a single ultra high-speed link, in order to reduce costs yet maintain all the other benefits of optics.

2.4 Attenuation Factors

For high frequencies (in and above the GHz range), electrical interconnect bandwidth is limited by the resistance, skin effect, and dielectric dispersion of wires [DKR+90]. The frequency-dependent dispersion of optical signals is much less than that of electrical signals over the range of distances used in interconnections because optics does not rely upon conduction. Absorption, scattering and misdirected light comprise the main loss factors in an integrated optical system. Figure 2-3 shows the attenuation

[4]The reason behind this phenomena is due to the intrinsic responsivity of Si, which is explained in more depth in Chapter 3.

[5]For instance, GaAs liftoff of lasers and flip-chip bonding techniques can be employed.

as a function of frequency for various types of optical and electrical transmission mediums [CO90].

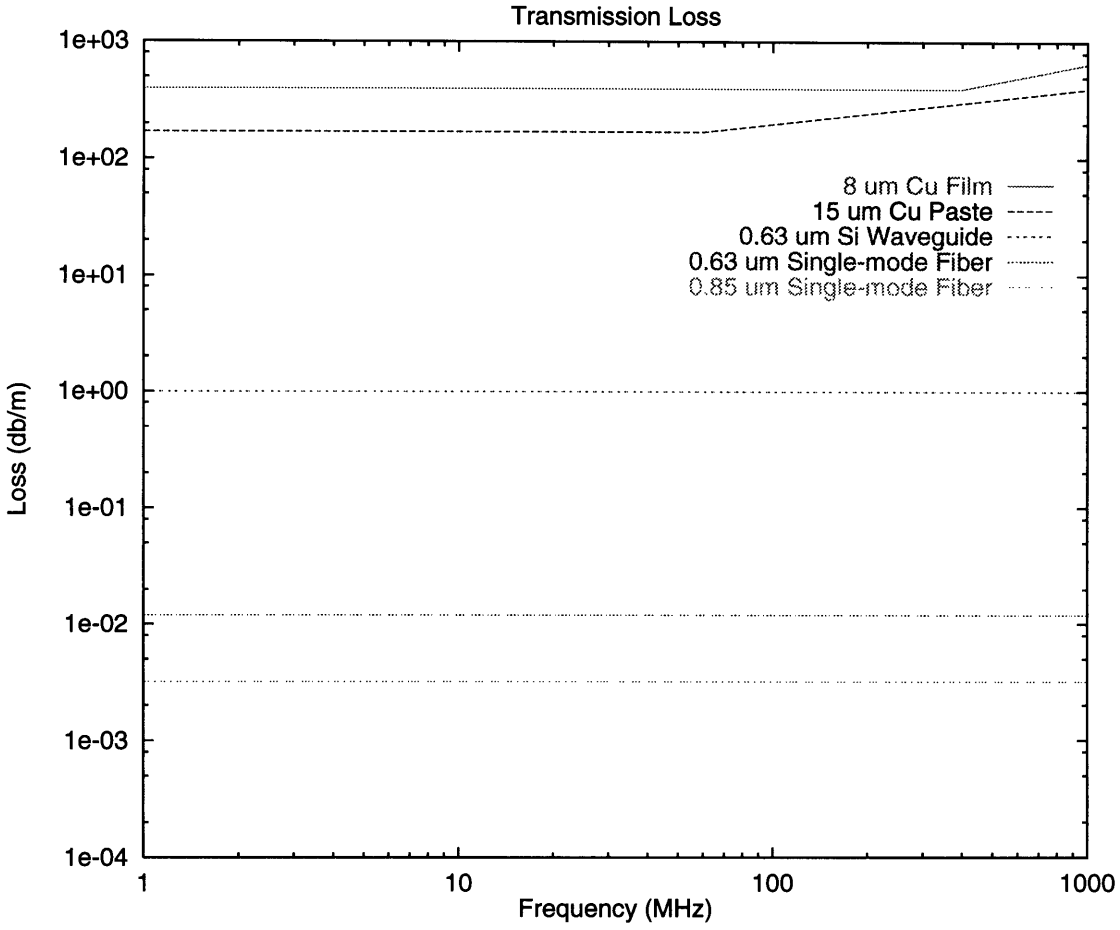


Figure 2-3: Attenuation of Optical and Electrical Interconnections

2.5 Fanin and Fanout Limitations

The fanin and fanout of optical signals is huge because optical signals require very little power and can cross without interference, unlike their electrical counterparts. In fact, in a $1mm^2$ area, an estimated 10^4 optical signals can coexist;[6] 10^4 parallel wires in such an area is impractical.

[6]This assumes a 2-D array (100 X 100) of 10,000 5-micron diameter beams each carrying a 1 Gb/s signal.

In an actual implementation of an optoelectronic receiver system, a photodetector-receiver combination converts an optically modulated signal back to its electrical form. Typically, these components limit the bandwidth of the entire link. Therefore, in order to achieve high bandwidth, it is often necessary to integrate the photodetector and receiver on the same chip. Additionally, for monolithically integrated optoelectronic interconnects, the area of the photoreceptors and amplifiers may severely limit the interconnect density; therefore, it is vital to keep this area small.[7]

2.6 Impedance Mismatch Noise

Free-space optical interconnects do not require precise impedance matching. Fiber-optic links, however, require more careful impedance matching so that reflected optical energy will not feed back into the source laser and degrade its performance [CO90]. Electrical interconnects are limited by the fact that they must drive inherently low-impedance transmission lines — typically around 50 Ω. The low impedance drivers required for such a transmission line consume considerable chip area and power. Optoelectronic devices can be made small, with relatively low impedances, and they will efficiently convert to and from the free-space optical impedance of 377 Ω. In other words, an optoelectronic interconnect is a special type of optical isolator. It functions as an ideal transformer, isolating two separate voltages in circuits and performing well even down to DC frequencies [Mil90]. In this manner, it can handle arbitrarily long strings of "one" or "zero" in a non-return to zero (NRZ) digital system, though most systems require some means of coding to facilitate clock recovery on the receiver end.

As Knight [T. 89] pointed out, electrical interconnections tend to be one-to-one in nature to achieve high speeds and easy termination of transmission lines. High speed transmission lines can be broadcast to multiple points (multidrop line), however, termination of such lines to match impedances can be extremely difficult, especially

[7] Keeping areas small also ensures detector and other parasitic capacitances are small which lowers required communication energy.

when signals go off-chip and parasitic package inductances as well as capacitances must be considered. Free-space or planar imaging techniques using microlens arrays or holograms [Jah90, Fel90] can be employed to broadcast an individual optical signal to many destinations with low interarrival skews due to matched transmission distances and the low noise properties of optical propagation. Moreover, free-space optical interconnects do not experience back-loading effects, and are, thus, limited only by the available power and detector area that must be covered. In fact, thousands of interconnections can be easily accommodated with a single lens [Mil90].

These combined facts encourage using optical interconnections to solve many problems with high speed clock distribution. Figure 2-5 depicts a monolithic optical approach to synchronous clock distribution in a digital system. In such a configuration, redundancy of emitters and receivers can increase reliability and reduce mechanical alignment problems. Another example of integrated optics comes from the work of Jahns [Jah90]. Figure 2-4 depicts an integrated optical imaging arrangement with 2-D microlens arrays.

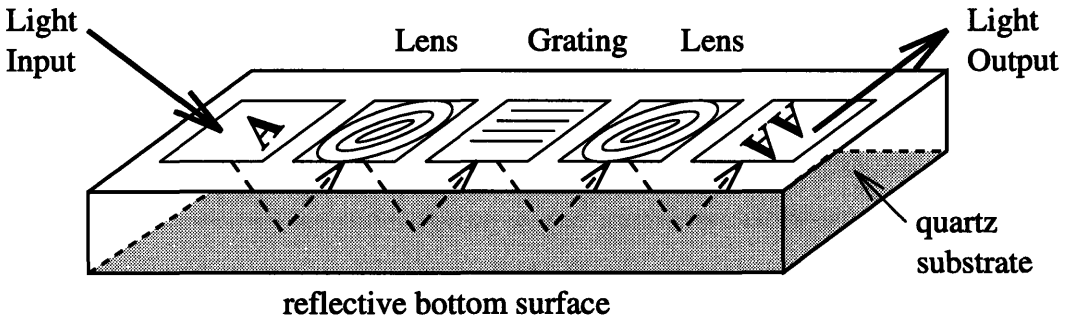


Figure 2-4: Integrated Optical Imaging System

2.7 Switching Noise

Another effect which can degrade the performance of electrical interconnect systems is noise caused by high-frequency switching transients. This noise voltage can be modeled as $NL\frac{di}{dt}$ where N is the number of simultaneous switching elements, L is the effective inductance, and $\frac{di}{dt}$ is the slew rate of the driver. In order to minimize this

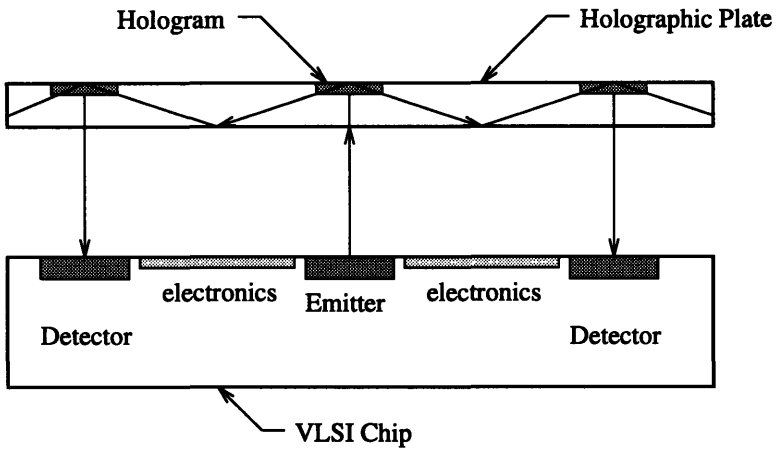


Figure 2-5: Synchronous Optical Clock Distribution

effect, the driver can be slowed, the number of simultaneous switching elements can be minimized, static and dynamic ground loops — which arise out of the difficulty in maintaining a constant ground potential throughout a system — can be broken, or decoupling capacitors can be used. Slowing the output drivers degrades system performance considerably; reducing the number of simultaneous switching elements requires special partitioning and system designs; breaking ground loops requires special circuit isolation techniques; using decoupling capacitors may consume considerable area. In typical low-noise electrical systems, the last three recommendations are usually employed.

2.8 Crosstalk Noise

Electrical signals are also degraded by interference due to stray capacitances and mutual inductances that exist between nearby transmission lines. Crosstalk noise only becomes worse for higher frequencies and denser circuits [Rub90a]. Optical systems, however, only experience small evanescent wave coupling between adjacent waveguides, neglecting scattering due to surface roughness and waveguide interfaces.

2.9 EMI Immunity

Because optics does not rely upon transduction except at the electronic interfaces, it is immune to EMP susceptibility, crosstalk, frequency dependent loss, and impedance mismatches, unlike electronics, which experiences all of the above effects. These factors lead to the high reliability of optical systems.[8]

2.10 Mutability of Connections

One of the major advantages of optics over electronics - the ability to make and break connections easily - stems from the fact that no physical contact is required to make optical interconnections, unlike electronics.[9] Electronic interconnects, on the other hand, require mechanical contacts and bonding / debonding. Bonding 10^5 connections is unwieldy for electronics, and is no concern for optics, however, mechanical alignment of the optical beams may be difficult.

Some of the direct consequences of the mutability of optical interconnections include the following [Cau90]:

- connections between removable objects
- connections between remote objects
- large number of connections
- "smart" or selective connections: with the merger of optics and complex electronic logic, connections can adapt to meet system signal-to-noise requirements, to take advantage of redundant hardware for fault-tolerance, to monitor and output vital link statistics (such as Bit Error Rate, BER, and available bandwidth), and to take advantage of electronic error detection and correction (EDC) circuitry.[10]

[8]Laser lifetimes have improved, and spatial modulators provide an alternative solution to the laser isolation problem.

[9]Even capacitively coupled techniques require precise control of inter-plate separation distances.

[10]Optics is great for regular patterns like a "perfect shuffle", but electronics is much better at

2.11 Reliability Concerns

The reliability of a large system made up of many identical components can be approximated by the reliability of one of the components divided by the number of such components in the system. The approximation holds true for random failures (which as a result of being a Poisson process, have exponential first-order interarrival times), not those due to component wear out (which have a log-normal distribution) [Ols93]. For example, a system composed of 1000 lasers each having a random failure time of 10^5 hours will have a mean time between failure (MTBF) of about 100 hours. Wear out failures will only increase this rate.

Many aspects of electrical connections can lead to system reliability problems. Electrical connectors, being mechanical in nature, can make lossy connections, or simply wear out from excessive make-break cycles. Moreover, VLSI systems can experience mechanical bonding problems due to poor alignment or thermal expansions of dissimilar materials, and electromigration[11] or component failure can also cause system failure. All these effects lower overall system reliability.

Optical interconnections also suffer from alignment problems, but none of the other mechanical limitations of electrical connections. The lifetime of monolithic lasers used in optical interconnections follows the traditional "bathtub" curve, where in the early stages of use, the rate of laser attrition due to component failures is high, then the rate decreases for quite some time, and finally, the laser extinction rate increases again due to wear out [Ols93]. The wear out of a semiconductor laser results in an increased threshold current,[12] making driver feedback circuits necessary in order to adjust the laser bias current to achieve a constant BER over time (before final failure).

Finally, electrically wiring two electronic gates in parallel to improve system yield and reliability will not work. If the gates were connected optically, however, then the

irregular connections: some hybrid combination may be needed to produce low area and efficient logic utilization.

[11] Cu deposited by chemical vapor deposition (CVD) can be used to increase resistance to electromigration and increase conductivity over traditional Al wiring [Shr90, DKR+90].

[12] Threshold current also increases with laser temperature.

setup would work. Fault-tolerance through optical interconnections and redundant electronics is easy to implement because of the unidirectional nature of free-space optical interconnections versus the bidirectional character of electrical wires [Hua91].

2.12 Communication Energy

Electrical interconnections can be modeled in two distinct manners [Kim91]: for short distances, a lumped RC model will suffice; for longer distances, a distributed transmission line model should be used. For the lumped RC case, the entire line must be charged to a final voltage and then discharged. Thus, the required energy is

$$E_e = C_t V^2 \ , \tag{2.1}$$

where C_t is the total line capacitance, including output driver and input receiver capacitances. For inter-chip communication, the input and output bonding pad capacitances would also be included. In the case of a distributed transmission line, only the length of line corresponding to the pulse width must be charged. The required energy follows the same equation as for the RC case, except that E_e is independent of interconnect length.[13] The length of interest (pulse width) is

$$L_{pw} = \frac{c}{\sqrt{\epsilon_r}} \tau_{pw} \ , \tag{2.2}$$

where c is is speed of light, τ_{pw} is the pulse width of the signal, and ϵ_r is the relative dielectric permittivity of the material ($\frac{\epsilon_{mat}}{\epsilon_o}$) which is 3.9 for SiO_2. Moreover, all high speed transmission lines must be terminated. The additional power that is dissipated due to a series-terminated line in the round-trip time of the line can be expressed as

$$P_t = \frac{V^2}{2R} \ , \tag{2.3}$$

[13]However, the resistive losses in a long line may require the use of repeater amplifiers, hence, increased required energy.

where R, the termination resistance, is matched to the characteristic impedance, Z_o, of the line. Figure 2-6 shows such a series-terminated transmission line.

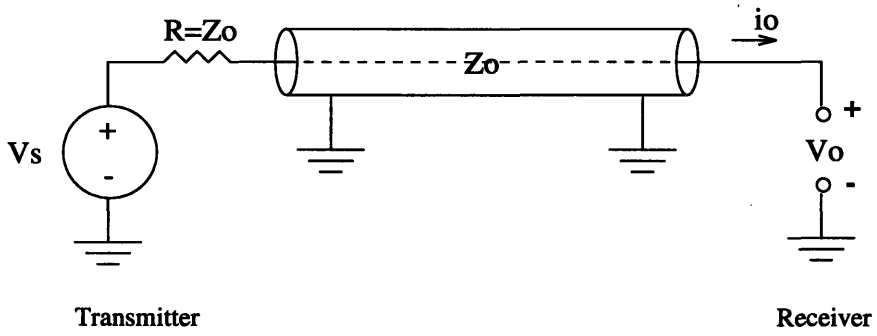


Figure 2-6: Series-Terminated Electrical Transmission Line

For an optical interconnect link, the total capacitance that must be driven is the photoreceiver plus amplifier input capacitances. The energy required to drive the laser can be expressed as

$$E_l = 2 \times 3\tau P_L \,, \tag{2.4}$$

where the laser driver is assumed to be 50 percent effective in delivering power to the laser, P_L is the average laser output power required, and τ is the rise time of the signal which is assumed to equal the fall time; the total pulse time is estimated at 3τ. An expression for τ comes from the characteristic capacitance relation:

$$\tau = C_t \frac{V}{I_{PD}} \,, \tag{2.5}$$

where C_t is the total driving point capacitance as mentioned above, V is the voltage swing needed at the receiver input, and I_{PD} is the photodetector current. The photocurrent follows the relation:

$$I_{PD} = \frac{q(P_L - P_{th})\eta}{h\nu} \,, \tag{2.6}$$

where q is the electron charge, P_{th} is the required power to bias the laser at threshold, η is the overall link efficiency (which is the product of the external laser quantum efficiency, the efficiency of the light guide material, and the photodetector quantum

efficiency), h is Planck's constant, and ν is the operating frequency. The above expressions can be substituted into the original power expression to obtain the required power for the optical interconnect:

$$E_l = 6C_tV\left[\frac{h\nu P_L}{\eta q(P_L - P_{th})}\right] . \tag{2.7}$$

Figure 2-7 depicts a typical free-space optical communications link, where η_o is defined to be the impedance of free space, 377 Ω. The transmission end consists of a laser which is driven by a modulated voltage source (a simple switch is the modulator here); the receiver end is comprised of a photodiode and amplifier. The amplifier operates in the transimpedance mode with an open loop gain of -A. More details will be provided on this configuration in later chapters.

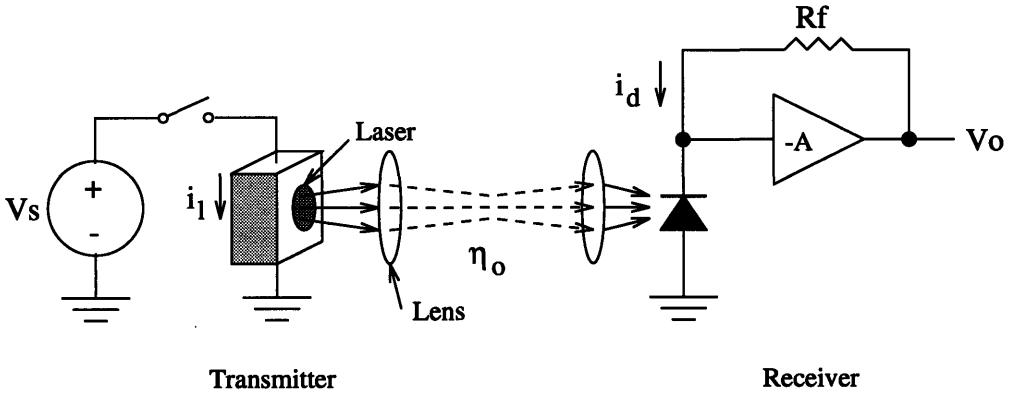


Figure 2-7: Free-Space Optical Communication Link

The high quantum efficiency of today's semiconductor lasers (around 80 percent) translates into about 1mW of optical power produced from 1mA of electrical current. Photodiodes with over 80 percent quantum efficiency are also manufacturable. Electrical drivers, on the other hand, require about 10-20mA switching current which is an order of magnitude higher than their optical counterparts [CO90].

Communication energy is actually only less for optics than for electronics when the distances for communication are greater than some critical length. Figure 2-8 based on Huang and Miller [Hua91, Mil90] depicts the critical communication distance where optics is more energy efficient than electronics for the same BER, ignoring cost

concerns. The critical distance of 200 μm shown would be closer to about 200 cm if today's costs are taken into account, however, with optics prices dropping (especially due to integration) and optical efficiencies increasing rapidly, this critical dimension will continue to decrease [Mil90].

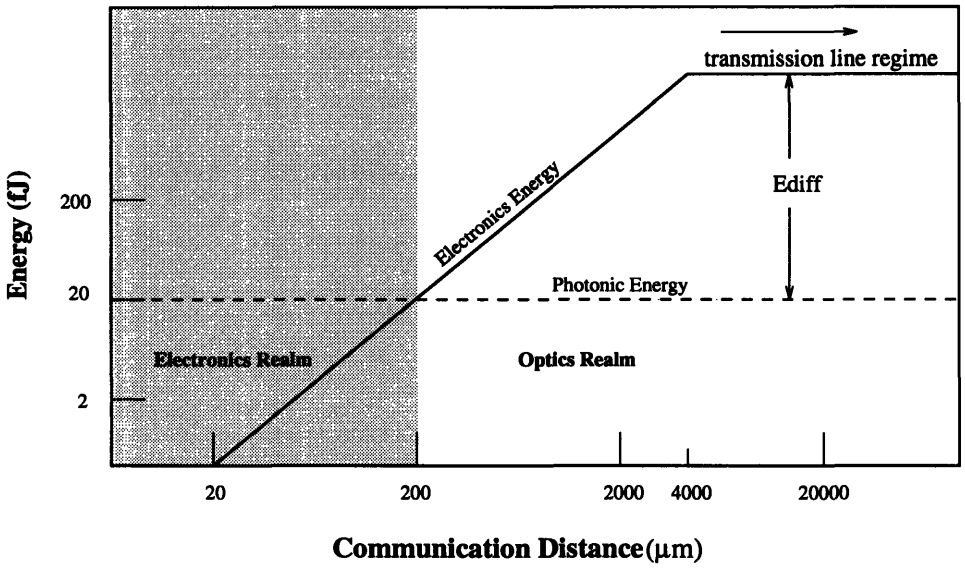


Figure 2-8: Communication Energy Versus Distance

Notice that for communication distances beyond 4 cm (4000 μm), the electronics energy curve remains constant. This phenomenon occurs because the electrical transmission medium is transformed into a transmission line, which when properly terminated, requires constant energy, independent of length.[14]

Figure 2-8 assumes that the communication channel rate is 1Gb/s, NRZ. An electrical wire can be modeled as a transmission line when the length of the wire is shorter than one-fourth of the signal's wavelength. The signal's wavelength can be calculated as follows:

$$\lambda_{sig} = \frac{v_{prop}}{f_{sig}}, \tag{2.8}$$

where v_{prop} is the characteristic velocity of the signal propagating down the wire which

[14]However, extremely long transmission lines require repeater amplifiers to compensate for resistive losses, hence their energy increase with increased length is not constant.

is assumed to be around 15 cm/ns (0.5 ft/ns), and f_{sig} is the frequency of the signal.[15]

For a 1Gb/s NRZ signal,

$$\lambda_{1Gb/s} = \frac{15cm/ns}{1ns} = 15cm. \tag{2.9}$$

Therefore, if the length of the electrical interconnect is longer than about 4 cm, it can be modeled as a transmission line. The energy difference, $Ediff$, should also be evaluated in order to determine if the additional cost of optics over electronics is justified.

[15] which corresponds to the highest possible transition rate of the signal

Chapter 3

Si Optical Sensors

One basic rule that most VLSI designers have in the back of their mind is:

> Use silicon... except when you can't.

3.1 Overview

This chapter begins with a description of photosensor operation including important device parameters and their inter-relationships. Next, the discussion turns to various designs of monolithic Si photodetectors which are compatible with standard CMOS processes. Mathematical expressions (derived in Appendix A) governing important device parameters are used to analyze possible design tradeoffs to optimize device performances under various constraints. These tradeoffs are incorporated into the *photogen* CAD program of Chapter 4.

Silicon comprises over 70 percent of the earth's crust. The material properties of silicon include such desirable traits as visible light sensitivity, room-temperature operation, and relative ease of manufacturing. These facts make it an ideal material for many photodetector applications. In the 200 to 1100 nm regime — which includes the visible spectrum — silicon is most sensitive, making it the most popular choice for optical sensor material. Also, many vendors offer competitive silicon detector solutions in high-speed packaging styles [Ham90, Sie90, Tex90, Adv90, Mot88, Opt90].

3.2 Photosensor Types

The three main categories of photosensors are depicted top-to-bottom in Figure 3-1. The various photodiode and phototransistor types all operate on the same basic principle. Because of carrier concentration gradients at a P-N semiconductor junction, electrons in the N-side diffuse across the boundary to the P-side, and holes diffuse in the opposite direction from the P-side to the N-side. The diffusion process creates a counterbalancing electric field in equilibrium which appears across the junction and results in a built-in device voltage.[1] A small region around the junction is thus depleted of mobile carriers, and hence, is termed the depletion region.

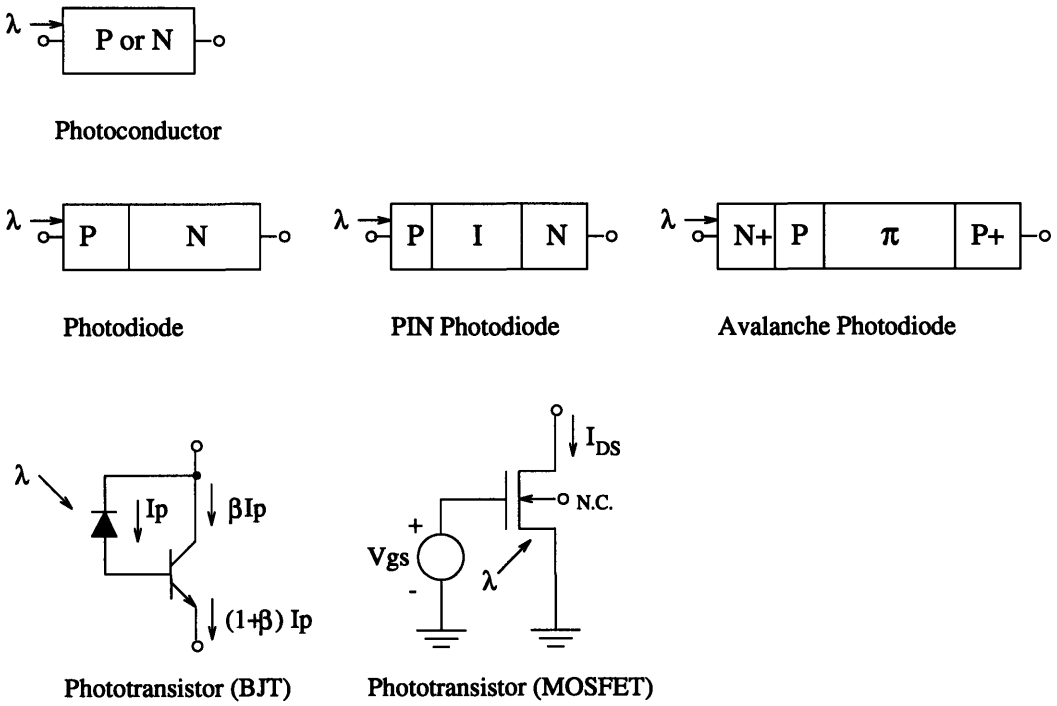


Figure 3-1: Photodetector Types

A photon from an incident light beam creates an electron-hole pair in the material only if its energy[2] is greater than or equal to the bandgap energy, E_G, of the material. E_G is 1.12ev for silicon; the wavelength corresponding to the bandgap energy of the

[1] This voltage may be nullified at the device terminals by proper choice of metal contacts which create metal-silicon diodes

[2] $E_p = h\nu$ where h is Planck's constant and ν is the frequency of light

material is termed the critical wavelength,[3] λ_c. This critical wavelength represents the longest optical wavelength which will generate electron-hole pairs in the material.

Electron-hole pairs produced within a diffusion length of the depletion region are eventually separated by the electric field and the carriers are swept across the junction to produce a current flow in an external circuit. Carriers outside this region recombine and thus, do not contribute to the photocurrent.

3.2.1 Photoconductor

The photoconductor is the simplest optical sensor that can be designed. It consists of a slab of semiconductor material which is typically doped N-type or P-type to tightly control its characteristics. When photons with energies above the bandgap of the material are absorbed in the material, electron-hole pairs are produced[4] which increase the material's conductivity. If an external voltage is applied across the device, the measured current will be proportional to the light intensity. In a real device, there are many processes that can degrade performance. Carrier trapping, recombination, and thermal generation of carriers can cause non-linearities [LHA+93]. Also, absorption may occur for photon energies less than the band gap because of lattice defects or impurities.

3.2.2 Photodiode

Figure 3-2 depicts an equivalent circuit model for a photodiode [Ham90]. I_L represents the photo-generated current which is proportional to the incident light intensity, I_D is the diode junction current, C_J is the junction capacitance, R_{sh} is the shunt resistance, R_s is the series resistance, I_{sh} is the shunt current, V_D is the diode voltage, and I_o is the output current.

Given the aforementioned circuit model, and the Shockley ideal diode equation,

[3] $\lambda_c = \frac{hc}{E_G} = 1.1\mu$m for Si

[4] which is the same as an electron being promoted from the valence band to the conduction band, leaving a hole in the valence band

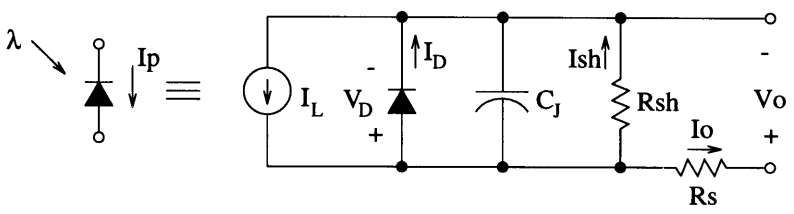


Figure 3-2: General Photodiode Circuit Model

one can solve for the output current, which is

$$I_o = I_L - I_D - I_{sh} .$$

Upon making the Schottky diode equation substitution, this becomes:

$$I_o = I_L - I_s(e^{\frac{qV_D}{kT}} - 1) - I_{sh} , \tag{3.1}$$

where I_s is the diode saturation current. The open circuit output voltage, V_{oc} is also easily to derive from the Schottky diode equation and the fact that $I_D = I_L - I_{sh}$:

$$I_D = I_S\,(e^{\frac{qV_D}{kT}} - 1)$$

$$\ln\left(\frac{I_D}{I_S} + 1\right) = \frac{qV_D}{kT}$$

$$V_{oc} = \frac{kT}{q}\ln\left(\frac{I_L - I_{sh}}{I_s} + 1\right) . \tag{3.2}$$

Equation 3.2 clearly illustrates the logarithmic dependence of the open-circuit output voltage on light intensity when I_L is large compared to I_{sh}. V_{oc} also varies considerably with temperature, making this photovoltaic diode mode difficult to use for light intensity measurements.

Finally, the short-circuit output current, I_{sc} can be calculated:

$$I_{sc} = I_{sh}\left(\frac{R_{sh}}{R_s + R_{sh}}\right)$$

$$I_{sc} = \left(\frac{R_{sh}}{R_s + R_{sh}}\right) I_L - \left(\frac{R_{sh}}{R_s + R_{sh}}\right) I_S \left(e^{\frac{qI_{sc}R_s}{kT}} - 1\right) . \quad (3.3)$$

In Equation 3.3, the diode nonlinearities arise from the exponential in the second term. If the value of R_s is made very small (under 1 Ω) and the value of R_{sh} very large (tens of MΩ or higher), then I_{sc} will be very close to I_L. Typically, a photodiode is operated with a very low load resistance, or even a reverse-bias voltage applied ($V_o < 0$) to increase response linearity and speed. This last, reversed-biased condition is termed the photoconductive or photocurrent[5] mode of a diode.

Figure 3-3 shows the reduced circuit model for a photodiode used in the photocurrent mode, where the diode has a reverse-bias voltage applied. R_{sh} is the parallel shunt resistance, C_J is the parallel junction capacitance, and R_s is the series resistance of the device. Notice that the direction of photo-induced current flow is in the opposite direction of conventional current. A photodiode can also be used in the photovoltaic mode, where no reverse-bias voltage is applied, and a photo-generated voltage appears from anode to cathode when incident light strikes the diode junction. A solar cell is an example of a photodiode operated in this mode. The photocurrent mode offers performance advantages over the photovoltaic mode in terms of its linearity, response speed, capacitance, stability, and temperature coefficient; therefore, unless otherwise stated, all photodiode designs to follow are assumed to operate in the photocurrent mode. For photodiodes operating in the photocurrent mode, typical range of values for R_s is 10 to 100 Ω, R_{sh} is 10^9 to 10^{10} Ω, and C_J is highly dependent upon processing and area [Mil86].

The PIN photodiode enhances the simple PN photodiode's characteristics by producing the same amount of current for a lower device capacitance. As a consequence,

[5]This name is a misnomer because the photodiode is always a current source, with or without an applied reverse bias.

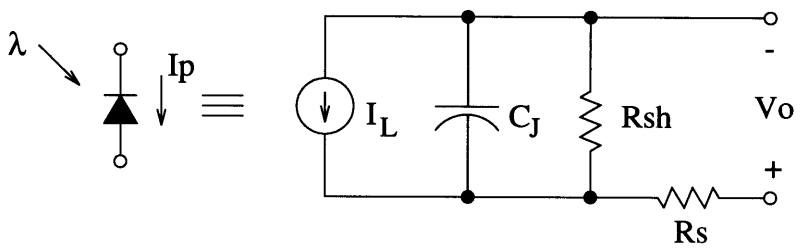


Figure 3-3: Photocurrent-mode, Photodiode Model

it offers a faster response as well. An additional intrinsic, (high resistance) region, *I*, sandwiched between the P and N regions. This intrinsic region is depleted in normal operation along with the P and N regions, providing a much larger collection region for electron-hole pairs than a simple PN diode, and greater resistance to breakdown voltage and leakage. Moreover, the intrinsic region thickness can be tailored to optimize responsivity and response time. Typical PIN photodiodes have been fabricated to have quantum efficiencies at or near 100 percent [Ham90, Adv90]. The capacitance of a PIN photodiode is lowered significantly because the effective "plates" of the capacitor have been spread apart by the length of the *I* region.

The avalanche photodiode requires a huge[6] reverse-bias voltage to create the high electric field required for avalanche multiplication. Optically generated carriers are swept into this high field region and accelerated. There, electrons and holes impact with other atoms and cause secondary electrons and holes to be produced, and these new carriers cause yet more carriers to be produced, and so on, effectively multiplying the number of carriers. The gain produced is on the order of 50 to 500 without a cost in speed, so this device is great for sensing low light levels. However, the large bias voltage required, the noise generated by the random carrier multiplication process,[7] and the temperature compensation needed for stability further complicate their use in monolithic integrated optoelectronic circuits today.

As depicted in Figure 3-4, MSM (metal-semiconductor-metal) diodes consist of closely-spaced interdigitated metal electrodes atop a semiconductor material. Volt-

[6] 150V to 300V is typical

[7] the receiver sensitivity can be improved by using a APD, but only as long as the multiplication noise is less than the noise of the amplifier

ages of opposite polarity are applied to adjacent electrodes to produce strong electric fields between the electrodes. These high electric fields sweep optically-generated carriers out of the exposed, underlying semiconductor material, allowing the carriers to be collected at the electrodes. The most common type of MSM photodetectors used today in optoelectronic applications are of the Schottky barrier variety, which have rectifying contact(s) from the metal to the semiconductor layer. The use of rectifying contacts guarantees that the dark current will be dominated by the thermal injection of electron-hole pairs in the semiconductor region where they are generated [JP86]. MSM photodetectors have advantages of planarity, simplicity, and compatibility with FET processing technology, and are used in many GaAs-based receivers. However, the Schottky barrier potentials on simple structures are often low, and in order to limit the dark current in such devices, special contact metals, extra semiconductor layers, and processing steps are often required to boost the Schottky barrier levels [SS90]. The MSM detector design is not compatible with an unmodified Si CMOS process, and thus was not considered an effective Si detector for this thesis investigation.

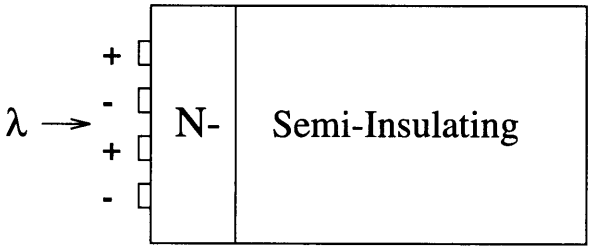


Figure 3-4: Schottky MSM Photodiode

3.2.3 Phototransistor

A bipolar phototransistor can be modeled as a bipolar transistor with a photocurrent injected between the collector and base. The transistor amplifies the photocurrent by the beta of the device, producing a much greater current at the emitter or collector. However, the performance penalty is speed — the photocurrent must charge up the base-emitter capacitance which increases rise and fall times drastically. Moreover, beta variation translates into uncertainty in sensitivity.

A light-sensitive MOSFET can be created by fabricating an N-channel transistor in a P-well process, applying a positive gate-to-source voltage to ensure the transistor is not in the cutoff region, and leaving the well connection floating [KDJS87]. MOSFETs can be rendered light-insensitive by tying their P-well (for an nFET) to ground, or N-well (for a pFET) to V_{DD}, or by shielding the device with a metal layer over the device.[8]

In the linear regime, the device's current response has been shown [KDJS87] to vary logarithmically with light power according to the following relation:

$$I_{DS} = \beta \left\{ [V_{GS} - V_{th0} - \kappa_1 (\sqrt{2\phi_F - V_T \ln(1 + \mathcal{R} P_{opt}/I_0)} - \sqrt{2\phi_F})] V_{DS} - \frac{{V_{DS}}^2}{2} \right\} , \quad (3.4)$$

where:

$$\beta = \frac{W}{L} \frac{\mu_n \epsilon_0 \epsilon_{ox}}{t_{ox}} , \quad (3.5)$$

$$V_T = \frac{kT}{q} , \quad (3.6)$$

$$\phi_F = \frac{kT}{q} \ln \left[\frac{N_{P-well}}{n_i} \right] , \quad (3.7)$$

$$\kappa_1 = \frac{t_{ox}}{\epsilon_{ox}} \sqrt{2qN_{P-well} \frac{\epsilon_{si}}{\epsilon_o}} . \quad (3.8)$$

I_0 is the junction reverse-bias saturation current, P_{opt} is the light power in μW, and $\mathcal{R}$ is the responsivity which is defined in Section 3.3. V_{th0} is the nonilluminated threshold voltage, and N_{P-well} is the donor concentration per cm^3 in the P-well.

At the P-well and substrate P-N junction, optically generated holes are swept into the P-well region, which causes electrons to flow across the source and P-well

[8] Shielding all non-photonic devices near the area where light impinges on the chip is a good idea in order to lower minority carrier injection into the substrate — current which can cause latchup in CMOS circuits.

N-P junction in order to regain equilibrium. Most injected electrons flow into the channel region if the source area is small, increasing the drain current; otherwise, many electrons get pulled down across the P-well and substrate P-N junction, creating a parasitic[9] vertical bipolar phototransistor mode of operation.

If the MOSFET's source area is too large, this parasitic bipolar device can degrade optical response of the device, and, for high enough light intensities, may lead to CMOS latchup. Hence, the photo-MOSFET has an inherent size constraint which may limit its application usefulness, especially in the niches of strip detectors and alignment-tolerant systems. Also, for large V_{DS} values, impact ionization (avalanche multiplication) has been shown to occur in the pinch-off region of the channel near the drain. This effect dominates over photo-induced currents; therefore, the device should be operated with low V_{DS} voltages to obtain optimal photo responsivity. Because of these limitations, the optical MOSFET will not be considered as candidate detector for the CMOS photodetector CAD tool in Chapter 4.

3.3 Important Photodiode Parameters

This section describes parameters which are vital for proper design or selection of photodiodes which operate in the photoconductive mode. For mathematical derivations of the parameters discussed below, please refer to Appendix A. A useful table containing important device design parameters may be found in Appendix B.

3.3.1 Active Area

Two different photodetector diffusion types [LHA+93] are shown in Figure 3-5. For the mesa type, the PN junction emerges on the sides of the device, so it is difficult to passivate this junction area with oxide. Therefore, it is often exposed and vulnerable to contamination and atmospheric degradation. Also, the top area of the device,

[9]This parasitic N+/P-well/N-sub device can be used by itself as a CMOS process-compatible bipolar phototransistor as long as care is taken to limit minority carrier injection into the substrate and to isolate these devices far from CMOS circuits.

which is unshielded by any protective or routing metal layers, is defined as the device active area. In a planar type detector, on the other hand, the PN junction emerges at the surface of the device, around its perimeter, where it usually is passivated with an oxide layer. This makes it much more reliable. For an applied reverse-bias voltage, a depletion region forms across the diode's PN junction — including the surface junction. This surface depletion region, if unshielded, effectively extends the active device area beyond the area defined by the lithographic masks. If the geometry is carefully planned, this surface depletion region can be used to maximize device responsivity through the use of the vertical PN junction feature, at the expense of photodetector response uniformity.[10] The *photogen* tool in Chapter 4 has an option to generate photodetectors which rely upon this principle.

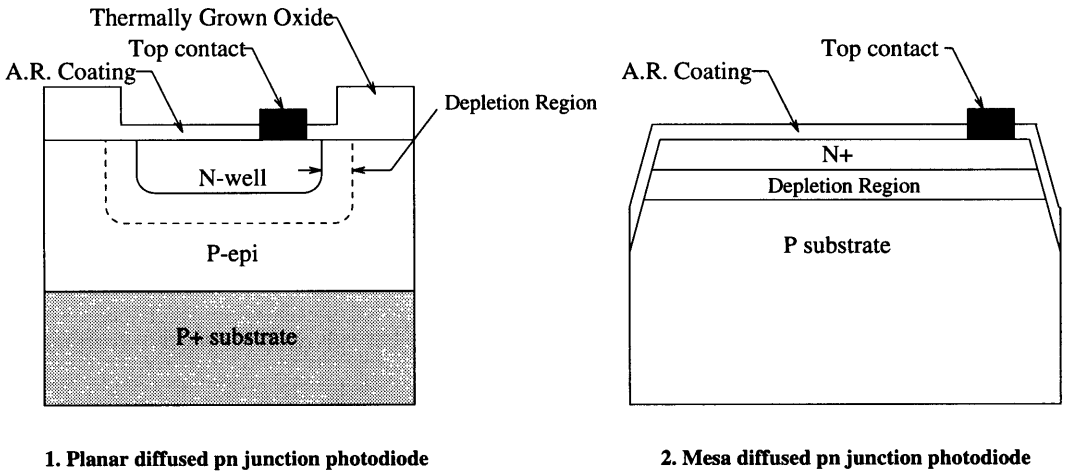


Figure 3-5: Photodiode Junction Diffusion Structures

3.3.2 Breakdown Voltage (reverse bias)

When the internal electric field of a reverse-biased photodiode exceeds about 3×10^5 V/cm,[11] a high-field effect known as impact ionization is initiated. This is the same effect which causes carrier multiplication in an avalanche photodiode. The

[10]It does this by absorbing photons along the tub edge, which extends deeper into the substrate than any other diffusion feature.

[11]This number depends upon the doping levels, and can range from 2×10^5 V/cm to 8×10^5 V/cm for Si.

external reverse-bias voltage which must be applied to the device to support such a high internal electric field is termed the breakdown voltage of the device. For a simple abrupt junction PN diode, the maximum electric field — which occurs at the junction edge — can be expressed as

$$|\mathcal{E}_{max}| = \sqrt{\frac{2q}{\epsilon_{Si}}\left(\frac{N_A N_D}{N_A + N_D}\right)(V_{bi} + V_{rev})}\,. \tag{3.9}$$

Assuming that $V_{rev} \gg V_{bi}$ and solving for V_{rev}, we get:

$$V_{br} = V_{rev} = \left(\frac{\mathcal{E}_{max}^2\,\epsilon_{Si}}{2q}\right)\left[\frac{N_A + N_D}{N_A N_D}\right]\,. \tag{3.10}$$

Assuming the diode's built-in voltage is 0.6V, and the ionized donor concentration is 10^{15}, the applied reverse voltage, V_{rev} must be around 584V to support a maximum electric field of 3×10^5 V/cm.

For a PIN diode, the breakdown voltage, V_{br}, can be expressed simply as

$$V_{br} = |\mathcal{E}_{max}|\,W - V_{bi}\,, \tag{3.11}$$

where W is the width of the intrinsic (depletion) region.

To support a maximum electric field intensity of 3×10^5 V/cm, the PIN diode must have a reverse voltage of about 300V applied.

Because voltages in typical integrated circuits are nowhere near this high breakdown voltage, the avalanche mode of a typical photodiode cannot be utilized by monolithic photodiodes in optoelectronic circuits. This parameter is often ignored in monolithic photodiode designs, except in the case of providing electrostatic discharge (ESD) protection for small devices.

3.3.3 Capacitance

The junction (depletion-layer) capacitance of a silicon PN diode is simply the capacitance of a parallel plate capacitor with plate separation W and stored charge equal to the charge stored in the PN junction depletion region (also known as the space

charge layer). This implies:

$$C_J = \frac{\epsilon_{Si} A}{W} , \tag{3.12}$$

where A is the cross-sectional area of the diode, W is the depletion region width, and ϵ_{Si} is the permittivity of silicon. The depletion widths of an abrupt-junction PN diode and a linear-graded PN diode can be expressed as

$$W_{aj} = \sqrt{\left(\frac{2\epsilon_{Si}}{q}\right)\left(\frac{N_A + N_D}{N_A N_D}\right)(V_{bi} + V_{rev} - 2V_t)} , \tag{3.13}$$

$$W_{gj} = \sqrt[3]{\frac{12\epsilon_{Si}}{qa}(V_{bi} + V_{rev} - 2V_t)} , \tag{3.14}$$

where:

$$V_{bi} = V_t \ln\left(\frac{N_A N_D}{{n_i}^2}\right) , \tag{3.15}$$

$$V_t = \frac{kT}{q} . \tag{3.16}$$

N_D is the ionized donor concentration in the N-material, N_A is the ionized acceptor concentration in the P-material, V_{bi} is the built-in voltage (usually around 0.6V), V_t is the thermal voltage, (0.026V for room temperature [300K]), a is the doping grading coefficient, V_{rev} is the applied reverse-bias voltage, and n_i is the intrinsic carrier concentration of silicon.

If $N_A \gg N_D$, as is the case in a one-sided abrupt junction, Equation 3.13 reduces to

$$W_{aj_1} = \sqrt{\left(\frac{2\epsilon_{Si}}{qN_D}\right)(V_{bi} + V_{rev} - 2V_t)} . \tag{3.17}$$

The capacitance per-unit-area expression for a reverse-biased abrupt-junction PN diode is

$$C_{aj} = \sqrt{\frac{q\epsilon_{Si}N_A N_D}{2(N_A+N_D)(V_{bi}+V_{rev}-2V_t)}}\,. \tag{3.18}$$

For an abrupt, one-sided junction, assuming $N_A \gg N_D$, Equation 3.18 reduces to

$$C_{aj_1} = \sqrt{\frac{q\epsilon_{Si}N_D}{2(V_{bi}+V_{rev}-2V_t)}}\,. \tag{3.19}$$

For typical (small) optoelectronic photodetector device sizes and technologies, the junction capacitance is also usually small. The diode's RC time constant[12] limits its speed of response. Since R_s and R_{sh} can be ignored most of the time, the detector circuit-limited frequency response can be expressed as

$$\frac{V_{in}(f)}{I_{ph}(f)} = \frac{R}{1+j2\pi f\, R_L\, C_j}\,. \tag{3.20}$$

However, for PIN structures with wide depletion regions, the diode's bandwidth is usually limited by the transit time of photo-generated carriers across the depletion region. Moreover, when the diode is actually connected to a circuit, its total capacitance adds to the front-end (receiver) capacitance which can significantly impact receiver sensitivity.

3.3.4 Crosstalk (for arrays)

Crosstalk between adjacent detector array elements is defined as the ratio of the photocurrent generated in a dark detector to the photocurrent generated in an adjacent, 60-80 percent illuminated detector [Ham90]. The intensity of illumination is adjusted to achieve a specified output current in the illuminated detector. This parameter is extremely important for arrays because strong crosstalk can significantly degrade the S/N ratio on a particular channel.

[12] $\tau_{RC} \approx R_L\, C_t$ where R_L is the load resistance of the circuit to which the diode is connected, and C_t is the total capacitance of the diode.

3.3.5 Dark Current

Dark current is the current generated by a reverse-biased diode in the absence of incident light. The dark-current component of shot noise is the major source of noise for the diode in sub-GHz operation [Lin89]. It has four contributing factors: trap-related thermally-generated carriers in the depletion region, diffusion current, tunneling, and surface leakage.

Thermally generated minority carriers in the non-intrinsic region outside the depletion region can diffuse into the depletion region and contribute a diffusion current component to the dark current. The diffusion current comes from the Shockley equation [Sze69].

$$I_{diff} = I_s \left[e^{\left(\frac{qV}{kT}\right)} - 1 \right] , \tag{3.21}$$

where I_s is the diode saturation current, V is the applied voltage, k is Boltzmann's constant, and T is the diode temperature.

For complete ionization of the impurities, the saturation current, I_s can be expressed [Sze69] as

$$I_s = qA{n_i}^2 \left[\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A} \right] , \tag{3.22}$$

where A is the junction area, D_n and D_p are minority carrier diffusion constants, L_n and L_p are minority carrier diffusion lengths, and N_D and N_A are the ionized donor and acceptor concentrations, respectively.

Traps in the depletion region near the center of the bandgap which emit carriers are the cause of the trap-related component of the dark current. An electron in the valence band is absorbed by a trap. Later, due to thermal excitation, the electron is promoted from the valence to the conduction band. In thermal equilibrium these processes would balance out, but in the high-field depletion region, the emitted carriers are swept away by the electric field before they can be recaptured by traps. This component of the current can be modeled [Lin89] as

$$I_{gd} = \frac{qn_i AW}{\tau_e}\left[e^{\left(\frac{qV}{2kT}\right)} - 1\right] , \tag{3.23}$$

where W is the depletion width and τ_e is the effective minority carrier lifetime.

When a large reverse-bias voltage is applied to a photodiode, many empty conduction bands on the N-side of the junction line up directly with filled valence bands on the P-side of the junction. If the barrier is small,[13] electrons can tunnel across the junction to fill the vacant sites. Quantum tunneling occurs primarily for direct bandgap material under high (> 20V) reverse bias, so this factor can be neglected for indirect bandgap materials, like silicon. Finally, surface leakage currents can be limited by passivation oxide material.

Equations 3.21 through 3.23 give light to the fact that dark current increases for increased reverse-bias[14] and device active area. Therefore, with proper device geometry and conservative reverse-bias, dark current can be limited.

3.3.6 Detectivity

Specific detectivity, $D*$, measures photodetector signal-to-noise ratio (S/N), normalized to a unit active device area.

$$D* = \frac{\sqrt{A\Delta f}}{NEP} \qquad \left[\frac{cm\sqrt{Hz}}{W}\right] , \tag{3.24}$$

where A is the device area, Δf is the bandwidth, and NEP is the Noise Equivalent Power defined in Section 3.3.7.

3.3.7 Noise

Noise is represented by random changes in the output signal which did not originate from changes in the corresponding input signal. In a reverse-biased photodiode, there are two main sources of noise: shot noise and Johnson, also known as thermal noise.

[13] i.e. large reverse bias and low enough bandgap energy

[14] The device depletion width, W, is proportional to the square-root of applied reverse voltage via Equation 3.13.

Shot noise is due to the fact that photo-generated carriers occur in discrete numbers, and create noise because of their random arrival times across the junction. It is modeled mathematically as

$$I_{shot} = \sqrt{2qI_{dk}\Delta f} \qquad [A] \ , \tag{3.25}$$

where I_{dk} is the dark current from Section 3.3.5, and Δf is the bandwidth of interest.

Thermal (Johnson) noise is caused by the random motion of carriers in a resistor. For the case of the photodiode model in Figure 3-3, the diode's thermal noise is generated by the shunt resistance, R_{sh}.

$$I_{therm} = \sqrt{\frac{4kT\Delta f}{R_{sh}}} \qquad [A] \ , \tag{3.26}$$

where k is Boltzmann's constant, and T is temperature. If I_L, the photo-generated light current is much greater than the dark current, I_{dk}, then the shot component of noise in Equation 3.25 is replaced by

$$I_{shot} = \sqrt{2qI_L\Delta f} \qquad [A] \ . \tag{3.27}$$

The total photodiode noise current, for a bandwidth Δf, is simply:

$$I_{nt} = \sqrt{(I_{shot})^2 + (I_{therm})^2} \ . \tag{3.28}$$

If the photodiode is not reverse-biased, the shot component of the noise can be ignored; if the reverse-bias is greater than about 2V such that $I_{shot} \gg I_{therm}$, the thermal noise component can be ignored.

The Noise Equivalent Power, NEP, is defined as the total photodetector noise, divided by the photodiode's responsivity, $\mathcal{R}$. It represents the light power required to produce a photocurrent equal to the noise current; in other words, for the S/N ratio to equal 1.

$$NEP = \frac{I_{nt}}{\mathcal{R}} \qquad [W] \ , \tag{3.29}$$

where I_{nt} is the total noise current and $\mathcal{R}$ is the responsivity. The NEP and $D*$ parameters are useful figures of merit for photodiode sensitivity. However, in actual applications, the receiver amplifier design usually contributes much more total noise to the system, so the responsivity parameter of a photodiode is often a better indicator of performance [LHA+93].

3.3.8 Quantum Efficiency

The quantum efficiency of a photodetector is defined to be the fraction of incident photons which are absorbed and eventually become current-producing carriers. If the detector is designed such that most photons are absorbed in the high-field intrinsic region, the quantum efficiency, η, can be simply expressed as

$$\begin{aligned} \eta &= \frac{I_p/q}{E_o/h\nu} \\ &= (1-R)(1-e^{-\alpha d}) \, . \end{aligned} \tag{3.30}$$

where I_p is the photo-generated current, E_o is the incident optical intensity, R is the Fresnel surface reflectivity (fraction of optical power reflected at the semiconductor-air interface), and d is the absorption depth, which approximately equals the depletion width.

Assuming normal incidence, the Fresnel reflection coefficient, R, and the transmissivity, T, can be expressed [Wya91] as

$$\begin{aligned} R &= \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 \\ T &= (1-R) \end{aligned} \tag{3.31}$$

$$= \frac{n_2}{n_1}\left(\frac{2n_1}{n_1+n_2}\right)^2$$

$$= \frac{4n_1n_2}{(n_1+n_2)^2}\,, \tag{3.32}$$

where n_1 and n_2 are the refractive indexes of the incident and transmission mediums, respectively. For a SiO_2 - air interface, $R \approx 0.33$.

The absorption coefficient, α, is a measure of how good the material is at absorbing light. If the non-intrinsic (top) material thickness is significant, it can absorb sufficient light power to lower the overall quantum efficiency. Equation 3.30 then becomes:

$$\eta = (1-R)(1-e^{-\alpha d})e^{-\alpha l}\,, \tag{3.33}$$

where l is the thickness of the non-intrinsic layer that the light must penetrate before reaching the depletion layer. If the absorption within one-diffusion length of the depletion region is taken into account [Sze69], Equation 3.33 can be rewritten as

$$\eta = (1-R)\left(1-\frac{e^{-\alpha d}}{1+\alpha(L_P+L_N)}\right)e^{-\alpha(l-L_N)}\,, \tag{3.34}$$

assuming the top diode layer is N-type. L_P is the extrinsic Debye hole diffusion length, and L_N is the extrinsic Debye electron diffusion length. L_P (and similarly, L_N) can be expressed in terms of the hole diffusivity constant, D_P, and the hole lifetime, τ_P:

$$L_P = \sqrt{D_P\,\tau_P} \tag{3.35}$$

$$L_P = \sqrt{\frac{\epsilon_{Si}kT}{q^2N_A}}\,. \tag{3.36}$$

Similar relations exist for L_N as Equations 3.35 and 3.36 with the P's and N_A replaced, respectively, with N's and N_D.

The Einstein relation in two forms is shown below:

$$D_P = \mu_P \frac{kT}{q} \tag{3.37}$$

$$D_N = \mu_N \frac{kT}{q}, \tag{3.38}$$

where μ_P is the hole mobility, and μ_N is the electron mobility can be used along with Equations 3.35 and 3.36 to obtain expressions for τ_P and τ_N in terms of known parameters:

$$\tau_P \left(\mu_P \frac{kT}{q} \right) = \frac{\epsilon_{Si} kT}{q^2 N_A}$$

$$\tau_P = \frac{\epsilon_{Si}}{\mu_P q N_A} \tag{3.39}$$

$$\tau_N = \frac{\epsilon_{Si}}{\mu_N q N_D}. \tag{3.40}$$

Figure 3-6 plots values of the absorption coefficient, α, over a range of wavelengths for different semiconductor materials [Sib90]. The plot clearly shows the critical absorption wavelength[15] corresponding to the bandgap energy of the material at which the material becomes transparent. For this reason, Si and GaAs are not suitable detector materials for wavelengths above 1 μm. The variation in the slopes of the curves with wavelength is due to the different densities of energy levels in the valence and conduction bands of the material.

For a direct band gap material like GaAs, there is no constraint on the electron-hole creation process and the absorption coefficient increases rapidly for wavelengths less than the critical absorption wavelength [Gow84]. For indirect band gap materials. like Si and Ge, carrier generation must be accompanied by interactions with the crystal lattice, so α increases only gradually for wavelengths just short of the critical

[15] $\lambda_c = \frac{hc}{E_G}$ varies with temperature because E_G and the band edges are dependent upon temperature.

absorption wavelength. At still shorter wavelengths, α increases more rapidly because excitation across a direct band gap becomes possible.

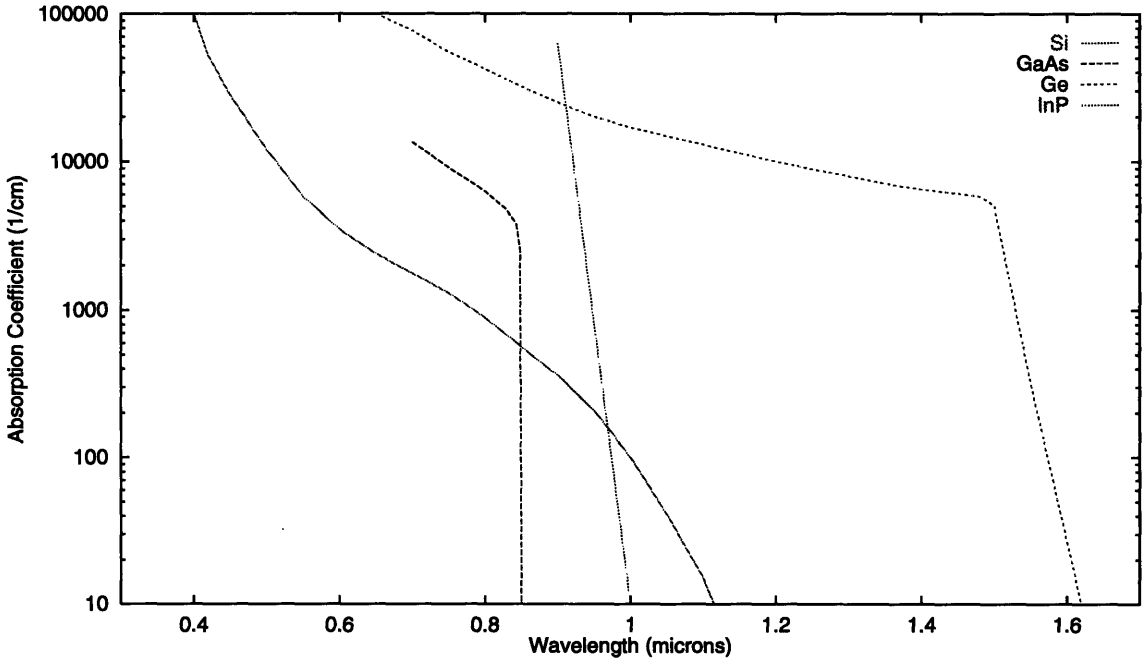


Figure 3-6: Absorption Coefficient Versus Wavelength for Various Materials

The penetration depth — the distance into the device where most of the light is absorbed[16] — is defined as follows:

$$P_{depth} = \frac{1}{\alpha} \times 10^4 \ [\mu m] \ . \tag{3.41}$$

From Figure 3-6 and Equation 3.41, it is easy to calculate the penetration depth for Si irradiated with 850 nm light to be about 10 μm. Because of shallow ($< 2\mu$m) diffusion depths in a modern, unmodified CMOS VLSI process, it is difficult to make Si photodetectors with high quantum efficiencies in such a process.[17]

An alternate way of achieving a high quantum efficiency for a medium depletion width is to use sideways illumination on a lateral junction, rather than using normal light incidence. The design of monolithic edge detectors using lateral[18] edge junctions to increase efficiencies has been explored [Kim91], but difficult packaging issues

[16]Specifically, the intensity drops to $I_o e^{-1}$, where I_o is the surface intensity.

[17]In fact, more responsive detectors can often be fabricated in older CMOS technologies!

[18]One can control the length dimension to make a junction which is capable of absorbing most

have yet to be ironed out before these detectors can compete with monolithic surface detectors. Back-illumination techniques have also been explored [Leh86], resulting in higher quantum efficiencies due to minimized surface recombination, and no reflections due to top-surface wiring. Another method to achieve a high responsivity without a large depletion width for top-surface illuminated photodiodes is to provide a reflective coating on the back side of the die to return unabsorbed carriers back to the depletion region. A serious limitation with this idea is that the response time of the device can be increased significantly, especially for large substrate depths.

An optimal detector design requires the quantum efficiency to be as high as possible. This requirement translates into low surface reflectivity[19] and a material with a large enough α for the wavelength(s) of interest to assure a reasonable penetration depth which can be matched by the intrinsic region depth in either a PIN or reverse-biased PN diode, yet not so high that all the photo-generated carriers are absorbed near the surface. Low carrier recombination is also desirable, as well as low absorption outside one diffusion length of the intrinsic region, which demands thin non-intrinsic material.[20]

3.3.9 Response Linearity

The response of silicon photodiodes is usually linear within a few tenths of a percent over the range of power from minimum detectable power to tens of milliwatts. The largest non-linearities inherent in the device stem from the nonlinear dependence of responsivity to the applied reverse bias. To improve linearity, the reverse bias can be increased, the series resistance, R_s, decreased, and the effective load resistance, R_L, decreased.

On the other hand, a very high-speed device requires a small depletion region (for short carrier transit times) which, in turn, requires a low applied reverse-bias voltage, which often creates a high device series resistance, R_s. When the bias voltage

side-ways incident photons within the penetration depth, yet requiring only a modest depletion width.

[19] made possible through use of anti-reflective ($\frac{\lambda}{4}$) dielectric surface coatings

[20] yet thick enough to support the depletion region in the material

is low, the effective internal diode bias voltage across the space charge layer can modulate drastically for large photocurrents, since most of the bias voltage will be dropped across the series resistance of the diode. The modulation in the bias voltage with photo-generated current varies the depletion region thickness which degrades responsivity, causing a substantial non-linearity in diode sensitivity. To reduce the nonlinearities, the series resistance should be minimized, and the reverse-bias voltage increased to the point where the slowest allowable transit time is reached — which raises quantum efficiency and lowers device capacitance.

3.3.10 Response Uniformity

For most applications, a uniform photodetector response is desired for light impinging upon any portion of the exposed detector area. Uniformity depends upon device surface features including surface resistance distribution, top surface wiring placement and exposed junction types (for a planar process), the quality of the package window or fiber connection, and the photodiode bottom surface reflectivity. For wavelengths less than 800 nm, the conditions of the diode's front surface greatly influence the response uniformity. Here the sensitivity of the uniformity to surface features is inversely proportional to the illuminated area. For wavelengths greater than 800 nm, the photons can penetrate further into the substrate, and for longer wavelengths, the reflectivity of the back surface may have a considerable effect.

3.3.11 Responsivity

This parameter is one of the most important characterizing parameters of a photodiode. It is defined to be the fraction of incident light power that results in photocurrent at the device terminals. The responsivity, $\mathcal{R}$, is defined as

$$\mathcal{R} = \frac{I_p}{P_{opt}}$$

$$= \frac{\eta\lambda q}{hc}$$

$$= \frac{\eta\lambda}{1.24 \times 10^{-6}} \left[\frac{A}{W}\right], \tag{3.42}$$

where I_p is the generated photocurrent (in Amperes), P_{opt} is the incident optical power (in Watts), η is the quantum efficiency, λ is the wavelength (in meters), q is the electron charge, h is Planck's constant, and c is the speed of light in a vacuum. The responsivity is also commonly expressed in units of mA/mW. If every incident photon produced an electron-hole pair, and all of the photo-generated carriers were collected by the photodetector (100 percent quantum efficiency), for a wavelength of 850 nm, the responsivity would be around 0.685 A/W.

Figure 3-7 shows the variation of the responsivity of a typical silicon photodiode over a broad range of light wavelengths, for a constant light intensity. Apparent in this figure is the extreme responsiveness of silicon photodiodes to wavelengths around 850 nm, making them prime candidates for near-infrared optical interconnect applications. For long wavelengths, the curve also shows a rapid decrease in responsivity in agreement with the critical wavelength for silicon. There is also a low wavelength cutoff point which occurs because the penetration depth of short-wavelength light in a material is much less than light of longer wavelengths. See Figure 3-6 and Equation 3.41 for further clarification.

Essentially, fewer carriers are optically generated within a diffusion length of the depletion layer, contributing to the photocurrent. More carriers are generated in the non-intrinsic top layer[21] before reaching the depletion region, and simply recombine there, contributing nothing to the photocurrent. Thus, the quantum efficiency of detectors decreases for short wavelengths, especially when the thickness of this non-intrinsic region is comparable to the absorption depth of the material. Additionally, because optical windows and fiber connections are sources of reflections and absorption, geometry and packaging concerns play a very important role in the design of

[21]The light must first penetrate this region before reaching the intrinsic region, assuming top-surface, not side-illumination of the diode junction.

high-responsivity photodetectors.

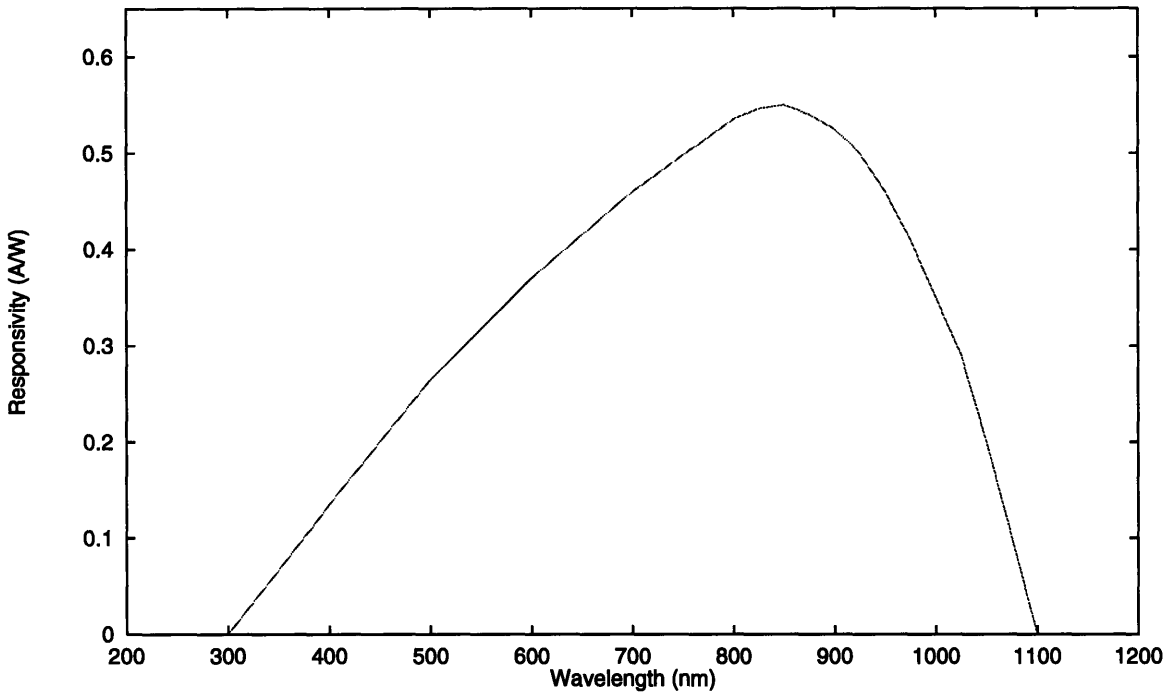


Figure 3-7: Spectral Responsivity of a Typical Silicon Photodiode

3.3.12 Response Speed

The response speed of typical photodiodes is usually specified in terms of the rise and fall times[22] of the photocurrent. These values essentially reflect the bandwidth of the photodiode and include the transit time of carriers through the depletion region (due to drift current), diffusion time of carriers to be collected at the edge of the depletion region, and the intrinsic RC time constant of the device. Out of these three components, the ultimate limit to response speed is the carrier transit time. Small area PN diodes operating in the photoconductive mode can have response times as low as 1ns; PIN photodiodes can have response times much smaller than this number.

In the transit time arguments which follow, it is assumed that the electric field in the depletion region is sufficient to accelerate the photo-generated carriers there to their saturation velocities. For example, in a PIN photodiode, the proper reverse-bias

[22]These times are measured from the 10 and 90 percent points of the photocurrent waveform.

voltage can be applied to make the electric field greater than or equal to the required saturation field intensity, $\mathcal{E}_{sat}$. This required voltage is simply:

$$V_{rev} > |\mathcal{E}_{sat}| W \ , \tag{3.43}$$

where W is the depletion width.

Equation 3.43 relies upon the fact that in a PIN diode, the electric field intensity is fairly constant in the intrinsic region as long as it is very lightly doped or relatively short; however in a simple abrupt-junction PN diode, the electric field reaches its maximum value at the PN junction, then decreases linearly on either side, reaching 0 V/m at the depletion region edges. For this reason the electric field in a PN diode will not maintain the carriers at their saturation velocities for the entire depletion region width even when the maximum electric field exceeds the saturation field intensity. However, as shown later in this chapter, the maximum field intensity for similarly-doped and similarly-biased PIN and PN diodes is higher in the PN case than in the PIN type. Therefore, if the carrier velocities in a PIN diode are saturated, so too, will be the carrier velocities in the PN diode for a large portion of the depletion region.

For a simple case illustrating transit-time limited response, assume N_o electron-hole pairs are generated at the (top) P-edge of a PIN diode. The holes are collected here, while the electrons must traverse the drift region and be collected at the N-edge of the depletion region. The time this takes is simply:

$$t_{trn} = \frac{\ell}{v_{sn}} \ , \tag{3.44}$$

where $\ell \approx W$, the depletion width, and v_{sn} is the saturation velocity of electrons. For Si, v_{sn} is about 10^5 m/s, and for a depletion thickness of 20 μm, $t_{trn} = 0.2$ ns. The drift current is simply:

$$i(t) = \frac{N_o q}{t_{trn}} = \frac{N_o q v_{sn}}{\ell} \ , \tag{3.45}$$

for $0 < t < t_{trn}$.

If the photo-generated carriers were produced only in the middle of the depletion

region, the holes would drift to the P-side of the region, and the electrons would drift to the N-side, each traveling only half the total distance. In reality, the photo-generated carrier distribution follows the exponential absorption profile with depth. The transit time is increased by the lower saturation velocity of holes over electrons, and to a big extent by photo-generated carriers outside the depletion region which diffuse into the depletion region to be collected. This diffusion current component contributes a slow "tail" to the diode's response due to excess absorption outside the depletion region. Since the minority carrier lifetimes can easily be over 10 ns, in most materials [Lin89], to minimize the rise and fall times, it is important to generate as many of the carriers as possible in the high-field, depleted region of the diode where the carrier velocities are fast, and to minimize the number of slow carriers generated outside this region. This constraint also implies that a thin non-intrinsic top region (which also helps increase quantum efficiency) is needed. A depletion width matched to the absorption depth is desirable to maximize quantum efficiency, however a smaller depletion width will decrease transit time, increase device capacitance, and increase the device RC time constant.

The transit-time limited frequency response of a PIN diode assuming high-frequency operation, low light penetration depth, and generation of carriers all one end of the depletion layer has been shown [Lin89, Gow84] to be

$$\frac{I(f)}{I(0)} = \frac{\sin(\pi f\, t_{tr})}{\pi f\, t_{tr}} \,. \tag{3.46}$$

For the case where $t_{tr} = 0.2$ ns, the 3 dB down-point in output power occurs when $f \approx 2.21$ GHz.

For a given device material and monochromatic light wavelength (which determines α) and desired responsivity, the geometry and reverse-bias voltage can be adjusted to minimize RC time constant and transit-time effects. In conventional VLSI photodiode designs, the geometry constraint is often not exercised in order to improve overall photodetector performance. The *photogen* tool detailed in Chapter 4 performs such optimizations.

3.3.13 Series Resistance

The series resistance of the device, R_s, depends mainly upon material bulk resistivities (small yet finite), ρ, metal-silicon (ohmic) contact resistance, and also upon geometry, including area. The effective series resistance appears as R_s in the diode model shown in Figure 3-3. A low value of R_s is required for fast response (lower RC constant) and good diode linearity.[23] The series resistance can be minimized through low-resistance (many) metal-diffusion contacts, low interconnect wiring resistance, shallow diffusion depths, and high device doping densities. However, high doping densities reduce the depletion width[24] for a given reverse-bias voltage through Equation 3.17 and, therefore, also increase device capacitance.

The N+N/P photodiode style in Figure 3-8 has a unique advantage of allowing a low resistance top contact (N+) over the entire device area without degrading device depletion width as in the N+/P design. This top diffusion region effectively lowers R_s.

3.3.14 Shunt Resistance

Shunt Resistance,[25] R_{sh}, is the resistance of a dark photodiode with 0 volts[26] of applied bias. This effective resistance plays an important role in determining the lower limit to the sensitivity of the device because the thermal noise depends primarily upon the value of this resistance.

3.4 Monolithic VLSI Photodetector Designs

Two conditions must be met by inexpensive monolithic photodetectors. They must meet electrical and optical specifications, and they must also be compatible with the IC technology used in chip fabrication. The types of detector junctions available in

[23] A low value of R_s lowers modulation of the depletion width with generated photocurrent

[24] by the square-root of the doping concentration

[25] R_{sh} is also commonly termed, source impedance, R_{so}

[26] some photodiode manufacturers specify the shunt resistance to be the slope of the dark current with respect to the applied voltage when the reverse-bias voltage is around 10mV [Ham90].

a standard AT&T CMOS technology (0.9 μm silicon gate length) are illustrated in Figure 3-8. The protective nitride and thick oxide layers can cause reflections and internal interference patterns which degrade overall responsivity of the diode. Optical response can be enhanced by removing the protective nitride over photodiode surfaces and adjusting the oxide thickness. With an additional processing step the thickness of the oxide layer[27] only over the photodetector areas could be controlled to minimize interference effects. The option of adjusting oxide thickness was not explored in this thesis because it would have required a simple, nonetheless special, modification to the standard CMOS shuttle process, resulting in more expensive runs.

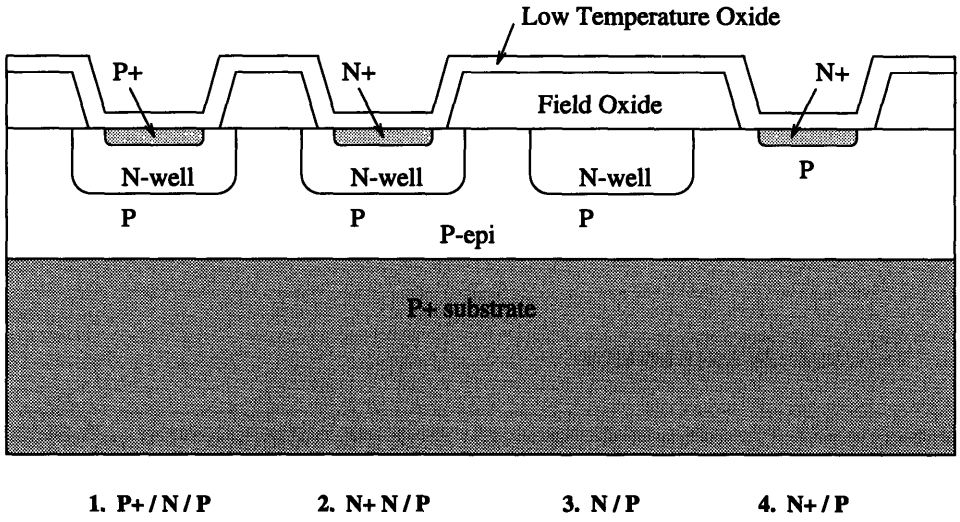


Figure 3-8: AT&T CMOS Process-Compliant Si Photodiodes

An evaluation of these same types of photodiodes, except for a P-well process, can be found in the literature [AOV88]. Their findings are reported in Table 3.1 for a 4-μm P-well technology. I_{rev} is the reverse-bias photocurrent, and C_J is the diode depletion capacitance. The P/N and P+P/N (graded junction) diode capacitances were found to follow a $1/V^3$ relationship as a function of the applied reverse voltage because of their extended depletion regions; the P+/N and N+/P (abrupt junction) diodes shared a $1/V^2$ dependence of the capacitance on the applied voltage. The P/N and P+P/N diodes also have larger photocurrents (higher responsivities) by about

[27]This oxide layer also serves to passivate surface leakage currents of the top-surface PN junction edge.

Parameter	N+/P/N	P+ P/N	P/N	P+/N	Units
I_{rev} (0.5V_{rev})	0.9	10	10	1.35	pA/mm^2
C_J	335	50	45	195	pF/mm^2

Table 3.1: Characteristics of CMOS-compatible Photodiodes

an order of magnitude over the abrupt junction types for the same reason of extended depletion regions.

For the AT&T process, the N+N/P diode type was chosen as the primary style to be implemented in the *photogen* program of Chapter 4 actually before outside work [AOV88] was known. Please refer to Figure 3-8 for the arguments which follow. First, in the P+/N/P diode style the photocurrent-producing junction is the top P+/N junction; in a typical CMOS process, the bottom N/P junction is back-biased to provide adjacent tub isolation.[28] Because the doping concentration of the P+ region in the N-well is higher than the N-well concentration, most of the depletion region appears on the N-well side of the diode. The depletion region thickness in the P+/N-well diode is much smaller than the depletion region which occurs in the N-well/P-substrate junction because the P-material doping concentration is much greater in the former case. Therefore, the responsivity of the P+/N diode is much less than the N+N/P and the N/P diodes, and the junction capacitance is much higher. However, the carrier transit time is lower, which can provide increased bandwidth for small detector areas.

The N+N/P diode has both low capacitance and high responsivity. The capacitance of the N+N/P design is lowered by the graded-junction between the N-well and the P-substrate, and the lower doping densities over P+/N and N+/P (abrupt-junction) designs. Additionally, the N+ layer can be spread over the entire diode active area, significantly reducing the effective series resistance, R_s of the diode, improving linearity and increasing device bandwidth. At short wavelengths, however, this N+ layer can lower diode sensitivity because of increased surface recombination.

As integrated circuit processes continue to advance, the minimum feature sizes

[28]The P-substrate is tied to V_{SS}, and the N-well is tied to V_{DD} in a typical CMOS process

continue to decrease. Not only are lateral dimensions reduced, but also many vertical distances, such as diffusion depths shrink as well. The reduction in many dimensions requires a corresponding reduction in the voltages used, assuming a constant $\mathcal{E}$-field scaling constraint is imposed [All88]. In order for depletion widths of transistor channels to decrease proportionately, the doping in the bulk near the source/drain junctions must increase by the same factor.

The CMOS process-compliant PN diode designs all rely upon the scaled doping densities, and hence, all diode depletion widths will decrease as well. The P+/N/P and N+/P designs are most sensitive to technology changes because they will invariably suffer the depletion width reduction required by the native MOSFET devices for an unmodified, advanced CMOS process. The N+N/P and N/P designs rely upon the well-substrate junction characteristics, and the depletion width here can be controlled by other processing steps, and will not significantly effect circuit density if it is allowed to be large. In most unmodified, advanced CMOS processes, however, it will reduce as well.

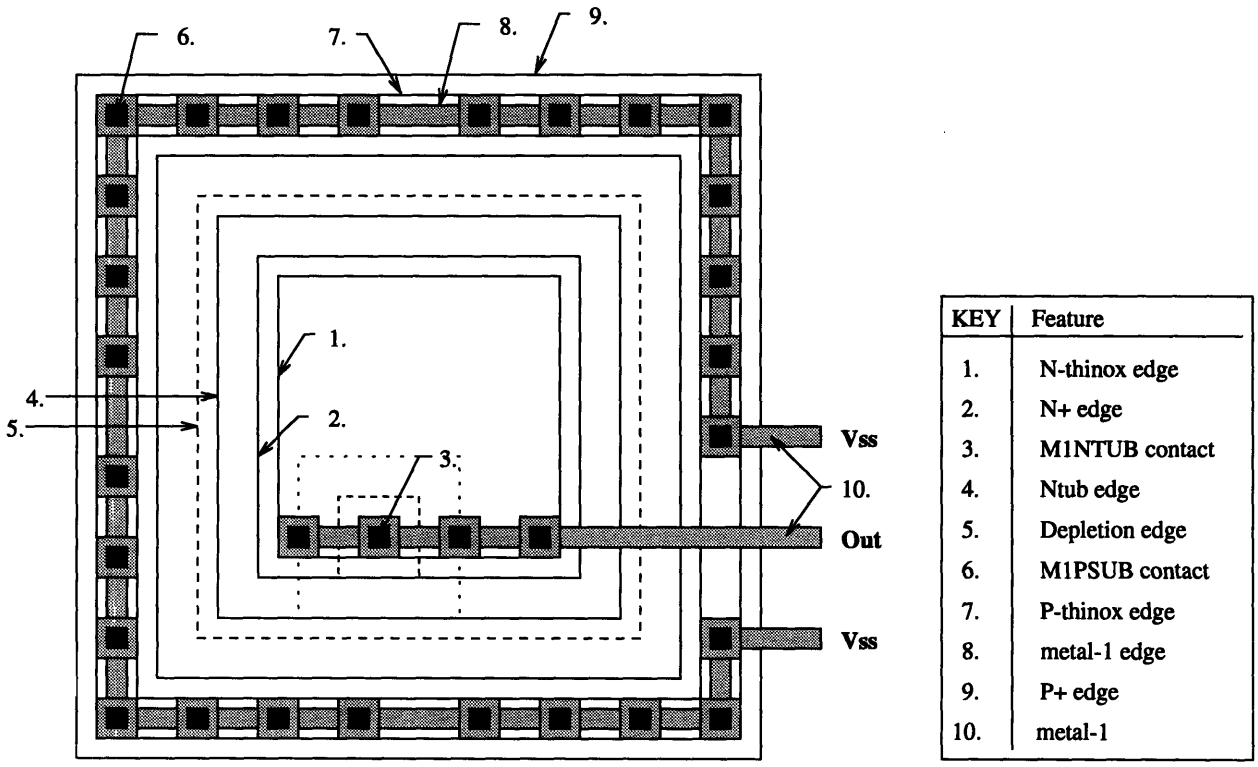
The N+N/P and N/P designs use the well-substrate junction for this purpose, and the well depth is the deepest mask-controlled feature in any standard CMOS process. Advanced CMOS optoelectronic processes may need an extra ion implantation step in order to preserve or tightly control this vertical photodiode junction depth as the well depth in advanced processes becomes too small to attain a high enough responsivity-to-capacitance ratio to make this geometry desirable over simpler designs.

3.4.1 Monolithic Photodetector Geometry Examples

Three different photodetector geometries are presented. Each of these photodiode geometries were designed to meet photolithography constraints of a typical CMOS VLSI process, and to allow easy CAD tool generation of photodiode arrays. All geometries presented are supported by the photodetector synthesis and analysis CAD tool in Chapter 4.

Figure 3-9 shows a simple rectangular photodetector geometry for the N+N/P style. The N+N/P style was chosen for its good response and low capacitance for

arguments explained earlier. The V_{SS} tie around the perimeter of the reverse-biased diode serves as the anode connection as well as a guard ring to prevent minority carriers from interfering with adjacent detector signals — which would cause crosstalk — and from causing latchup in local CMOS electronics. This detector geometry provides good uniformity of response and low junction capacitance, however, its responsivity is limited by the depletion region thickness between the well and the substrate, and this dimension cannot be near the penetration depth of Si for 850 nm light[29] because of the limited reverse-bias voltage available[30] in a typical CMOS process. In any case, this depletion region thickness must be supported in the region between the N-well edge and P+ edge of the V_{SS} guard ring for the device to operate properly; this constraint is indicated by the presence of the depletion region dashed line (5.) in Figure 3-9.



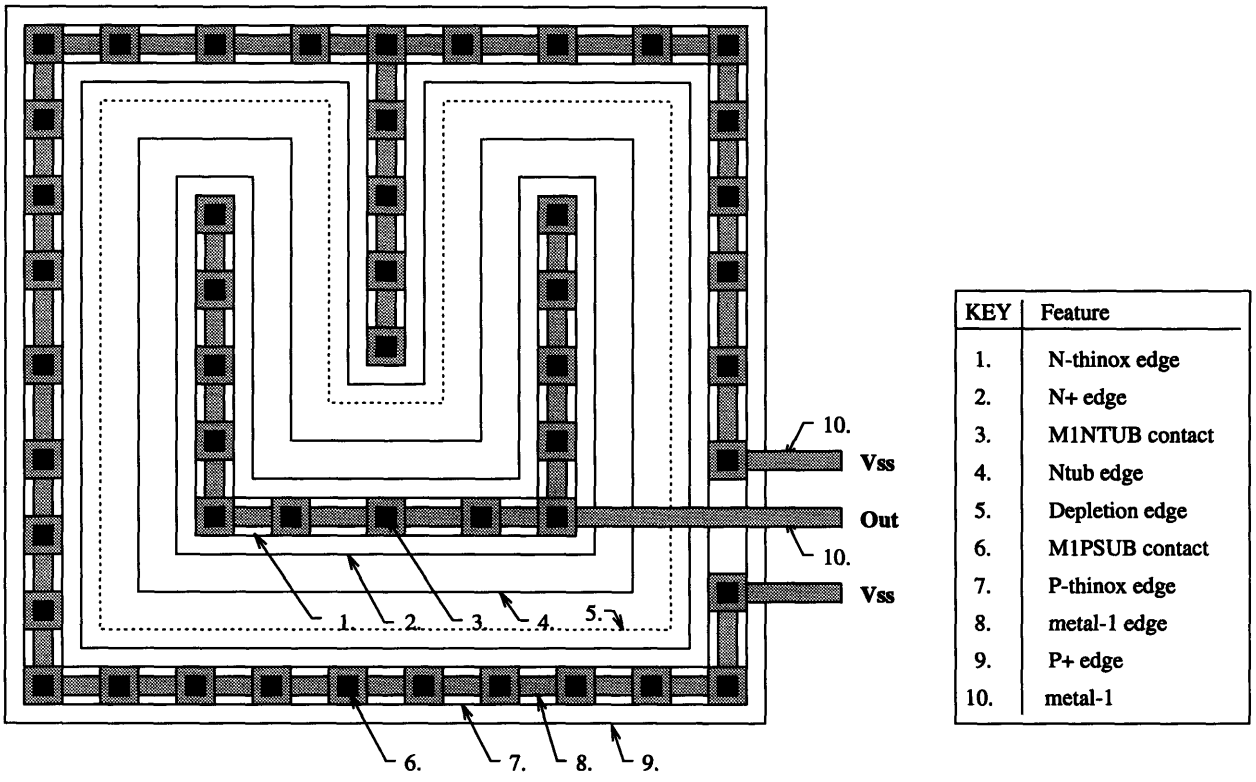
KEY	Feature
1.	N-thinox edge
2.	N+ edge
3.	M1NTUB contact
4.	Ntub edge
5.	Depletion edge
6.	M1PSUB contact
7.	P-thinox edge
8.	metal-1 edge
9.	P+ edge
10.	metal-1

Figure 3-9: Simple N+N/P Rectangular Photodetector Geometry

[29]Still, it is larger than the depletion widths for the P+/N/P and N+/P styles presented earlier.

[30]Typically, in a single-well process, the maximum bias voltage is $-V_{DD}$ without using a back-bias generator, but for practical circuit biasing concerns, it is usually around $\frac{-V_{DD}}{2}$.

Figure 3-10 depicts a simple interdigitated finger geometry for the N+N/P style. This geometry is designed to maximize the diode's responsivity at the expense of increased device capacitance and dark current, and decreased response uniformity (for small light spots). The responsivity increases dramatically for areas of the the diode where a vertical surface PN junction is present, and it decreases substantially for other areas, such as the large areas of interconnect. If the interdigitated fingers are small (many) compared to impinging light source, local extrema in the responsivity tend to average out to obtain a more uniform response to light position. The interdigitated style has a higher capacitance because of denser PN junction surface area, and a correspondingly higher dark current level, but very high quantum efficiency because it uses primarily the depth of the vertical PN surface edge junctions to collect light.



KEY	Feature
1.	N-thinox edge
2.	N+ edge
3.	M1NTUB contact
4.	Ntub edge
5.	Depletion edge
6.	M1PSUB contact
7.	P-thinox edge
8.	metal-1 edge
9.	P+ edge
10.	metal-1

Figure 3-10: N+N/P Interdigitated Finger Photodetector Geometry I

Figure 3-11 illustrates a larger interdigitated finger geometry for the N+N/P style, which utilizes the commonly used VLSI technique of mirroring cells to share common power or signal rails. Please refer to the key in Figure 3-10 for labels to the various

device features. Notice how the geometry has changed from the "comb-style" in Figure 3-10 to the "back-to-back comb-style" in Figure 3-11. This new geometry reduces the light-blocking disadvantages of the guard ring and signal wiring, providing better responsivity for larger detectors. The *photogen* program in Chapter 4 is able to determine which photodetector style(s) can be generated based upon the required size of the photodetector, bias conditions, etc. It then automatically synthesizes mask layouts, and extracts device parasitics for SPICE simulations.

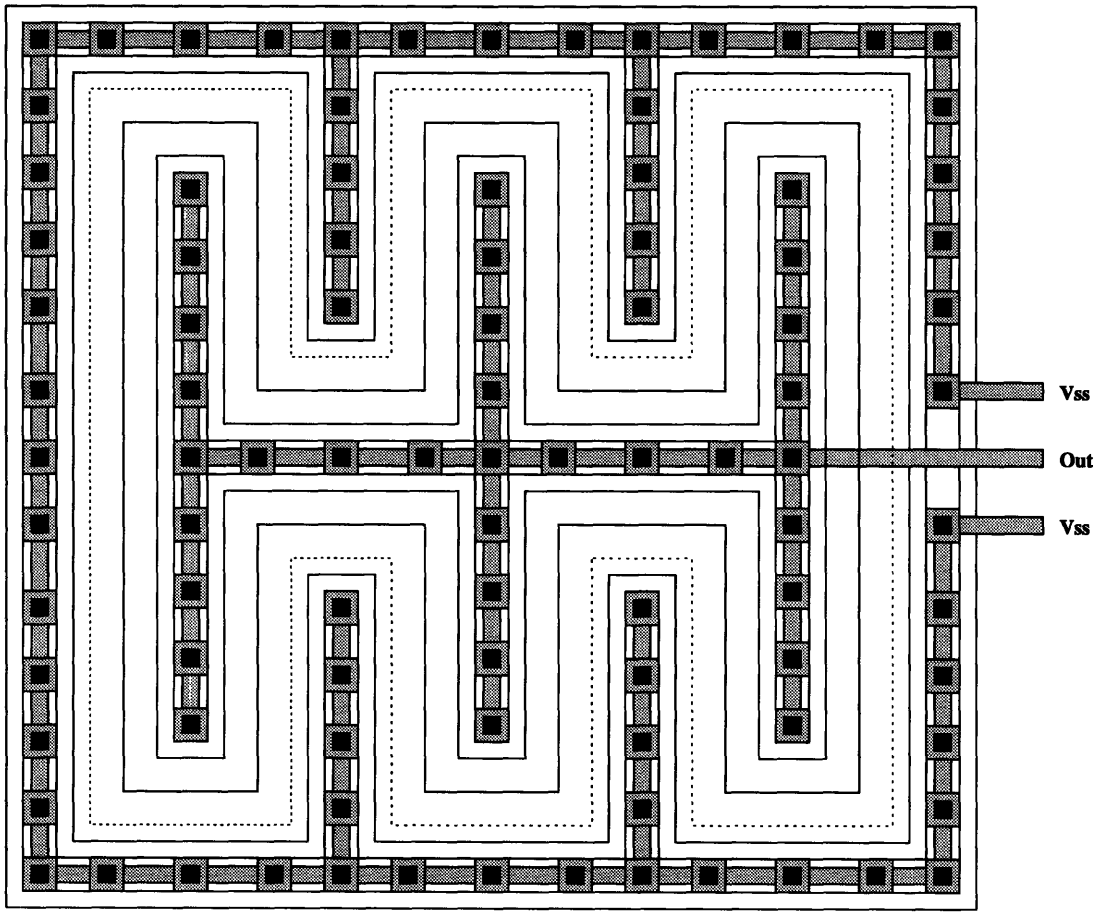


Figure 3-11: N+N/P Interdigitated Finger Photodetector Geometry II

Each photodetector geometry was carefully planned so that synthesizing photodiode arrays can be as simple as overlapping diode instances, sharing guard rings and conserving light sensitive area. The leaf cells can also be pitch-matched to other dimensions without sharing guard rings in order to improve crosstalk isolation. These features allow the *photogen* tool in Chapter 4 to save considerable time in make large

arrays. Figure 3-12 shows three simple 1-D (1X4) arrays made for each of the three photodiode geometries presented earlier. Larger arrays, especially 2-D ones, are simply made by arraying 1-D blocks which have additional routing lines to allow easy access to signals when arrayed.

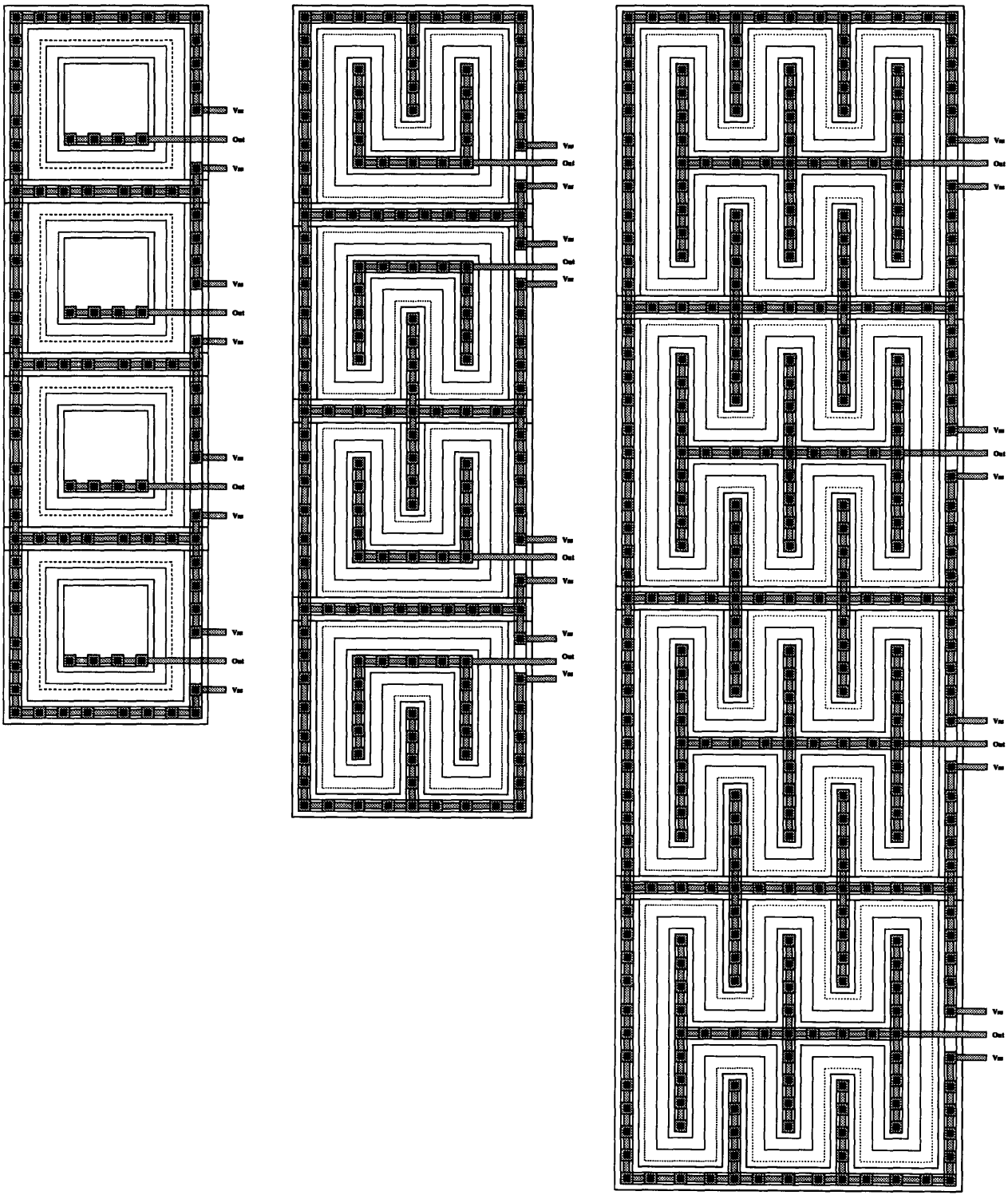


Figure 3-12: N+N/P 1-D (1X4) Photodetector Arrays

Chapter 4

Photodetector CAD Tool Design

4.1 Overview

The complexity of optoelectronic designs is continually increasing, and few research design teams today can possess all the detailed knowledge about free-space optical system setups, optoelectronic interconnection technology, VLSI fabrication, computer architecture and packaging mechanics required to design and build custom optoelectronic interconnect systems [KFC+90]. In order to facilitate a fast turn-around time of optoelectronic designs, standard design tools must be developed. Designs concerns could then shift from managing complexity toward improving costs and more developed testing procedures. A similar trend has occurred in traditional VLSI design tools in the past decade [All88].

The motivation for writing new CAD tools will be discussed in the context of AT&T's MULGA[1] VLSI design system. The required and desired features of a photodetector design and analysis tool include hardware platform independence, process technology independence, and output format independence.[2] Moreover, in an optoelectronic CAD tool, and in particular a photodetector design tool, many options should be user-selectable: photodetector styles compliant to process constraints, synthesis of mask levels including array wiring, and parameter extraction for SPICE

[1]MULGA stands for **M**odular **UNIX**-system based **L**ayout and **G**raphic **A**id

[2]Some common mask formats include Caltech's CIF, Mentor Graphic's L, and AT&T's *xymask*.

modeling. A photodetector design tool, written in C, that meets all the stated criteria is presented in this chapter.

4.2 CMOS VLSI Photodetector Design Tradeoffs

The design tradeoffs mentioned below rely heavily upon the interdependencies of various device parameters mentioned in Chapter 3, which are derived in Appendix A. Please refer to these sections as needed.

4.2.1 Responsivity Versus Bandwidth

The tradeoff between responsivity and bandwidth is the most fundamental design tradeoff of a photodetector because it is inherently coupled to the classical gain-bandwidth interplay from circuit theory.

For most photodetector designs, a high responsivity is desired so that low power can be used to achieve a given signal-to-noise, S/N, ratio which translates into an effective Bit Error Rate, BER. However, a high enough photodetector bandwidth is also needed so that electronic speeds are accommodated.[3] These two desired features, conflict because a high responsivity implies a longer carrier transit time. Transit-time limited response occurs mainly in PIN photodiodes with large intrinsic regions; regular PN photodiodes are generally limited by the photodiode RC time constant because of the small depletion widths, and hence, larger junction capacitance value.

Referring back the Equation 3.42, the only variables under control are the quantum efficiency, η, and the wavelength, λ. It appears as if a larger λ is desirable, however, the wavelength also affects α, the absorption coefficient which, in turn, affects the value of η through Equation 3.34. From this equation, a lower value of λ will maximize the value of η. This is a more mathematical argument for the peaking of the responsivity of a Si photodetector around 850 nm as shown in Figure 3-7 than given in Chapter 3, for values of λ less than the critical wavelength. A high quantum

[3] Actually, the photodetector bandwidth should be at least 2–3 times the minimum bandwidth required in order to prevent signal distortion.

efficiency depends upon low surface reflectivity, R, shallow non-intrinsic absorption depth, l, high depletion width, W, and high absorption coefficient value, α.

Also, it is essential to have low Debye lengths to limit diffusion currents which requires (from Equation 3.35) that the doping level in at least the top layer be high. This constraint is quite beneficial because a heavily-doped top region not only lowers diffusion currents where absorption is significant, but also reduces the value of R_s which improves linearity and frequency response. A shallow top layer is also beneficial to further lower R_s and non-intrinsic absorption — especially for short wavelengths — as well. However, the depletion region on this side of the diode will be diminished which lowers the responsivity and increases the capacitance slightly. As long as the other side of the diode is lightly doped, a large enough depletion region can be supported in this layer to achieve a reasonable responsivity.

A large depletion width, W, requires low doping densities, and high reverse-bias conditions or a large intrinsic, I, region for a PIN photodiode. This constraint poses no difficulties as long as the bottom part of the diode is lightly doped so it supports most of the depletion region, and a moderate bias voltage is applied. Having a large W lowers C_J which is exactly the effect desired for a RC-limited simple PN photodiode.[4] The depletion region in a PIN photodiode is regulated mainly by the I region thickness, which is tailored to meet both responsivity and speed requirements. In conclusion, most of the features of a typical planar VLSI process complement the design of integrated PN photodiodes.

4.2.2 Responsivity Versus Process Complexity

The responsivity of a photodiode can be further enhanced by controlling existing processing steps or adding new ones, requiring more masks. For instance, removing the nitride layer and adjusting oxide layer thickness over the photodetector areas helps reduce surface reflectivity. Ion implantation can be used to control photodetector tub depths and dopings more closely. Also, more efficient PIN photodiodes can be

[4]The depletion region in a simple PN photodiode is typically small enough that transit-time limits are not reached.

produced.

4.2.3 Geometry Versus Responsivity and Noise

Integrated VLSI photodetector geometries like the simple rectangular type shown earlier in Figure 3-9 have high response uniformity, low responsivity,[5] and low noise for small device areas. More complex geometries like the interdigitated styles shown in Figures 3-10 and 3-11 have better uniformities as long as the impinging light spot size is large compared to feature sizes, and a higher responsivity because they take advantage of the high density of vertical (deep) surface PN junctions. However, these two styles produce higher noise due to increased dark currents (because the effective junction area has increased), and it follows that they have larger junction capacitances.

4.2.4 Geometry Versus Responsivity and Series Resistance

High contact and wiring densities reduce optical power reaching the photodiode junction which lowers responsivity; yet at the same time R_s is lowered which increases bandwidth and improves linearity. For all the photodiode designs presented in Chapter 3, minimum wire widths were used throughout because the photocurrents are so small that they will not cause significant voltage drops to appear across interconnect wiring.

4.3 MULGA: AT&T's VLSI CAD tool set

MULGA is AT&T's integrated MOS VLSI symbolic design and analysis tool set [ACW87]. It uses a symbolic intermediate circuit description language, *icdl*, which relieves the designer from having to know the process design rules. *Icdl* uses a virtual grid for design viewing, and it understands the following objects: devices (transistors), wires, pins (labels), and contacts. Once a design is entered, it can be compacted to a

[5] because of the small W for typical CMOS processes and operating voltages

particular technology, and then simulated. Because the *icdl* format is symbolic, the designer can focus on higher level issues such as topology and testability.

A flowchart of the typical cell design process is shown in Figure 4-1. An outline of the module design process is highlighted in Figure 4-2. MULGA files are located in special directories corresponding to function (e.g. all *icdl* files in the **icdl** directory). An example of the hierarchy used in a design directory is shown in Figure 4-3. Most MULGA commands are executed from the top of the design directory.

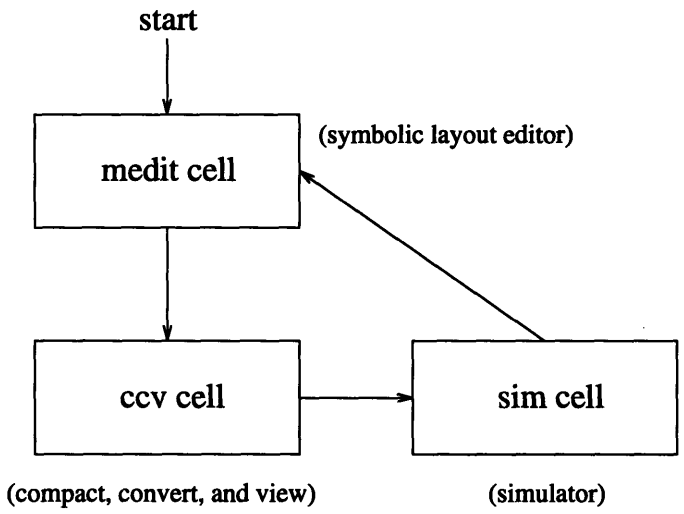


Figure 4-1: Cell Design Loop

4.3.1 Cell Design Loop

The cell is the lowest level in the hierarchy of a chip design. Following Figure 4-1 each cell is iteratively designed and simulated until it functions properly. In MULGA, cells are either designed procedurally or interactively through the *medit icdl* editor. In *icdl*, all symbols are explicitly defined in order that obvious design rule violations — like inadvertent transistors created by placing poly over diffusion — can be caught. In *icdl*, symbols can only be placed at grid intersections on a virtual, topological net. The grid is virtual because only during compaction is the actual grid spacing determined, in order to meet design rules. This feature allows the designer to focus on connectivity information and topology without being concerned with design rules. However, this added flexibility is attained with some considerable costs. For instance,

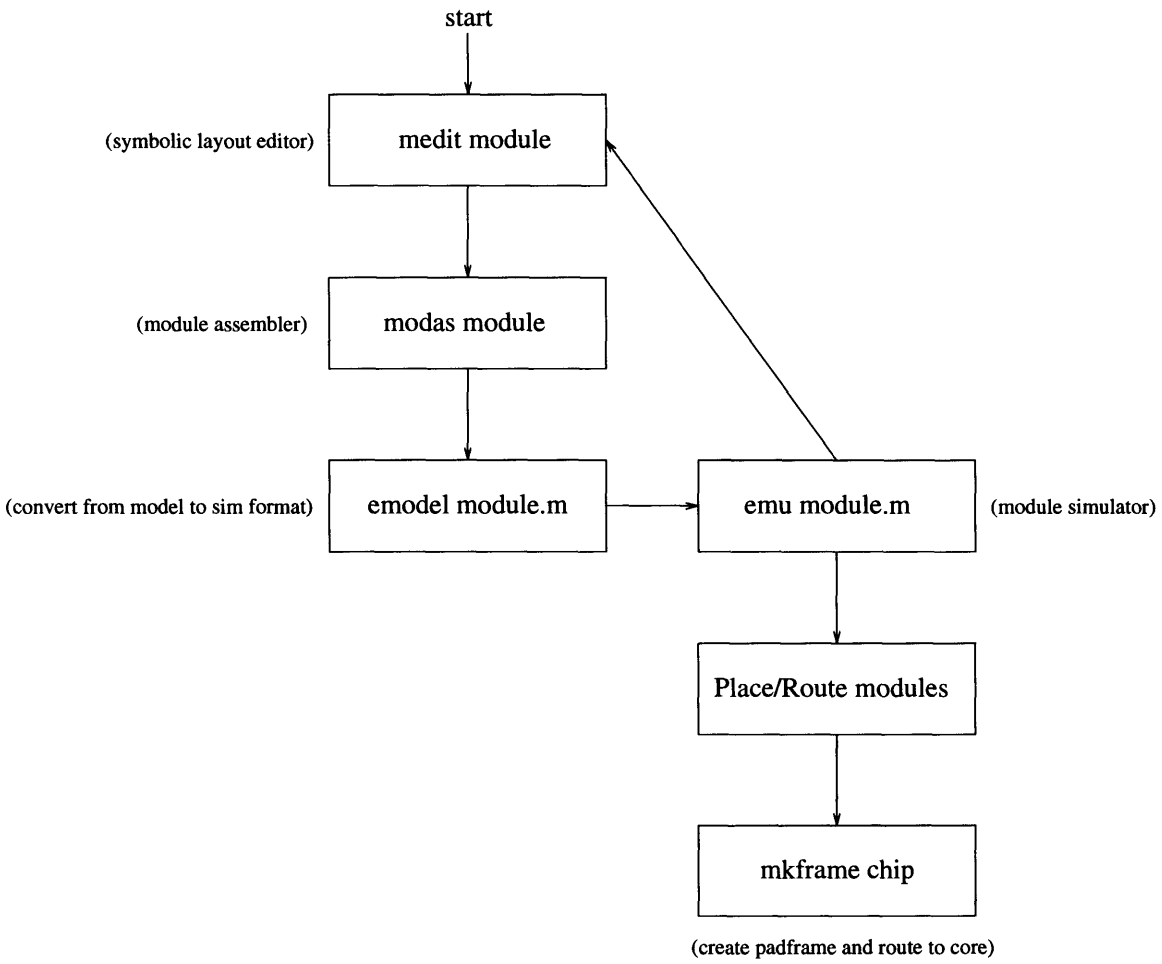


Figure 4-2: Module Design Loop

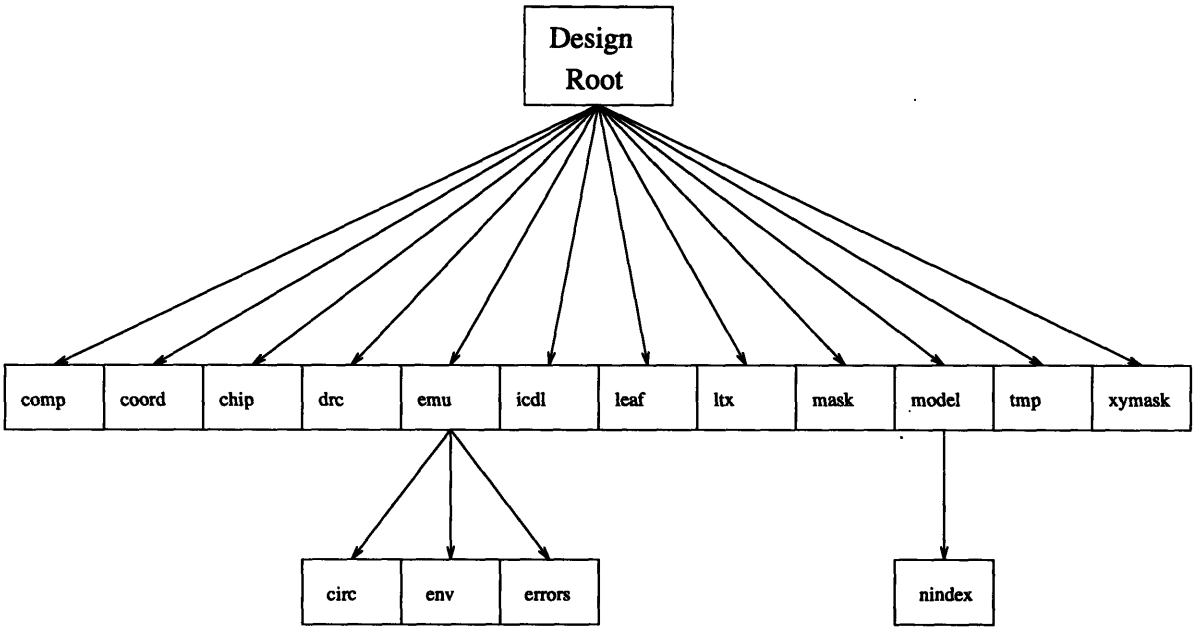


Figure 4-3: MULGA Directory Tree

very intelligent compactors are needed in order to produce area-efficient layouts, and a designer may spend considerable time and effort discovering how to finesse a not-so-intelligent compactor to place devices and wires in compact spaces. Other drawbacks of a symbolic circuit editor will be described shortly.

The *ccv* utility is used to interactively change the layout of a circuit to find its optimal area. This command compacts a design using *comp* to satisfy process design rules. Next, the design is converted from *icdl* format to *mask*, a geometric layout format which can be displayed by *medit*.

4.3.2 Module Design Loop

A complete chip consists of one or several modules that are either abutted or routed together; each module is composed of several leaf cells that are abutted. The utility, *modas*, compacts a symbolic *icdl* design into a *mask* file, also pitch-matching adjacent cells. To save processing time, it only compacts those *icdl* designs which have changed since the last compaction. The place-and-route program is an optional step which allows modules to be automatically placed according to constraints and routed. The

mkframe command places the I/O buffers for the specified AT&T process[6] around the perimeter of the design and routes signals from these buffers to the core circuitry. The final chip *mask* design is then converted to *xymask* format for fabrication.

4.3.3 Limitations of MULGA

The MULGA system has been the major CAD package used by AT&T MOS VLSI designers since the early 1980's. Since that time, it has gone through many revisions, but there still are limitations to the system which cannot be avoided without a serious redesign of the entire package or the addition of separate tools. Most AT&T researchers have opted for adding new programs to the system rather than attempting an entire system redesign. Several powerful event-driven simulators have been added to the suite of tools in the past few years, for example.

The MULGA system is composed of separate, yet interrelated UNIX programs with support for various window managers. For most application suites, if one program is changed or added, yet still supports the same database formats,[7] no additional work is needed to make the program compatible with the entire tool suite. However, the amount of work required to update the existing programs in the MULGA package to handle new optoelectronic components, such as lasers, photodetectors, and waveguides is great because of limitations in the designs of the tightly-coupled program modules. Less time and debugging effort is required to design separate programs that can easily integrate into the MULGA environment (and talk to each other with their own database formats) and not be tied down to MULGA-specific constraints. The optoelectronic "front-end" would also be portable to other CAD platforms as long as their respective database formats are supported as well

A serious limitation in the *icdl* netlist format is that the symbols (transistors, wires, pins and contacts) are hard-coded, so a language extension would be required to handle optoelectronic components. To further complicate matters, the symbolic

[6]Standard pads are available for each AT&T process, and fixed padframe sizes are used in the various shuttles.

[7]including circuit and simulator netlist formats and mask layout formats

compactor (*modas*) doesn't recognize photodetector or other optoelectronic geometries in the *icdl* design file. It would simply flag design violations or completely compact away the optoelectronic components. Moreover, the extraction utilities would not create the proper models or even know what elements should be included in the simulation netlist. Design Rule Checking, DRC, programs may also need to be modified to handle new exceptions to the design rules caused by the various devices.

Another drawback of the MULGA system is that *icdl*, *mask*, and *xymask* file formats are not completely interconvertible. The *xymask* format is used for actual chip submissions only — it can only be viewed and edited by a primitive tool, *GRED*; *mask* is mainly used for viewing via the *medit* tool, and only RECTangle objects can be edited. This means that typically, final chip designs (in *xymask*) are converted to *mask* for viewing. However, not all *xymask* constructs are supported in the conversion from *xymask* to *mask*,[8] so if changes are made to a converted *mask* file with *medit*, the design can be converted back to *xymask*, but extensive editing of the new *xymask* will be required. Moreover, a *mask*-to-*icdl* converter is not available, so if a designer is given only a *mask* file (or a CIF file from an outside vendor), it is extremely difficult to make design changes. Figure 4-4 shows the mask file conversions possible in MULGA.

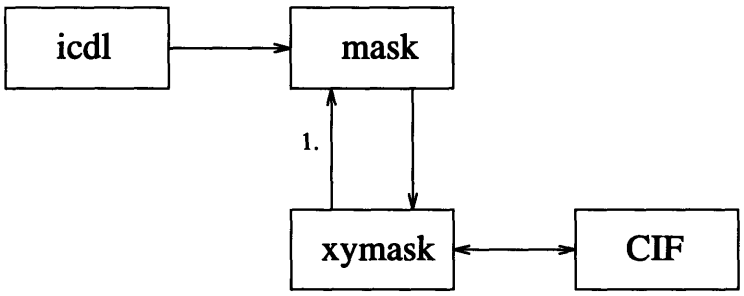


Figure 4-4: MULGA Mask Format Interconversions

Having several possible styles of photodetectors which comply to typical CMOS process constraints compounds the complexity. Adding capability for integrated waveguides and lasers (in an enhanced optoelectronic technology) to the tools re-

[8]This is a long-known, but unfixed bug!

quires knowledge of issues like substrate minority carrier injection via light absorption, laser current leakage, material reflectivities and transmissivities, critical angles of waveguides, etc. which current tools do not consider.[9]

4.4 *Photogen* — A Photosensor CAD Tool

This section describes the design and implementation of a photodetector synthesis and analysis CAD tool, named *photogen*. In Section 4.2, relevant photodetector design tradeoffs were discussed in light of what design constraints are imposed by a standard CMOS VLSI process technology. Now, important features required in such a tool are outlined, then the implementation of each feature is explained in more detail.

4.4.1 Required Features of a Photodetector CAD Tool

As mentioned earlier in Section 4.3.3, there is clearly a need for a separate photodetector synthesis and analysis tool that can be integrated with *any* tool set, not just the AT&T MULGA tool suite, with little or no extra effort, until commercial optoelectronic tools become available.

The program should be written in a common programming language, such as C, in order to obtain hardware platform independence and to ease making changes or enhancements by developers. Compatibility with the UNIX operating system is a major advantage because most VLSI CAD tools are designed for UNIX workstations. Furthermore, modular code is essential to ease program maintenance and enhancements. New features can simply be implemented in separate modules, then re-linked into a final executable with small effort.

A CAD tool that does not depend upon process parameters particular to a given technology (process independence) is also a very important asset. In particular, having a tool that supports the standard MOSIS VLSI shuttle service would greatly reduce design fabrication costs, placing the option of custom photodetector design

[9] CMOS tools could easily estimate minority carrier substrate injection currents in order to prevent latchup and recommend body plug placements.

into the hands of even small companies, universities, and individuals.

CAD tool independence, hardware platform independence, and process technology independence features are highly desirable, but unfortunately come at a substantial cost in program complexity. This complexity, however, can be organized into manageable sections of code (modules) which help improve program maintenance and upgrades.

4.4.2 Integration into Commercial CAD Systems

The *photogen* tool easily integrates into the AT&T MULGA environment by sharing the MULGA directory tree structure of Figure 4-3. A subdirectory, **photogen**, is added to the MULGA working directory to contain all the *photogen*-related technology files and source code. Auto-install shell scripts are provided which ease the installation from a single compressed *tar* distribution file. Separate *Makefiles* are provided which ease compilation. Compiler and machine-specific options can also be set in convenient *Makefile* environment flags. Finally, a simple test script is provided which can be used to verify proper compilation by comparing actual generated files with expected output.

Output File Formats

Since this tool was originally designed to work with MULGA, it has program output modules included which generate AT&T *mask* and *xymask*-compatible output files. Note that from previous arguments, *icdl* files are not — and should not — be produced with this tool because the AT&T compactor tools are not designed to handle new photodetector (symbolic) geometries. AT&T's ADVICE file format is also supported for simulation.

The majority of other commercial CAD tools can input mask files encoded in the CIF file format, therefore, the *photogen* tool has incorporated a CIF output format module. Because of the modular nature of *photogen*, adding another output

file format, like GDS, is very straightforward.[10] An output module which produces Berkeley's SPICE netlist files and another module which produces AT&T's ADVICE netlist output files for simulation are included in the *photogen* package. Again, other simulation output file formats are possible.[11]

4.4.3 Internal Database Format

In order for the *photogen* program to attain CAD tool independence, it must maintain its own internal database of all process design rules, process layer information and process parameters so it can synthesis contacts, wires, other device features, and terminals and text for circuit connectivity. Moreover, it must be able to extract device parasitic information for simulation. Because it deals with generic parameter names which are in the internal database, it need not rely upon any one particular CAD tool database format. As will be described shortly, this internal database format is extremely flexible and straightforward, so little work is required to setup the program to handle any CMOS technology.

Hash Table Implementation

The *photogen* database is implemented as a simple hash table of process design rules, available routing layers, and process parameters [KR78]. The *photogen* program performs a hash table initialization when the program starts and a technology is selected, reading various process text files and mapping logical parameter names (which are strings) as indices or keys to their associated values (which can be strings or numerical values). When the *photogen* tool needs the process-specific value of a parameter, it simply looks up the corresponding value in the database. The hash table is quite small (under 250 entries) for a typical technology, so accesses are very fast.

The choice of repeated hash table lookups over simply initializing internal variable names to values read from text files at startup, then accessing variable values from then on may seem like overkill at first, but the technique has many advantages. If a

[10] CIF-to-GDS file format translators are also freely available.
[11] The LSIM format used in Mentor Graphic's GDT tool set is also common.

writer of a *photogen* technology file failed to initialize an important parameter, the hash table module would catch the error when it tried to lookup the value, and report which parameter is ill- or undefined using one central error handling routine for string and numeric constants.[12] In contrast, for an internal variable approach, variables would need to be pre-initialized to some illegal value before being initialized from the file, so that undefined variables can be flagged. This hash table implementation choice organizes the tool's data resources, greatly simplifies the error detection mechanisms of the CAD tool, makes catching design rule bugs easier, and allows developers easy access to the database format without having to add new program variables; new parameters can easily be added for yet more geometry synthesis modules. All these benefits are bought for the small price of additional software overhead, and a minor reduction in operating speed.

Technology Files

The process parameters, design rules, and technology layers text files reside in the **PARAMS**, **RULES**, and **LAYERS** subdirectories, respectively, of the *photogen* tree. These files contain the information used in initializing the internal database mentioned earlier. For the scalable MOSIS technologies, there exists a single layer file, **mosis.cif**, in the **LAYERS** subdirectory which contains definitions for all the CIF layers. Only a single layers file is needed for all the MOSIS layers because the CIF mask output file format is always produced for MOSIS SCMOS designs.[13]

Another file, **mosis.scmos**, in the **RULES** subdirectory encapsulates the MOSIS generic SCMOS rules presented in Appendix C.[14] These design rules are simply scaled by the value of λ[15] to arrive at the generic (over-conservative) design rules for that technology. Finally, separate process parameter files should be provided (in the **PARAMS** subdirectory) for specific MOSIS vendor technologies; otherwise syn-

[12]Out-of-bounds checking is, of course, done in a local routine where the variable is used.

[13]If another common output format module, like GDS, were incorporated into the tool, a separate **LAYERS** file, like **mosis.gds** would be needed.

[14]Vendor-specific design rules can be used instead of these generic SCMOS rules in order to further optimize detector design.

[15]λ is defined to be one-half the minimum feature size for the technology.

thesized layouts will be sub-optimal, and extracted parasitic models will not be very accurate. Essentially, if more tightly compacted geometries and the highest responsivity photodetector designs are desired, then vendor-specific **RULES** and **PARAMS** files should be included.

For vendor-specific non-MOSIS designs, three separate files (one in each of the **PARAMS**, **RULES**, and **LAYERS** subdirectories) must be provided. For example, the files **LAYERS/att0.90**, **RULES/att0.90**, and **PARAMS/att0.90** were produced for the AT&T 0.9 μm technology used to fabricate a simple photosensor array described in Chapter 5. Actually, only important parameters, layers, and design rules need to be specified, so these files are not excessively difficult to create, to enhance features, or to modify. Until a standard database format can be agreed upon among CAD tool designers, this simple hash table database approach seems to be the best solution for CAD tool independence.

4.4.4 MOSIS Compliance

By providing full MOSIS support, very inexpensive optoelectronic receiver chips including photodetector arrays can fabricated. The *photogen* program requires either user entry of the value of λ for the selected (scalable) technology, or the name of a MOSIS-compatible vendor-specific technology, for which **PARAM**, **LAYER**, and **RULE** files have been created. The latter case, will of course, create more efficient layouts and more accurate simulation models.

4.4.5 Operation

Figure 4-5 is a simple flow of operations for the *photogen* program showing its two distinct modes of operation. First, it can simply take command-line options, and run in what is known as batch mode, where it reaches its final output without further user intervention. The other mode is interactive, where the design and analysis loop can be iterated many times until estimated device performance meets required optical and electrical specifications. Finally, the synthesis and parasitic extraction steps are

performed.

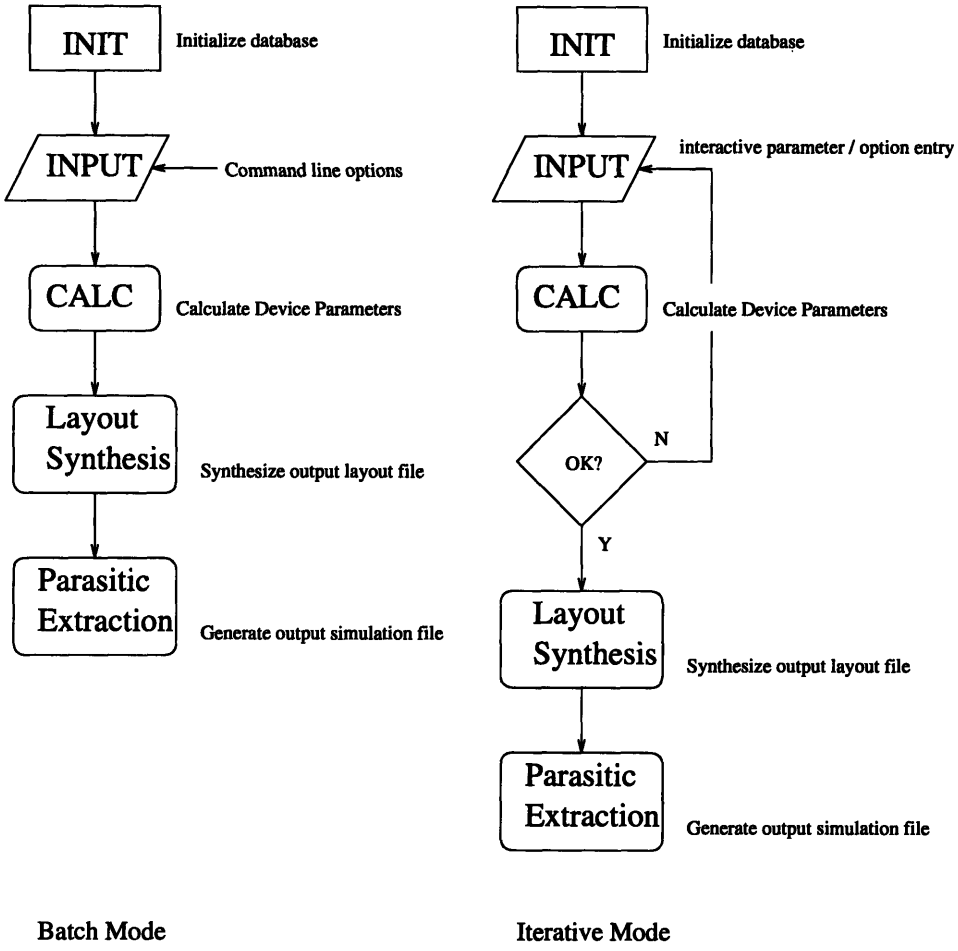


Figure 4-5: *Photogen* Process Flow

User Input

Before the *photogen* tool can synthesize photodetector layout and circuit model files, there are several parameters and options which must be specified. First, the following optical link parameters must be specified:

- CMOS technology: MOSIS SCMOS, AT&T 0.9 μm, etc.
- Layout file format: CIF, *mask*, *xymask*, etc.
- Simulator file format: ADVICE, SPICE, etc.
- Laser light wavelength

- Laser power output
- Optical link efficiency[16]
- Optical spot to detector size ratio
- Photodiode reverse-bias voltage

Next, there are many options which may be selected in order to fine-tune detector performance:

- Photodetector geometry style (as shown in Figures 3-9, 3-10, and 3-11).
- Detector periphery dimensions (L x W)
- Detector contact density
- Detector wiring density
- Detector metal-2 wiring (when available)
- Arrays: 1D or 2D (1x1 — or single detector — is the default)
- Array spacings (x-dimension and y-dimension pitches)
- Array wiring style: straight 1D, top 2D, or both 2D
- Desired Responsivity: [mA/mW]
- Desired Response Time
- Maximum Noise power

[16] η_{link}, the efficiency from laser output to chip surface

4.4.6 Layout Synthesis Algorithms

Three separate layout synthesis program modules exist for each of the CMOS photodetector geometric styles presented in Chapter 3. These modules all call three high-level mask synthesis subroutines which generate RECTangles in particular process layers, and CONTACTs between the various layers. The RECT() subroutine requires the bottom-left and top-right coordinates and the process layer name as arguments. The CONTACT routine requires the contact type (instance) name and the center coordinate. The WIRE() subroutine requires wire length and width, starting and ending points, and the layer name. It simply calls RECT() with the calculated coordinates. The TRANSISTOR primitive works in a similar way as the CONTACT() call: an instance name and center location must be specified.

Before the layout synthesis begins, each contact type used in the detector is first generated out of constituent RECTangles. Figure 4-6 shows the contacts used in the AT&T 0.9 μm CMOS N-well process. This contact generation process produces a library of contacts which can later be instanced in the final mask file to greatly reduce layout file length and layout algorithm complexity. The option to generate various transistor instances is also provided for later CAD tool expansion to include electronic amplifier synthesis as well.

The RECT() and WIRE() mask synthesis subroutines, in turn, look up layer names in the database, and call the appropriate output file module in order to produce layout mask of the correct format. The individual RECTangles and CONTACTs are placed with regard to the photodetector style, process design rules, selected contact and wire densities, and estimated depletion width for the given bias conditions. If a detector of a given geometric style fails to meet all the geometric specifications,[17] then the next smaller style will be attempted.

If no initial geometry is specified, and a moderate responsivity, low-noise detector is required, the style attempted first by default will be the simple rectangular design shown in Figure 3-9. If higher responsivity is required, but more noise and capacitance

[17]for instance, if the depletion width is not being supported or other features are smaller than minimum sizes set by the process design rules

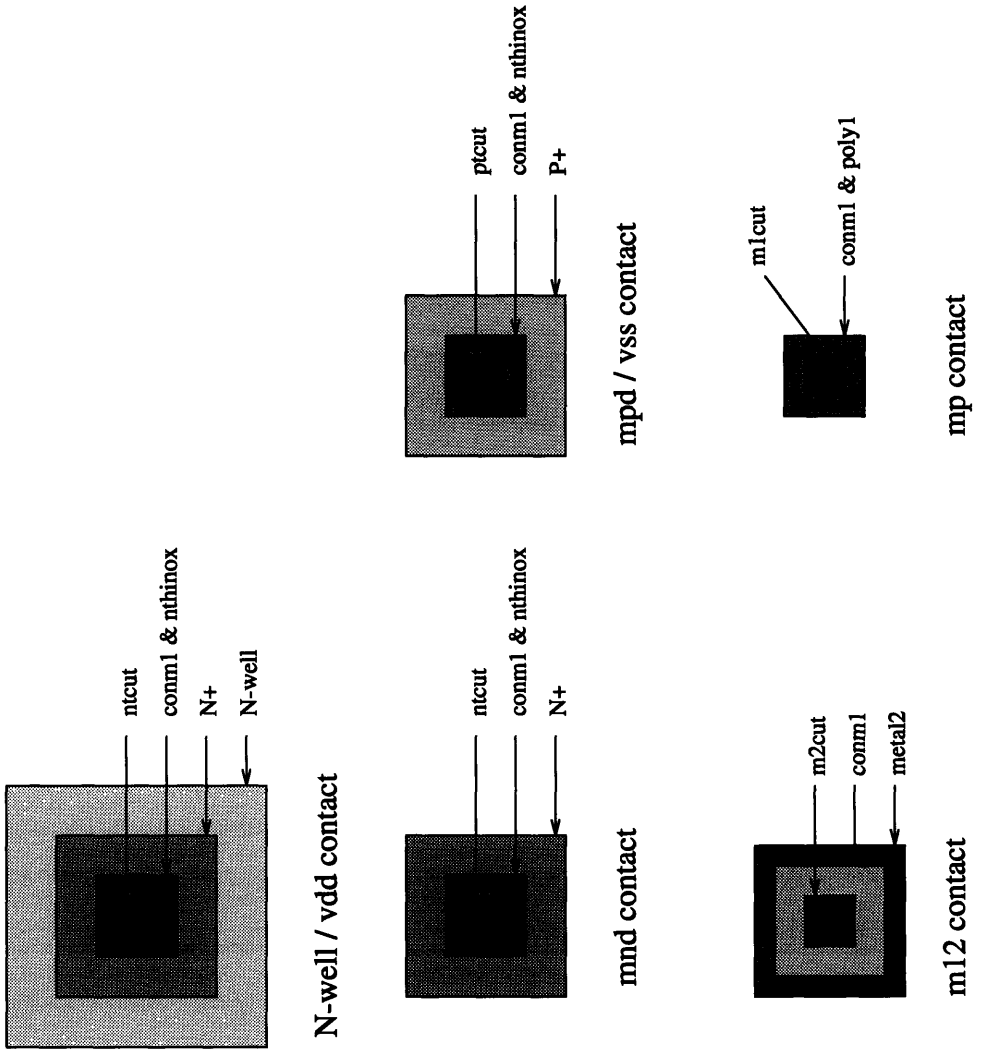


Figure 4-6: Mask Contact Types for N-well Process

can be tolerated, then the style shown in Figure 3-10 is used. Finally, if the highest responsivity is required at the expense of large capacitance and higher noise, then the style in Figure 3-11 is selected initially. Once a single detector cell is produced, it is simply instantiated at the correct pitch in order to produce 1-D arrays. The wiring to the 1-D array is automatically routed from the individual detectors to either the left or to the right or half to either left or right sides. An array of simple photodetectors for the latter wiring style is shown in Figure 4-7.

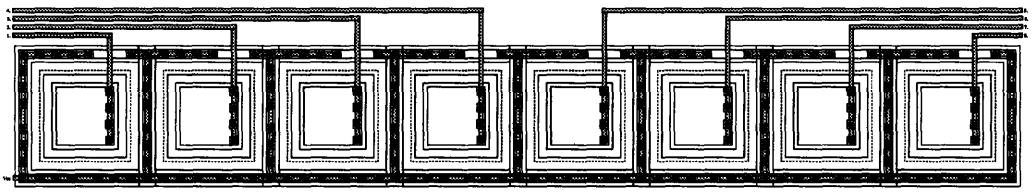


Figure 4-7: 1-D (1x8) Wired Photodetector Array

To make 2-D arrays, a previously-created 1-D (wired) array is simply instantiated forming adjacent 1-D arrays as depicted in Figure 4-8. This wiring approach works fine for large vertical inter-pixel spacing; it will provide small inter-pixel vertical pitches only when the number of horizontal pixels is small because of the wiring space requirements. To achieve higher density 2-D arrays at the expense of bandwidth, true 2-D x-y (row-column) addressing support can be added, with pixel selection electronics incorporated into each detector. This technique can be used to produce large 2-D imaging arrays.

4.4.7 Parasitic Extraction Algorithms

As the layout synthesis is proceeding, important diode parameters are being calculated (or estimated) including the periphery and area junction capacitances, series resistance (from contact density, reverse-bias voltage, bulk material resistivity and geometry including device area), and shunt resistance. From the process parameters, a diode model card is created which contains estimates for all the nonlinear capacitive parasitics. Alternatively, a linear capacitor model with capacitance value equal to the junction capacitance for the given area and reverse-bias conditions can be substituted

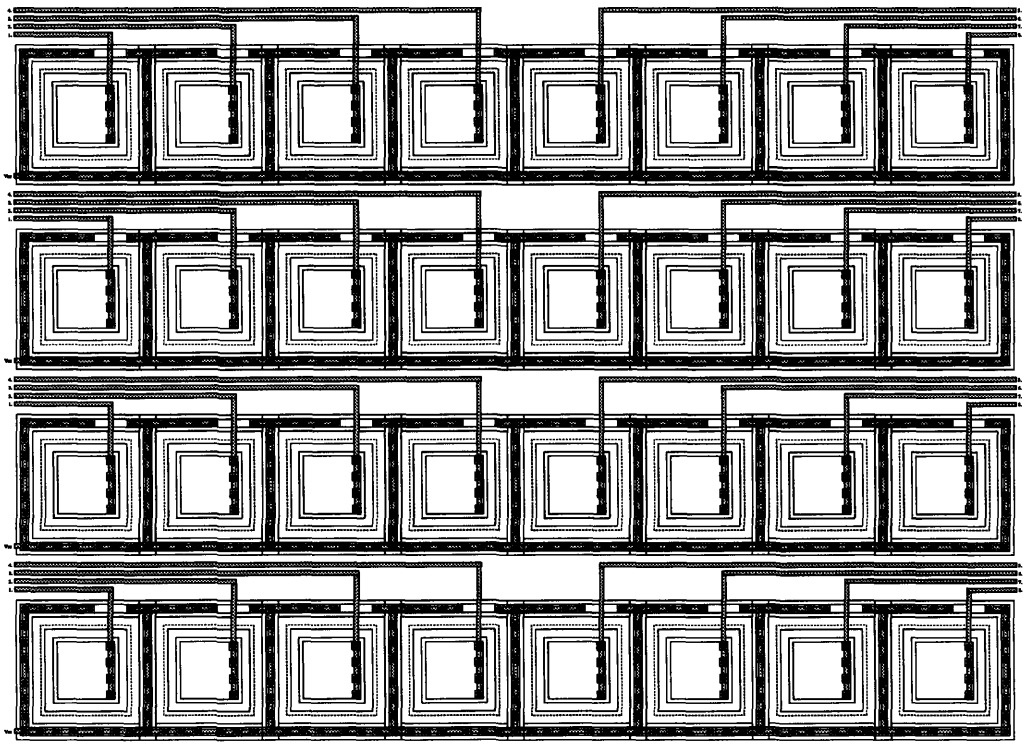


Figure 4-8: 2-D (4x8) Wired Photodetector Array

for the nonlinear diode capacitance. This linear model will be accurate as long as the reverse-bias voltage remains fairly constant, as in the case where the diode is connected to a transimpedance amplifier which attempts to maintain a constant diode bias voltage through negative feedback.

Chapter 5

Results

Figure 5-1 shows a linear microlaser array fabricated by AT&T. This laser array emits light at 850 nm wavelength, and was used to test fabricated photodiode arrays. Notice how one of the bonding wires has inadvertently snapped off the bonding pad causing local damage to the chip. The mechanical nature of bonding along with the electrical parasitics associated with off-chip interconnects are two driving forces behind monolithic optoelectronic systems.

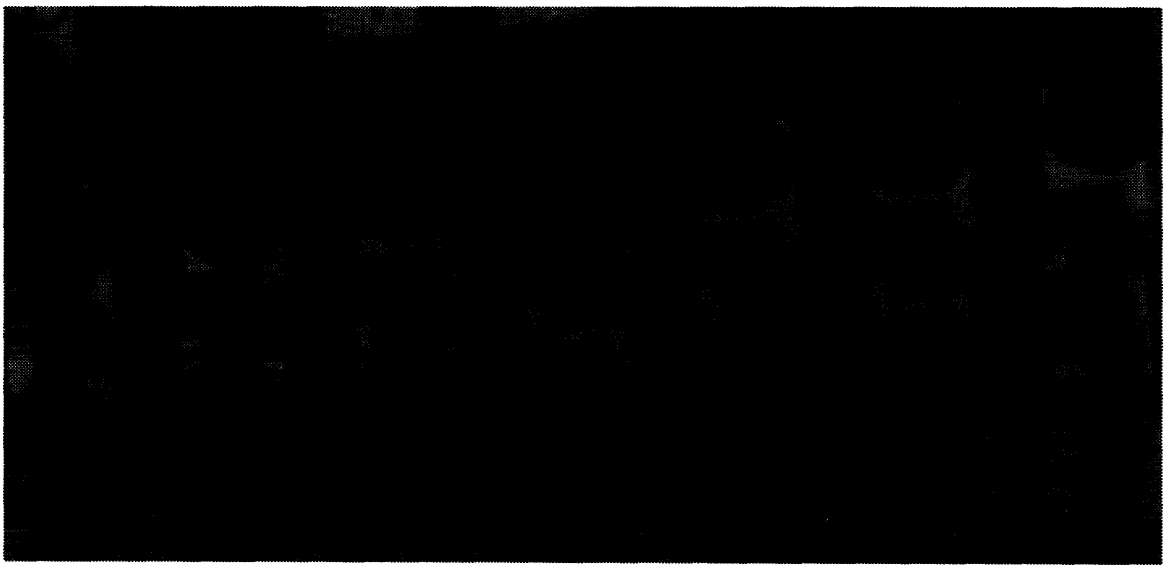


Figure 5-1: AT&T GaAs Microlaser Array

5.1 MULGA Mask Files

A 1 x 2 array of simple photodiodes was generated with the *photogen* program of Chapter 4 for the AT&T 0.9 μm CMOS process. Figure 5-2 shows a *medit* top-level view of the array including the bonding pads. Figure 5-3 depicts a view of the circuit at the lowest hierarchy, and Figure 5-4 is a closeup of a section of the photodiode array. Referring back to Figure 5-3, from left to right the bonding pads are electrically connected to V_{SS}, V_{DD}, and the two photodetector signals. The last two signal pads have integrated ESD protection.

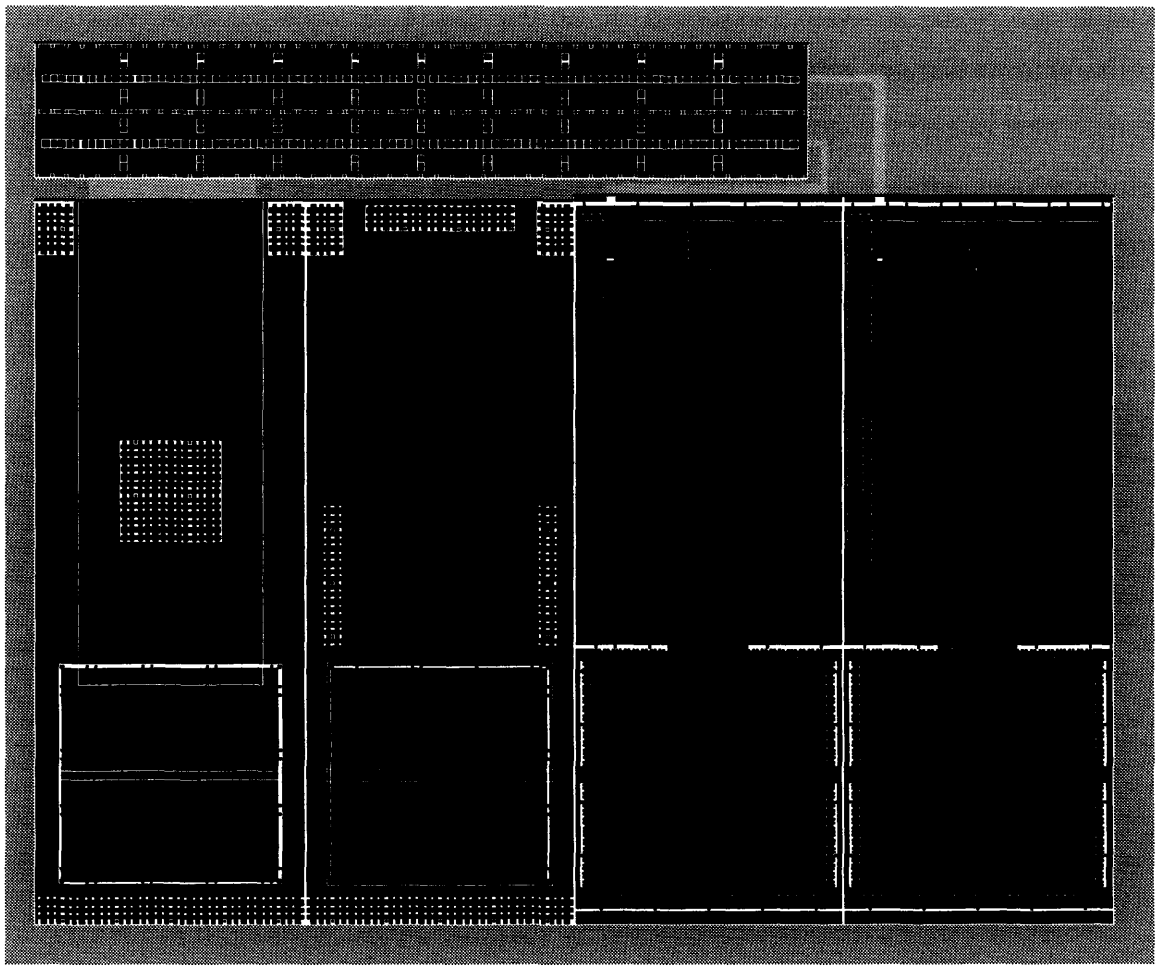


Figure 5-2: Medit View of Photodetector Array Top Levels

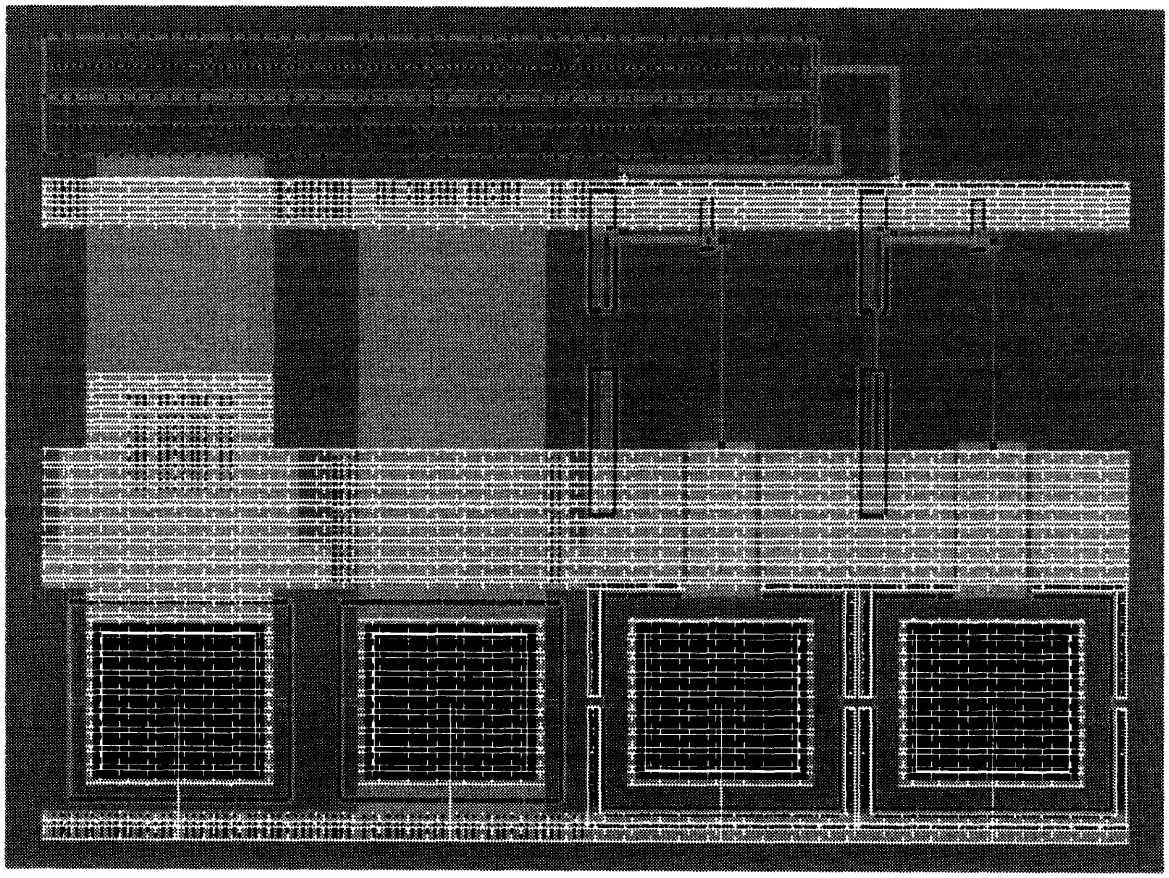


Figure 5-3: Medit View of Entire Photodetector Array

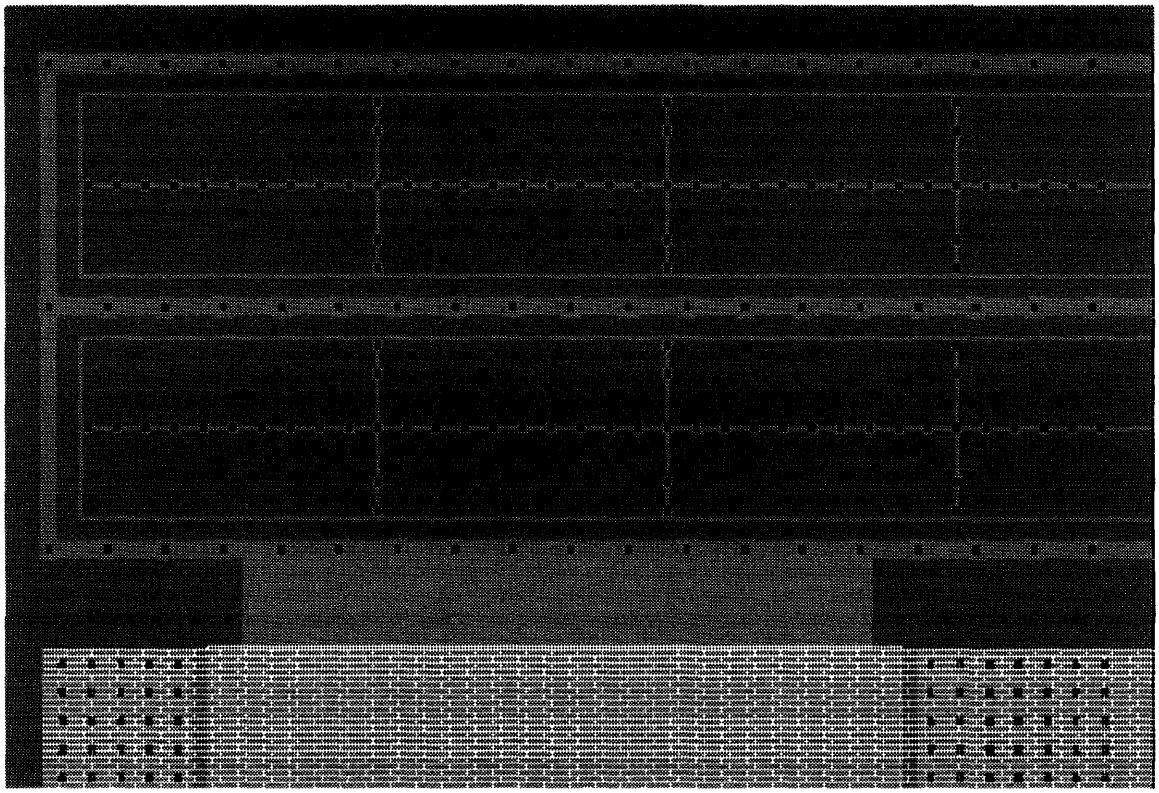


Figure 5-4: Closeup of Photodetector Array

5.2 Fabrication and Bonding

Because of limited departmental funding and time constraints, only one chip design was submitted to a shuttle run.[1] Many other small projects shared the run with another, larger design, so available area was also limited. In fact, the simple array design submitted used all allotted (600 μm X 400 μm) area. After the wafers were fabricated, wafer probe tests revealed some problems with the large project, so the die were never packaged. With the help of other AT&T processing engineers, in one week the rest of the packaging steps were finished.

First, photoresist was spun on the wafer for protection. Then a diamond wafer saw was used to cut the wafer into die. The photoresist was then removed. Finally, the die were mounted substrate-side-down into windowless packages using conductive epoxy to provide a highly-conductive substrate contact. Using a wire bonding machine, the

[1]The MOSIS support for the *photogen* program was not completely implemented at this time.

die were connected to the packages. Figure 5-5 shows the fabricated circuit before wire bonding. Figure 5-6 illustrates a bonded array. Figure 5-7 is a closeup of the ESD protection (two diodes and a resistor) circuitry of an input pad.

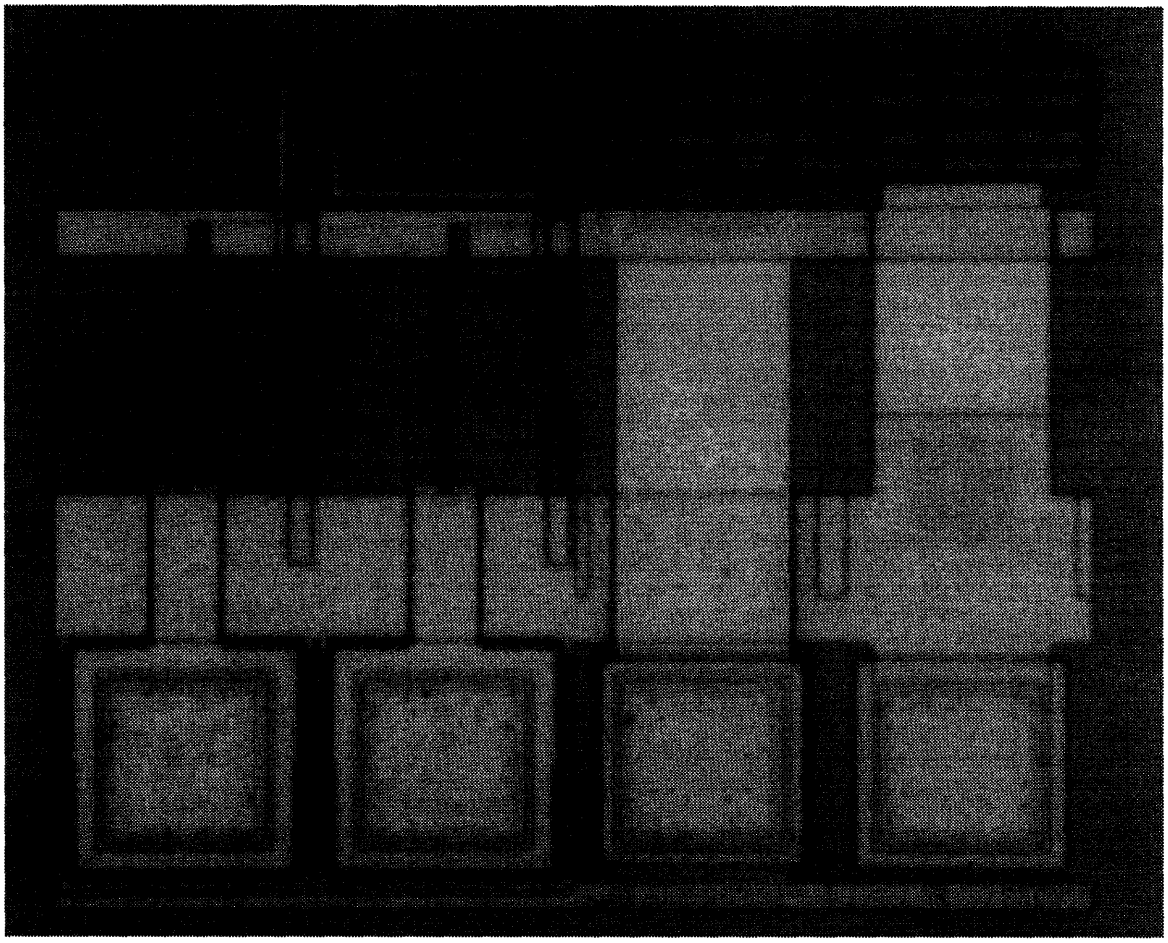


Figure 5-5: AT&T 0.9μm CMOS Photodetector Array

5.3 Optical and Electrical Measurements

Next, the packaged arrays were tested. Monochromatic light with 850 nm wavelength was passed through a neutral density filter in order to adjust the light intensity on the photodiode surface. The three distinct filters used and the corresponding light intensity reaching the surface of the photodiode for an initial laser output power of 1mW is given in Table 5.1.

Table 5.2 shows the measurement results for the simple photodiode arrays for 0V

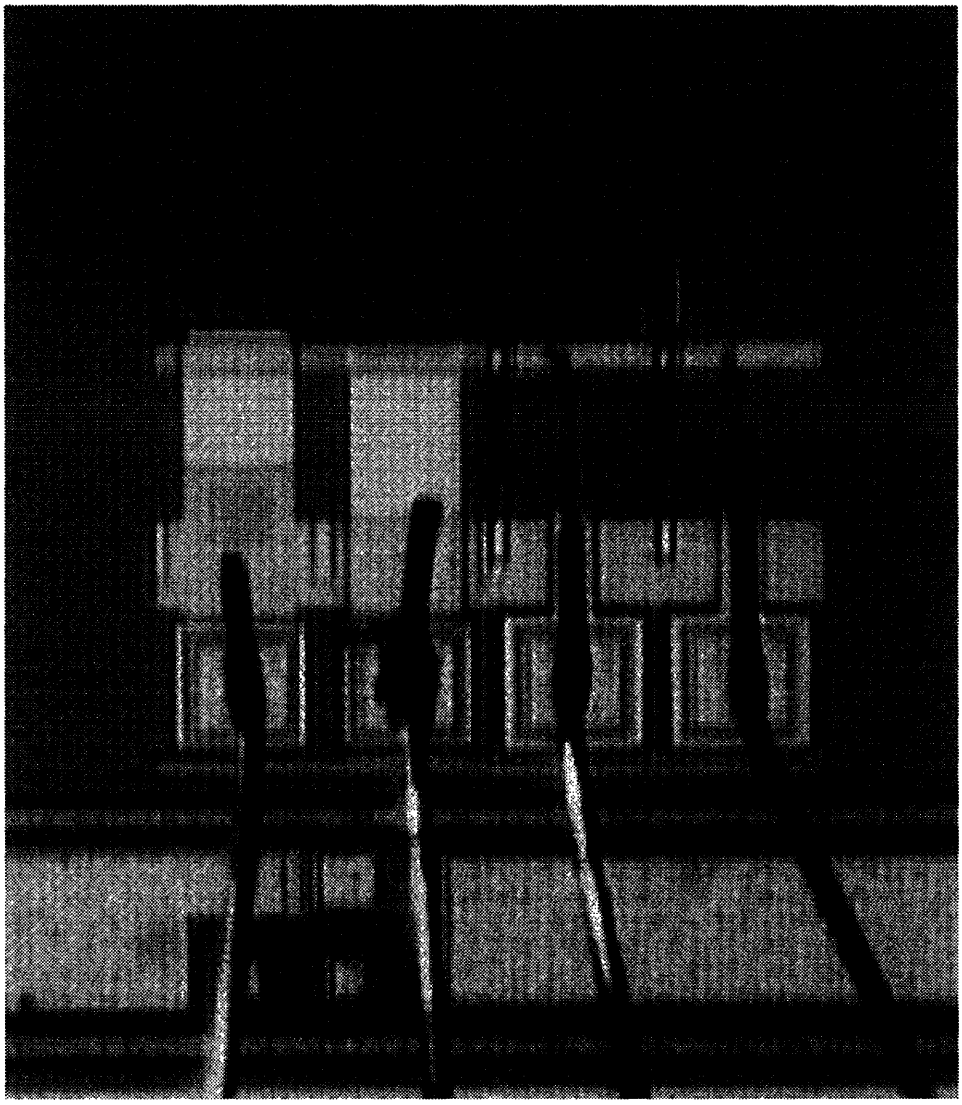


Figure 5-6: AT&T 0.9μm CMOS Photodetector Array Bonding

ND filter	Power
150	0.43 mW
90	109 μW
210	5.3 μW

Table 5.1: Neutral Density Filter Powers

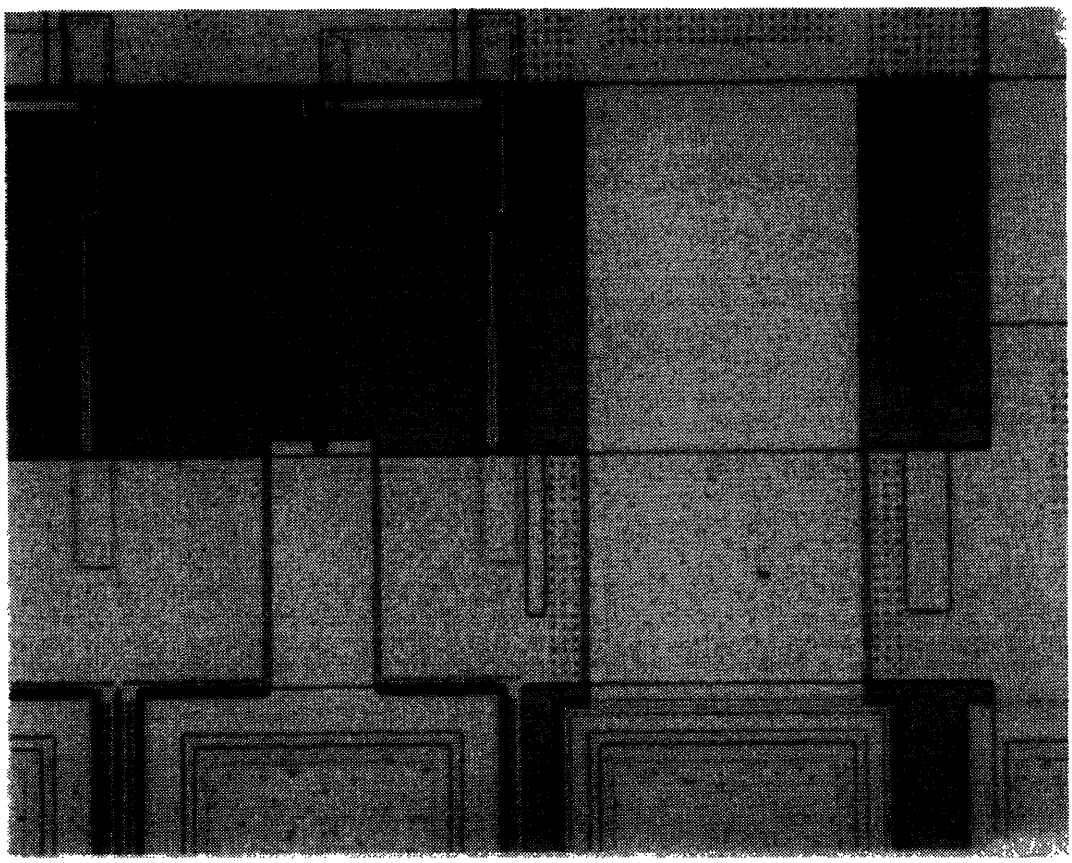


Figure 5-7: AT&T 0.9μm CMOS Photodetector Array Pads

ND filter	Measured I_{ph}	Responsivity
150	47.7 μA	0.11
90	12.1 μA	0.11
210	0.6 μA	0.11

Table 5.2: Photodiode Surface Edge-Junction Response for 0V bias

ND filter	Measured I_{ph}	Responsivity
150	3.3 μA	0.01
90	0.8 μA	0.01
210	0.04 μA	0.01

Table 5.3: Photodiode Center (Tub) Response for 0V bias

applied bias where the light spot impinges upon the surface edge junction. Table 5.3 shows the responsivity for the simple photodiode arrays for 0V applied bias where the light spot strikes the central tub area. These results are in agreement with estimated detector performance and the discussions in Chapters 3 and 4. For the 0V bias condition with hardly any light on the photodiode, however, the dark current was measured to be about 61 nA/mm^2 which seemed rather large. The estimated junction capacitance was around 160 pF/mm^2.

Table 5.4 shows the measurement results for the simple photodiode arrays for 2.5V applied bias where the light spot is incident upon the surface edge junction. Table 5.5 shows the responsivity for the simple photodiode arrays for 0V applied bias where the light spot impinges upon the central tub area. These results are in agreement with estimated detector performance and the discussions in Chapters 3 and 4. The responsivity of 0.55 measured in Table 5.4 is close to the theoretical optimal — which is also maximal — responsivity of 0.65 for silicon for a wavelength of 850 nm. The estimated junction capacitance for this reverse bias point was around 72 pF/mm^2.

ND filter	Measured I_{ph}	Responsivity
150	237.9 μA	0.55
90	59.3 μA	0.55
210	2.8 μA	0.55

Table 5.4: Photodiode Surface Edge-Junction Response for 2.5V reverse bias

ND filter	Measured I_{ph}	Responsivity
150	17.7 μA	0.04
90	4.3 μA	0.04
210	0.4 μA	0.04

Table 5.5: Photodiode Center (Tub) Response for 2.5V reverse bias

Chapter 6

Conclusions and Future Research

6.1 Overview

Monolithic optoelectronic integrated circuits are becoming popular today because of the limitations in electrical interconnects discussed in Chapter 2. By integrating photosensors and electronics on the same substrate, costs are reduced, circuit board real estate is decreased, electrical parasitics are decreased, increasing available bandwidth, and the overall system becomes much more reliable than a hybrid approach because of lower mechanical problems[1] and the use of proven technologies. Combining two such system components may strain a typical process technology which may, in turn, require some minor changes in order to optimize performance. However, with redundant components, and error-correcting electronics, the system can usually tolerate sub-optimal performance by using many lower-bandwidth communication channels instead of a few ultra-high speed ones.

In designing monolithic optoelectronic circuits including integrated photodetectors, it is very important to use a standard process technology — at least in the early stages of design. VLSI Si processes offer the advantages of cheap, easy-to-integrate photodiodes with little or no process modifications. Optoelectronic receiver chips can benefit from the low cost and high reliability of integrated CMOS electronics as well.

[1] such as those associated with wire-bonding, for instance

6.2 Conclusions

It has been shown in Chapter 3 that various types of photosensors are compatible with a standard Si CMOS process. The simple abrupt graded junction N+N/P diode was shown to have the best optical and electrical performance of all the types presented so it was used as a template for a photodetector analysis, synthesis and extraction CAD tool presented in Chapter 4. Using Si as a photodetector material is not the most efficient, nor does it allow design of the fastest photosensors, but it has been proven in Chapter 5 that one can make high responsivity, low dark current, inexpensive, mass-producible monolithic photodetectors in a standard, unmodified CMOS process technology. This approach is certainly a step in the right direction to providing low-cost optoelectronic solutions to high speed interconnect problems.

Current VLSI tools (AT&T's MULGA CAD tool especially) are difficult to modify to support photodetector designs. Optoelectronic components such as photodiodes and lasers often require foreign geometries, and new design rules are often needed to "stretch" process limits. There also exist many database limitations in most conventional CAD tools.

A new CAD tool, *photogen*, was written to address the need for computer automation in the area of photodetector design and analysis. By determining all device characteristics before fabrication, one can quickly attempt many different photodetector design geometries and even different process technologies before settling on a design. In the long run, design automation tools save tremendous amount of effort. They allow other engineers to design optoelectronic photodetectors without having to worry about all the physics. Once designs have been proven, they can be made into libraries of useful components for different applications.

6.3 Future Work

The *photogen* program described in Chapter 4 is only the beginning of possibilities for monolithic photodetector design. The sections below highlight future improvements.

6.3.1 Fabrication

Many more photodetector designs need to be fabricated and characterized. Different styles and geometries should be tried in other technologies in order to develop more accurate models.

6.3.2 Process Modifications

If process modifications are allowed, the photodetector responsivity could be greatly improved. Anti-reflection coatings can be used or better control of oxide thickness would also limit reflections. Also, if packaging constraints allowed back-illumination of the substrate, then other style detectors may be possible. Side-junction-illuminated devices allow high responsivities with low reverse-bias, but they also suffer from edge effects of the die. If this style were to be implemented, better techniques for passivating the junction edge (through oxide for instance) need to be developed.

6.3.3 CAD Tool Enhancements

Many enhancements to the *photogen* tool are possible. New photodetector geometries can be implemented as program modules. For instance, an octagonal photodetector geometry would be useful in reduce spreading resistance. It may also be possible to modify the interdigitated photodetector style to gain better response uniformity for smaller spot sizes, if that be a concern. Other enhancements are outlined below:

- Providing support for modified, or non-silicon process (such as GaAs) would allow easy generation of MSM and PIN photodetector types.
- Adding true 2-D array x-y addressing with integrated selector electronics. This would make generating pixel arrays for a camera very straightforward.
- Providing accurate modeling of substrate minority carrier injection to help determine crosstalk noise for arrays. This feature would help to determine optimal array spacing for isolation.

- Providing generators for receiver front-end electronic circuits — amplifiers in particular.

- Adding a common window interface (possibly X11) would be nice, but not essential as long as one can view mask files with the local CAD tool. It would, however, be an effective display for demos.

Appendix A

Derivations of Photodiode Equations

Important equations used in Chapter 3 are derived in this section.

A.1 PN Abrupt-Junction Photodiode

As shown earlier in Figure 3-1, a PN diode consists of a P-type region abutted next to an N-type region, forming a PN junction. Because of the carrier concentration gradients at the PN junction, electrons diffuse across the junction from the N-side to the P-side; similarly, holes diffuse from the P-side to the N-side. These diffusion currents expose underlying fixed donor and acceptor ions, causing an electric field to be formed across this region. This electric field, in turn, induces electron and hole drift currents which counterbalance the diffusion currents until an equilibrium is reached.

Because this high $\mathcal{E}$-field region is depleted of mobile carriers, it is termed the depletion region or space-charge layer. Far away from the PN junction, the semiconductor is not upset, and the material is quasi-neutral. In order to simplify the three-dimensional equations describing the physical nature of an actual device, a much simpler one-dimensional step junction model is assumed where all variables change in only one spatial dimension. Also, perfect ohmic contacts are assumed to exist at

the far ends of the device.

A.1.1 Carrier Concentrations

Figure A-1 depicts the carrier concentrations on either side of an step-junction PN diode for both log and linear scales. The diode is more heavily doped on the P-side, so it is essentially a P+/N junction.

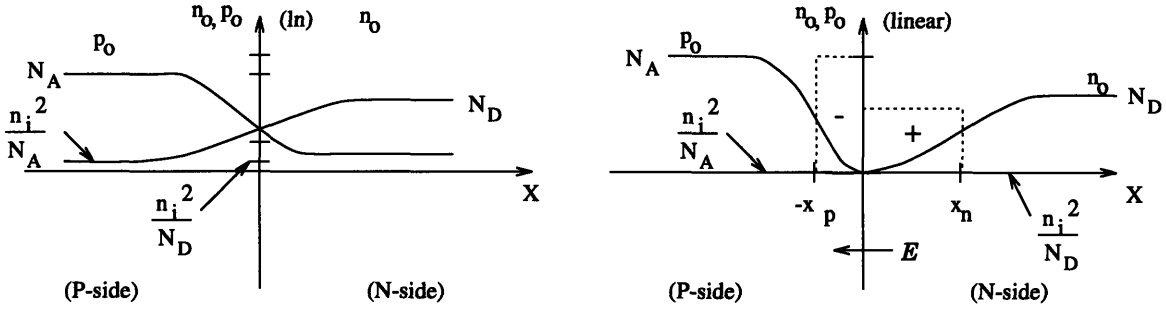


Figure A-1: Step-Junction PN Diode Carrier Concentrations

A.1.2 Charge Density

The resulting charge-density, $\rho(x)$, distribution due to this equilibrium condition is depicted in Figure A-2. The dashed lines represent a simplified model for the charge distribution known as the depletion approximation, whereby all the fixed charge is assumed to lie within the dashed boxes, $(-x_p < x < x_n)$ and the regions outside are quasi-neutral ($\rho = 0$). The total width of the depletion region, W, is simply defined to be $x_p + x_n$. This model vastly reduces calculation complexity for abrupt junction diodes, and is accurate only when the width of the space-charge region, W, is large[1] compared to the extrinsic Debye length on either the N-side or the P-side. This condition can be expressed mathematically for the N-side and P-side, respectively, assuming complete ionization as

[1] a factor of 5 is typically used

$$W \geq 5\sqrt{\frac{kT\epsilon}{q^2 N_D}} \tag{A.1}$$

$$W \geq 5\sqrt{\frac{kT\epsilon}{q^2 N_A}}\,. \tag{A.2}$$

These constraints are easily satisfied for an abrupt-junction PN diode. N_D and N_A are the ionized donor and acceptor ion concentrations on the N-side and P-side, respectively.

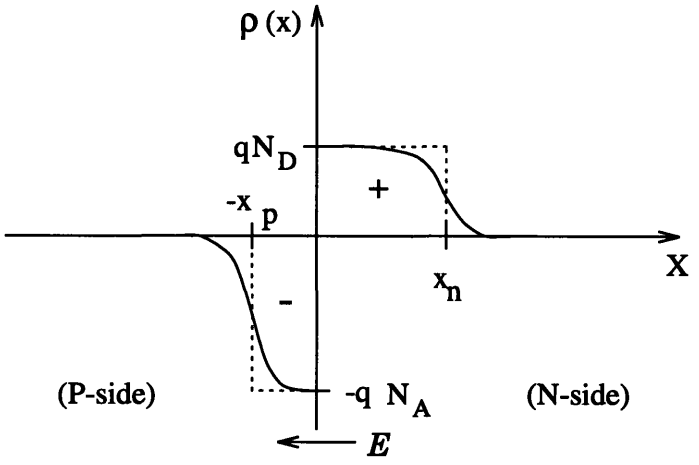


Figure A-2: Step-Junction PN Diode Charge Density Distribution

Using this depletion approximation, the charge distribution, $\rho(x)$, can be expressed as

$$\rho(x) = 0 \quad \forall (x < -x_p) \tag{A.3}$$

$$= -qN_A \quad \forall (-x_p \leq x < 0) \tag{A.4}$$

$$= +qN_D \quad \forall (0 < x \leq x_n) \tag{A.5}$$

$$= 0 \quad \forall (x > x_n)\,. \tag{A.6}$$

A.1.3 Electric Field Intensity

To determine the electric field from the charge distribution, we can make use of Poisson's equation:

$$\nabla^2\Phi = -\frac{\rho}{\epsilon}\,. \tag{A.7}$$

Substituting an expression for the electric field in terms of the potential,

$$\mathcal{E} = -\nabla\Phi\,, \tag{A.8}$$

into Equation A.7, for one (x-) dimension, we obtain:

$$\begin{aligned} \nabla\mathcal{E} &= \frac{\rho}{\epsilon} \\ \frac{\partial\mathcal{E}}{\partial x} &= \frac{\rho(x)}{\epsilon} \end{aligned} \tag{A.9}$$

To solve for the electric field for the three separate regions, we integrate Equation A.9 over the corresponding segments to obtain:

$$\mathcal{E}(x) = 0 \quad \forall\,(x < -x_p) \tag{A.10}$$

$$\begin{aligned} \int_{-x_p}^{x} dE &= \frac{1}{\epsilon}\int_{-x_p}^{x} -qN_A dx \\ \mathcal{E}(x) - \mathcal{E}(-x_p) &= \frac{-q}{\epsilon}\int_{-x_p}^{x} N_A dx \end{aligned}$$

$$\mathcal{E}(x) = \frac{-qN_A}{\epsilon}(x + x_p) \quad \forall\,(-x_p \leq x \leq 0) \tag{A.11}$$

$$\mathcal{E}(x) = \frac{qN_D}{\epsilon}(x - x_n) \quad \forall\,(0 \leq x \leq x_n) \tag{A.12}$$

$$\mathcal{E}(x) = 0 \quad \forall (x > x_n) \tag{A.13}$$

The resulting electric field distribution is shown in Figure A-3. The linear electric field distribution due the depletion approximation is shown in dotted lines; the actual $\mathcal{E}$-field is drawn in solid lines.

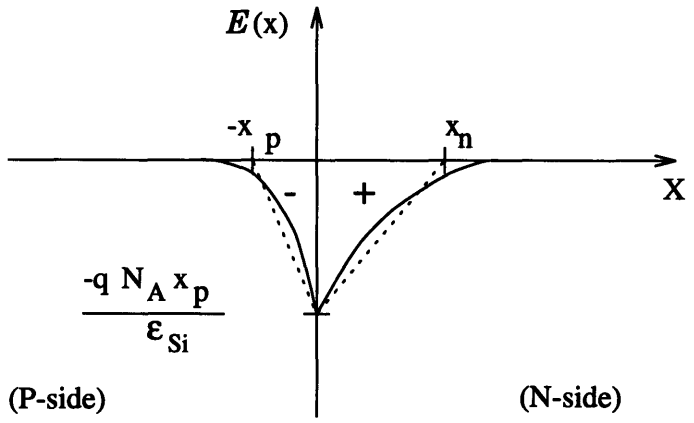


Figure A-3: Step-Junction PN Diode Electric Field Distribution

A.1.4 Potential Distribution

The values of the potential at both $x = +\infty$ and $x = -\infty$ are easily calculated from the exponential nature of the carrier concentrations.

$$N_A = n_i e^{\frac{-q\Phi(-\infty)}{kT}}$$

$$\Phi(-\infty) = \frac{-kT}{q} \ln \frac{N_A}{n_i} \tag{A.14}$$

$$N_D = n_i e^{\frac{-q\Phi(+\infty)}{kT}}$$

$$\Phi(+\infty) = \frac{kT}{q} \ln \frac{N_D}{n_i} \tag{A.15}$$

The built-in potential, V_{bi}, can be expressed simply as:

$$\begin{aligned} \Phi_{bi} &= \Phi_n - \Phi_p \\ &= \frac{kT}{q} \ln \frac{N_A N_D}{{n_i}^2} \end{aligned} \tag{A.16}$$

Next, the potential distribution, $\Phi(x)$, can be calculated throughout the PN diode using the one-dimensional form of Equation A.8 and from the electric field expressions found in Equations A.10 through A.13.

$$\Phi(x) = \Phi(-\infty) \quad \forall (x \leq -x_p) \tag{A.17}$$

$$\int_{-x_p}^{x} d\Phi = -\int_{-x_p}^{x} -\frac{qN_A}{\epsilon}(x + x_p)dx$$

$$\Phi(x) = \Phi(-\infty) + \frac{qN_A}{2\epsilon}(x_p + x)^2 \quad \forall (-x_p < x \leq 0) \tag{A.18}$$

$$\Phi(x) = \Phi(+\infty) - \frac{qN_D}{2\epsilon}(x - x_n)^2 \quad \forall (0 \leq x < x_n) \tag{A.19}$$

$$\Phi(x) = \Phi(+\infty) \quad \forall (x \geq x_n) \tag{A.20}$$

Figure A-4 shows the potential distribution for an abrupt-junction PN diode. Applying the principal of overall charge neutrality during equilibrium we obtain:

$$qN_A x_p = qN_D x_n \tag{A.21}$$

Next, continuity of Φ at $x = 0$ requires that

$$\Phi(-\infty) + \frac{qN_A}{2\epsilon} x_p^2 = \Phi(+\infty) - \frac{qN_D}{2\epsilon} x_n^2 \tag{A.22}$$

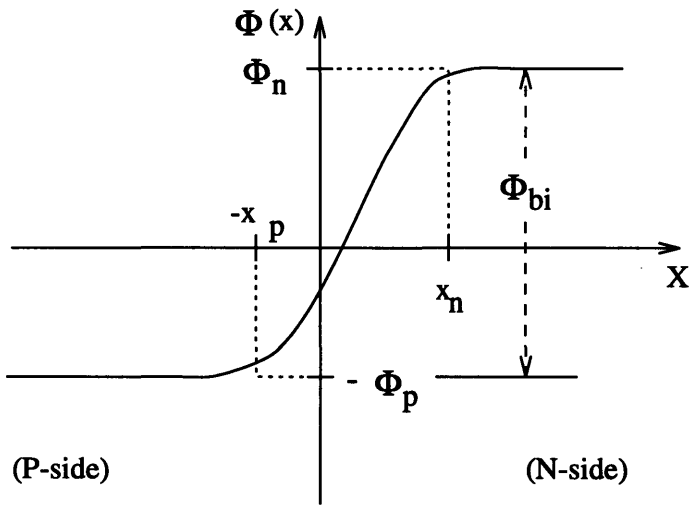


Figure A-4: Step-Junction PN Diode Potential Distribution

A.1.5 Depletion Width

Combining Equations A.21 and A.22 we arrive at expressions for x_p and x_n:

$$x_n = \sqrt{\frac{2\epsilon}{qN_D}\left(\frac{N_A}{N_A+N_D}\right)\Phi_{bi}} \tag{A.23}$$

$$x_p = \sqrt{\frac{2\epsilon}{qN_A}\left(\frac{N_D}{N_A+N_D}\right)\Phi_{bi}} \tag{A.24}$$

$$W = x_n + x_p$$

$$= \sqrt{\frac{2\epsilon}{q}\left(\frac{N_A+N_D}{N_AN_D}\right)\Phi_{bi}} \tag{A.25}$$

An applied reverse-bias voltage simply adds to the built-in potential assuming that all the bias voltage is dropped across the depletion region, increasing W by the square-root of the reverse-bias. Essentially, Φ_{bi} is replaced by $\Phi_{bi} + \Phi_{rev}$. To more correctly model thermal effects [Sze69], 2 V_t can be subtracted from Φ_{bi} resulting in the following expression for W:

$$W = \sqrt{\frac{2\epsilon}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)(\Phi_{bi} + \Phi_{rev} - 2\Phi_t)} \tag{A.26}$$

where the thermal potential, $\Phi_t = \frac{kT}{q}$.

A.1.6 Depletion Capacitance

Now the expression for the depletion capacitance can be found:

$$C_J = \frac{\epsilon A}{W} \tag{A.27}$$

$$= A\sqrt{\frac{q\epsilon}{2}\left(\frac{N_A N_D}{N_A + N_D}\right)\left(\frac{1}{\Phi_{bi} + \Phi_{rev} - 2\Phi_t}\right)} \tag{A.28}$$

Equation A.28 can be rewritten in terms of a zero-bias capacitance, C_{J0}.

$$C_{J0} = C_{J|V_{rev}=0} = \frac{\epsilon A}{\sqrt{\frac{2\epsilon}{q}\left(\frac{N_A+N_D}{N_A N_D}\right)\Phi_{bi}}} \tag{A.29}$$

Factoring out C_{J0} from Equation A.28 results in

$$C_J = \frac{C_{J0}}{\left(1 + \frac{\Phi_{rev}}{\Phi_{bi}}\right)^{1/2}} \tag{A.30}$$

If the one-half power in Equation A.30 is replaced by a variable m, then this expression for C_J can model PN junctions having various degrees of abruptness. Typically, m takes on values between one-third and one-half, corresponding to linear-graded and step-junction extremes.

$$C_J = \frac{C_{J0}}{\left(1 + \frac{\Phi_{rev}}{\Phi_{bi}}\right)^{m}} \qquad \forall\,(1/3 \leq m \leq 1/2; \Phi_{rev} \geq 0) \tag{A.31}$$

A.2 PN Linearly-Graded Junction Photodiode

A.2.1 Charge Density

Figure A-5 depicts the carrier concentrations and the charge density distribution for a linearly-graded junction PN diode. The negative of the grading coefficient, $-a$, is simply the slope of the linear doping profile.

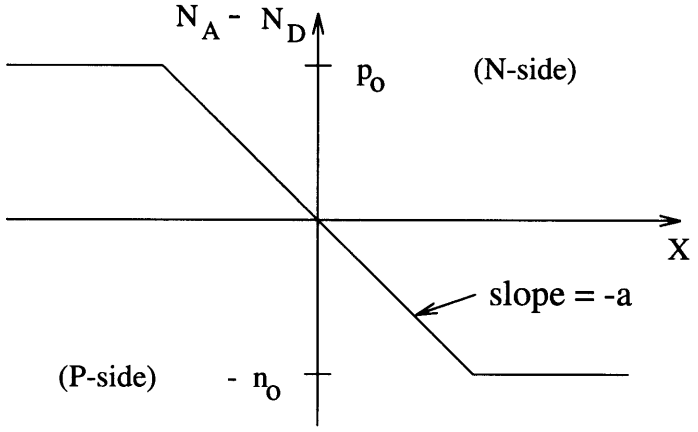


Figure A-5: Linear-graded PN Diode Carrier Concentration

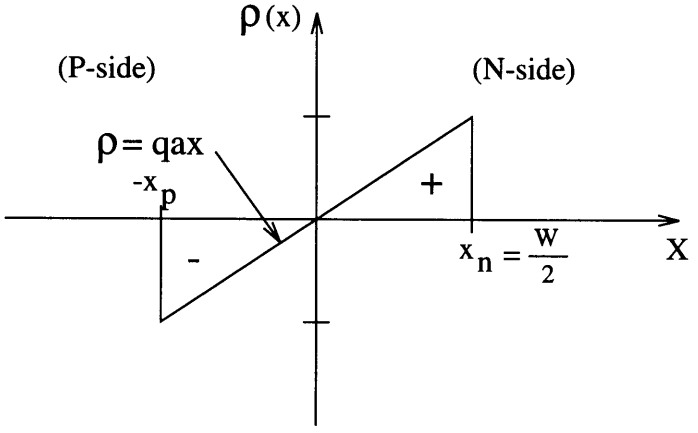


Figure A-6: Linear-graded PN Diode Charge Density Distribution

From the principle of charge neutrality, we have:

$$x_n = x_p = \frac{W}{2} \tag{A.32}$$

A.2.2 Electric Field Intensity

Solving for the $\mathcal{E}$-field, using Equation A.9, we obtain:

$$\frac{\partial \mathcal{E}}{\partial x} = \frac{qax}{\epsilon} \quad \forall (-x_p \leq x \leq x_n)$$

$$\int_{-x_p}^{x} d\mathcal{E} = \int_{-x_p}^{x} \frac{qax}{\epsilon} dx$$

$$\mathcal{E}(x) - \mathcal{E}(-x_p) = \frac{qa}{2\epsilon}(x^2 - x_p^2)$$

$$\mathcal{E}(x) = \frac{qa}{2\epsilon}\left[x^2 - \left(\frac{W}{2}\right)^2\right] \quad \forall (-x_p \leq x \leq x_n) \tag{A.33}$$

$$\int_{-\infty}^{x} d\mathcal{E} = 0 \quad \forall (x < -x_p)$$

$$\mathcal{E} = 0 \quad \forall (x < -x_p) \tag{A.34}$$

$$\mathcal{E} = 0 \quad \forall (x > x_n) \tag{A.35}$$

Figure A-7 depicts the electric field intensity for the linearly-graded junction PN diode.

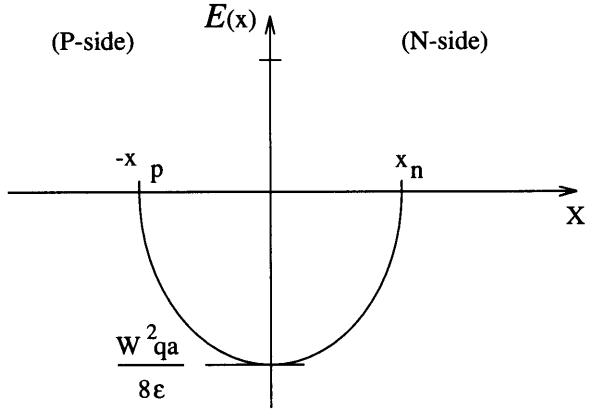


Figure A-7: Linear-Graded PN Diode Electric Field Distribution

A.2.3 Potential Distribution

Now to derive the potential distribution, $\Phi(x)$, we use Equations A.14, A.15 and A.8 and the previously-defined electric field distribution in Equations A.33 through A.35.

$$\begin{aligned}
\Phi(x) &= \Phi(-\infty) \quad \forall (x \leq -x_p) \\
&= -\frac{kT}{q} \ln \frac{p_o}{n_i} \\
&= -\frac{kT}{q} \ln \frac{qaWN_A}{2n_i} \quad \forall (x \leq -x_p)
\end{aligned} \tag{A.36}$$

$$\begin{aligned}
\Phi(x) &= \Phi(+\infty) \quad \forall (x \geq x_n) \\
&= \frac{kT}{q} \ln \frac{n_o}{n_i} \\
&= \frac{kT}{q} \ln \frac{qaWN_D}{2n_i} \quad \forall (x \geq x_n)
\end{aligned} \tag{A.37}$$

$$\begin{aligned}
\int_{-x_p}^{x} d\Phi &= -\int_{-x_p}^{x} \frac{qa}{2\epsilon} \left[x^2 - \left(\frac{W}{2}\right)^2 \right] dx \\
\Phi(x) - \Phi(-x_p) &= -\frac{qa}{2\epsilon} \left[\frac{x^3}{3} + \frac{x_p^3}{3} - \left(\frac{W}{2}\right)^2 x - \left(\frac{W}{2}\right)^2 x_p \right] \\
\Phi(x) &= \Phi(-x_p) + \frac{qa}{2\epsilon} \left[\left(\frac{W}{2}\right)^2 (x + x_p) - \frac{x^3}{3} - \frac{\left(\frac{W}{2}\right)^3}{3} \right] \\
\Phi(x) &= \Phi(-x_p) + \frac{qa}{2\epsilon} \left[\left(\frac{W}{2}\right)^3 + \left(\frac{W}{2}\right)^2 x - \frac{x^3}{3} - \frac{W^3}{24} \right]
\end{aligned}$$

$$\begin{aligned}\Phi(x) &= \Phi(-x_p) + \frac{qa}{2\epsilon}\left[\frac{W^3}{12} + \left(\frac{W}{2}\right)^2 x - \frac{x^3}{3}\right] \\ &\forall\,(-x_p \leq x \leq x_n)\end{aligned} \tag{A.38}$$

The built-in potential, Φ_{bi}, can be computed from Equations A.36 and A.37.

$$\Phi_{bi} = \Phi(+\infty) - \Phi(-\infty) = \frac{kT}{q}\ln\frac{q^2a^2W^2N_AN_D}{4{n_i}^2} \tag{A.39}$$

We now use the fact that $\Phi(x=0) = 0$ to determine $\Phi(-x_p)$:

$$\begin{aligned}\Phi(-x_p) - \frac{2a}{q\epsilon}\left[-2\left(\frac{W}{2}\right)^3\right] &= 0 \\ \Phi(-x_p) &= -\frac{\Phi_{bi}}{2}\end{aligned} \tag{A.40}$$

This result makes sense — since the junction is symmetrically doped, half the built-in voltage is dropped across each side.

A.2.4 Depletion Width

Setting Equations A.37 and A.38 equal to each other at $x = x_n$, we have:

$$\begin{aligned}\Phi(x_n) &= \frac{kT}{q}\ln\frac{qaWN_D}{n_i} \\ &= -\frac{kT}{q}\ln\frac{qaWN_A}{n_i} + \frac{qaW^3}{24\epsilon}\end{aligned}$$

$$\begin{aligned}\frac{qaW^3}{12\epsilon} &= \frac{kT}{q}\ln\frac{q^2a^2W^2N_AN_D}{4{n_i}^2} \\ &= \Phi_{bi}\end{aligned} \tag{A.41}$$

$$\begin{aligned}W_{gj} &= \sqrt[3]{\frac{12\epsilon\Phi_{bi}}{qa}} \\ &= 2x_n\end{aligned} \tag{A.42}$$

Equation A.41 must be iteratively solved for the built-in W in order to find the built-in potential, Φ_{bi}.

Accounting for the effect of the reverse-bias potential, Φ_{rev}, on W, we have

$$W_{gj} = \sqrt[3]{\frac{12\epsilon}{qa}(\Phi_{bi} + \Phi_{rev})} \tag{A.43}$$

For an abrupt-junction PN diode with $N_A = N_D = N$, we have from Equation A.25:

$$W_{aj} = \sqrt{\frac{4\epsilon}{qN}(\Phi_{bi} + \Phi_{rev})} \tag{A.44}$$

which is less than W_{gj}, as expected.

A.2.5 Depletion Capacitance

The larger depletion width of a graded-junction PN diode over an abrupt-junction PN diode and the fact that less charge is stored in the space charge layer for the same applied reverse-bias voltage, substantially lowers the capacitance of the graded-junction diode. Refer to Section A.1.6 for further details.

A.3 PIN photodiode

A PIN diode simply consists of a PN diode with an intrinsic layer sandwiched between the P and N layers. The analysis follows the same procedure as a step-junction PN diode. Figure A-8 depicts the characteristic charge density, electric field and potential distributions for a PIN diode having an intrinsic width of L.

The peak electric field for the PIN diode is much less than the peak electric field of a step-junction PN diode for the following arguments. First, the built-in potential, Φ_{bi}, is the same as for the PN case, $\frac{kT}{q} \ln \frac{N_A N_D}{n_i^2}$. To maintain this potential in the PIN diode, where there is a constant $\mathcal{E}$-field over the large intrinsic region requires that x_p and x_n become smaller than in the PN case, which effectively lowers the maximum electric field. This feature gives the PIN diode better immunity to reverse breakdown.

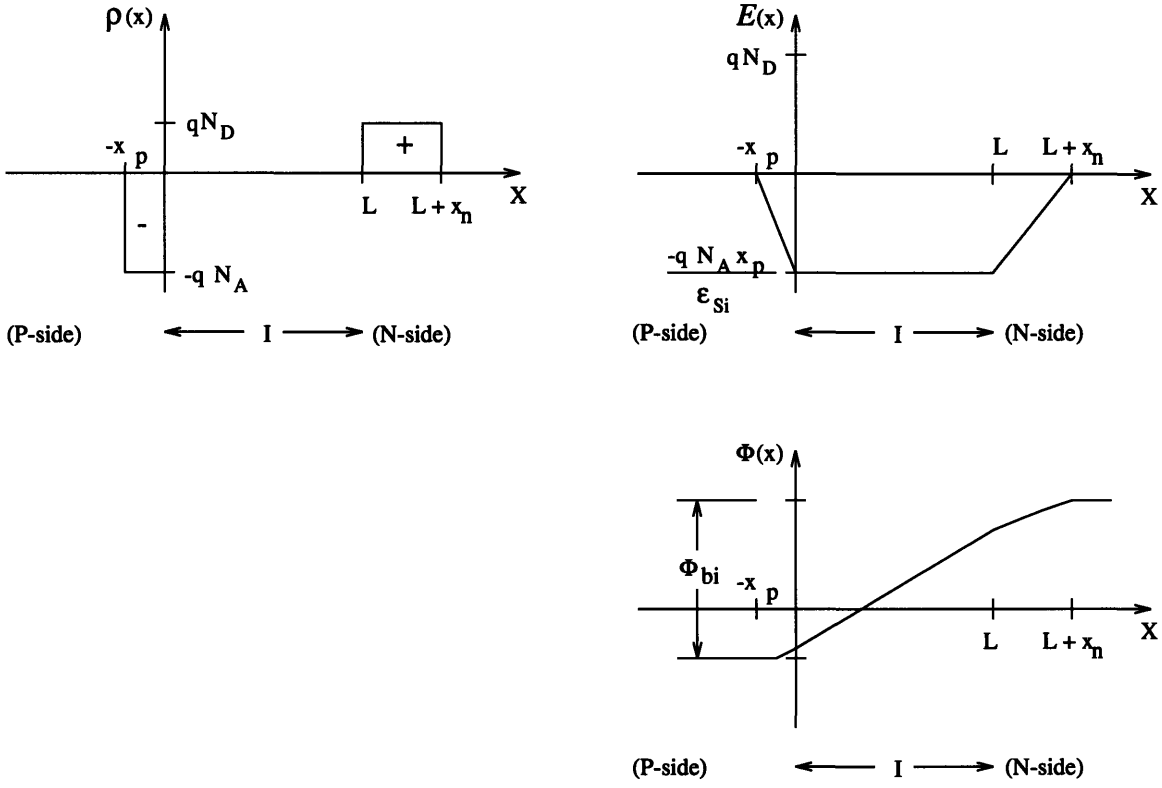


Figure A-8: PIN Diode Characteristic Distributions

Appendix B

Important Photodiode Parameters

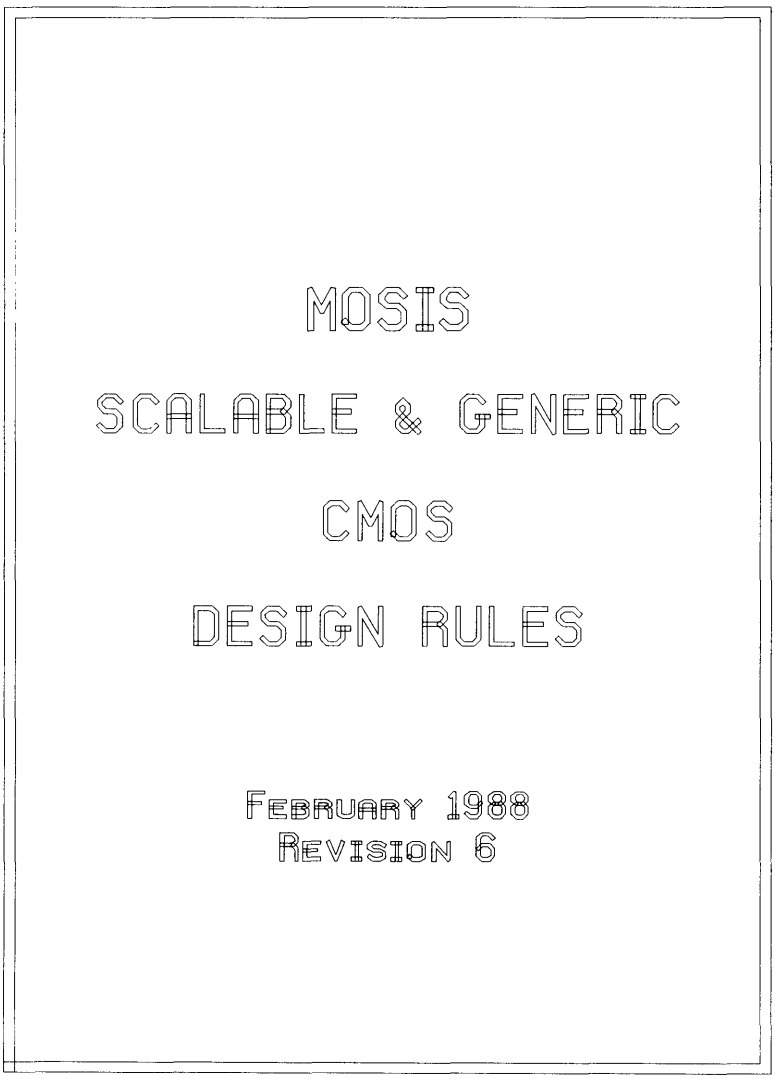
Table B.1 was developed from discussions in Chapter 3 and with the aid of the Photonics Handbook [LHA+93]. N_D and N_A are the ionized donor and acceptor concentrations in the N and P layers, respectively, $\mathcal{E}_{max}$ is the maximum electric field intensity which can be supported across the junction before breakdown occurs, R_L is the load resistance of the electronic receiver, I_L is the photo-induced light current, $\tau_{n,p}$ are the minority carrier lifetimes, Δf is the bandwidth of interest, v_{sat} is the saturated velocity of carriers, T is temperature in Kelvin, and ϵ is the dielectric permittivity of the material.

Symbol	Name	Dependencies	Design Value
A	Active Area	area, geometry, wiring density, W	small for low C_J and noise, large for high $D*$ and alignment tolerance
C_J	Capacitance (depletion)	A, $1/W$, ϵ	low for fast response
I_{dk}	Dark Current	A, W, $1/\tau_{n,p}$	low for low noise
I_{shot}	Shot Noise	$\sqrt{I_{dk}}$, $\sqrt{I_L}$, $\sqrt{\Delta f}$	low for low noise
I_{therm}	Thermal Noise	$\sqrt{T}$, $\sqrt{\Delta f}$, $1/\sqrt{R_{sh}}$	low for low noise
I_{nt}	Total Noise Current (rms)	I_{shot}, I_{therm}	low for low NEP and high $D*$
$D*$	Detectivity	$\sqrt{A}$, $\sqrt{\Delta f}$, 1/NEP	high for better S/N ratio
NEP	Noise Equivalent Power	I_{nt}, $1/\mathcal{R}$, $\sqrt{\Delta f}$	low for high $D*$
P_{opt}	Incident Optical Power	application	high for high S/N ratio, not too high for good linearity
η	Quantum Efficiency	R, geometry, design, α	high for high $\mathcal{R}$ and low NEP
R	Fresnel Reflectivity	ϵ's of materials	low for high $\mathcal{R}$
$\mathcal{R}$	Responsivity	wavelength (λ), η	high for sensitivity to low light levels
R_s	Series Resistance	A, ρ, geometry, wiring density	low for high bandwidth and linearity
R_{sh}	Shunt Resistance	$1/A$, $1/\sqrt{\rho}$	high for low thermal noise
ρ	Bulk Resistivity	process	high for low C_J, low for small t_{tr}
$\tau_{rise,fall}$	Response Speed	C_J, (R_s+R_L), W, $1/v_{sat}$, $\tau_{n,p}$	low for high bandwidth
τ_{RC}	Device Time Constant	C_J, (R_s+R_L)	low for high bandwidth (more important for PN)
t_{tr}	Transit Time	W, $1/v_{sat}$	low for high bandwidth (more important for PIN)
V_{bi}	Built-in Voltage	T, $\ln(N_A)$, $\ln(N_D)$	small for big W
V_{br}	Reverse Breakdown Voltage	W, $\mathcal{E}_{max}$	high for large applied V_{rev}
V_{rev}	Reverse-bias Voltage	application	low for low noise, high for high speed and better linearity
W	Depletion Width	I length (PIN), $1/\sqrt{N_A}$, $1/\sqrt{N_D}$, $\sqrt{V_{bi}}$, $\sqrt{V_{rev}}$	small for fast t_{tr} response but high C_J; big for slower t_{tr} response and low C_J
I	(PIN) Intrinsic thickness	process, application	small for fast t_{tr} response but high C_J; big for slower t_{tr} response and low C_J

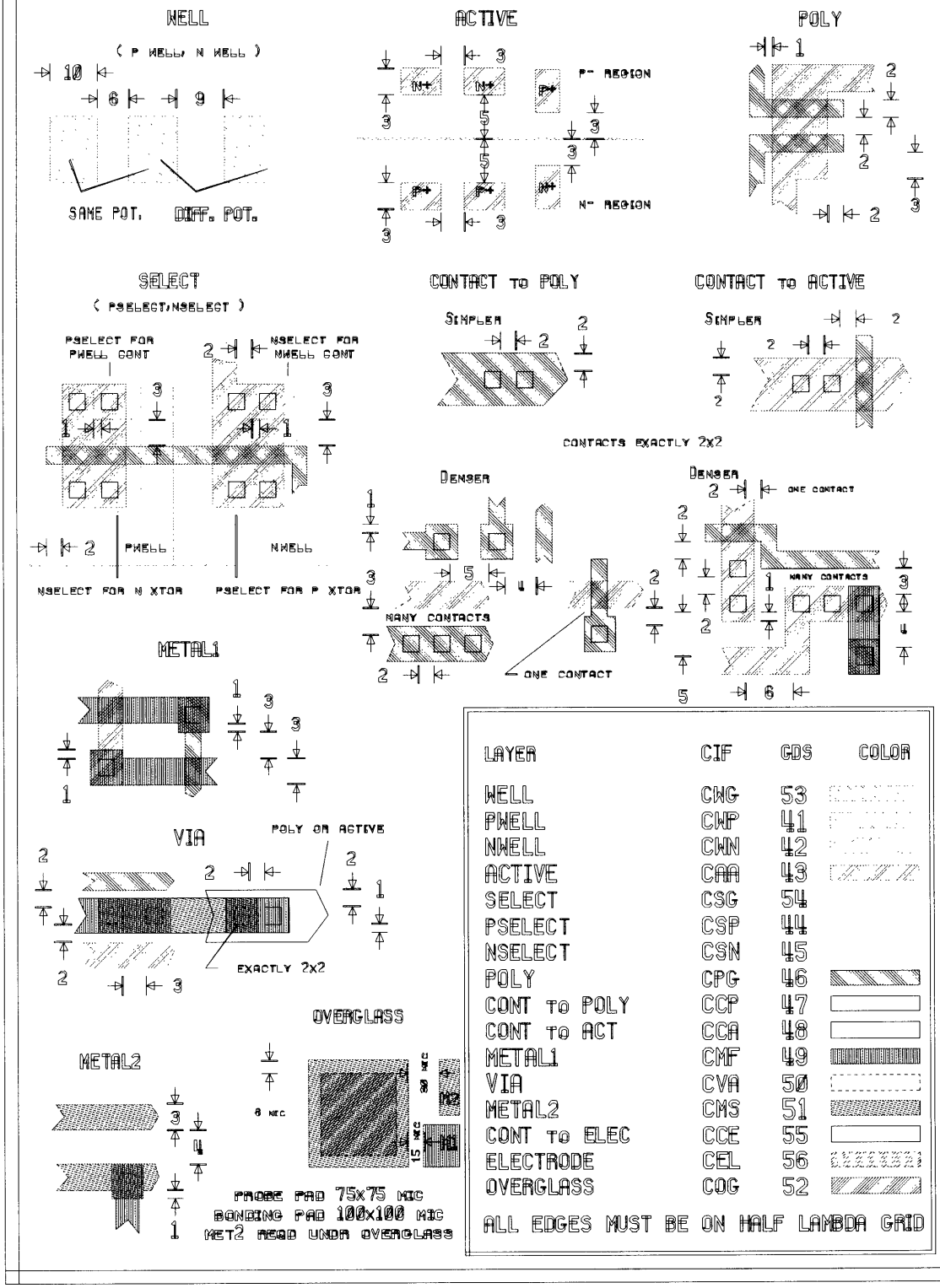
Table B.1:

Appendix C

MOSIS SCMOS Design Rules



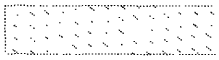
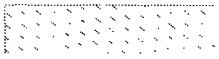
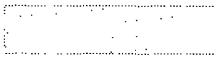
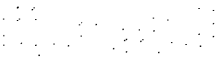
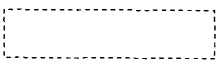
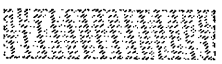
MOSIS CMOS SCALABLE RULES (REV6)



LAYER	CIF	GDS	COLOR
WELL	CWG	53	
PWELL	CWP	41	
NWELL	CWN	42	
ACTIVE	CAA	43	
SELECT	CSG	54	
PSELECT	CSP	44	
NSELECT	CSN	45	
POLY	CPG	46	
CONT TO POLY	CCP	47	
CONT TO ACT	CCA	48	
METAL1	CMF	49	
VIA	CVA	50	
METAL2	CMS	51	
CONT TO ELEC	CCE	55	
ELECTRODE	CEL	56	
OVERGLASS	COG	52	

ALL EDGES MUST BE ON HALF LAMBDA GRID

LAYER NAMES AND COLORS

LAYER	CIF	CALMA #	COLOR
WELL	CWG	53	
PWELL	CWP	41	
NWELL	CWN	42	
ACTIVE	CAA	43	
SELECT	CSG	54	
PSELECT	CSP	44	
NSELECT	CSN	45	
POLY	CPG	46	
CONT to POLY	CCP	47	
CONT to ACT	CCA	48	
METAL1	CMF	49	
VIA	CVA	50	
METAL2	CMS	51	
CONT to ELEC	CCE	55	
ELECTRODE	CEL	56	
OVERGLASS	COG	52	

TECHNOLOGIES AND REQUIRED LAYERS

PROCESS	TECHNOLOGY	REQUIRED LAYERS
PWELL AND N SUBS TWIN TUB	SCP	CWP CSN, CSP
N WELL AND P SUBS TWIN TUB	SCN	CWN, CSN, CSP
ALL *	SCG	CWG, CSG
ALL **	SCE	CWP, CWN CSP, CSN

* FOR A P WELL OR N SUBS TWIN TUB PROCESS, MOSIS SETS CWP=CWG AND CSP=CSG

FOR AN NWELL OR P SUBS TWIN TUB PROCESS, MOSIS SETS CWN=CWG AND CSN=CSG

** FOR A P WELL OR N SUBS TWIN TUB PROCESS, MOSIS IGNORES CWN

FOR AN NWELL OR P SUBS TWIN TUB PROCESS, MOSIS IGNORES CWP

INSTRUCTIONS FOR DESIGNERS

DRAW IN UNITS OF LAMBDA.

ON ANY LAYER NO SPACE OR FEATURE SIZE
MAY BE LESS THAN 2 LAMBDA WIDE.

ALL FEATURE EDGES
MUST BE ON A HALF LAMBDA GRID.

WHAT YOU DRAW WILL BE VERY
CLOSE TO WHAT YOU GET. MOSIS
WILL TELL YOU THE DIFFERENCES.

SCALE YOUR CIF OR GDS TO
METRIC UNITS. NEVER SUBMIT
A DESIGN IN CENTILAMBDA.

WHAT VALUES LAMBDA ?

LAMBDA=1.Ø MICRONS
FOR 2.Ø MICRON FABRICATORS

LAMBDA=Ø.8 MICRONS
FOR 1.6 MICRON FABRICATORS

LAMBDA=Ø.6 MICRONS
FOR 1.2 MICRON FABRICATORS

LAMBDA=Ø.5 MICRONS
FOR 1.Ø MICRON FABRICATORS

QUESTIONS TO FABRICATORS

MOSIS

MERGES CONTACT TO ACTIVE AND CONTACT TO POLY LAYERS TO GENERATE THE CONTACT MASK.

CAN APPLY BLOATS OR SHRINKS TO ALL LAYERS.

CAN ADJUST ACTIVE OVERLAP OF CONTACT INDEPENDENTLY OF BLOAT TO ACTIVE

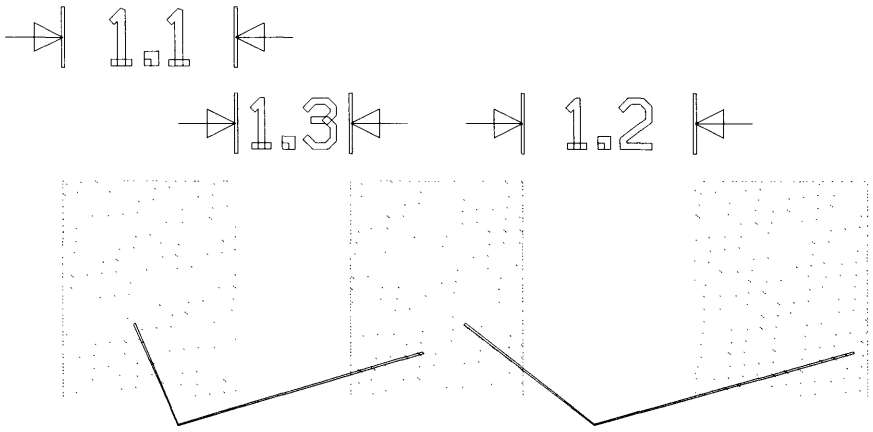
CAN ADJUST POLY OVERLAP OF CONTACT INDEPENDENTLY OF BLOAT TO POLY

FABRICATORS

WHAT VALUE OF LAMBDA CAN YOU SUPPORT? WHAT BLOATS OR SHRINKS DO YOU REQUIRE?

1. WELL
(NWELL.PWELL)

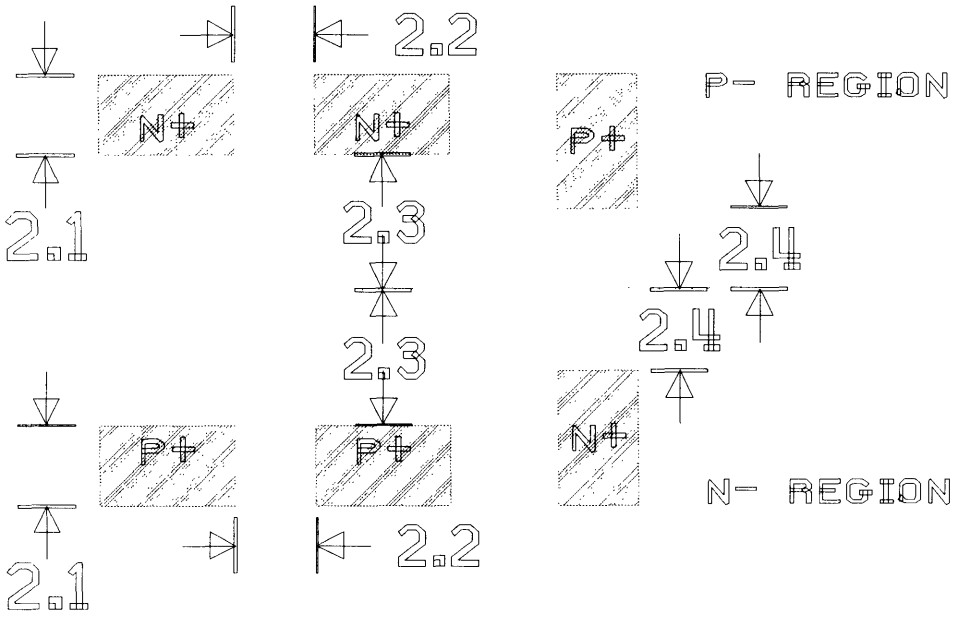
		LAMBDAS
1.1	WIDTH	10
1.2	SPACE DIFF. POT.	9
1.3	SPACE SAME POT.	0 OR 6



NOTE: IF BOTH P AND N WELLS SUBMITTED, THEY MAY NOT OVERLAP BUT THEY MAY BE COINCIDENT.

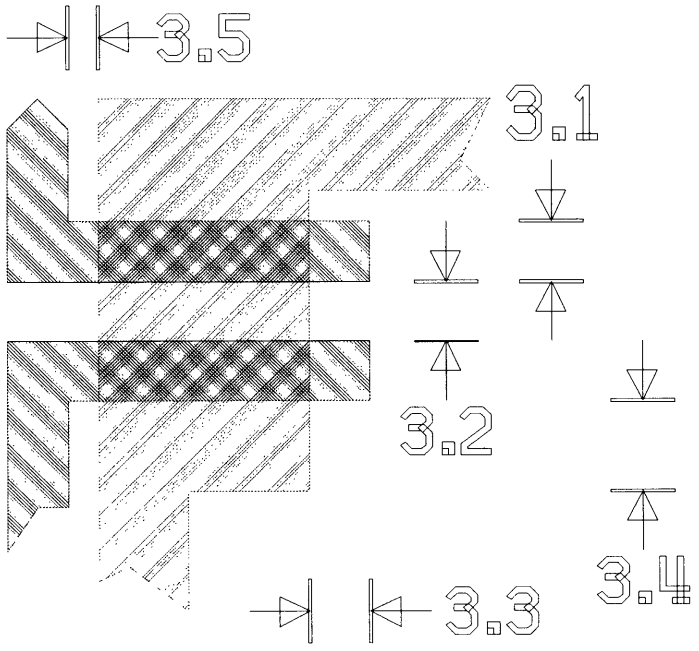
2. ACTIVE

		LAMBDAS
2.1	WIDTH	3
2.2	SPACE	3
2.3	SOURCE/DRAIN ACTIVE TO WELL EDGE	5
2.4	SUBS./WELL CONTACT. ACTIVE TO WELL EDGE	3



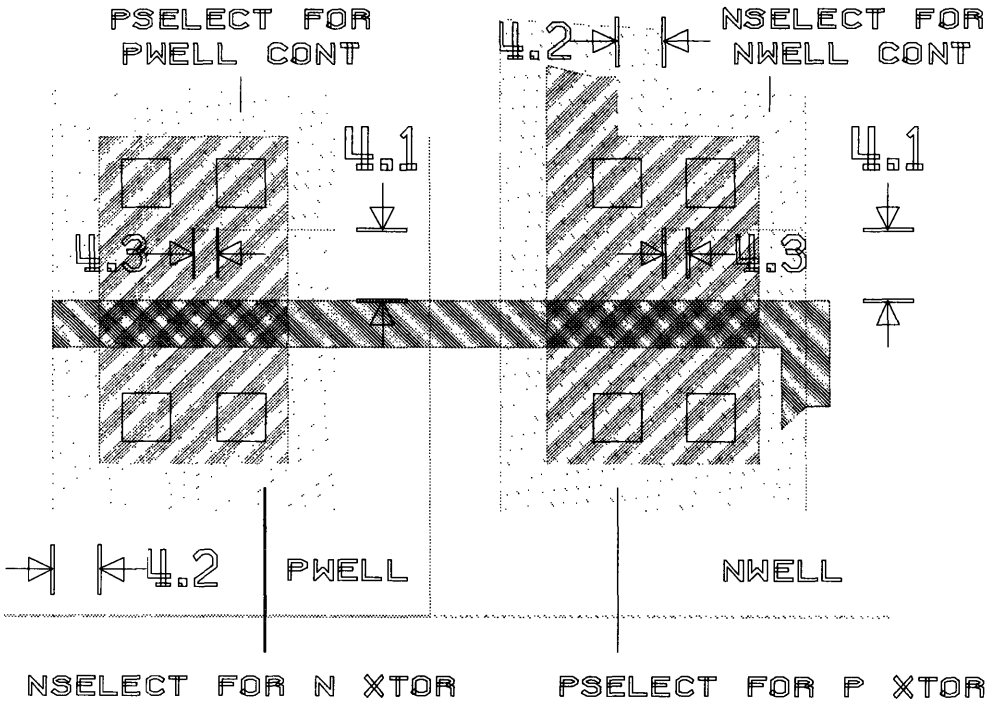
3. POLY

		Lambdas
3.1	Width	2
3.2	Space	2
3.3	Gate Overlap of Active	2
3.4	Active Overlap of Gate	3
3.5	Field Poly to Active	1



4. SELECT (PSELECT,NSELECT)

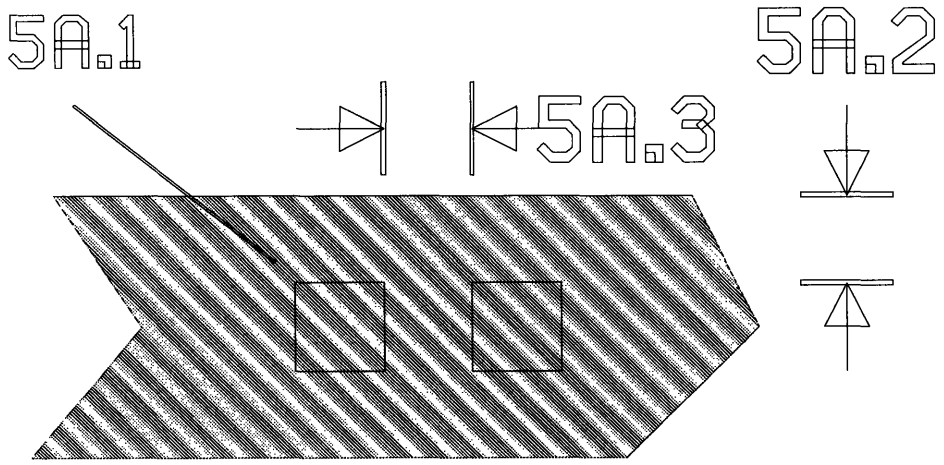
		LAMBDAS
4.1	SELECT SPACE (OVERLAP) TO (OF) CHANNEL TO ENSURE ADEQUATE SOURCE/DRAIN WIDTH	3
4.2	SELECT SPACE (OVERLAP) TO (OF) ACTIVE	2
4.3	SELECT SPACE (OVERLAP) TO (OF) CONTACT TO WELL OR SUBSTRATE	1
4.4	MIN WIDTH AND SPACE	2



NOTE: PSELECT AND NSELECT MAY BE COINCIDENT, BUT MUST NOT OVERLAP

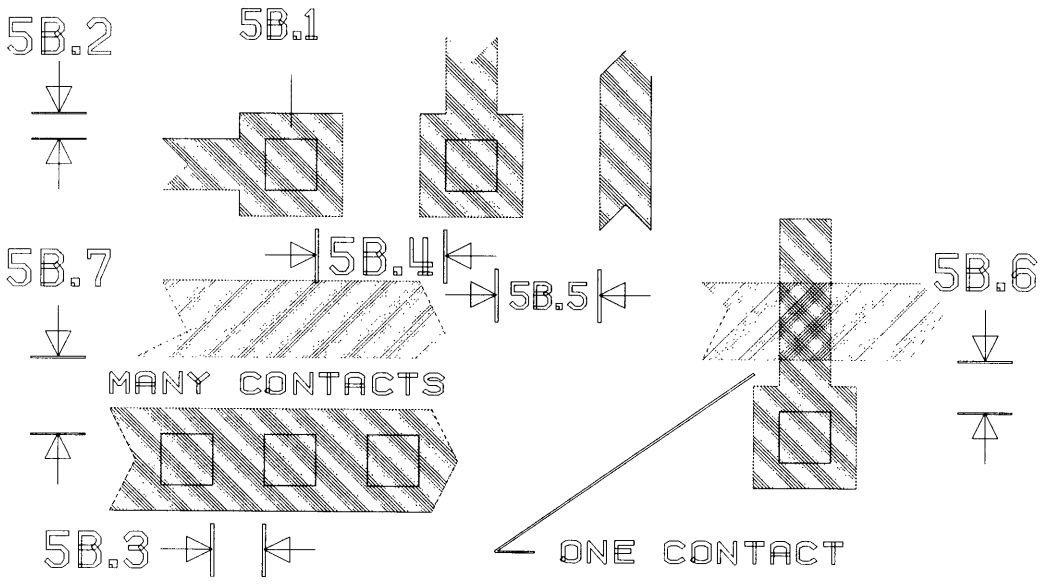
5A. SIMPLER CONTACT TO POLY

		LAMBDAS
5A.1	CONTACT SIZE EXACTLY	2x2
5A.2	POLY OVERLAP	2
5A.3	SPACING	2



5B. DENSER CONTACT TO POLY

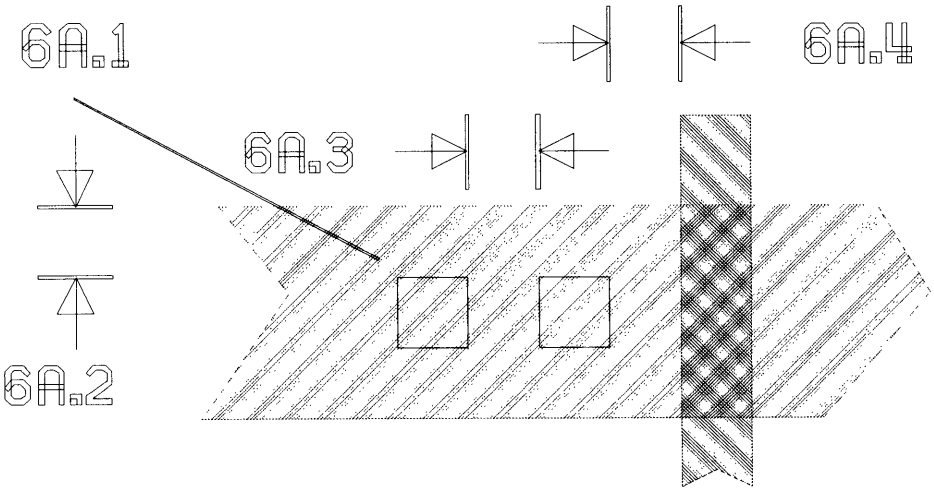
		LAMBDAS
5B.1	CONTACT SIZE, EXACTLY	2X2
5B.2	POLY OVERLAP OF CONTACT	1
5B.3	SPACING ON SAME POLY	2
5B.4	SPACING ON DIFF POLY	5
5B.5	SPACE TO OTHER POLY	4
5B.6	SPACE TO ACT, ONE CONTACT	2
5B.7	SPACE TO ACT, MANY CONTACTS	3



NOTE: YOUR ASSOCIATING CONTACTS WITH POLY OR ACTIVE ALLOWS MOSIS TO INDEPENDENTLY BLOAT THE LAYER AND THE LAYER OVERLAP OF THE CONTACT

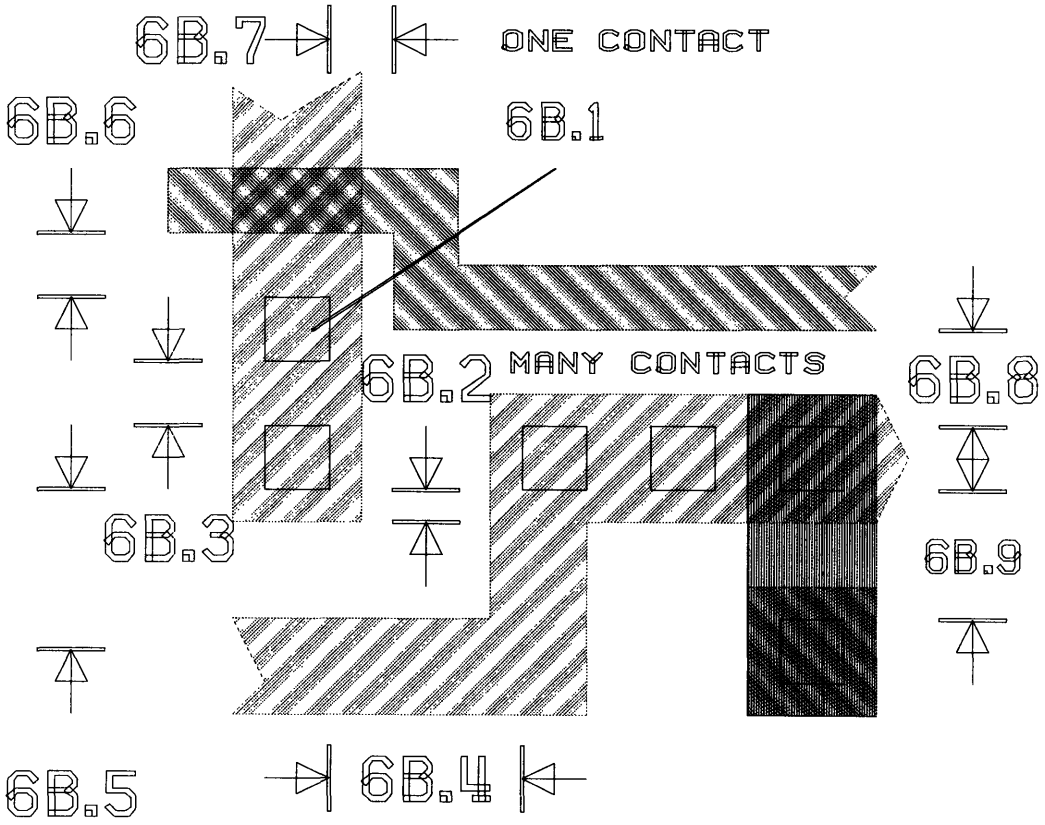
6A. SIMPLER CONTACT TO ACTIVE

		LAMBDAS
6A.1	CONTACT SIZE EXACTLY	2x2
6A.2	ACTIVE OVERLAP	2
6A.3	SPACING	2
6A.4	SPACE TO GATE	2



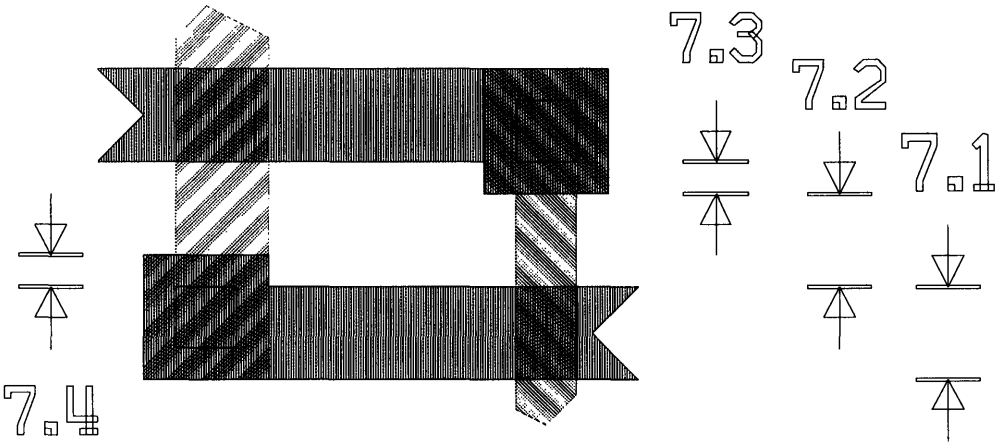
6B. DENSER CONTACT TO ACTIVE

		LAMBDAS
6B.1	CONTACT SIZE, EXACTLY	2x2
6B.2	ACTIVE OVERLAP	1
6B.3	SPACING ON SAME ACTIVE	2
6B.4	SPACING ON DIFF ACTIVE	6
6B.5	SPACE TO DIFF ACTIVE	5
6B.6	SPACE TO GATE	2
6B.7	SPACE TO FIELD POLY, ONE CONT.	2
6B.8	SPACE TO FIELD POLY, MANY CONT.	3
6B.9	SPACE TO CONTACT TO POLY	4



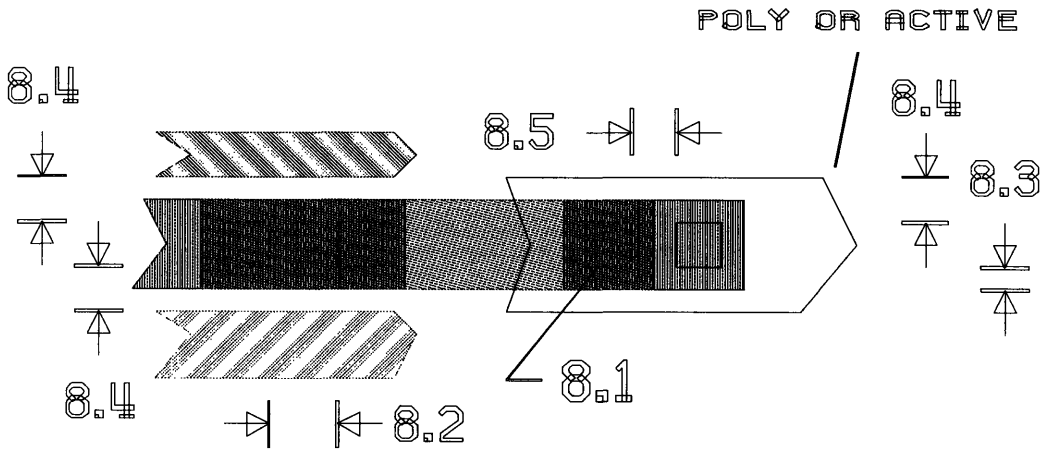
7. METAL1

		LAMBDAS
7.1	WIDTH	3
7.2	SPACE TO METAL1	3
7.3	OVERLAP OF CONTACT TO POLY	1
7.4	OVERLAP OF CONTACT TO ACTIVE	1



8. VIA

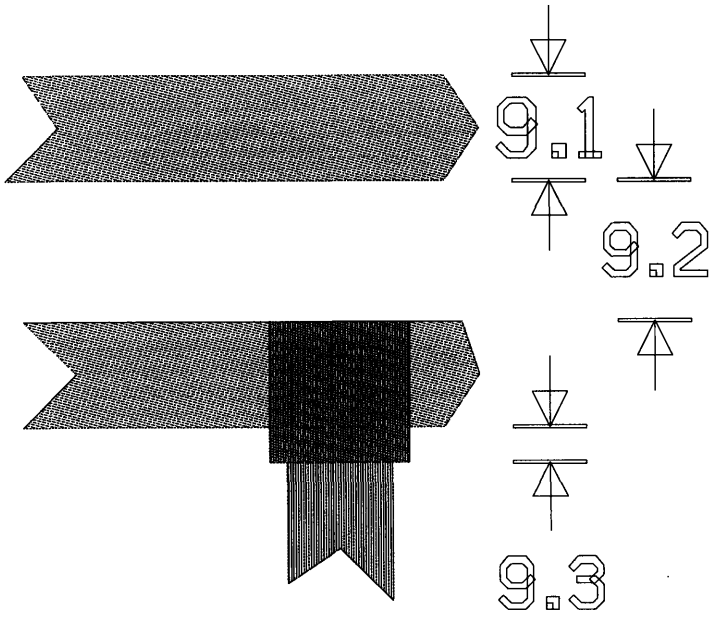
		LAMBDAS
8.1	SIZE, EXACTLY	2x2
8.2	SEPARATION TO VIA	3
8.3	OVERLAP BY METAL1	1
8.4	SPACE TO POLY OR ACTIVE EDGE	2
8.5	SPACE TO CONTACT	2



NOTE: OBJECTIVE IS VIA ON A FLAT SURFACE. VIA STACKED OVER CONTACT NOT ALLOWED.

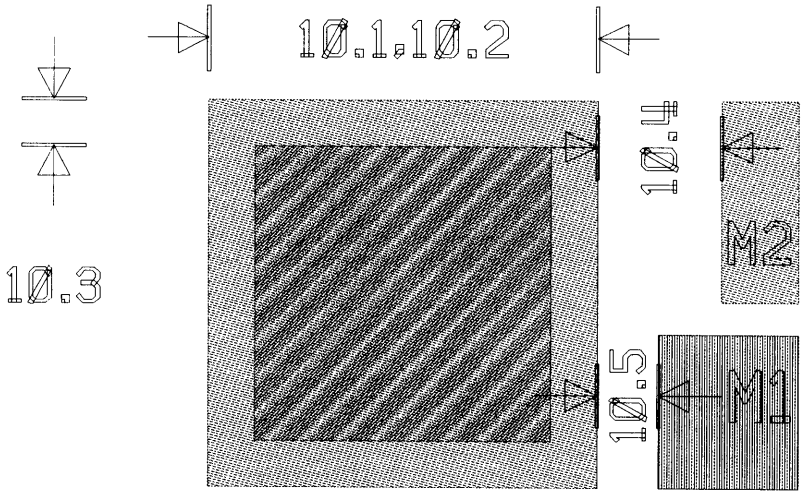
9. METAL2

		LAMBDAS
9.1	WIDTH	3
9.2	SPACE TO METAL2	4
9.3	OVERLAP OF VIA	1



10. OVERGLASS

		MICRONS
10.1	Bonding Pad	100x100
10.2	Probe Pad	75x75
10.3	Pad Overlap of glass	6
10.4	Pad Space to Unrelated Metal2	30
10.5	Pad Space to Unrelated Metal1, Poly or Active	15



Metal2 Required Under Overglass

Bibliography

[ACW87] Bryan Ackland, Bob Clark, and Neil Weste. *MULGA: A Symbolic VLSI Design System, Users Manual.* AT&T Bell Laboratories, Holmdel, New Jersey, 1987.

[AD88] B. D. Ackland and A. G. Dickinson. Interfacing differential optical signals to digital circuitry. Technical Memorandum 11357-880701-17TM, AT&T Bell Laboratories, Holmdel, New Jersey, 1 July 1988.

[Adv90] Advanced Optoelectronics, City of Industry, California. *Photosensor Product Catalog*, 1990.

[All88] Jonathan Allen. *Introduction to VLSI Design.* MIT Video Course Study Guide. MIT Center for Advanced Engineering Study, Cambridge, Massachusetts, fifth edition, 1988.

[AOV88] Philippe Aubert, Henri J. Oguey, and Raymond Vuilleumier. Monolithic optical position encoder with on-chip photodiodes. *IEEE Journal of Solid-State Circuits*, 23(2):465–473, April 1988.

[Cau90] H. John Caulfield. The unique advantages of optics over electronics for interconnections. In Alfred P. DeFonzo, Stuart K. Tewksbury, and John R. Carruthers, editors, *Microelectronic Interconnects and Packages: System and Process Integration*, volume 1390 of *International Symposium on Advances in Interconnection and Packaging*, pages 399–402, Boston, Massachusetts, November 1990. SPIE.

[CO90] George Chiu and Modest M. Oprysko. Overview of optical interconnect technology. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 364–374, Boston, Massachusetts, November 1990. SPIE.

[DKR+90] A. Deutsch, G. V. Kopcsay, V. A. Ranieri, J. K. Catalado, et al. Electrical characteristics of lossy interconnections for high-performance computer applications. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on*

Advances in Interconnection and Packaging, pages 161–176, Boston, Massachusetts, November 1990. SPIE.

[DPC+89] A. G. Dickinson, M. E. Prise, N. C. Craft, et al. A free space optical data link using GaAs multiple quantum well modulators and Si CMOS photodetectors. Technical Memorandum 11356-881013-25TM, AT&T Bell Laboratories, Holmdel, New Jersey, 31 July 1989.

[Fel90] Michael R. Feldman. Holographic optical interconnects for multichip modules. In Alfred P. DeFonzo, Stuart K. Tewksbury, and John R. Carruthers, editors, *Microelectronic Interconnects and Packages: System and Process Integration*, volume 1390 of *International Symposium on Advances in Interconnection and Packaging*, pages 427–433, Boston, Massachusetts, November 1990. SPIE.

[Gow84] John Gowar. *Optical Communication Systems*, chapter 12. Prentice-Hall International Series in Optoelectronics. Prentice/Hall International, Englewood Cliffs, New Jersey, 1984.

[Ham90] Hamamatsu Photonics K. K., Solid State Division, Hamamatsu City, Japan. *Photodiodes: Catalog*, September 1990.

[Hua91] Alan Huang. The evolving relationships between optics and electronics. In Booz, Allen & Hamilton, Inc., editor, *DARPA Optical Computing Workshop*, Arlington, Viginia, 16 May 1991. DARPA.

[Hut86] L. D. Hutcheson. High speed optical interconnect development. In O. Glenn Ramer and Paul Sierak, editors, *High Frequency Optical Communications*, volume 716 of *Proceedings of SPIE*, pages 32–39, Cambridge, Massachusetts, September 1986. SPIE.

[Jah90] Jürgen Jahns. Integrated packaging of optical backplanes. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 523–526, Boston, Massachusetts, November 1990. SPIE.

[JL+89] J. L. Jewell, Y. H. Lee, et al. Low-threshold electrically pumped vertical-cavity surface-emitting microlasers. *Electronics Letters*, 25(17):1123–1124, 17 August 1989.

[JML+89] J. L. Jewell, S. L. McCall, Y. H. Lee, et al. Lasing characteristics of gaas microresonators. *Applied Physics Letters*, 54(15):1400–1402, 10 April 1989.

[JP86] D. J. Jackson and D. L. Persechini. Monolithically integrable high speed photodetectors. In O. Glenn Ramer and Paul Sierak, editors, *High Frequency Optical Communications*, volume 716 of *Proceedings of SPIE*, pages 104–108, Cambridge, Massachusetts, September 1986. SPIE.

[KDJS87] S. D. Kirkish, J. C. Daly, L. Jou, and Shing-Fong Su. Optical characteristics of CMOS-fabricated MOSFET's. *IEEE Journal of Solid-State Circuits*, sc-22(2):299–301, April 1987.

[KFC+90] F. Kiamilev, J. Fan, B. Catanzaro, S. Esener, and S. H. Lee. The architecture of an integrated computer aided design system for optoelectronics. In Alfred P. DeFonzo, Stuart K. Tewksbury, and John R. Carruthers, editors, *Microelectronic Interconnects and Packages: System and Process Integration*, volume 1390 of *International Symposium on Advances in Interconnection and Packaging*, pages 311–329, Boston, Massachusetts, November 1990. SPIE.

[KHK90] Raymond K. Kostuk, Yang-Tung Huang, and Masayuki Kato. Multiprocessor optical bus. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 515–522, Boston, Massachusetts, November 1990. SPIE.

[Kim91] H. H. Kim. Can optical interconnects improve electrical interconnects of digital processors? Technical Memorandum 11126-910522-05TM, AT&T Bell Laboratories, Holmdel, New Jersey, 26 June 1991.

[KR78] Brian W. Kernighan and Dennis M. Ritchie. *The C Programming Language.* Prentice-Hall Software Series. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1978.

[Leh86] R. F. Leheny. Monolithically integrated pin/FET receiver technology: a review. In Davis H. Hartman, Robert L. Holman, and Doyle P. Skinner, editors, *Integration and Packaging of Optoelectronic Devices*, volume 703, pages 116–121, Cambridge, Massachusetts, September 1986. SPIE.

[LHA+93] Teddi C. Laurin, Patricia L. Hume, Kathleen A. Alibozek, et al. *The Photonics Design & Application Handbook.* Publishers of Photonics Spectra, Pittsfield, Massachusetts, 39th edition, 1993.

[LHH+90] Gail R. Lalk, Sarry F. Habiby, Davis H. Hartman, et al. Potential roles of optical interconnections within broadband switching modules. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 386–400, Boston, Massachusetts, November 1990. SPIE.

[Lin89] Chinlon Lin, editor. *Optoelectronic Technology and Lightwave Communication Systems*, chapter 14. Van Norstrand Reinhold, New York, 1989.

[LJ+89] Y. H. Lee, J. L. Jewell, et al. Room-temperature continuous-wave vertical-cavity single-quantum-well microlaser diodes. *Electronics Letters*, 25(20):1377–1378, 28 September 1989.

[McB90] Dennis J. McBride. CAD in new areas of the package and interconnect design space. In Alfred P. DeFonzo, Stuart K. Tewksbury, and John R. Carruthers, editors, *Microelectronic Interconnects and Packages: System and Process Integration*, volume 1390 of *International Symposium on Advances in Interconnection and Packaging*, pages 330–335, Boston, Massachusetts, November 1990. SPIE.

[Mil86] Eric Miller. Introduction to practical fiber optics. Application Note 244, National Semiconductor, Santa Clara, California, November 1986.

[Mil90] David A. B. Miller. Quantum well devices for optics in digital systems. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 496–501, Boston, Massachusetts, November 1990. SPIE.

[Mil91] David A. B. Miller. Future directions in optics for computing. In Booz, Allen & Hamilton, Inc., editor, *DARPA Optical Computing Workshop*, Arlington, Viginia, 16 May 1991. DARPA.

[Mot88] Motorola Technical Information Center, Phoenix, Arizona. *Optoelectronics Device Data*, 2 edition, 1988.

[Ols93] J. J. Olsen. Control and reliability of optical networks in multiprocessors. Technical Report 985, Lincoln Laboratory - Massachusetts Institute of Technology, Lexington, Massachusetts, 22 December 1993.

[Opt90] Optek Technology, Inc., Carrollton, Texas. *Optek Data Book*, 1990.

[Rub90a] Barry J. Rubin. Electrical characterization of the interconnects inside a computer. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 314–328, Boston, Massachusetts, November 1990. SPIE.

[Rub90b] Lawrence M. Rubin. CAE tools for verifying high performace digital systems. In Alfred P. DeFonzo, Stuart K. Tewksbury, and John R. Carruthers, editors, *Microelectronic Interconnects and Packages: System and Process Integration*, volume 1390 of *International Symposium on Advances in Interconnection and Packaging*, pages 336–358, Boston, Massachusetts, November 1990. SPIE.

[Shr90] U. A. Shrivastava. Design, simulation model and measurements for high density interconnections. In Alfred P. DeFonzo, Gnanalingam Arjavalingam, and James Pazaris, editors, *Microelectronic Interconnects and Packages: Optical and Electrical Technologies*, volume 1389 of *International Symposium on Advances in Interconnection and Packaging*, pages 122–137, Boston, Massachusetts, November 1990. SPIE.

[Sib90] M. J. N. Sibley. *Optical Communications.* Optical and Electro-Optical Engineering Series. McGraw-Hill, Inc., New York, first edition, 1990.

[Sie90] Siemens Components, Inc., Optoelectronics Division, Cupertino, California. *Optoelectronics Data Book 1990*, 1990.

[SS90] Julian B. D. Soole and Hermann Schumacher. Transit-time limited frequency response of InGaAs MSM photodetectors. *IEEE Transaction on Electron Devices*, 37(11):2285–2291, November 1990.

[Sze69] S. M. Sze. *Physics of Semiconductor Devices.* John Wiley & Sons, Inc., New York, 1969.

[T. 89] T. F. Knight, Jr. Technologies for low latency interconnection switches. In *Proceedings 1989 ACM*, Symposium on Parallel Algorithms and Architectures, pages 351–358. ACM, 1989.

[Tex90] Texas Instruments, Dallas, Texas. *Optoelectronics and Image Sensor Data Book*, 1990.

[Wil91] Richard C. Williamson. Optical interconnections and computing, integrated optoelectronic devices. In Booz, Allen & Hamilton, Inc., editor, *DARPA Optical Computing Workshop*, Arlington, Viginia, 16 May 1991. DARPA.

[Wya91] Clair L. Wyatt. *Electro-Optical System Design*, chapters 11, 19. Optical and Electro-Optical Engineering Series. McGraw-Hill, Inc., New York, 1991.