HIGH-SPEED PHOTODIODES IN STANDARD CMOS TECHNOLOGY

Saša Radovanović

.

Publisher:
Print Partners Ipskamp,
Bergselaan 169 B, 3037 BJ Rotterdam, The Netherlands,
www.ppi.nl

Cover design: Srdjan Nikolić-Peka

ISBN 90-365-2115-7

HIGH-SPEED PHOTODIODES IN STANDARD CMOS TECHNOLOGY

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. F. A. van Vught,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op woensdag 1 december 2004 om 16.45 uur

door

Saša Radovanović

geboren op 11 Augustus 1974
te Peć, Servië and Montenegro

Dit proefschrift is goedgekeurd door:

De promotor: Prof. dr. ir. Bram Nauta

De assistent promotor: dr. ir. Anne Johan Annema

Dit onderzoek is gesteund door het MESA$^+$ institute als onderdeel van het TeraHertz projekt.

... Naposletku, sve čim se ovaj naš život kazuje - misli, napori, pogledi, osmesi, reči, uzdasi - sve to teži ka drugoj obali, kojoj se upravlja kao ka cilju, i na kojoj tek dobiva svoj pravi smisao. Sve to ima nešto da savlada i premosti: nered, slut ili nesmisao. Jer sve je prelaz,most čiji se krajevi gube u beskonačnosti a prema kom su svi zemljini mostovi samo dečje igračke, bledi simboli.

A sva je naša nada sa one strane...

Ivo Andrić

... In the end, everything through which this life of ours is expressed - thoughts, efforts, glances, smiles, words, sighs - is all reaching out to another shore, as towards its aim, and only there will it be granted its true meaning. Everywhere there is something to overcome or to bridge: disorder, death, meaninglessness. Everything is a transition, a bridge whose ends are lost in infinity, beside which all the bridges of this earth are only children's toys, pale symbols.

And all our hope lies on the other side...

Ivo Andrić, Nobel Prize for literature 1961

Contents

CHAPTER 1

Project motivation

1.1 Motivation

In the last decades, the speed of microprocessors has been increasing exponentially with time and will continue to do so for at least another decade. However, the local computing power of the microprocessor alone does not determine the overall speed of a system. Equally important to the processor's bare computing power is the speed at which data can be distributed to and from the processor. That means that the speed of the data-input and -output channel must keep pace with the processor's computing power. For the future it is expected that these data-communication channels will become the speed-bottle-neck for the whole system.

For short and medium distance (centimeters up to hundreds of meters) the data communication channels are usually implemented as wired electrical connections. However, at high speeds major problems occur: poor impedance matching results in distorted signals, signal losses due to the skin-effect, significant Electro-Magnetic noise is generated which degrades the system performances. For this reason, a research project named "Fiber-to-the-chip" is started. The main goal is to increase the data-rate in short-haul communica-

tion by using optical fibers instead of wires. The main focus is given at the receiver side; the objective was to design a low-cost Gb/s receiver that can be easily integrated with the rest of electronic circuitry.

The electronics for the long distance channels is typically realized with expensive exotic technologies: Gallium-Arsenide High-Electron-Mobility-Transistors (GaAs HEMT) [1], [2], and Indium-Phosphide Hetero-Junction-Bipolar-Transistors (InP HBT) [3], [4]. The maximum bit-rate for these systems is around 100 Gb/s per channel. The first reason for adequacy of these expensive blocks is long distance links: the cost per length of the fiber is low. The second reason for the efficiency of this solution is that a large number of users share the links: the cost per user is low.

For the medium and short distances however, as well as for the small number of users (fiber-to-the-home or fiber-to-the-desk) the optical receivers and transmitters should not be expensive. This project aims at increasing the speed and lowering the costs of the fully integrated optical receivers in standard CMOS technology. These receiver chips (inside microprocessors for example) have integrated light-sensors and thus they are cheap and do not have wire-speed limitations. The light shines directly onto the CMOS! That means that microprocessors could be connected in networks by just using light. The result could then be a low-cost and high-speed fully integrated optical data communication system for distances ranging from chip-to-chip (cm range) up to up to hundreds of meters, typical for LAN environments.

1.2 Project phases and scientific challenges

The main goal of the project is to design a low-cost optical receiver in inexpensive CMOS technology. The bit rate should be up to the Gigabit/second range. The receiver alone can be used for future chip-to-chip, board-to-board communication (λ=850 nm), and in some other applications like fully integrated CD-ROM (λ=780 nm) and DVD pick-up units (λ=650 nm, λ=400 nm).

In order to lower the cost of the high data-rate communication, typically short wavelength optical components (λ<850 nm) are used in combination with silica multimode fibers [6]. The optical transmission in *long-distance* communication is typically based on long wavelengths (1300-1600 nm) and single-mode fiber, due to the optimal transmission parameters of the fiber in that region [7]. Both attenuation and dispersion are low such that signals can reach larger

distances without the need for repeaters. However, the components for these wavelengths are very expensive for short-haul application, mainly because of the light sources and the hybrid technology in the receiver module. In the latter, Si-based electronics has to be combined with the GaAsP photodiode. Therefore, the cost aspect forms a decisive argument regarding all the advantages of fiber-optic communication.

Combining the multimode fiber with the short wavelength components will make our system more economically attractive. Light Emitting Diodes (LEDs) and Vertical Cavity Surface Emitting Lasers (VCSELs) operating at 850 nm are less expensive than lasers operating at 1300 nm and 1550 nm. They also open the possibility of integrating a Si photodiode with the receiver front-end using straightforward CMOS technology.

1.3 Part of the TeraHertz project

This research work has been accomplished within the framework of the TeraHertz project, sponsored by the MESA+ research institute of the University of Twente. The aim of the project is to establish interconnect between superconducting electronics on the one side and semiconductor electronics on the other, for very high data rates. The superconductor circuits may operate at a clock speed up to a few THz [5], whereas opto-electronic conversion of CMOS circuits is in a few MHz range. Therefore, there is a six order of magnitude speed difference. The first attempt was to run the total system in the few GHz range, requiring drastical improvement in speed of optical receiver in standard CMOS technology. Besides the speed, a complicating factor is the difference in signal levels and in operating temperature between the two. The former requires temperatures running from 4-40 K, whereas CMOS technology operates mainly at the room temperature. The voltage level of the superconducting circuits ranges from 0.1-10 mV, whereas the CMOS circuits need at around 1 V.

The TeraHertz project combines the expertise of several research groups that are active within the MESA+ research institute:

- Low Temperature Division (LT)
- IC-Design (ICD)
- Testable Design and Testing (TDT)
- Lightwave Devices Group (LDG)
- Applied Analysis and and Mathematical Physics (AAMP)

These groups geared towards the acquisition, transportation and manipulation of information at very high data-rates. TeraHertz program's point of focus is the design, fabrication and testing of a demonstrator analogue-to-digital converter (ADC) suitable for wideband code division multiple access (WCDMA) communication. As such, it provides a solid basis for future software radio applications. Besides this demonstrator, it is expected that spin-off will be generated that is useful for other applications as well.

"*... You also see that people from very different disciplines are working together fruitfully, like superconductor-people and semiconductor-people. That is quite unique and it will help to establish the name and reputation of both Terahertz and MESA+.*"

From interview with Gerrit Gerritsma († 2002).

1.4 Outline of the thesis

This thesis consists of 7 chapters. The first one is the introductory chapter where motivation for the "Fiber-to-the-chip" project is given. The goal is to design monolitically integrated optical receiver in straightforward CMOS technology, for short-haul optical communication and bit-rates up to a few Gb/s. This achievement can minimize the total cost of the system.

The second chapter gives the main motivations for using optical communication on short distances in comparison with the straightforward wired (electrical) communication. The advantages and disadvantages of the optical communication system are presented. The following part of the chapter explains major characteristics of light: light spectra, frequency, wavelength, rays or photons. There the three key building blocks for optical communication system: light sources, optical fiber, and light detectors. These blocks are also explained in chapter 2. Light is typically *created* either using lasers or high-intensity LEDs. Main characteristics as well as advantages and disadvantages of lasers and LEDs are briefly discussed. Optical fibers as typical *"medium"* for optical communication are presented with their main characteristics. Finally at the *receiver* side, we will present photodetectors in general, together with their requirements and specifications. The important properties of light absorption will also be investigated such as absorption coefficient and absorption depth. Related to this, the quantum efficiency, responsivity as well as the wavelength sensitivity range for CMOS photodiodes will be analyzed.

Chapter 3 presents a detailed analysis of the time and frequency responses of photodiodes in CMOS technology for λ=850 nm. Physical processes inside a photodiode are thoroughly investigated using one particular CMOS technology; it is a standard 0.18 µm CMOS since that was the available fabrication technology. For every high-speed photodetector there are two main parameters that define their figure-of-merit: responsivity and bandwidth. The bandwidth is the main limiting factor for Gb/s optical detection. For easier comparison among the photodiodes, it is assumed that the active area of all analyzed diodes is the same providing the same maximal responsivity. There are actually two in nature different bandwidths of the photodiode: intrinsic (physical) and extrinsic (electrical). The first is inversely related to the time that excess carriers need to reach junctions and thus, to be detected at the output terminal. The second bandwidth is related to diode capacitance as well as the input impedance of the subsequent transimpedance amplifier. By approximation, the total bandwidth is the lowest between these two. These bandwidths will be separately analyzed in detail in chapter 3.

The calculated intrinsic bandwidth of typical photodiode in standard CMOS for λ=850 nm, is in the low MHz range. This is two orders of magnitude too low. Chapter 4 presents a solution for more than *4 times* bandwidth improvement in comparison with state-of-the art detectors in standard CMOS technology without sacrificing diode responsivity. This is achieved by using an inherently robust analog equalizer. Complex adaptive algorithms are not required. The proposed configuration is spread and temperature robust. Using this approach, a 3 Gb/s data-rate for λ=850 nm and 0.18 µm CMOS technology, is achieved. This is the highest speed ever achieved with the integrated photodetector in standard CMOS for the moderate responsivity of 0.4 A/W.

For very low wavelength λ=400 nm (blue light), the light penetration depth in silicon diode is very small (0.2 µm). Chapter 5 analyzes both time and frequency response of photodiode structures in CMOS. Excess carriers are generated close to junctions allowing large diffusion current bandwidth (hundreds of MHz up to a few GHz range). This diffusion current mainly determines the total diode photocurrent. The same chapter investigates polysilicon photodiodes designed using NMOS and PMOS gates. The measured bandwidth of the poly photodiode was 6 GHz, which figure was limited by the measurement equipment. This is the highest bandwidth among all photodiodes designed in standard CMOS technology. However, the quantum efficiency of this poly photodiode is low (<8 %) due to the very small light sensitive diode volume. The

diode active area is limited by a narrow depletion region and its depth by the technology.

Chapter 6 analyzes the frequency behavior of photodiode in standard CMOS in general, for the whole wavelength-sensitivity range 400 nm<λ<850 nm: a general device-layer/p-substrate diode is analyzed for various scales of the technology. The photodiode bandwidth can be improved using the analog equalization method independent on input wavelength,. This chapter analyzes the starting and the maximal equalization frequencies that depends on CMOS technology and input wavelength i.e. it depends on the roll-off in the diode frequency characteristics.

Chapter 7 presents the most important conclusions in the thesis and gives an overview of the publications originating from the work described in this thesis.

1.5 Main contribution of this thesis

The work presented in this thesis gives the original contribution to the Gb/s integrated optical receivers in standard CMOS. For the first time, an integrated photodetector and electrical amplifier for a Gb/s data-rate is designed in standard CMOS technology for λ=850 nm. The achieved 3 Gb/s data-rate is more than 4 times the data-rate of the state-of-the-art integrated CMOS optical detector, and *300 times* the speed of the conventional photodetector in standard CMOS.

References

[1] Joongho Choi; Sheu, B.J.; Chen: "A monolithic GaAs receiver for optical interconnect systems O.T.-C.", *Solid-State Circuits, IEEE Journal of*, Volume: 29, Issue: 3, March 1994, pp.328-331.

[2] C. Takano, K. Tanaka, A. Okubora; J. Kasahara: "Monolithic integration of 5-Gb/s optical receiver block for short distance communication", *Solid-State Circuits, IEEE Journal of* , Volume: 27, Issue: 10 , Oct. 1992, pp.1431-1433.

[3] M. Bitter, R. Bauknecht, W. Hunziker, H. Melchior: "Monolithic InGaAs-InP p-i-n/HBT 40-Gb/s optical receiver module", *Photonics Technology Letters, IEEE* , Volume: 12 , Issue: 1, Jan. 2000, pp.74-76.

[4] H.-G. Bach, A. Beling, G.C. Mekonnen, W. Schlaak: "Design and fabrication of 60-Gb/s InP-based monolithic photoreceiver OEICs and modules", *Selected Topics in Quantum Electronics, IEEE Journal of*, Volume: 8, Issue: 6, Nov.-Dec. 2002, pp.1445 - 1450.

[5] http://mesainfo.el.utwente.nl/terahertz/terahertz.html

[6] I Radovanović, W. van Etten and H. Freriks: "Ethernet Based Passive Optical Local Area Networks for Fiber-to-the-Desk Application", *Journal of Lightwave Technology*, the Special Issue on optical networks in November 2003, pp.1166-1170.

[7] Jeff Hetcht: "*Understanding Fiber Optics*", Prentice Hall, New Jersey, 1999.

CHAPTER 2

Introduction

2.1 Why optical interconnection?

For nearly forty years scientists are using light to "talk" over distance. The birth of optical communications occurred in the 1970s with two key technology breakthroughs. The first was the invention of the semiconductor laser in 1962 [1]. The second breakthrough happened in September 1970, when a glass fiber with an attenuation of less than 20 dB/km was developed [2], [3]. With the development of optical fibers with an attenuation of 20 dB/km, the threshold to make fiber optics a viable technology for telecommunications was crossed. The first field deployments of fiber communication systems used Multimode Fibers (MMFs) with lasers operating in the 850 nm wavelength band. These systems could transmit several kilometers with optical losses in the range of 2 to 3 dB/km. The total available bandwidth of standard optical fibers is enormous; it is about 20 THz. A second generation of lasers operating at 1310 nm enabled transmission in the second window of the optical fiber where the optical loss is about 0.5 dB/km in a Single-Mode-Fiber (SMF). In the 1980s, the telecom carriers started replacing all their MMFs operating at 850 nm. Another wavelength window around 1550 nm was developed where a standard SMF has its

minimum optical loss of about 0.22 dB/km.

From this small history of fibers it can be concluded that the main research focus was on long-distance communication. Chapter 1 described that the electronics for the long distance channels is typically realized with expensive exotic technologies such as GaAs or InP. The bit-rate for these systems is large, around 100 Gb/s per channel, with the low cost per length of the fiber and for a large number of users.

Replacing electrical wires with optical fibers for short distances for a small number of users is still challenging. The goal is to have low cost but high (Gb/s) bit-rates of the system. However, the important question is should we use light (fibers) to directly connect silicon chips and why?

A large study about this issue is published in the literature and some of the results will be briefly presented further in this introductory chapter. In [4], [5], [6], D. Miller tried to stress the practical benefits of optical interconnects and drawbacks of electrical systems for very high-speed communication (>10 GHz). His approach was to analyze the similarities and differences in optical and electrical systems, which will be briefly investigated in the following subsections.

2.1.1 Electrical and Optical Interconnection - Similarities

At the most basic level, optical and electrical physics are very closely linked. In practice, in both the electrical and optical cases, it is the electromagnetic wave that carries a signal through a medium (see Figure 2.1).

It is important to stress that in high-speed communication, it is not electrons that carry the signals in wires or coaxial cables; actually the signal is carried by electromagnetic wave [4]. It is also good to note that signals in wires propagate at the velocity of light (or somewhat lower than light velocity if coaxial cables are filled with a dielectric). Hence it is generally incorrect to say that signals propagate faster in optics. In fact, signals typically travel slightly slower in optical fibers than they do in coaxial cables because the dielectric used in cables has a lower dielectric constant than glass.

In case of electrical interconnection lines on chips, the signals do move at a lower speed, but this speed is determined by the overall resistance (R) and capacitance (C) of the interconnect line [7].

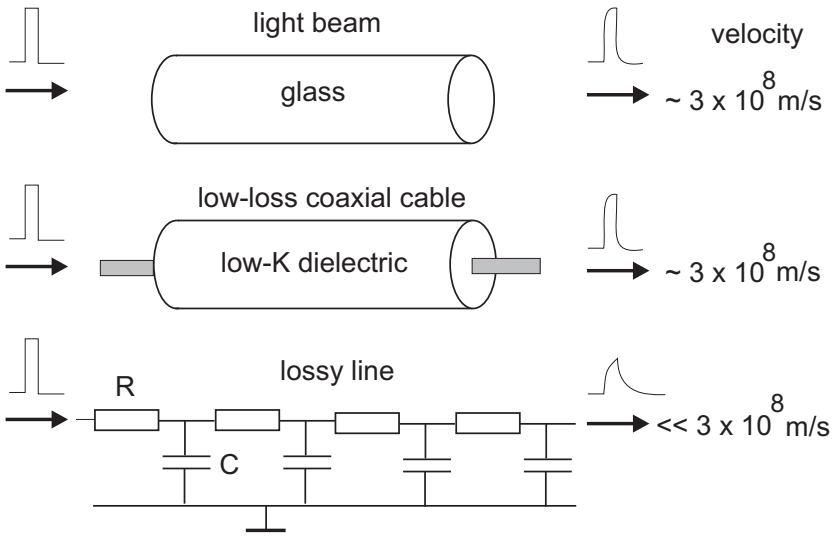


Figure 2.1: Types of optical and electrical propagation and their velocity. One possible model of the lossy line is presented.

2.1.2 Electrical and Optical Interconnection - Differences

Apart from large similarities, there are important basic differences between optical and electrical physics. The most important one is the higher (carrier) frequency and therefore a large photon energy.

The higher carrier frequency (shorter wavelength, typically in 1 µm range) allow us to use optical fibers to send optical signals without high loss [8]. There are small "wavelength windows" where the loss in the fibers (both singlemode and multimode) is small (<1 dB/km). The dispersion in singlemode and multimode fibers used in short distance communication is small too. In this way it is possible to avoid the major loss phenomena that in general limits the capacity of electrical interconnects on high frequencies: signal and clock distortion and attenuation. The optical generation and detection for interconnection is in principle quantum mechanical (e.g., counting photons). This is in contrast to a classical source/detection of voltages and currents; for example, detection of light in practice involves counting photons, not measuring electric field amplitudes. Two practical consequences are that all optical interconnections provide voltage isolation (used in opto-isolators), and optics can offer lower powers for interconnects: it can solve the problem of matching high-impedance low-power

devices to the low impedance (and/or higher capacitance) of electromagnetic propagation. Due to the optical interconnection, there are no inductive voltage drops on input/output pins and wires that exist in the electrical connection.

A signal propagating down an electrical line may start with sharply rising and falling "edges". However, these edges will gradually decrease because the loss-related distortion and dispersion, as illustrated in Figure 2.1. This "softening" of the edges makes precise extraction of timing information more difficult. For the same communication distances, optical systems have relatively little problem with such variations. The dispersion and loss in optical fiber are typically smaller than in electrical wires, which is explained in section 2.3.2. The optics becomes thus increasingly attractive at high bit rates but also in higher interconnect densities (e.g., high density edge connectors for boards, or even very high density connections of chips), and arguments for optics become increasingly strong as the number of lines on the board increases. However, the disadvantage of optics is in the systems with optical connectors, because the connector size is much larger than a fiber diameter.

Optics also offers several additional opportunities that have essentially no practical analogy in the electrical case, including use of short pulses for improved interconnect performance [9]. A very important advantage of optical fibers is that they can be deployed in environments with large electromagnetic interference (EMI) and radio-frequency interference (RFI), such as airports, factories, military bases etc.

In total, the advantages of optical interconnection in comparison with the straightforward electrical connection are summarized below, [4]:

- Immune to noise (electromagnetic interference and radio-frequency interference
- Signal Security (Difficult to tap)
- Nonconductive (does not radiate signals) - electrical isolation
- No common ground required
- Freedom from short circuit and sparks
- No inductive voltage drops on pins and wires
- Reduced size and weight cables (but not connectors)

- Ability to have 2-D interconnects directly out of the area of the chip rather than from the edge
- Resistant to radiation and corrosion
- Less restrictive in harsh environments
- Low per-channel cost [2]
- Lower installation cost in future (Wavelength Division Multiplexing [10])

Despite the many advantages of fiber optic systems, there are some disadvantages. Because of the relative newness of the technology, fiber optic components are still expensive even though the prices decrease dramatically in a last couple of years. Fiber optic transmitters (but not the receivers[1]) are still relatively expensive compared to electrical interfaces. The lack of standardization in the industry has also limited the acceptance of fiber optics. Many industries are more comfortable with the use of electrical systems and are reluctant to switch to fiber optics. However, a huge bandwidth advantage of the optical interconnection will move the future industries more towards optics. Nevertheless, the signal processing is typically electrical, so the interface between the electrical and optical signals will always be required and the total speed in the system will be limited by the electronics.

2.2 Characteristics of light

The working methods of optical communication and optical fibers depend on basic principles of optics and the interaction of light with matter. From a physical standpoint, light can be seen either as electromagnetic waves or as photons. Both view points are valid and valuable, but the simplest view for a fiber transmission is to consider light as rays travelling in straight lines and for a light detection to see the light as a number of incident photons on the photodetector surface.

Light is only a small part of the electromagnetic (EM) spectrum. The difference in radiation in different parts of EM spectrum is a quantity that can be measured: length of wave/frequency of EM-field and energy of photons. In

[1] A 3 Gb/s data-rate optical receiver in inexpensive CMOS technology is presented in chapter 4 of this thesis.

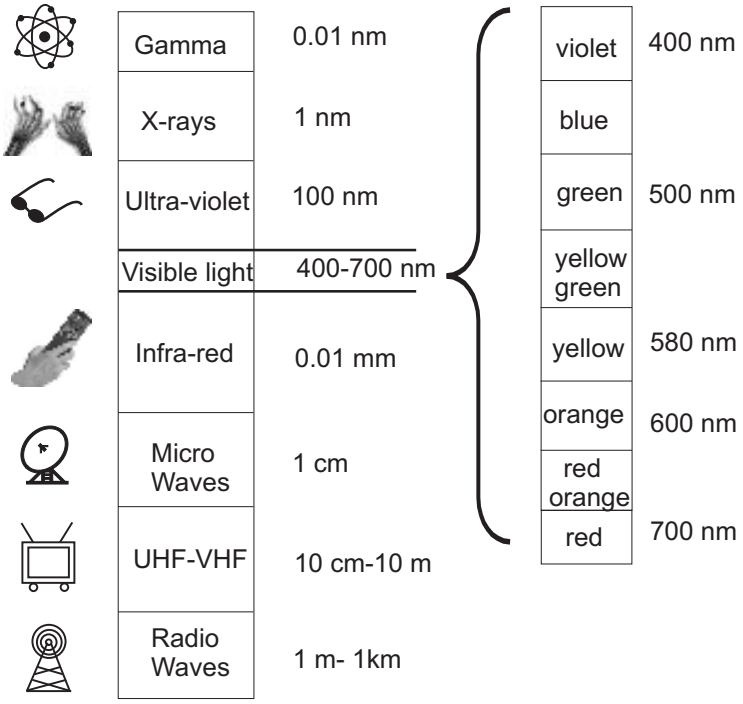


Figure 2.2: Electromagnetic spectra with applications in the certain frequency range.

some parts of the spectrum, frequency is used the most; in others wavelengths and photon energies. In Figure 2.2 the EM spectrum is presented with typical applications in certain spectral ranges.

In the optical world most common used light quantity is a wavelength, measured in micrometers or nanometers. It is inversely proportional to frequency f:

$$\lambda = \frac{c}{f} \tag{2.1}$$

where c is the speed of light in the transporting medium.

2.3 Optical fiber types

Optical fibers are characterized in general by the number of modes that propagate along the fiber. Basically, there are two types of fibers: single-mode fibers and multi-mode fibers. The basic structural difference is the different core size.

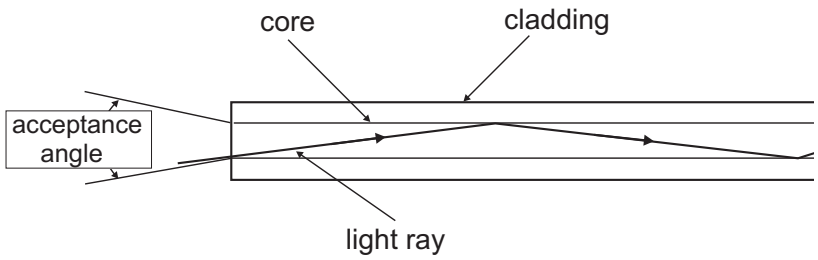


Figure 2.3: Single-mode optical fiber with typically one light ray (small core diameter 9 µm).

2.3.1 Single-mode fibers

Single-mode fibers have lower signal loss and higher information capacity (bandwidth) than multimode fibers. They are capable of transferring higher amounts of data due to low fiber dispersion [2]. These fibers illustrated in Figure 2.3, are mainly used for long-haul optical communication because of their typical loss that is less than 0.2 dB/km.

2.3.2 Multimode Fibers

As their name implies, multimode fibers propagate more than one mode. This is illustrated in Figure 2.4. The number of modes M_{n} propagated depends on the core size and numerical aperture (NA) and can be approximated by:

$$M_{\mathrm{n}} = \frac{V^2}{2} \quad \text{and} \quad \frac{V^2}{4} \tag{2.2}$$

for step index fiber and gradient index fiber, respectively. V is known as the normalized frequency, or the V-number, which relates the fiber size, the refractive index, and the wavelength. The V-number is:

$$V = \left[\frac{2\pi a}{\lambda}\right] \times \mathrm{NA} \tag{2.3}$$

NA is closely related to the acceptance angle and it is approximately [8]:

$$\mathrm{NA} = \sqrt{n_0^2 - n_1^2} \approx n_0 \sin\Theta_c \tag{2.4}$$

[2]Basically, dispersion is the spreading of light as light propagates along a fiber. This causes intersymbol interference i.e. an incorrect bit detection at the fiber's output.

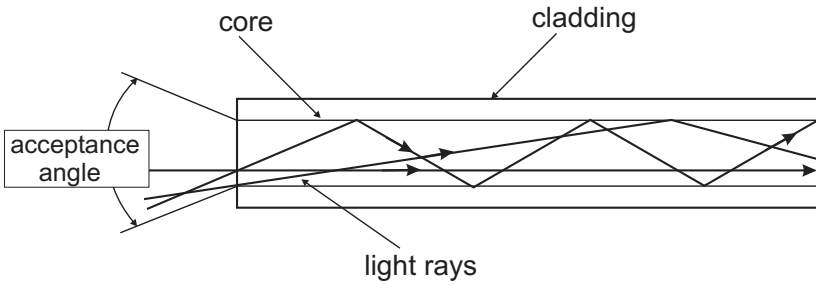


Figure 2.4: Multimode-mode optical fiber with multiple light rays (large core diameter 50 µm). The angles of the light rays are refracted at the air/fiber interface according to Snell's law.

where n_0 and n_1 are refractive index of the core and cladding respectively, and Θ_c is the confinement angle in the fiber core. As the core size and NA increase, the number of modes increases. Typical values of fiber core size and NA are 50 µm to 100 µm and 0.20 to 0.29, respectively.

A large core size and a higher NA have several advantages. Light is launched into a multimode fiber with more ease. Higher NA and larger core size make it easier to make fiber connections: during fiber splicing, core-to-core alignment becomes less critical. Another advantage is that multimode fibers permit the use of light-emitting diodes (LEDs). Single mode fibers typically must use laser diodes due to their small diameter (< 10 µm). LEDs are cheaper, less complex, and last longer and they are preferred for the large number of applications [8].

Nevertheless, multimode fibers have some disadvantages. As the number of modes increases, the effect of modal dispersion increases. Modal dispersion (intermodal dispersion) is important because, as the pulses spread, they can overlap and interfere with each other, limiting data transmission speed. Typical dispersion values for fiber are measured in nanoseconds per kilometer of fiber. These can be translated into an analog bandwidth limit in the transmission.

For instance, if one ray travels straight through a multimode fiber and another bounce back-and-forth at the acceptance angle Θ_c through the same fiber, the second ray would travel further for:

$$l_1 = l\left(\frac{1}{\cos\Theta_c} - 1\right) \qquad \text{[m]} \tag{2.5}$$

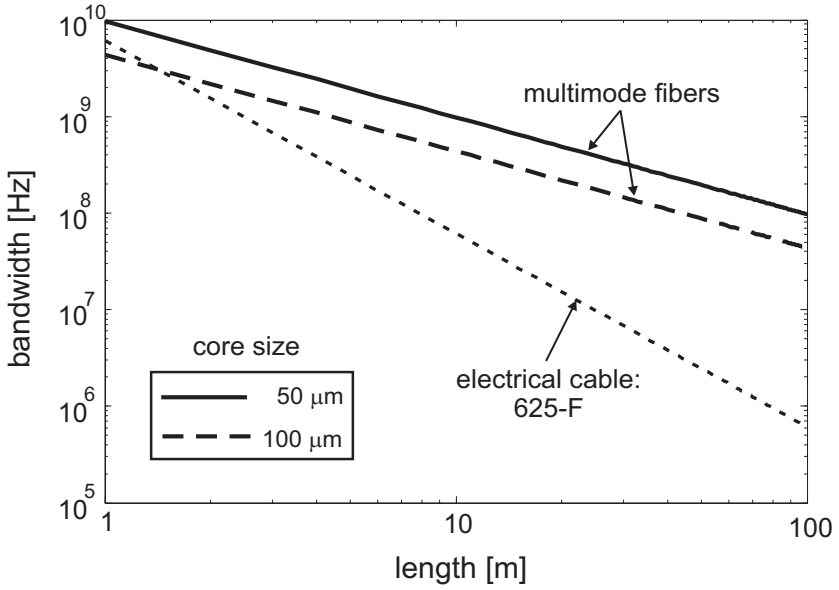


Figure 2.5: The analog bandwidths of the multimode fibers having core diameters of 50 µm and 100 µm and electrical cable 625-F vs. fiber/cable length.

where l is the length of the multimode fiber. The ray that goes down the center of the fiber with speed v will reach the output τ_{r} seconds before the the ray that bounces at the acceptance angle:

$$\tau_{\mathrm{r}} \approx \frac{l_1}{v}\left(\frac{1}{\cos\Theta_c} - 1\right) \tag{2.6}$$

Thus, an instantaneous pulse at the start will spread out τ_{r} seconds at the end. The analog bandwidth of the multimode fiber is inversely proportional to the pulse spread.

For a typical NA values of multimode fibers of 0.20 to 0.29, the acceptance angle calculated using equation (2.4) ranges from 11.5° to 17°. If we take the speed of the ray in optical fiber to be about $2 \cdot 10^8$ m/s [11], the dispersion t_{r} can be calculated from equation (2.6). The analog bandwidth of the multimode fiber, inversely proportional to dispersion is calculated and presented in Figure 2.5.

As far as electrical cables are concerned, the attenuation A_{tt} in dB is pro-

portional to the length of the cable and square-root of the frequency [12], [13]:

$$A_{\mathrm{tt}} = e^{-3k_1 l\sqrt{f}} e^{-3k_2 lf} \tag{2.7}$$

where f is the frequency expressed in megahertz, k_1 and k_2 are parameters defining the electrical cable type and l is the cable length expressed in kilometers. The first exponential term is due to a skin-effect and the second exponential term is due to the dielectric loss. One should notice that the additional advantage of optical fibers is that the fiber-loss is independent on signal frequency over their normal operating range [11].

For a very small attenuation cable 625-F [13], $k_1 = 0.6058$ and $k_2 = 0.0016$. Since $k_1 \gg k_2$, the bandwidth frequency of the cable f_{cab} can be approximated as:

$$f_{\mathrm{cab}} = \frac{1}{400 k_1^2 l^2} \tag{2.8}$$

The 625-F cable bandwidth is illustrated in Figure 2.5. For larger transmission distances, the bandwidth of the electrical cable drops significantly in comparison with the bandwidth of the multimode fibers.

2.3.3 Plastic optical fibers

Multimode fibers made entirely of plastic have higher losses than silica fibers. Therefore, they have long been outweighed, especially for communication. However, they have also the advantage of being lighter, inexpensive, flexible, and easy to handle. Since the single-mode fibers are proven unsuitable for LAN installations (high connectors cost and costly technical expertise) plastic fibers appear to be a possible solution: the physical characteristics meet the same challenges as copper and glass. It has ability to withstand a bend radius of 20 mm with no change in transmission, an 1 mm bend without breaking or damaging the fiber.

The main disadvantage of plastic fibers is their high loss. The best laboratory fibers have loss around 40 dB/km. At 650-nm wavelength (for communication using red LED) plastic fibers have loss of about 150 dB/km. Unlike glass-fibers, the loss of plastic fibers is lower at shorter wavelength and is much higher in the near infrared, as illustrated in Figure (2.6). As a result, plastic optical fibers have only limited application. They are used mainly for flexible bundles for image transmission and illumination, where light does not need to go far. In communication, plastic fibers are used for short links, like within the office

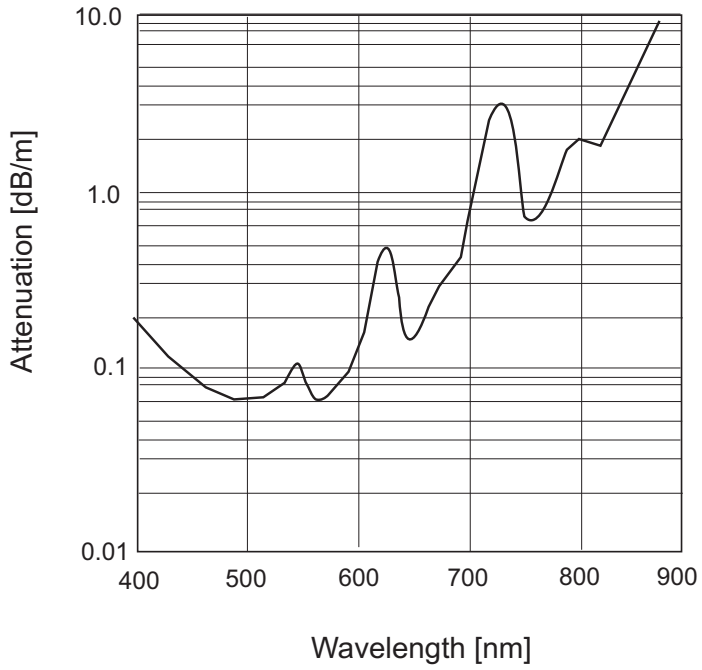


Figure 2.6: Attenuation versus wavelength for a commercial plastic multimode step-index fiber [11]. It typically decreases with wavelength while for the single-mode fibers it increases.

building or cars.

Another important concern is long term degradation at high operation temperatures. Typically, plastic fibers can not be used in applications where the temperature ranges up to 85°C. This leaves only a little margin with engine compartments of car which can get hotter. Some other heat-isolators have to be explored and used. Plastic fibers are designed similar to glass- fibers. High index cladding (see Figures 2.3, 2.4) is around the low-index core. Commercial plastic fibers are usually multimode because the material's high losses makes long-distance transmission impossible.

2.4 High intensity light sources

Light source in the fiber-optic communication system converts an electrical input signal into an optical signal. The important parameters of the source are:

- the dimension of the light-emitting area and the radiation pattern of the

optical bundle

- the efficiency
- the lifetime
- the effect of temperature on its transfer characteristics

Typical high-intensity light sources are lasers and LEDs. Both sources operating at 850 nm, which is typical wavelength used for short distance communication, are economically attractive, since they are less expensive than lasers operating at 1300 nm and 1550 nm.

2.4.1 Lasers

A vertical cavity laser (VCSEL), formed in a single epitaxial growth, is realized by sandwiching a light-emitting semiconductor diode between multi-layer crystalline mirrors. The technologies used for VCSEL fabrication are typically InGaN or AlGaAs. Unlike edge-emitting lasers, which require a larger wafer area and power consumption, the laser output from a VCSEL is emitted from a relatively small area (5-50 μm^2) on the surface of the chip, directly above the active region. The VCSEL is illustrated in Figure 2.7. The VCSELs physical structure yields numerous inherent advantages including: compact size and surface area, high reliability, flexibility in design, ability to efficiently test each die while still in the wafer state, low current requirements, efficient fiber coupling, high speed modulation, and the ability to build multiple lasers on a single semiconductor. A big advantage of VCSELs is that they can be modulated with very high frequencies (>50 GHz).

2.4.2 Light Emitting Diodes (LEDs)

The working principle of the LED is based on emission of photons due to recombination of holes and electrons in a semiconductor device. The number of carriers present in the active LED region is proportional to the forward current through the LED. The dimensions of the emitting area of an LED are similar to the core diameter of the multimode fiber.

In most LEDs the light is not completely monochromatic i.e. show relatively broad spectra. The visible light from an LED can range from red (at a wavelength of approximately 700 nanometers) to blue-violet (about 400 nanometers). Some LEDs emit infrared (IR) energy (850 nanometers or longer).

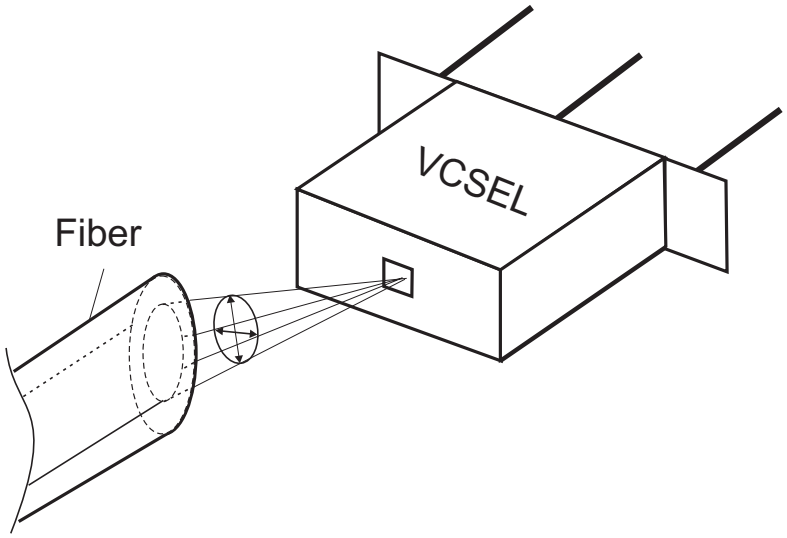


Figure 2.7: VCSEL structure with light emitted from the surface of the chip. Possible coupling with both the single-mode and multimode optical fibers.

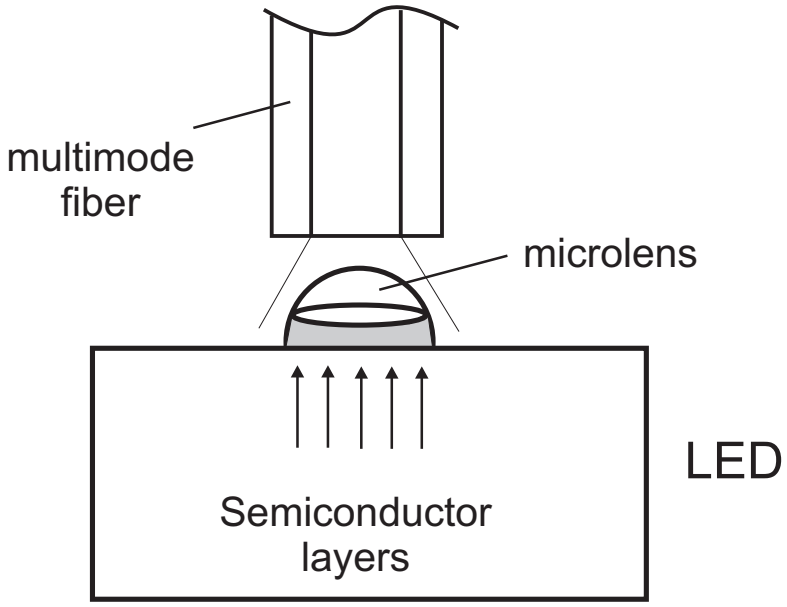


Figure 2.8: A LED coupled to a multimode fiber.

2.5 Photodetectors - introduction

A silicon photodetector is in general a solid state transducer used for converting light energy into electrical energy. The following subsections present the main photodetector characteristics.

2.5.1 Ideal photodetector

In the ideal case, the photodetector should meet the following requirements:

* detect all incident photons,
* has the bandwidth larger than the input signal bandwidth,
* not to introduce additional noise, apart from the quantum shot-noise from the received signal.

In most practical applications, additional requirements can be defined. The photodetector should be small, reliable, its characteristics should not be affected by age and environment and very important is to be cost-effective.

The requirements for ideal photodetectors are very hard to meet in reality, and the photodetectors usually have limited bandwidths with finite response time. They introduce unwanted noise and the efficiency of detecting incident photons is less then 100%. The lifetime is usually limited and some detectors degrade unacceptably as they age.

Most of the photodetectors used in the today's communications are photon-effect based i.e. they directly generate the photocurrent from interactions between the photon and the semiconductor material. There is huge number of atoms in the material, about N=10^{23} cm^{-3}. When light penetrates into the detector material, there is a high probability that the semiconductor will interact with an incident photon creating an electron-hole pair. Photodetectors are grouped into four categories: photo-multipliers, photoconductors, photodiodes and avalanche photodiodes. In this thesis the main focus will be on photodiodes. The limitations of photodiodes in standard CMOS in their quantum efficiency and in the bandwidth will be discussed in the following chapters.

2.5.2 Absorption of light in silicon

Once the light is absorbed in the photo-detector, electrons in the valence band become mobile carriers in the conduction band. This process leaves behind

mobile holes in the valence band. Basically, these two types of carriers are seen as a photocurrent at the photodiode terminals.

Light is absorbed exponentially versus depth inside the silicon [8]. The light intensity I inside the silicon can be approximated as:

$$I \propto e^{-\alpha x} \tag{2.9}$$

where α is the wavelength dependent absorption coefficient while x is the depth in silicon. The absorption coefficient for silicon can be approximated with the following formula [14]:

$$\alpha = 10^{13.2131 - 36.7985\lambda + 48.1893\lambda^2 - 22.5562\lambda^3} \quad 1/[\text{cm}] \tag{2.10}$$

The wavelength λ of the input light signal is given in [µm].

The photodiodes in CMOS technology are sensitive only for a particular wavelength range. The photon energy $h\nu$ is wavelength dependent and it should be larger than the band-gap of the semiconductor material (in this case silicon) [15]. For larger wavelengths λ>950 nm, photon energy is not high enough to create an electron-hole pair, and no electrical signal can be detected. For lower wavelengths on the other hand, $\lambda<$ 400 nm, excess carriers are generated very close to a photodiode surface. The surface recombination rate is high and only a small part of them contribute to the photocurrent. As a result, the wavelength sensitivity range of CMOS photodiodes is $\lambda \in [400 - 850]$ nm.

For best performance e.g. the highest speed and responsivity, the photodiode should be designed to allow the largest number of photons to be absorbed in depletion regions; in ideal case photons should not be absorbed until they have penetrated as far as the depletion region, and should be absorbed before penetration beyond it. The relative depth to which photon penetrates is a function of its wavelength (see chapters 3, 4 and 5). Short wavelength light (around blue and violet) are absorbed close to the photodiode surface while those with longer wavelength (infrared) may penetrate 10ths of micrometers deep in the substrate.

The calculated values of the absorption coefficient are shown in Figure 2.9. From this figure we conclude that the difference in absorption coefficient for the two boundaries is very large: $\alpha = 7.5 \times 10^2 \div 5.5 \times 10^4 \quad \text{cm}^{-1}$. This absorption coefficient is directly related to the absorption depth inside the CMOS

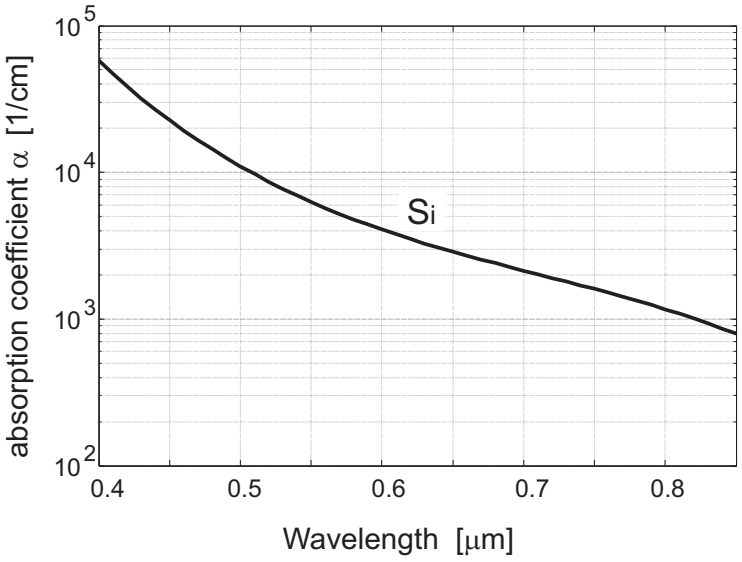


Figure 2.9: The absorption coefficient α for silicon photodiodes versus input wavelength of the light signal λ.

photodiode, as illustrated in Figure 2.10.

The light intensity drops exponentially inside silicon:

$$\frac{\partial I}{\partial x} \propto \alpha e^{-\alpha x} \tag{2.11}$$

The more light is absorbed in the photodiode, the more excess carriers are generated. We define a parameter $G(x)$ which is the *carrier generation rate* as a result of the incident light in the unity of time often modelled as:

$$G(x) = \Phi_0 \alpha e^{-\alpha x} \tag{2.12}$$

where Φ_0 is the *photon flux* at the silicon surface generated by a monochromatic optical source and can be further expressed as:

$$\Phi_o = \frac{P_{\text{in}}}{h\nu}(1 - R_{\text{f}}) \tag{2.13}$$

P_{in} is the input optical power density (W/cm^2), $h\nu$ is the photon energy and R_{f} is the reflection coefficient due to the different index of reflections of the "outside world" on the top of the silicon and the silicon itself [8]. During each

unit of time, $P_{\text{in}}/h\nu$ photons arrive with a frequency ν. The number of generated carrier pairs is $\sim \eta P_{\text{in}}/h\nu$ resulting in a photocurrent of $\sim \eta e P_{\text{in}}/h\nu$ [8] (where e is electron charge); this is often referred to as photodiode *responsivity*. It is defined as the average photocurrent per unit of incident optical power:

$$R = \frac{e\eta}{h\nu} \tag{2.14}$$

The parameter η is quantum efficiency. The quantum efficiency is often defined as the average number of (primary) generated electron-hole pairs per incident photon. For every photodetector there are typically four quantum efficiency components:

1. efficiency of light transmission to the detector (fraction of incident photons that reach the silicon surface)

2. efficiency of light absorption by the detector (fraction of photons reaching the silicon surface that produce electron-hole (EH) pairs)

3. quantum yield (number of EH pairs produced by each absorbed photon)

4. charge collection efficiency of the photo-detector (fraction of generated minority carriers by presence of light, that cross the pn junction before recombining).

However, during the calculations of the available output photocurrent, typically only the first and the fourth quantum efficiency components are taken into account. The other two components are taken to be equal to one. Typical value of the quantum efficiency in a CMOS photodiode is about 40%-70%.

The maximum possible responsivity varies with photon energy. For $\eta = 1$, the maximal responsivity can be simplified as: $R_{\text{max}} = \lambda/1.24$, where λ in [µm]. For the wavelength sensitivity range of CMOS photodiodes 400 nm$<\lambda<$850 nm, the maximum responsivity is in the range 0.32 A/W$<R_{\text{max}}<$0.64 A/W.

The responsivities of a typical Si photodiode, Ge photodiode and InGaAsP photodiode as a function of wavelengths are illustrated in Figure 2.11. The maximum responsivity is also depicted on the same figure ($\eta = 1$).

In the short-wavelength region (λ = 400 nm), the value of R_{max} decreases more rapidly than λ; this is caused by increased surface recombination for the shallow absorbtion depth. For large wavelengths ($\lambda>$850 nm) the responsivity of the CMOS photodiodes also declines; minority carriers are generated deep in the substrate and they are recombined with majority carriers. The highest responsivity is for 600 nm$<\lambda<$800 nm since the carriers are generated closer to the photodiode junctions.

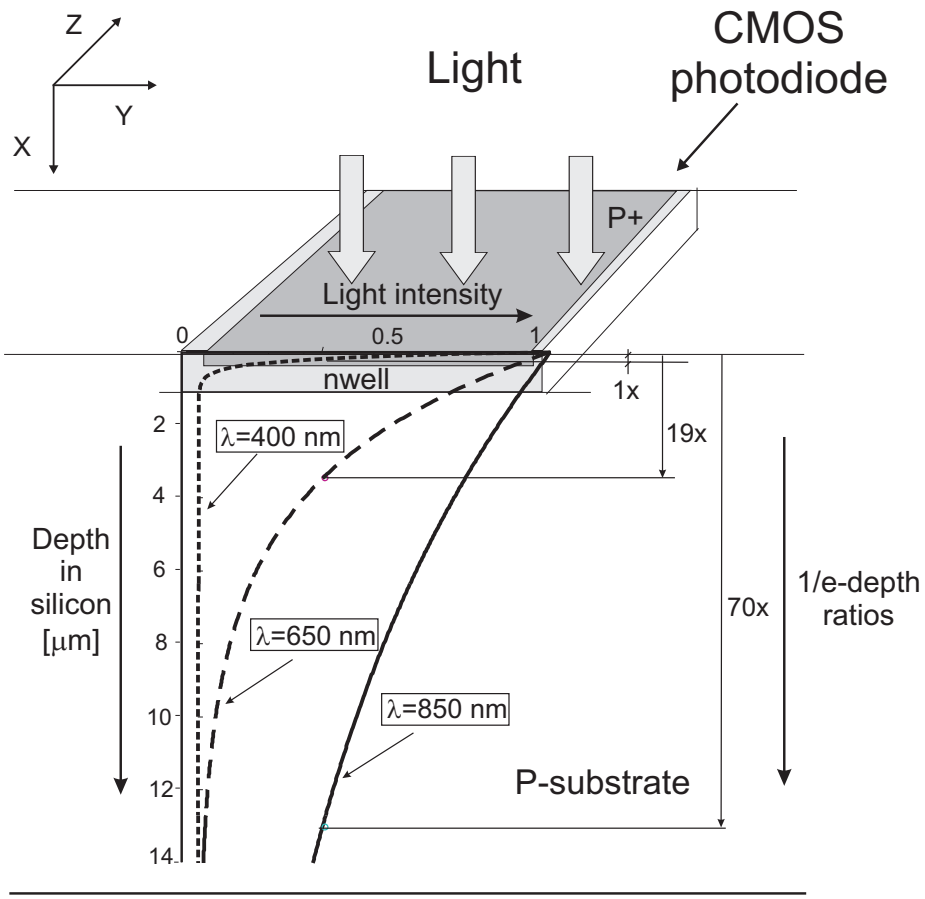


Figure 2.10: The absorption of light inside photodiode in standard CMOS technology. The difference between 1/e-absorption depth among λ = 400, 650 and 850 nm) is large; There is a causal relation between the photodiode responsivity and the bandwidth.

Figure 2.11 shows that silicon photodiodes are not useful in the longer wavelength region λ>950 nm. Other materials have the advantage of a smaller bandgap and higher mobility providing thus higher responsivity and higher bandwidths. However, silicon photodiodes can be integrated with the rest of the electronic circuitry which provides low-cost solution for high-speed optical detection.

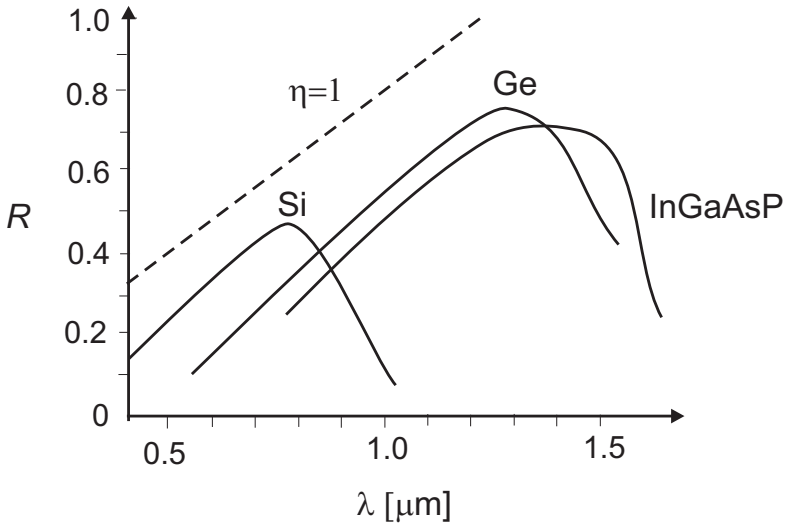


Figure 2.11: Responsivity of a Si photodiode, a Ge photodiode and a InGaAs photodiode as a function of the wavelength

2.6 High-speed optical receivers in CMOS for $\lambda = 850$ nm-literature overview

This section presents a brief overview of high-speed optical receivers in CMOS technology reported in the literature for $\lambda = 850$ nm. Only a few solutions for optical receivers are presented in standard CMOS (data-rates up to 700 Mb/s). The rest are designed using modified CMOS technology and high-voltage solutions (data-rates up to 1 Gb/s).

2.6.1 Using standard CMOS technology

A high-speed optical detection is typically achieved in two manners: the first one is a using "smart" photodiode design, which includes technology exploration and the layout design. The second one is using standard photodiodes with insufficient speed but doing signal processing afterwards in the subsequent electronic circuitry.

CMOS technology with feature size of 1µm

In [16], a data-rate of 622 Mb/s is achieved in 1-µm CMOS technology with a voltage supply of 5 V; the input wavelength is $\lambda = 850$ nm. The reported sensitivity of the detector is -15.3 dBm for a bit error rate (BER) of 10^{-9}. As comparison, the specified sensitivity for Gigabit Ethernet Standard is -17 dBm for the same BER, [17]. Chapter 6 of this thesis describes that using this technology with 0 V supply voltage, the calculated bandwidth of the photodiode is about 100 MHz. With the supply voltage included (5 V), the depletion region area increases almost three times resulting in the increased photodiode bandwidth. In [18] it was shown that the maximum bit-rate is typically beyond the photodiode bandwidth.

Very important features of 1 µm CMOS processes compared to a 0.18 µm CMOS processes, that strongly determine the photodiode bandwidth and the responsivity are:

1. the depth of the nwell is 3-4 times larger than in modern CMOS technology (4 µm). The fast diffusion current inside the nwell and fast drift current strongly determines the overall photocurrent. In modern 0.18 µm CMOS technology for example, the overall current is mainly determined by the slow substrate current.

For $\lambda = 850$ nm, a large portion of light (1/3) is absorbed in nwells, in comparison with newer CMOS technologies where most of the light is absorbed inside the substrate. The fast diffusion response inside the nwell dominates here more to the overall photocurrent. As a result, a photodiode bandwidth is larger.

2. the supply voltage is almost three times higher (5V/1.8V); as a result the depletion region width (depth) is about 50% higher.

3. the disadvantage of 1 µm CMOS is that it is hard to implement electronic circuits in the GHz range.

Together with the depletion region that has a couple of µm depth inside the epi-layer, the amount of the carriers that are generated deep in the substrate is 5 times lower than in modern CMOS technologies[3]. For comparison, the calculated photodiode bandwidth for a modern CMOS process (0.18 µm) is only 5 MHz for $\lambda = 850$ nm (see chapter 3).

[3]Slow diffusion of the substrate carriers that limit the photodiode bandwidth is tremendously reduced (exponential light absorbtion). This will be discussed in detail in chapter 3.

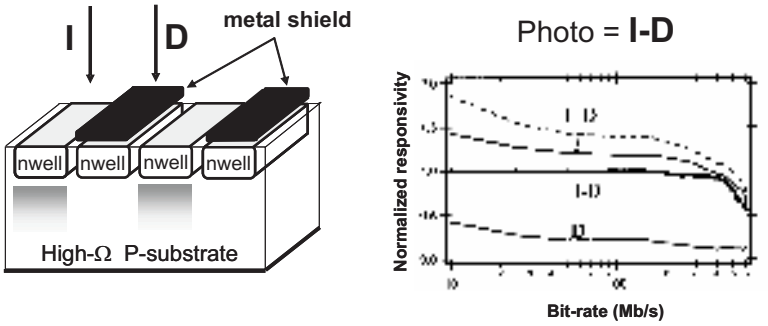


Figure 2.12: Spatially modulated light detector.

SML detector exploiting layout design

One solution in standard 0.25 µm CMOS technology where 700 Mb/s data-rate is achieved is presented in [18],[19]. The effect of the slowly diffusing carriers cancelled by subtracting two diode responses: immediate and deferred diode responses; this however results in lower responsivity and hence lower sensitivity.

The principle of the SML-detector allows one to cancel the effect of the substrate carriers with the cost of lower responsivity. The SML-detector consists of a row of rectangular p-n junctions (fingers) alternatingly covered and non-covered with a light blocking material, as shown in Figure 2.12. The masked fingers connected together form the *deferred* (D) detector. The other fingers connected together form the *immediate* (I) detector.

The slow tail in the time-response of both detectors is very similar, since approximately the same number of the substrate carriers diffuse towards the detectors. The fast total photodiode response is achieved by subtraction of the two diode responses. This however results in lower responsivity (about 75% of the input signal is lost) and hence lower sensitivity. For 300 Mb/s data-rate and BER=10^{-9} the reported sensitivity was -18 dBm. The detector responsivity for 700 Mb/s [19] was not reported; typically the optical power of the input signal is even higher since the noise in the circuit is increased for higher speeds.

2.6.2 CMOS technology modification

Very high-resistance substrate

A solution for 1 Gb/s optical detection is presented in [20]. An integrated receiver is designed in NMOS technology with a special high-resistive substrate

which behaves as a diode intrinsic (I) region. This PIN photodiode is used as a detector designed using n+ and p+ layers inside high-resistive n-substrate. A large intrinsic region ensures both the high speed and the high quantum efficiency of 82%. However, the supply voltage is -32 V. This is unrealistic biasing in modern CMOS processes where typical supply voltage is around 1 V.

Buried oxide layer

In order to increase the photodiode bandwidth, the dominant slow substrate diffusion current [18] can be cancelled by introducing an *buried oxide layer*. The working principle is similar with silicon-on-isolator (SOI) photodetectors. The biggest disadvantage of this technique is a reduced responsivity. A large portion of the excess carriers generated in the substrate do not contribute to the overall photocurrent. In [21], a bandwidth of 1 GHz is reported with the cost of very low[4] responsivity of 0.04-0.09 A/W, corresponding to the sensitivity of 2 dBm to -5 dBm. As a result, the input optical power should be at least 13 dB higher than required in Gigabit Ethernet Standard [17].

References

[1] G. P. Agrawal, N.K. Dutta: "*Long-Wavelength Semiconductor Lasers*", Van Nostrand Reinhold, New York, 1986.

[2] G. P. Agrawal: "*Fiber-Optic Communication Systems*", John Wiley and Sons, New York, 1997.

[3] R. Goff: "*Fiber Optics Reference Guide*", Focal Press, Third Edition, 2002.

[4] David A. B. Miller: "Physical reasons for Optical interconnection", *Int. J. Optoelectronics* 11, pp. 155-168, 1997.

[5] David A. B. Miller and H. M. Ozaktas:" Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture", *Journal of parallel Distrib. Comput.*, vol 41, p4242, 1997.

[6] David A. B. Miller: "Rationale and challanges for optical interconnects to electronic chips", *Proc. IEEE*, vol. 88, 2000, pp.728-749.

[4] Typically, responsivity of the photodiode is > 0.3 A/W corresponding to $>40\%$ quantum efficiency.

[7] H. B. Bakoglu: "*Circuits, interconnections, and packaging for VLSI*", Addison-Wesley, New York, 1990.

[8] W. Etten and J.vd Plaats: "*Fundamentals of Optical Fiber Communications*", Prentice-Hall, 1991.

[9] M. R. Feldman, S. C. Esener, C. C. Guest and S. H. Lee: "Comparison between electrical and optical interconnect based on power and speed consideration", *Appl. Optics*, vol. 27, pp. 1742-1751, 1998.

[10] E. A. de Souza, M. C. Nuss, W. H. Knox and D. A. Miller: "Wavelength-Division Multiplexing with femtosecond pulses", *Optics Letters*, vol. 20, pp. 1166-1168, 1996.

[11] Jeff Hetcht: "*Understanding Fiber Optics*", Prentice Hall, New Jersey, 1999.

[12] A. Bakker: "An Adaptive Cable Equalizer for Serial Digital Video Rates to 400 Mb/s", *ISSCC 1996, Digest of technical papers*, San Francisco, USA, pp. 174-175.

[13] W. Chen: "*Home Networking Basis: Transmission Environments and Wired/Wireless Protocols*", Prentice Hall PTR., July 2003.

[14] Wei Jean Liu, Oscal T.-C. Chen, Li-Kuo Dai and Far-Wen Jih Chung Cheng: "A CMOS Photodiode Model", *2001 IEEE International Workshop on Behavioral Modeling and Simulation*, Santa Rosa, California, October 10-12, 2001.

[15] S. M. Sze: "*Physics of semiconductor devices*", NewYork: WileyInterscience, 2-nd edition, p. 81, 1981.

[16] H. Zimmermann and T. Heide: "A monolithically integrated 1-Gb/s optical receiver in 1-µm CMOS technology", *Photonics technology letters*, vol. 13, July 2001, pp. 711-713.

[17] IEEE 10 Gigabit Ethernet Standard 802.3ae.

[18] D. Coppée, H. J. Stiens, R. A. Vounckx, M. Kuijk: "Calculation of the current response of the spatially modulated light CMOS detectors", *IEEE Transaction Electron Devices*, vol. 48, No. 9, 2001, pp. 1892-1902.

[19] C. Rooman, M. Kuijk, R. Windisch, R. Vounckx, G. Borghs, A. Plichta, M. Brinkmann, K. Gerstner, R. Strack, P. Van Daele, W. Woittiez, R. Baets, P. Heremans: "Inter-chip optical interconnects using imaging fiber bundles and integrated CMOS detectors", *ECOC'01*, pp. 296-297.

[20] C. L. Schow, J. D. Schaub, R. Li, and J. C. Campbell: "A 1 Gbit/s monolithically integrated silicon nmos optical receiver", *IEEE Journal Selected Topics in Quantum Electron.*, vol. 4, Nov.Dec. 1999, pp. 10351039.

[21] M. Ghioni, F. Zappa, V. P. Kesan, and J.Warnock: "VLSI-compatible high speed silicon photodetector for optical datalink applications", *IEEE Trans. Electron. Devices*, vol. 43, July 1996, pp. 10541060.

CHAPTER 3

High-speed integrated photodiodes for $\lambda = 850$ nm

This chapter presents frequency and time domain analyses of photodiode structures designed in standard CMOS technology, for $\lambda = 850$ nm. For clear explanation and illustration of the physical processes inside photodiode, one particular CMOS technology is analyzed in detail. A standard 0.18 μm CMOS is chosen since that was the available fabrication technology. Choosing another CMOS technology does not fundamentaly change the behaviour of the photodiode in general[1]*. The photodiodes are first analyzed as stand-alone detectors i.e. without subsequent electronic circuitry. This allows the analysis of the intrinsic photodiode behavior. The intrinsic behavior is related to the movement (drift and diffusion) of the generated carriers inside the diode. Based on the frequency analysis, the intrinsic (physical) diode bandwidth is determined. In the second part of this chapter, the diode is investigated as an "in-circuit" element, integrated together with the subsequent electronics. The electrical bandwidth of the photodiode is determined by the diode capacitance and the input impedance of the subsequent amplifier. These two bandwidths determine the total diode bandwidth which is by approximation the lower of the two. Further, the influence*

[1]The impact of the technology on photodiode behavior is discussed in detail in chapter 6.

of the diode layout (nwell, n+, p+ finger sizes) in general, on the intrinsic, the extrinsic and the total bandwidth are investigated; as an example, two nwell sizes of the finger diode structure are analyzed: the minimal nwell width (2 μm) and the nwell width much larger than minimal (10 μm).

3.1 Introduction

Depending on the wavelength of the input optical signal (400 nm≤λ≤850 nm), there are several applications for optical detectors:

• $\lambda = 850$ nm: 10 Gigabit/sec Fiber Ethernet (standard 802.3ae, [4]), short-haul communication (chip-to-chip, board-to-board), high-speed opto-couplers [1].

• $\lambda = 780$ nm: CD players and recorders

• $\lambda = 650$ nm: DVD players and recorders

• $\lambda = 400$ nm: DVD - blue ray disc

• 400 nm≤λ≤700 nm: CMOS image sensors This chapter shows that the bandwidth of the integrated CMOS photodetectors is wavelength dependent, structure dependent and layout dependent. It is important to notice that the technology used in thesis is a standard CMOS i.e. there are no technology modifications. *The depth* of the photodiode regions where light is absorbed is related to the photodiode structure; in this chapter various photodiode structures are studied in detail:

• nwell/p-substrate

• n+/p-substrate

• p+/nwell/p-substrate

• p+/nwell The size of the lateral depletion region in photodiodes depends on the well-technology. Thereby, the total depletion region contribution to the overall photocurrent is also different. There are two well technologies that will be analyzed in this chapter:

• *twin-well with adjoined wells* and

• *triple-well with separate wells* technology. For λ=600-850 nm, the light penetration depth is larger than 6 μm (90% of the absorbed light, **[1]**. For modern CMOS technologies, for example with 0.18 μm feature size, the deepest junctions are located close to the photodiode surface (typically 1-2 μm away from the surface). Bellow the junction, there is a substrate where a large portion of the light is absorbed. Thereby, the substrate current response is very important

in the total diode response. Two different kinds of p-substrate are analyzed:

- high-resistance substrate
- low-resistance substrate

The width of the photodiode regions located close to the surface (nwell, n+, depletion regions) can in general be optimized for the best photodiode performance (maximal bandwidth and responsivity). The diode can comprise a number of nwells, n+ fingers, or it can be designed as a single photodiode i.e. with the maximal nwell/n+ width. This is illustrated in Figure 3.1. The influence of the nwell/n+/p+ geometry on photodiode bandwidth will be calculated in general. Two different geometries will be discussed in detail:

- minimal nwell/n+ width L_{ymin}. Typically for standard CMOS, a minimal width is twice the nwell/n+ depth. For 0.18 µm, L_{ymin}=2 µm.
- nwell/n+ width much larger than the nwell/n+ depth: L_y=10 µm.

where L_y is the width of the nwell/n+/p+ regions. The designed photodiodes should achieve maximal bandwidth while still satisfying the required responsivity[2].

A typically used figure-of-merit for photodetectors is the responsivity-bandwidth product [8]. For easier comparison among photodiodes, the following assumptions are taken:

∗ fingered photodiodes are considered in general, to have the nwell/n+ stripes with different sizes. The nwell/n+ regions are placed at *minimal distance* defined by the technology [3]. In this manner, the junction area is maximized for the fixed width of the stripes.

∗ the active (light sensitive) area of all diodes is identical. The active area corresponds to the illuminated silicon area (with no light covers). Therefore, the absorbed input optical power is the same, as well as the maximal possible responsivity for all structures. As a result, the photodiodes performance is compared using their *bandwidth*.

∗ for simplicity of calculation, the photocurrent generated inside pwells that are chosen to have always the minimum size, will not be calculated separately. Instead, the calculated current inside nwells is taken to be larger for the ratio between the wells: $I_{\mathrm{nwell}} = I_{\mathrm{nwell}}(1 + \mathrm{nwell}_{\mathrm{size}})/\mathrm{pwell}_{\mathrm{size}}$.

∗ all junctions in photodiodes are step-junctions

[2]The photocurrent value is related to the diode responsivity as well as the amplification of the subsequent electronic circuitry [8].

[3]Larger distances between nwells decrease a carrier-gradient of the excess carriers inside p-regions i.e. decrease the diffusion speed of these carriers and limits the total diode bandwidth.

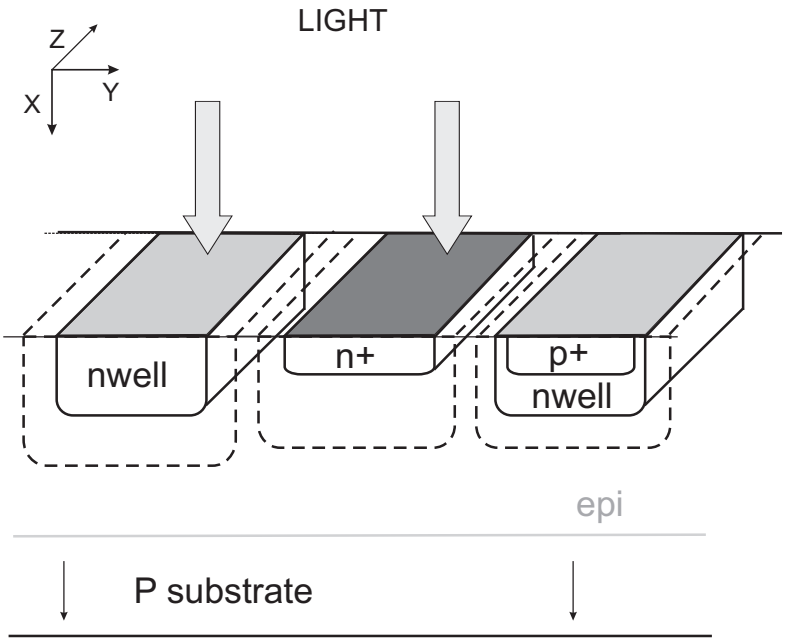


Figure 3.1: Photodiode structures in standard CMOS technology.

∗ the diode parameters such as doping concentrations, carrier lifetime etc. are taken from the standard Philips 0.18 µm CMOS process.

∗ nwell/p-substrate photodiode with adjoined wells and low-resistance substrate is the reference for the other analyzed photodiode structures. For easier comparison of the photodiode performances, their frequency response is normalized with a DC photocurrent density (J_{DC}) of the reference photodiode.

This chapter is organized as follows. In the first part of the chapter, a photodiode is analyzed as a stand-alone device. The intrinsic (physical) behavior of CMOS photodiodes for the aforementioned diode structures and layouts is analyzed. The analysis uses the calculated frequency and time responses of the diode with Dirac pulse as the input optical signal. Both responses will be thoroughly investigated. The calculations of the drift current profile in the depletion region and the diffusion current profiles in the remaining n- and p-regions are presented. The overall photocurrent is the sum of the drift and the diffusion currents. The frequency behavior of the overall photocurrent gives insight into the maximum intrinsic bandwidth of the photodiode.

In the second part of the chapter, the photodiode is investigated as an "in-

circuit" element, integrated together with the subsequent transimpedance amplifier (TIA). The diode capacitance and TIA's input capacitance together with the TIA's input resistance gives an extrinsic bandwidth. This bandwidth will be here referred to as *electrical* diode bandwidth. For the constant input resistance of the TIA, the larger the diode capacitance the smaller the electrical bandwidth. This capacitance is directly related to the diode layout i.e. related to the nwell size. However, larger nwell size decreases the intrinsic diode bandwidth, which will be presented further in this chapter. This chapter will show the trade-off in the diode layout design for the maximal *total* photodiode bandwidth.

3.2 Bandwidth of photodiodes in CMOS

A main topic of the chapter is the frequency response of photodiodes designed in 0.18 µm CMOS. In this way it is possible to predict the diode behavior in the frequency range of interest. Thus, the most suitable photodiode can be chosen for a certain application prior to the circuit fabrication and testing. In this manner, the total cost of the designed system can typically be minimized.

First, a nwell/p-substrate diode with high-resistive substrate and separate-well technology is analyzed in the frequency domain and in the time domain. The high-resistive substrate is chosen for simplicity of calculation in comparison with low-resistance substrate. After this, the frequency and time responses of nwell/p-substrate photodiode with low-resistance substrate will be calculated. This low-resistance substrate is used for a photodiode fabrication, and for this reason the later diode structure will serve as a reference for the other analyzed photodiode structures.

3.2.1 Intrinsic (physical) bandwidth

The intrinsic diode characteristic is related to the behavior of the excess carriers inside the photodiode. These carriers are moving inside the photodiode either by drift (inside depletion regions) or by diffusion (outside depletion region). In general, the photodiode response is the sum of the drift current I_{drift}, and the diffusion currents I_{diff_k}:

$$I_{\text{int}_{\text{total}}} = \sum_k I_{\text{diff}_k} + I_{\text{drift}}, \text{for} \quad k \in [\text{nwell}, \text{n+}, \text{p+}, \text{p-substrate}]. \tag{3.1}$$

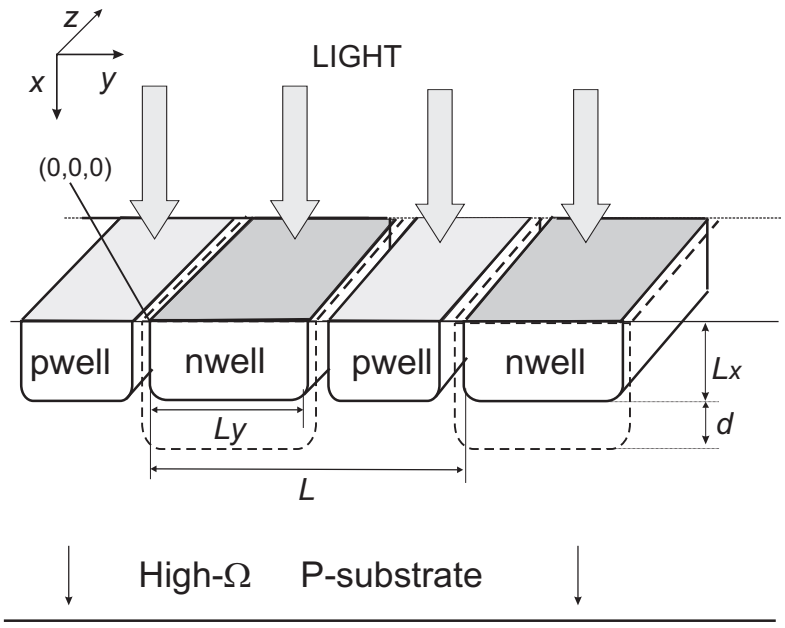


Figure 3.2: Finger nwell/p photodiode structure with *high resistance* substrate and twin-wells in standard CMOS technology.

For better understanding of the total diode response, the frequency response of every current component will be separately presented. The excess carrier profiles and further the currents of the different photodiode regions, are calculated by taking the Laplace transform of the diffusion equations in the time domain, [1]. This provides insight in the frequency domain behavior of CMOS photodiodes.

Nwell/p-substrate photodiode with high-resistance substrate in twin-well technology

This section presents a frequency analysis of the finger nwell/p-substrate photodiode, shown in Figure 3.2. The number of fingers N, is determined by the photodiode dimension in y-direction Y (see Figure 3.2), and the technology (minimal nwell/pwell finger width $L_{\min}$). This number can take every value in the range:

$$N = \left\lfloor \frac{Y}{L} \right\rfloor \quad \text{where} \quad L \in [L_{\min}, Y] \tag{3.2}$$

Nwell diffusion current

The nwell diffusion current shown in equation (3.1), is solved analytically for an impulse light radiation, in two-dimensions using a method similar to that in [1]. From the hole carrier profile, the current density is calculated at the border of the depletion region since the excess holes are collected there as a photocurrent.

The transport of the diffusive holes inside the photodiode is described by the diffusion equation [2]:

$$\frac{\partial p_{\mathrm{n}}(t,x,y)}{\partial t} = D_{\mathrm{p}}\frac{\partial^2 p_{\mathrm{n}}(t,x,y)}{\partial x^2} + D_{\mathrm{p}}\frac{\partial^2 p_{\mathrm{n}}(t,x,y)}{\partial y^2} - \frac{p_{\mathrm{n}}(t,x,y)}{\tau_{\mathrm{p}}} + G(t,x,y) \quad (3.3)$$

where $p_{\mathrm{n}}(t,x,y)$ is the excess minority carrier concentration inside the nwell, D_{p} is the diffusion coefficient of the holes in the n-doped layer and τ_{p} is the minority-carrier lifetime. Using equation 2.12 the hole generation rate $G(t,x,y)$ can be expressed as:

$$G(t,x,y) = \alpha\Phi_0(t)e^{-\alpha x}|_{x\in[0,L_x]} \quad (3.4)$$

where α is given in equation (2.10), and Φ_0 in equation (2.13). The two-dimensional (x,y) calculation of the hole-profile is carried out because the depth of the nwell region is comparable with its width. There are four boundary conditions for the hole-profile: two in the x-direction and two in the y-direction, as shown in Figure 3.3.

For the *first* boundary condition, the photodiode surface is assumed to be reflective i.e. the normal component of the gradient of the carrier density is zero. This is because the surface recombination process is far slower than a 1/MHz or a 1/GHz time-responses considered in this thesis. On the *other three* boundaries with depletion region, the electron densities are assumed to be zero:

$$\frac{\partial p_{\mathrm{n}}}{\partial x}|_{x=0} = 0 \quad p_{\mathrm{n}}|_{x=L_x} = 0 \quad (3.5)$$

$$p_{\mathrm{n}}|_{y=0} = 0 \quad p_{\mathrm{n}}|_{y=L_y} = 0 \quad (3.6)$$

The equation 3.3 is partial differential equation in time (t) and space domain (x,y). The carrier profile p_{p} is calculated first by taking the Laplace transform of the diffusion equation (3.3). In this manner the carrier profile is transformed in the frequency domain s. The obtained diffusion equation will be further

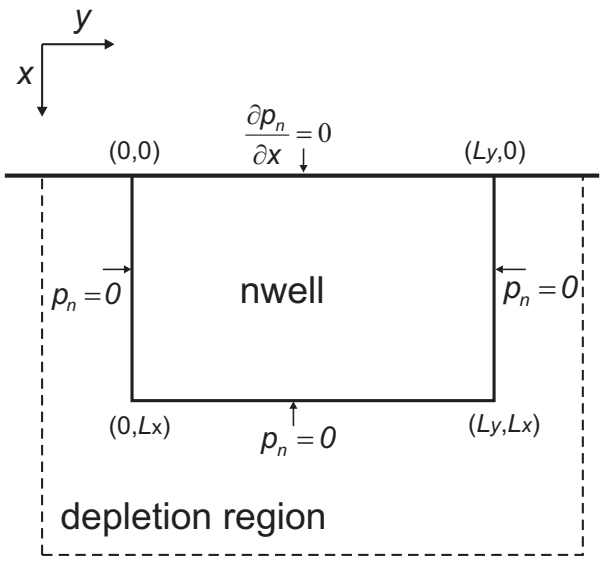


Figure 3.3: The boundary conditions for hole densities inside nwell.

solved in the space domain (x, y). In order to solve this equation analytically, the most suitable method is to use Discrete Fourier series in the space domain. This is certainly valid for y-direction where nwell/pwell represents light/no-light periodic function. The carrier distribution function $p_{\rm p}$ and the carrier generation function G are rewritten as a product of two Fourier series, one of a square wave in the x-direction (with index n) and the other of a square wave in the y-direction (with index m). Each of these decomposed terms of $G(s)$ drives one of the terms of decomposed of $p_{\rm n}$:

$$\begin{aligned}\frac{p_{\rm n}(s,x,y)}{\Phi_0(s)} &= \frac{16\alpha L_y L_{\rm p}^2}{l\pi D_{\rm p}} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{(2n-1)\pi(-1)^{\frac{n-1}{2}} e^{-\alpha L_x} + 2\alpha L_x}{4\alpha^2 L_x^2 + (2n-1)^2\pi^2} \\ &\times \frac{\sin(\frac{(2m-1)\pi y}{L_y})\cos(\frac{(2n-1)\pi x}{2L_x})}{(2m-1)\left(\frac{(2n-1)^2\pi^2 L_{\rm p}^2}{L_y^2} + \frac{(2m-1)^2\pi^2 L_{\rm p}^2}{4L_x^2} + s + 1\right)} \end{aligned} \quad (3.7)$$

The Fourier series is composed of odd sine and cosine terms. The even Fourier terms (full sines and cosines) do not contribute to the nwell current response

because the total area below the curves is zero. The odd sine and cosine terms are truncated sine and cosine functions, and the area below these curves is non-zero. This is illustrated in Figure 3.4.

Once the carrier profile is calculated, the hole-current frequency response can be determined for each set of indexes n and m (from equation (3.7)). The total contributed current is the integral of the current through the two side-walls and the bottom layers. The final expression for the nwell diffusion current is:

$$\begin{aligned}\frac{J_{\text{nwell}}(s)}{\Phi_0(s)} &= 32\frac{eL_{\text{p}}^2\alpha}{l\pi^2}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{(2n-1)\pi e^{-\alpha L_x}+(-1)^{\frac{(2n-1)-1}{2}}\alpha L_x}{4\alpha^2L_x^2+(2n-1)^2\pi^2} \\ &\times \frac{\frac{2L_x}{L_y}\frac{1}{2n-1}+\frac{L_y}{2L_x}\frac{2n-1}{(2m-1)^2}}{\frac{(2n-1)^2\pi^2L_{\text{p}}^2}{4L_x^2}+\frac{(2m-1)^2\pi^2L_{\text{p}}^2}{L_y^2}+1+s\tau_{\text{p}}}\end{aligned} \tag{3.8}$$

The total nwell response is the double sum of the n and m one-pole responses with the wavelength-dependent amplitudes. The nwell amplitude response is illustrated in Figure 3.5 and the nwell phase response in Figure 3.6. The total response is shown in Figure 3.7. The former Figure illustrates that the amplitude of higher Fourier terms decreases with n and m while the poles are placed further on the frequency axes. The sum of all components gives the total nwell response with an unusually *slow roll-of* (∼10 dB/decade). This is an feature of the nwell diffusion process that will be explored more in chapter 4 of this thesis.

The bandwidth of the nwell diffusion current is determined using equation (3.8) and introducing certain simplifications as given in [1]. The slowest and the dominant contribution to the nwell current corresponds to the case $n = m = 1$. The amplitude of the other contributions decrease quadratic with n and m. From equation (3.8), the -3 dB bandwidth frequency can be approximated with:

$$f_{3\text{dB}} \simeq \left(\frac{\lambda}{\lambda_{850}}\right)^{1/3}\frac{\pi D_{\text{p}}}{2}\left(\left(\frac{1}{2L_x}\right)^2+\left(\frac{1}{L_y}\right)^2+\left(\frac{1}{L_{\text{p}}}\right)^2\right) \tag{3.9}$$

The only difference in the equation above in comparison with the bandwidth equation given in [1] is that wavelength (λ) dependence is introduced. Equation (3.9) shows that the bandwidth of the nwell current is directly proportional to the diffusion constant of holes D_{p}, which is further proportional to the mobility of holes [15]. The higher the mobility, the faster the holes can reach the edges

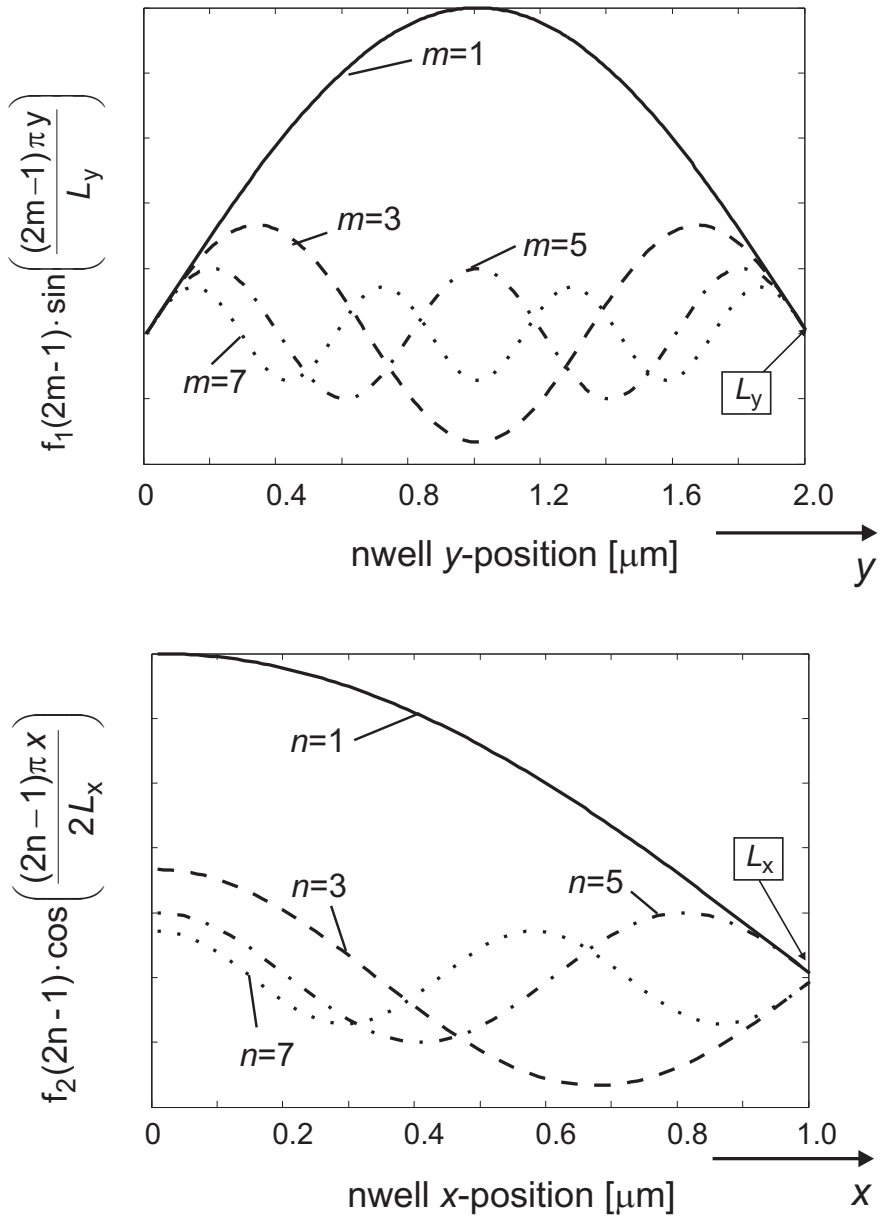


Figure 3.4: Minority carrier profile inside nwell in a) x-direction and b) y directions.

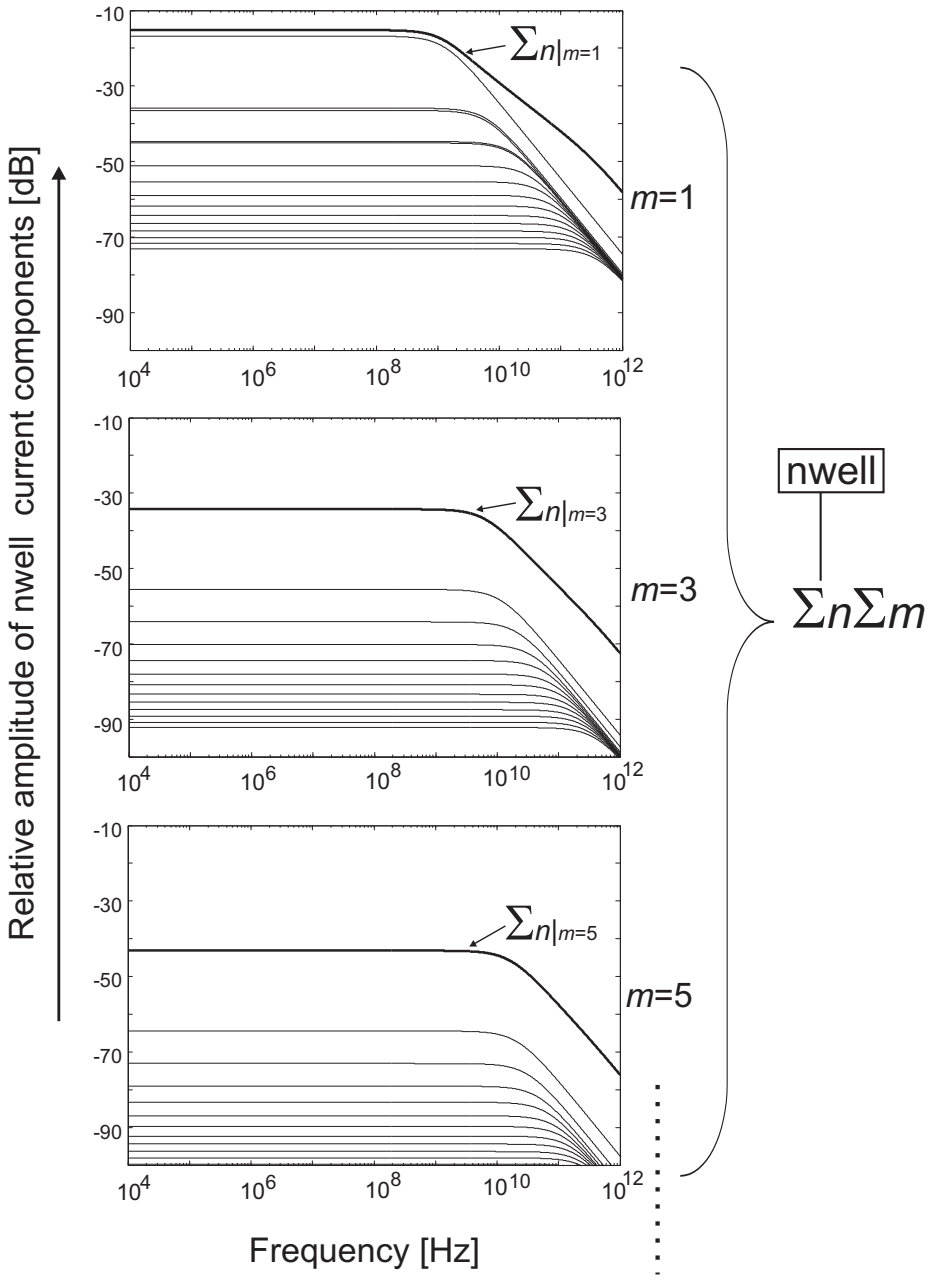


Figure 3.5: Amplitude of the double Fourier series of the nwell diffusion current for the vertical and the lateral nwell direction. A total nwell diffusion current is the sum of all terms with indices m and n, respectively.

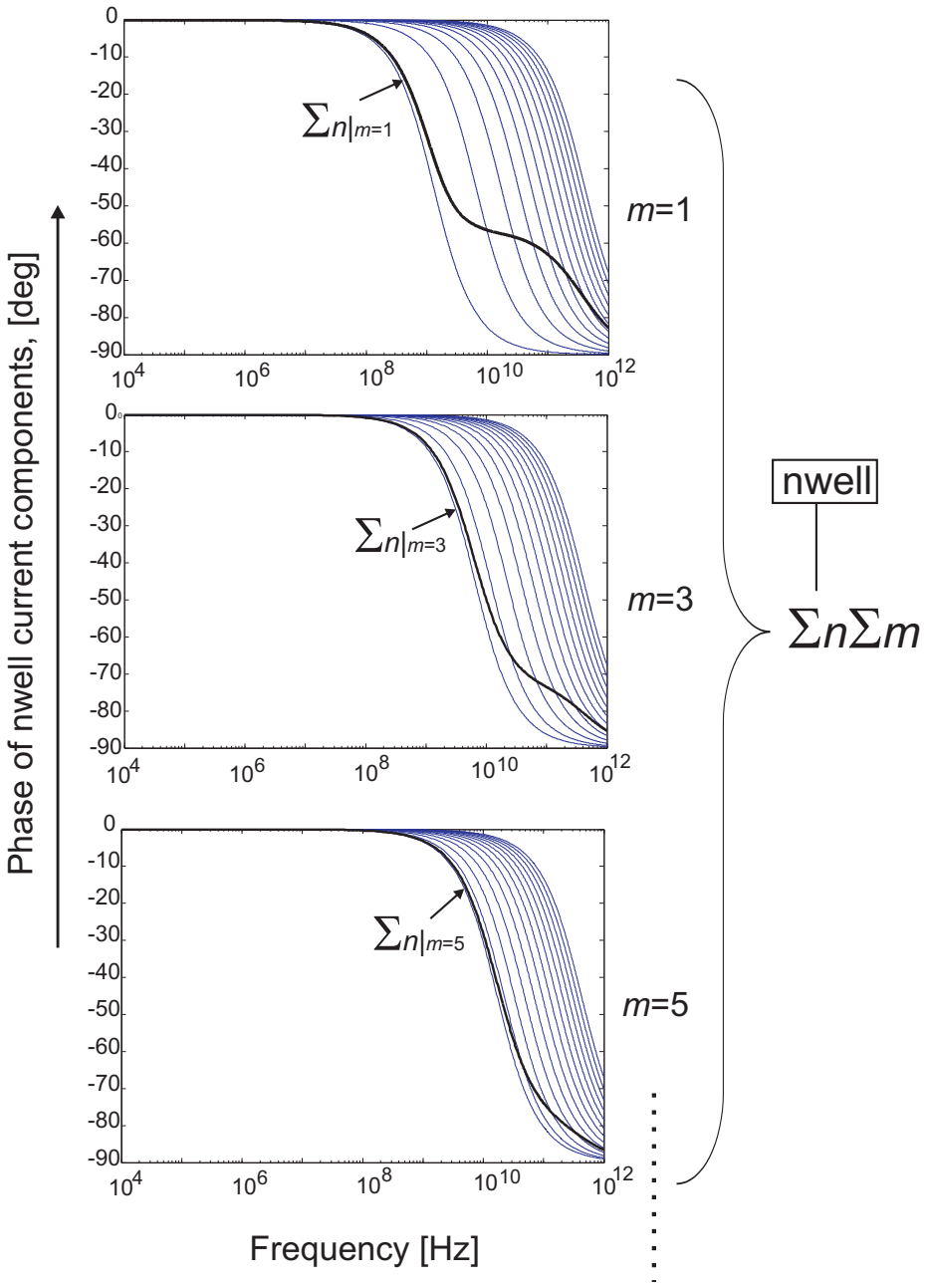


Figure 3.6: Phase of the double Fourier series of the nwell diffusion current for the vertical and the lateral nwell direction. A total nwell diffusion current is the sum of all terms with indices m and n, respectively.

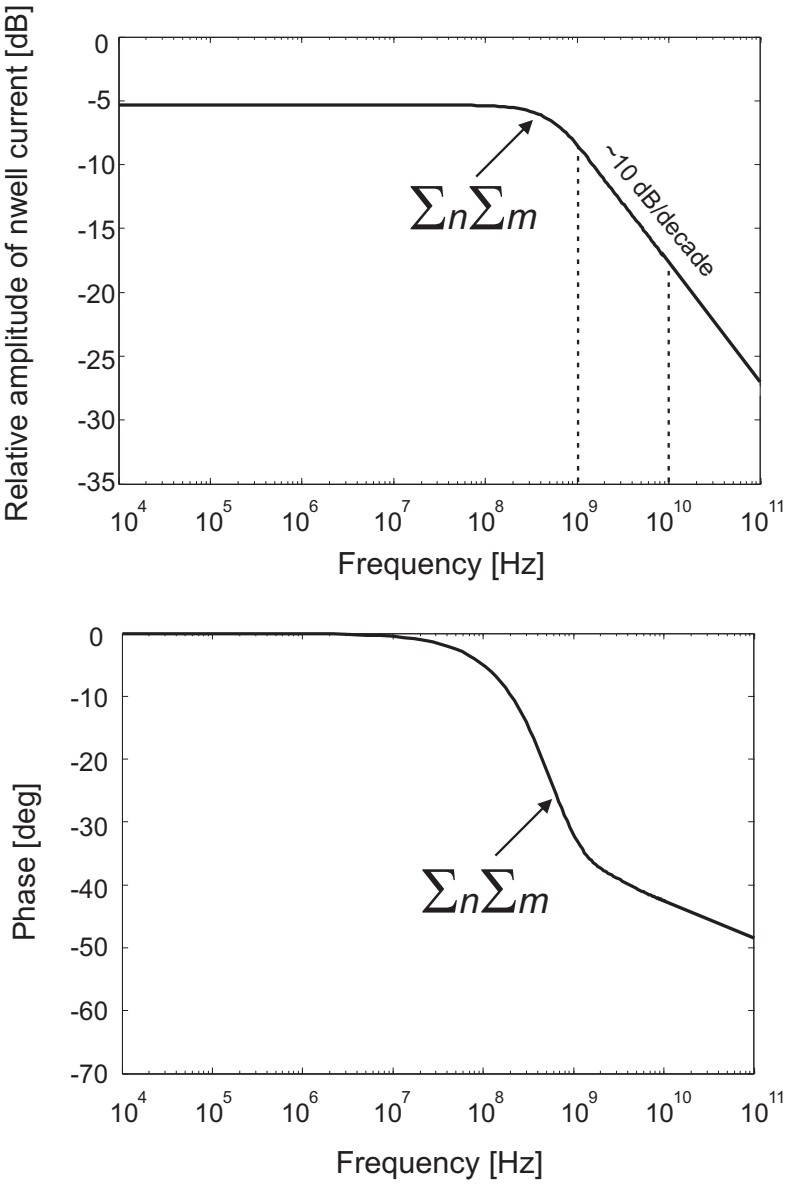


Figure 3.7: The total amplitude and phase of the nwell diffusion current.

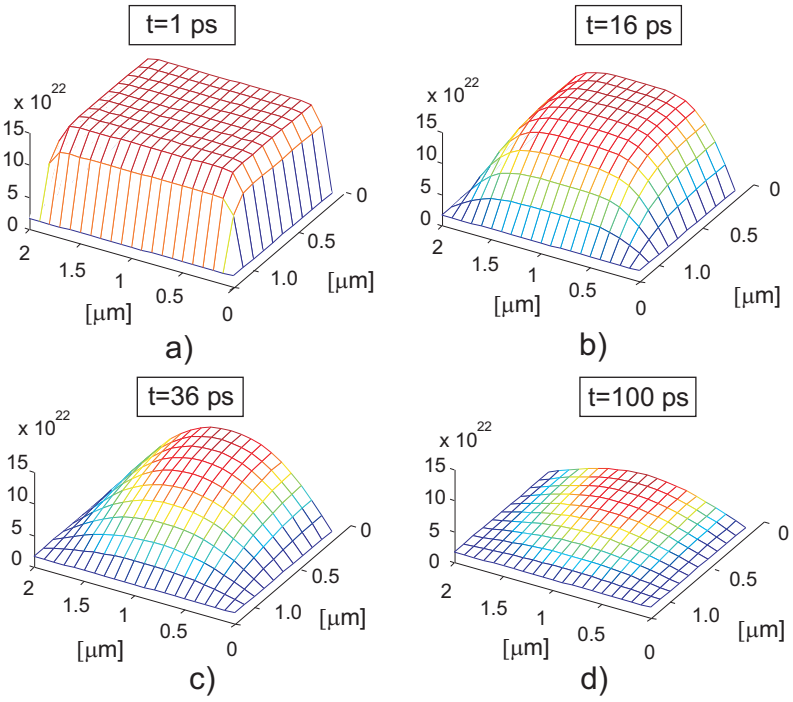


Figure 3.8: The calculated hole diffusion profile inside nwell with 2µm size, under incident light pulse (10 ps pulse-width). This profile is calculated after 1 ps, 16 ps, 36 ps, 100 ps.

of the depletion region. Thereby, the response is faster i.e. nwell diffusion bandwidth is higher.

The terms in the bracket concerning the depth L_x and the width L_y can be explained using Figure 3.8 and Figure 3.9. From the calculated frequency response, the time response of the nwell region is calculated using the inverse Laplace transform of equation (3.8).

The holes diffuse towards junctions due to the gradient of the hole concentration. The gradient is maximum over the minimum distance to the junctions. Thereby, the holes choose "minimal paths" towards the junctions. If L_x=$2L_y$ the hole in the top-middle position of the nwell can diffuse left, right or down with the equal probability since they are all "minimal paths". The nwell size in y-direction is twice the size in x-direction and for that reason the first bracket term is with $1/(2L_x)$.

The third term inside the brackets corresponds to the diffusion length of the

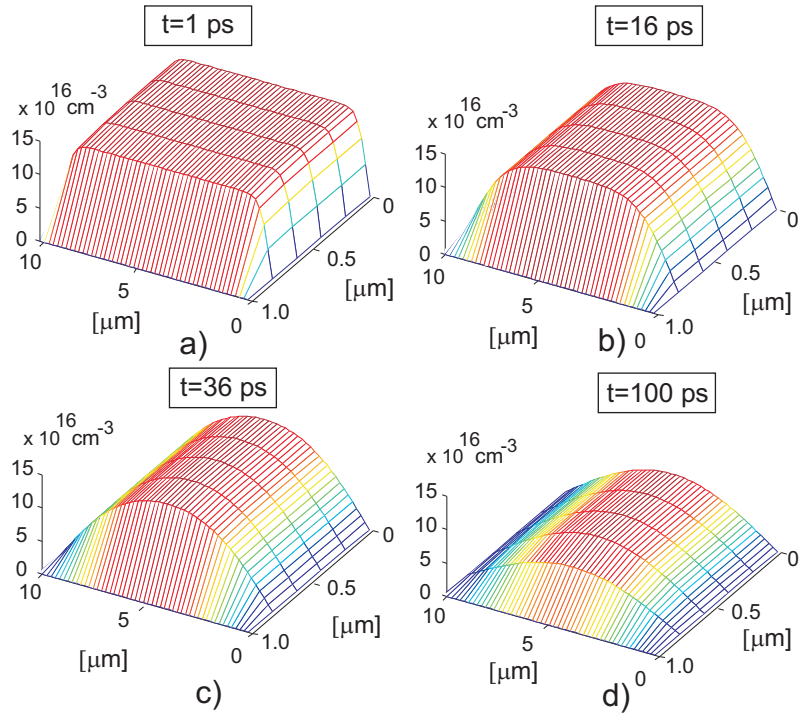


Figure 3.9: The calculated hole diffusion profile inside nwell with 10 µm size, under incident light pulse (10 ps pulse-width). The lateral nwell dimension is obviously less important for the diffusion process.

holes L_p. Typically in CMOS technology, the diffusion length is much larger than the minimum side of the nwell and its contribution in the equation (3.9) is small. It is clear that both the *layout* (lateral size) and the *technology* (related to the nwell depth and the doping concentration) are very important and determine the nwell diffusion bandwidth.

High-resistance substrate current

The second photocurrent component that is analyzed in this chapter is the substrate current. The frequency and the time responses of this current are presented further in this section.

A diffusion process of electrons generated in substrate below the depletion region areas is different in comparison with the diffusion of electrons generated outside these areas. However, the designed photodiodes in this thesis have the

minimal technological distance between the nwell/n+ and for the simplicity of the substrate current calculation, it is assumed that the depletion region is everywhere above the substrate; thus, the diffusion process is the same for all minority electrons.

The substrate current results from the excess minority electrons that diffuse either towards upper allocated junctions (with the nwell or n+) or deeper in the substrate where they are recombined.

The substrate current frequency response on a Dirac light pulse is solved using one-dimensional (vertical) diffusion equation [15]:

$$\frac{\partial n_p(t,x)}{\partial t} = D_{\mathrm{n}} \frac{\partial^2 n_{\mathrm{p}}(t,x)}{\partial x^2} + \frac{n_{\mathrm{p}}(t,x)}{\tau_{\mathrm{n}}} + G(t,x) \tag{3.10}$$

where $n_p(x,t)$ is the excess electron concentration inside the substrate, D_{n} is the diffusion coefficient of the electrons and τ_n being the minority-carrier lifetime. The electron generation rate $G(t,x)$ using equation (2.12) is:

$$G(t,x) = \alpha \Phi_0(t) e^{-\alpha(L_x+d)} e^{-\alpha x}|_{x\in[0,L_{\mathrm{fnt}}]} \tag{3.11}$$

where L_{fnt} is the finite substrate depth where the light is almost completely absorbed (99%), and d is the depletion region depth (see Figure 3.2).

There are two boundary conditions for the minority electrons in the x-direction; the first boundary is at the substrate top and the second at L_{fnt}. Both boundary conditions are taken to be zero since the carriers are either removed by the junctions or they are recombined:

$$n_{\mathrm{p}}|_{x=0} = 0 \tag{3.12}$$

$$n_{\mathrm{p}}|_{x=L_{\mathrm{fnt}}} = 0 \tag{3.13}$$

For $\lambda = 850$ nm the chosen bottom boundary $L_{\mathrm{fnt}} = 60$ µm, where 99% of the light is absorbed. Larger values for L_{fnt} leads to a slower response [18].

In order to solve equation (3.10), the Laplace transform of the equation is taken first. This is similar to the solving procedure given in [1]. The carrier profile $n_{\mathrm{p}}(t,x)$ is transformed to $n_{\mathrm{p}}(s,x)$ in the frequency domain. The carrier profile function n_{p} and the carrier generation function $G(s)$ are rewritten as the product of a Fourier series of a square wave in the x-direction (with index n). Each of these decomposed terms of $G(s)$ corresponds to the one of the terms of

decomposed of $n_{\mathrm{p}}(s,x)$. For each set of indexes n a carrier profile is calculated and expressed as:

$$n_{\mathrm{p}}(s,x) = 2\alpha\Phi(s)\pi \sum_{n=1}^{\infty} \frac{n \sin\left(\frac{n\pi x}{L_{\mathrm{fnt}}}\right)}{(\alpha^2 L_{\mathrm{fnt}}^2 + n^2\pi^2) D_{\mathrm{n}} \left(\frac{s}{D_{\mathrm{n}}} + \frac{1}{L_{\mathrm{n}}^2} + \frac{m^2\pi^2}{L^2{}_{\mathrm{fnt}}}\right)} \tag{3.14}$$

The carrier profile is a sum of n-sine signals that differs in amplitude for a factor $n^2/(\alpha^2 L_{\mathrm{fnt}}^2 + n^2\pi^2)$.

Further, the substrate current is determined from these carrier profiles. The total contributed current is the current through the upper depletion region:

$$\begin{aligned} J_{\mathrm{subs}}(s) &= eD_{\mathrm{p}} \frac{\partial n_{\mathrm{p}}(s,x)}{\partial x}|_{x=0} \\ &= \sum_{n=1}^{\infty} \frac{2\alpha e\Phi(s) e^{-\alpha(L_x+d)} n^2\pi^2}{L_{\mathrm{fnt}}(\alpha^2 L^2{}_{\mathrm{fnt}} + n^2\pi^2) \left(\frac{s}{D_{\mathrm{n}}} + \frac{1}{L_{\mathrm{n}}^2} + \frac{n^2\pi^2}{L_{\mathrm{fnt}}^2}\right)} \end{aligned} \tag{3.15}$$

The total substrate response is the sum of n Fourier terms with amplitudes that depend on the wavelength (α). The higher the indic n, the further the pole of the corresponding Fourier component. This is illustrated in Figure 3.10. The sum of all components gives the total substrate response with a slow roll-of (∼10 dB/decade).

The calculation of the substrate current response can be simplified by taking the infinite substrate depth, as given in [1]. This solution corresponds to the worst-case solution i.e. with the lowest substrate-current bandwidth. The calculated photocurrent is expressed as:

$$J_{\mathrm{subs}}(s) = e\alpha L_{\mathrm{n}} e^{-\alpha L_x} \frac{1}{\sqrt{1 + s\tau_{\mathrm{n}}} + \alpha L_{\mathrm{n}}} \tag{3.16}$$

This equation also shows a *low* current response decay with 10 dB/decade.

The inverse Laplace transform of the equation (3.15) is used to calculate the substrate current impulse response as a function of time. The result is illustrated in Figure 3.11. The diffusion process is slow and electrons need time (tens of ns) to reach the junctions located close to the photodiode surface. A certain number of carriers will also diffuse deeper into the substrate and do not contribute to the overall photocurrent.

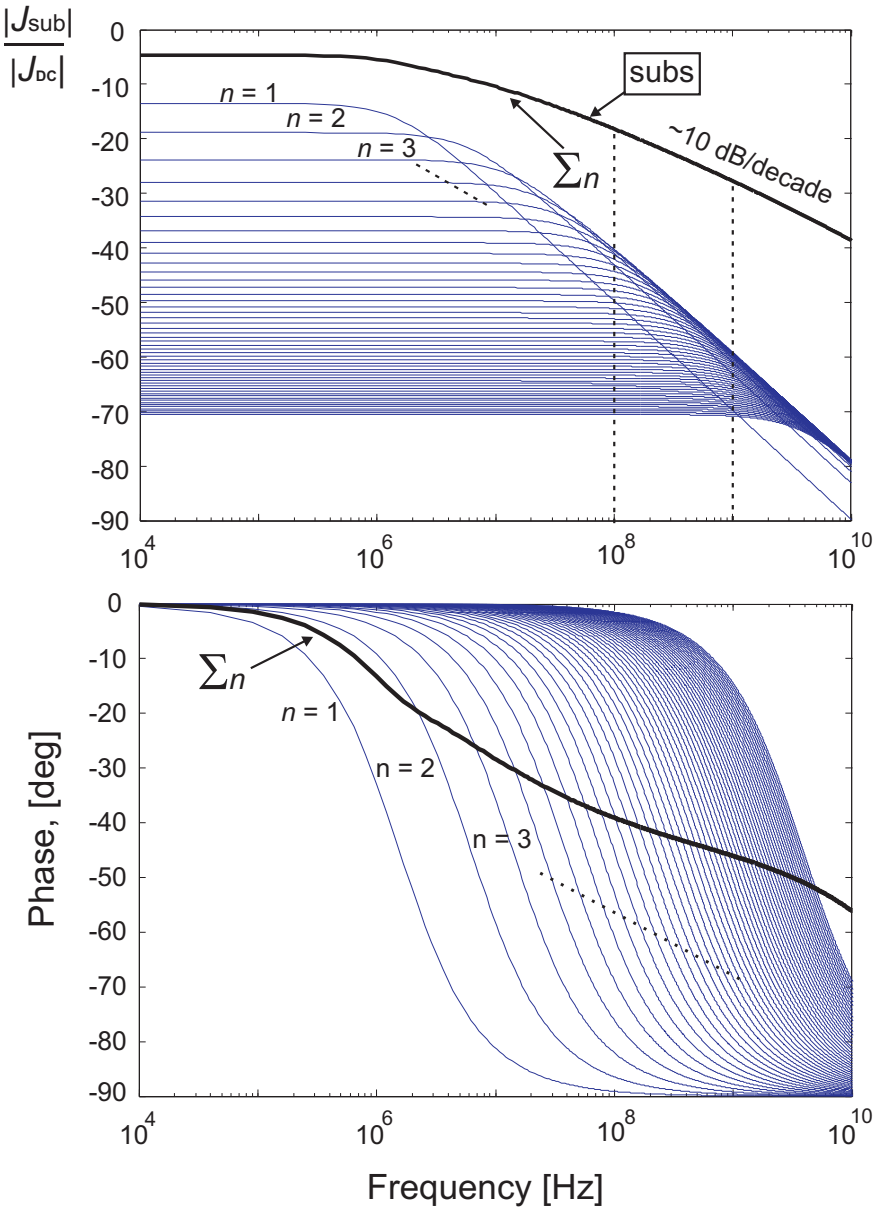


Figure 3.10: Frequency response of the substrate diffusion current. It is a sum of n one-pole sine-Fourier components.

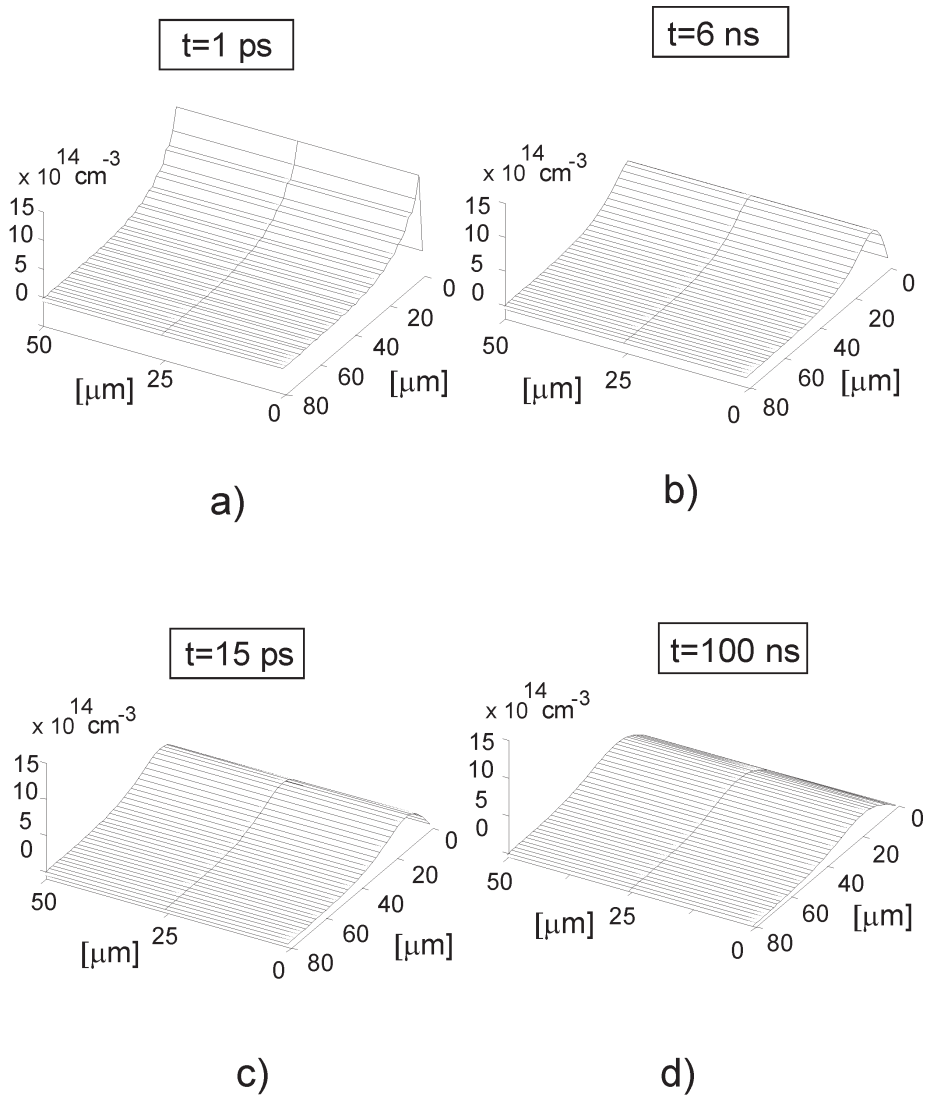


Figure 3.11: The calculated substrate diffusion profile inside p-substrate (depth 80 µm), under incident light pulse (10 ps pulse-width). This profile is calculated after 1 ps, 6 ns, 15 ns, 100 ns.

Depletion region response (drift response)

The third photocurrent component in equation (3.1), is the drift current inside the depletion regions in the vertical and lateral directions. The drift current is directly related to the depletion volume in which carriers are generated:

$$J_{\text{dep}} = \Phi e \left([e^{-\alpha L_x} - e^{-\alpha(L_x+d)}] \frac{A_{\text{total}}}{A_{\text{eff}_{\text{lat}}}} + [1 - e^{-\alpha(L_x)}] \frac{A_{\text{total}}}{A_{\text{eff}_{\text{ver}}}} \right) \tag{3.17}$$

where $A_{\text{eff}_{\text{lat}}}$ and $A_{\text{eff}_{\text{ver}}}$ are the effective lateral and vertical depletion region areas in comparison with the total photodiode area A_{total}. For the twin-wells technology with the adjoined wells, the vertical depletion region is much smaller than the lateral one, so for simplicity of calculations in this thesis it will be neglected. Typically, the drift current bandwidth is much larger than the diffusion current bandwidth; hence, for simpler calculations the drift current is taken to be independent on frequency. This is certainly valid for frequencies up to f_{3dBdrift}=0.4 v_{s}/W, [8], where v_{s} is saturation velocity of charge carriers and W is the depletion region depth. For 0.18 µm CMOS technology, this frequency is f_{3dBdrift}=8 − 10 GHz.

Total photodiode intrinsic characteristics

The sum of all diffusion and drift current components in the previous sections presents the total intrinsic photocurrent response of the nwell/p-substrate diode. Figure 3.12 shows the calculated responses of the two finger nwell/p-substrate diodes. The first response is for minimal nwell width, which is typically twice the nwell depth in CMOS technology; for 0.18 µm CMOS it is 2 µm. The second response is for the nwell width much larger than its depth i.e. for the nwell width of 10 µm. Both responses are calculated for $\lambda = 850$ nm. The values for the parameters in the analytical expressions were directly obtained from the process technology parameters for a fully standard 0.18 µm CMOS process.

For $\lambda = 850$ nm, the substrate current typically dominates the overall photocurrent frequency behavior up to a few hundreds of MHz (illustrated in Figure 3.12). Generated electrons need a few microseconds time to diffuse towards the junctions and limit the substrate current bandwidth (-3 dB) to about 5 MHz. The nwell diffusion current has a larger bandwidth that is mainly determined by the length of the shortest side of the nwell. For the narrow nwell with $L_y = 2$ µm, the shortest sides are both lateral and vertical dimensions. The calculated band-

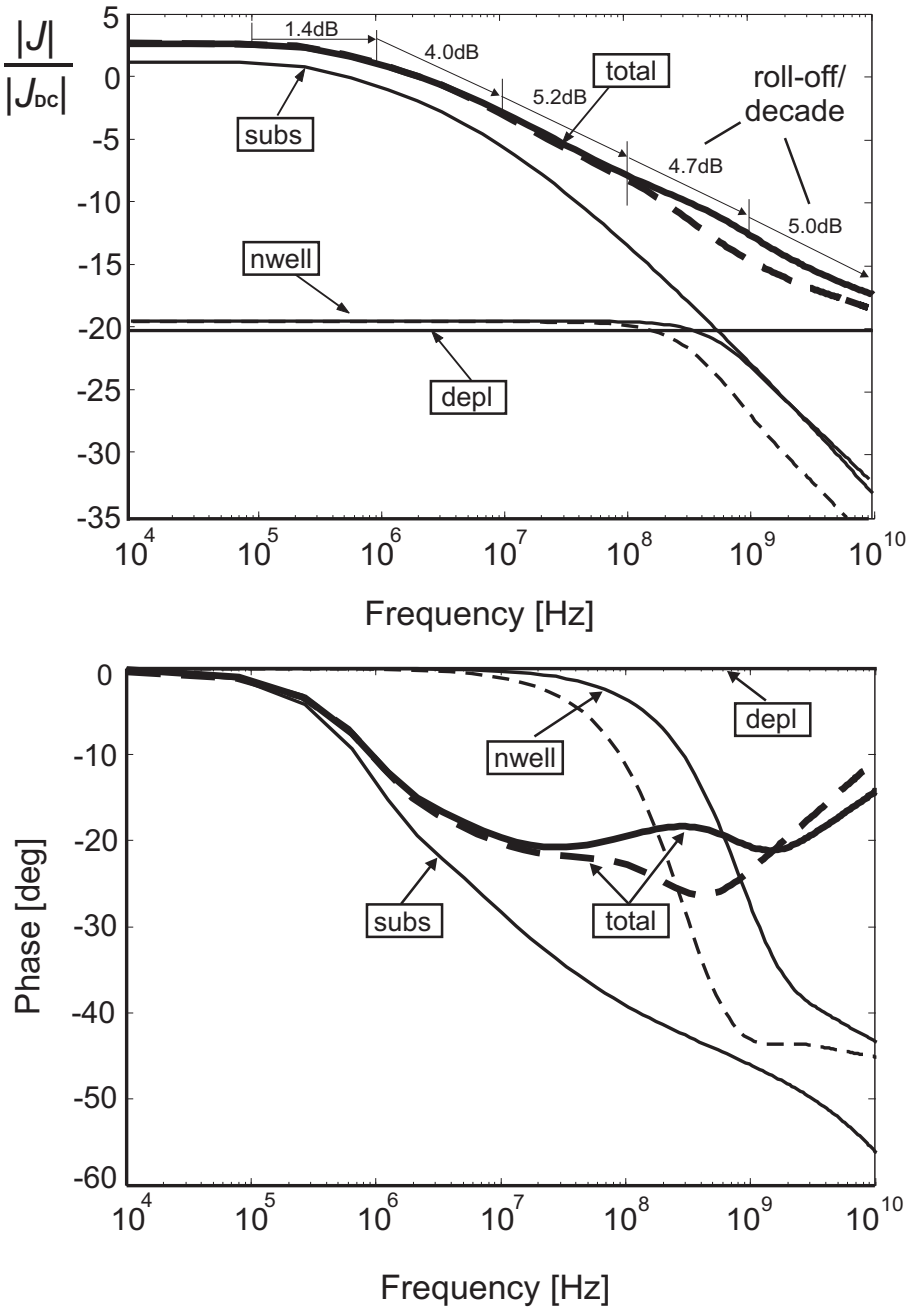


Figure 3.12: The calculated total photocurrent response of nwell/p-substrate photodiode with *high-resistance* substrate in a *twin-well* technology: 2 µm (solid lines) and 10 µm nwell size (dashed lines) for $\lambda = 850$ nm.

width is $f_{3\text{dBnwell}} = 930$ MHz. For wide nwell with $L_y = 10$ µm, the shortest side is the nwell depth only, and the charge gradient is lower than in the previous case. The calculated bandwidth in this case is $f_{3\text{dB}_{\text{nwell}}} = 450$ MHz. Thus, the larger the nwell width L_y, in comparison with its depth L_x ($L_y > 2L_x$), the lower the influence of the nwell-width on its bandwidth. The overall maximal intrinsic bandwidth is 5 MHz. This bandwidth is almost independent on the nwell geometry due to the dominant and size independent substrate current contribution: the fast diffusion response in the nwells and the fast drift response are overshadowed by the large substrate current.

Roll-of in the frequency characteristics

The overall intrinsic photodiode response shows a slow decay due to the combination of the three current components. It follows the dominant component in the frequency domain, as shown in Figure 3.12.

The roll-off of the total photocurrent response in the beginning (around the -3 dB point) follows the one of the substrate current response. For higher decades, the total roll-of is smaller in comparison with the substrate roll-of, due to the larger influence of the fast nwell and the depletion region currents. The maximal roll-off value for the frequencies between the -3dB frequency and the lower GHz range is about 5.7 dB/decade for $L_y = 10$ µm and 5.2 dB/decade for $L_y = 2$ µm, as illustrated in Figure 3.12. In the low-GHz range, the roll-off is lower (about 4.7 dB/decade) and it decreases with the frequency since the "flat" depletion region response dominates the overall photocurrent.

Nwell/p-substrate photodiode with low-resistance substrate

In standard CMOS processes, the circuit designers can typically choose between "high" or "low"-resistance substrate. This section analyses the photodiode frequency behavior of the nwell/p-substrate with a low-resistance substrate illustrated in Figure 3.13.

The only difference between this diode and the previously analyzed photodiode is in the substrate current response. This response is solved again using one-dimensional (vertical) diffusion equation. The two "p" layers are placed at the top of each other (see Figure 3.13) and the movement of the minority electrons in both layers is described with two diffusion equations:

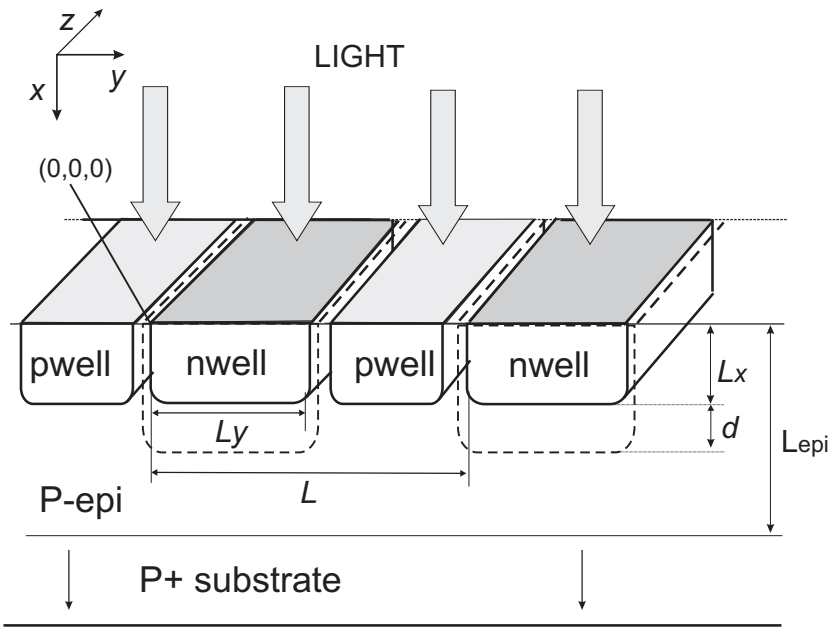


Figure 3.13: Finger nwell/p photodiode structure with *low-resistance* substrate and twin-wells in standard CMOS technology.

$$\begin{aligned} \frac{\partial n_{\mathrm{p1}}}{\partial t} &= D_{\mathrm{n}}\frac{\partial^2 n_{\mathrm{p1}}}{\partial x^2} - \frac{n_{\mathrm{p1}}}{\tau_{\mathrm{p1}}} + G_1(t,x) \\ \frac{\partial n_{\mathrm{p2}}}{\partial t} &= D_{\mathrm{n}}\frac{\partial^2 n_{\mathrm{p2}}}{\partial x^2} - \frac{n_{\mathrm{p2}}}{\tau_{\mathrm{p2}}} + G_2(t,x) \end{aligned} \tag{3.18}$$

where the electron generation rate at $t = 0$, in the top substrate layer $G_1(t,x)$ and in the bottom substrate layer $G_2(t,x)$ can be expressed as:

$$\begin{aligned} G_1(t,x) &= \alpha\Phi_0(t)e^{-\alpha(L_x+d)}e^{-\alpha x)}|_{x\in[0,L_{\mathrm{epi}}]} \\ G_2(t,x) &= \alpha\Phi_0(t)e^{-\alpha L_{\mathrm{epi}}}e^{-\alpha x)}|_{x\in[0,\infty]} \end{aligned} \tag{3.19}$$

where d is the depth of the depletion region, and L_{epi} the depth of the p-epi layer.

In order to calculate the substrate current response in the frequency domain (s), once again the Laplace transform of the diffusion equation (3.18) is taken. Between the two substrate layers there is a boundary condition related to both the current density and the minority carrier concentration [9]. Due to the continuity of currents, the current densities are equal between the two layers:

$$-qD_{\mathrm{p1}}\frac{\partial n_{\mathrm{p1}}(s,x)}{\partial x}|_{x=L_{\mathrm{epi}}} = -qD_{\mathrm{p2}}\frac{\partial n_{\mathrm{p2}}(s,x)}{\partial x}|_{x=L_{\mathrm{epi}}} \tag{3.20}$$

The second boundary condition is related to the continuity of the concentration of the minority carriers:

$$n_{\mathrm{p1}}(s, L_{\mathrm{epi}}) = n_{\mathrm{p2}}(s, L_{\mathrm{epi}}) \tag{3.21}$$

The other two boundary conditions for both electron densities at the bottom of the depletion region, $x = 0$, and at the infinite substrate depth, $x = \infty$, are taken to be zero. The infinitely large substrate is taken in order to avoid long and complex calculations [1].

The electrons generated deep in the low-resistance substrate have a higher probability of recombination than in the high-resistance substrate due to the higher doping concentration. The recombined carriers do not contribute to the overall photocurrent. On one side the photo-responsivity decreases, but on the other side, the diffusion current bandwidth increases [4].

Following the procedure described previously in this section, the total current response is calculated and the result is presented in Figure 3.14. In comparison with high-resistance substrate photodiode, more carriers diffuse towards the substrate bottom resulting in a lower diode DC responsivity. Therefore, the DC current is lower, but the overall bandwidth is higher. The calculated normalized amplitude of the overall photocurrent is 3.5 dB lower but with 2.3 times higher cutoff frequency: 8 MHz. The photodiode geometry has again almost no influence on the bandwidth due to the substrate current domination.

The roll-of in the total frequency response for all decades after -3 dB frequency is **1-2** dB larger (see Figure 3.14) in comparison to low-resistance substrate diodes. The maximal calculated roll-off value for the frequencies between the -3dB frequency and the lower GHz range is about 5.6 dB/decade for L_y=10 µm and 4.7 dB/decade for L_y=2 µm, as illustrated in Figure 3.14.

3.2.2 Comparison between simulations and measurements

A finger nwell/p-substrate photodiode with 2 µm nwell width was fabricated in a standard 0.18 µm twin-well CMOS technology. The chip-micrograph is shown in Figure 3.15[5]. The overall photodiode area is 50×50 µm^2. The nwells

[4] There is no need to wait a long time for this carriers to be collected as a photocurrent.
[5] The technology has five metal layers and one polyscilicon layer available.

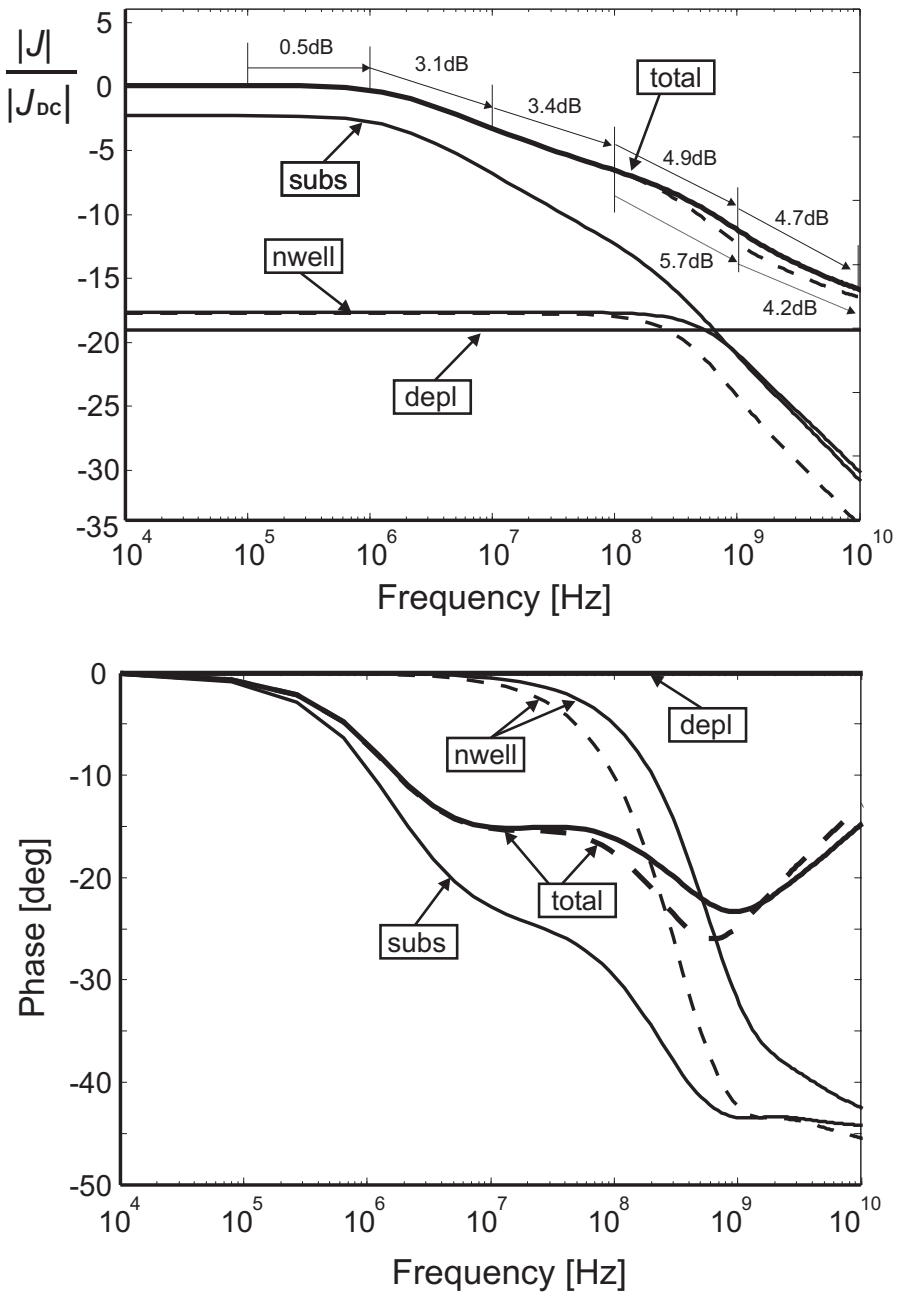


Figure 3.14: The calculated amplitude response of nwell/p-substrate photodiode with *low-resistance substrate* in a *twin-well* technology: 2 µm (solid lines) and 10 µ nwell size (dashed lines) for $\lambda = 850$ nm.

are connected with metal-2 and pwells with metal-1. The total metal area is about 13% of the total photodiode area meaning that the input light signal is decreased for 13%.

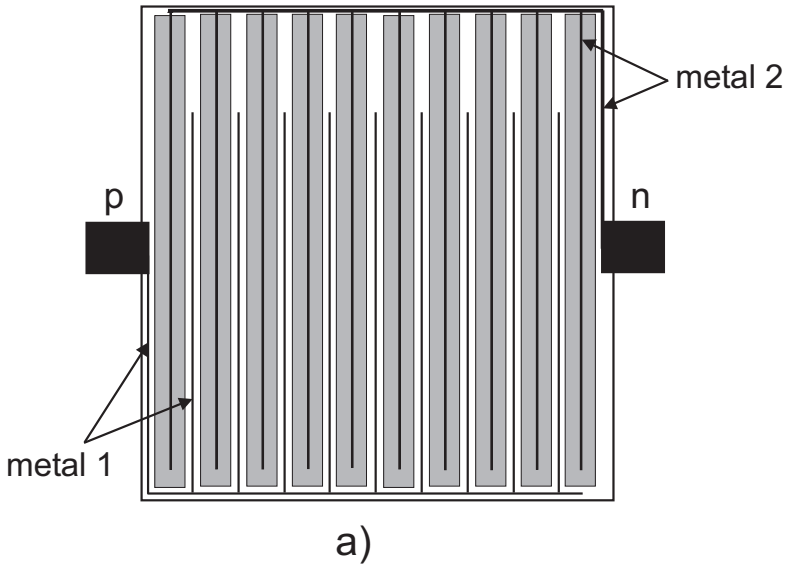
The responsivity of the photodiode shown in Figure 3.15 is measured first with the DC optical signal. The photoresponsivity and the frequency response are measured using on-chip measurements[6]. The diode is connected to a 1 V DC-supply using a bias-tee. A semiconductor parameter analyzer (SPA) HP4146B is used as the supply. The voltage and current compliances were set in order to avoid incorrect biasing. Using the SPA it was possible to monitor the diode characteristics thus, to check the correct contacting between the probes and bondpads.

The calculated photoresponsivity of the diode using equation (2.10), including the metal-coverage and without the light reflection R_{f} (defined in chapter 2) was 0.56 A/W. The exact amount of reflection depends on the thickness of the dielectric stack of layers between the silicon and the air. In the best case, these layers are transparent for the 850nm light and eliminate the reflection completely by forming an antireflection coating $d = \lambda/4\ n_{\mathrm{top}}$ [15], where n_{top} is the refractive index of the top transparent layer. In the worst case the reflection is not removed and effective amount of optical power incident to a photodiode is $P_{\mathrm{eff}} = 2/3 \cdot P_{\mathrm{in}}$. Due to this uncertainty, the photocurrent values for a different wavelengths can vary by one-third, as shown in Figure 3.16. The expected responsivity values range between 0.4 A/W-0.56 A/W.

An 850-nm VCSEL is used as a light source for testing the responsivity. The light was coupled from the laser to the photodiode using multimode fiber with 50 µm core-diameter. The fiber length is 1 m. The optical power is chosen in the range from 10µW (-20 dBm)[7], up to 120 μW (-9.2 dBm). The DC optical power at the fiber's output is measured using HP 8153A Lightwave Multimeter. With the responsivity of 0.4 A/W, and the metal area of 13% the calculated photocurrent is in the range from $I_{\mathrm{photo}} = 3.5$µA-42µA and with 0.56 A/W, the photocurrent is $I_{\mathrm{photo}} = 4.9$ µA-59 µA. The measured photocurrent is shown in Figure 3.16. This measured current values correspond more to the maximal light reflection. However, the diode responsivity is possibly reduced due to the positioning misalignment.

For the frequency response measurements, the insertion loss of RF cable is measured first in order to compensate for this loss. Then, the frequency

[6]The fabricated nwell/p-substrate photodiode was not packaged.
[7]-17 dBm sensitivity is specified as the Gigabit Ethernet standard.



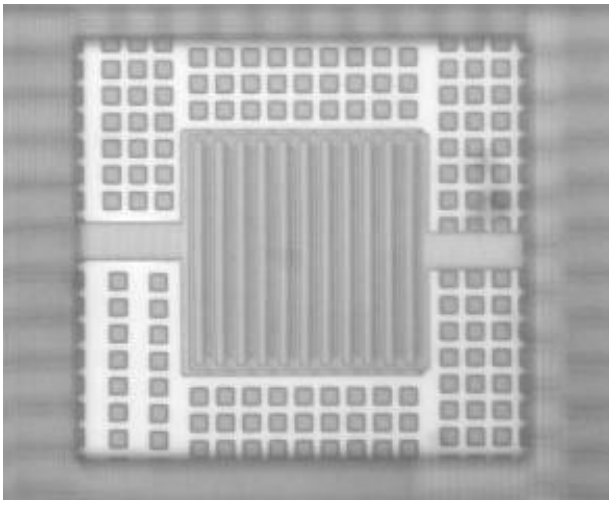


Figure 3.15: a) Layout of nwell/p-substrate photodiode with 2 µm nwell-width in standard CMOS technology b) chip-micrograph.

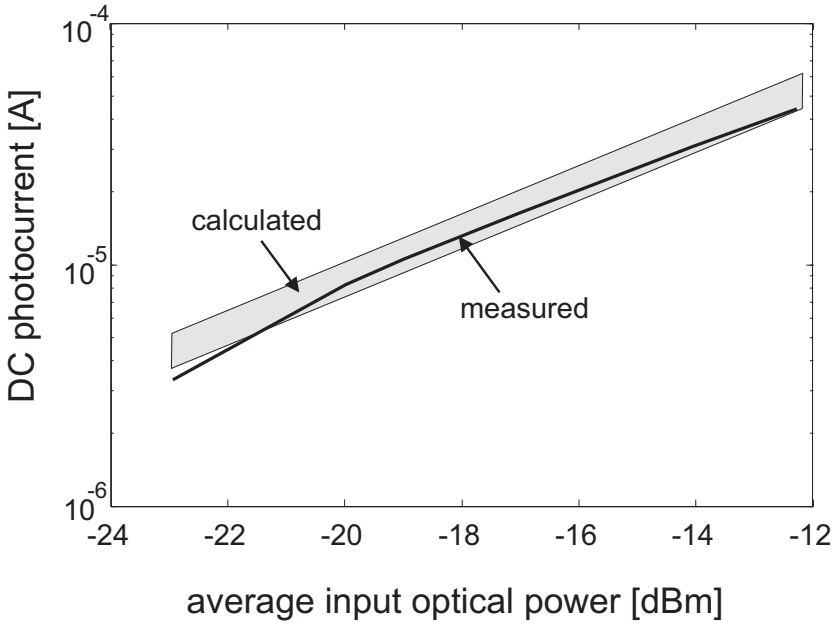


Figure 3.16: The measured DC photocurrent of nwell/p-substrate photodiode for $\lambda = 850$ nm.

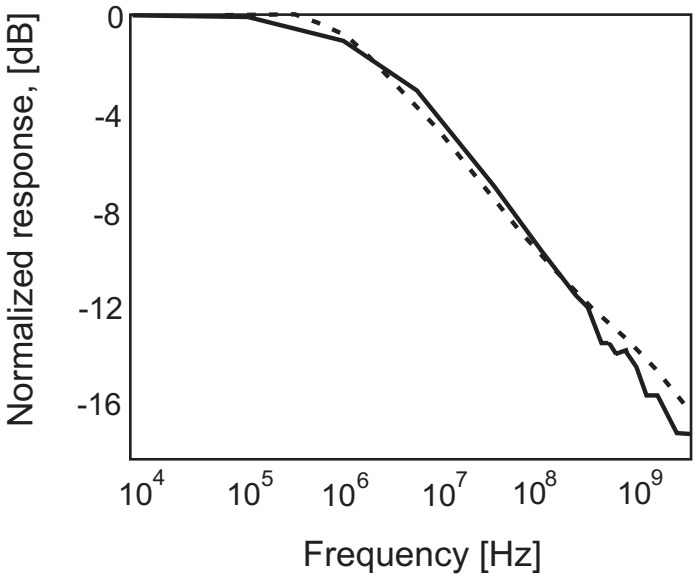


Figure 3.17: The measured (line) and calculated (dashed line) responses of a finger nwell/p-substrate photodiode with 2 µm nwell-size, $\lambda = 850$ nm.

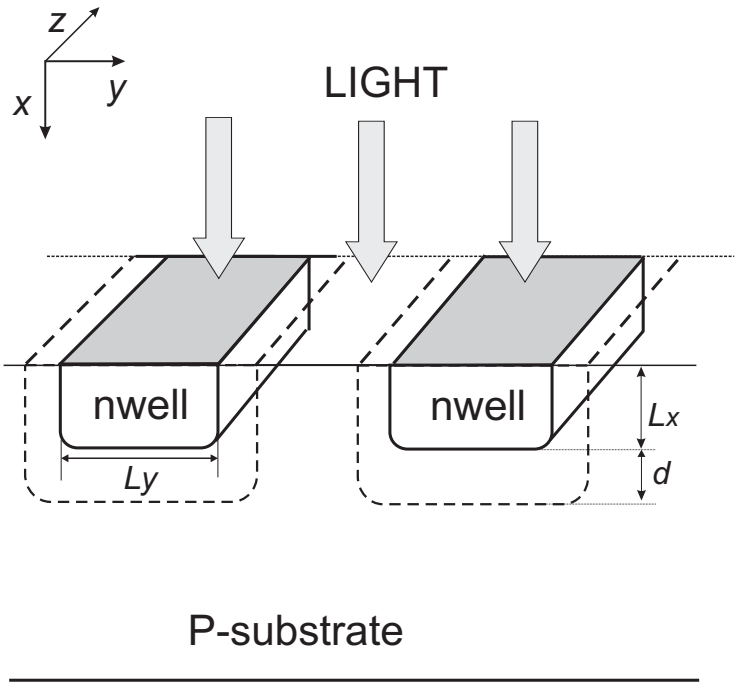


Figure 3.18: A finger nwell/p-substrate photodiode structure with *high resistance* substrate and *separate-wells* standard CMOS technology.

dependence of the output laser signal is compensated up to the frequency of interest (2 GHz). Further, the frequency response of the diode for $\lambda = 850$ nm is measured from 10 kHz up to 2 GHz using an E4404E Spectrum Analyzer. The measured and the calculated frequency responses are presented in Figure 3.17. Both responses show a slow decay after the -3 dB point which was described in the previous sections of this chapter. The calculated results using the process technology parameters comply well with the measurements.

Separate wells and high-resistance substrate

A separate well-CMOS technology combined with the high-resistance substrate is also used [1]; this photodiode structure is presented in Figure 3.18. The lateral depletion region between nwells is significantly increased. Therefore, the calculated amplitude of the drift current response in the depletion regions is higher in comparison to the adjoined-well diode. The overall photodiode bandwidth remains about 5 MHz because of the dominant substrate current

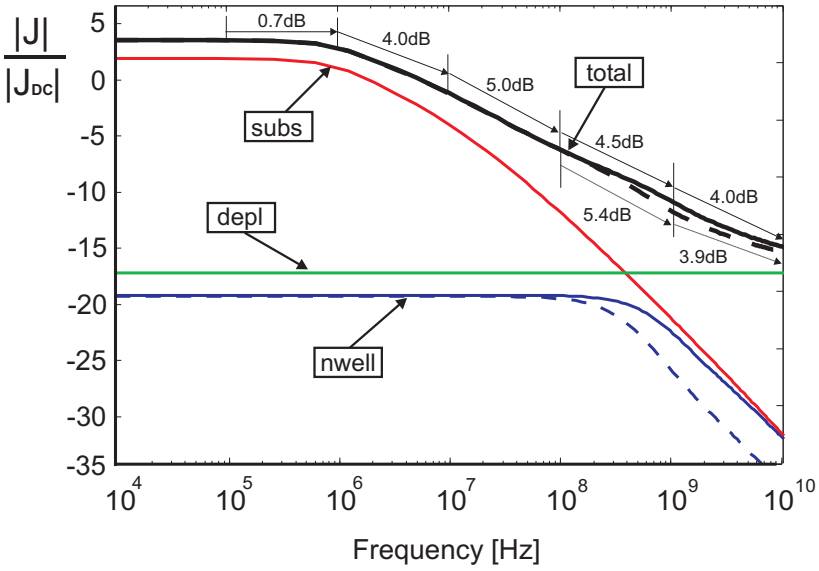


Figure 3.19: The calculated amplitude response of nwell/p-substrate photodiode with *high-resistance substrate* in a *separate-wells* technology: 2 µm (solid lines) and 10 µm nwell size (dashed lines) for $\lambda = 850$ nm.

contribution. On the other side, the depletion region response is dominant in higher decades and the roll-of in the total intrinsic diode characteristic is 1-2 dB lower in comparison with one in the twin well technology. The total photodiode intrinsic response in shown in Figure 3.19.

The maximal roll-off value in the photocurrent response is about 5.4 dB/decade for L_y=10 µm and 5.0 dB/decade for L_y=2 µm, as illustrated in Figure 3.19. In the low-GHz range, the roll-off is lowest among the previously discussed diodes ($\leq$4 dB/decade).

3.2.3 N+/p-substrate diode

This photodiode structure resembles a scaled-down of the nwell/p-substrate junction and it is presented in Figure 3.20. Therefore, the photocurrent response of this diode is similar with the one calculated for the nwell/p-substrate photodiode in the previous sections.

The response of the n^+ region is obtained by replacing the diffusion length L_p with the L_{p1} in the equation 3.8 as well as diffusion coefficient D_n with D_{n1} (corresponding to the doping of the n+ region). Since the doping concentration

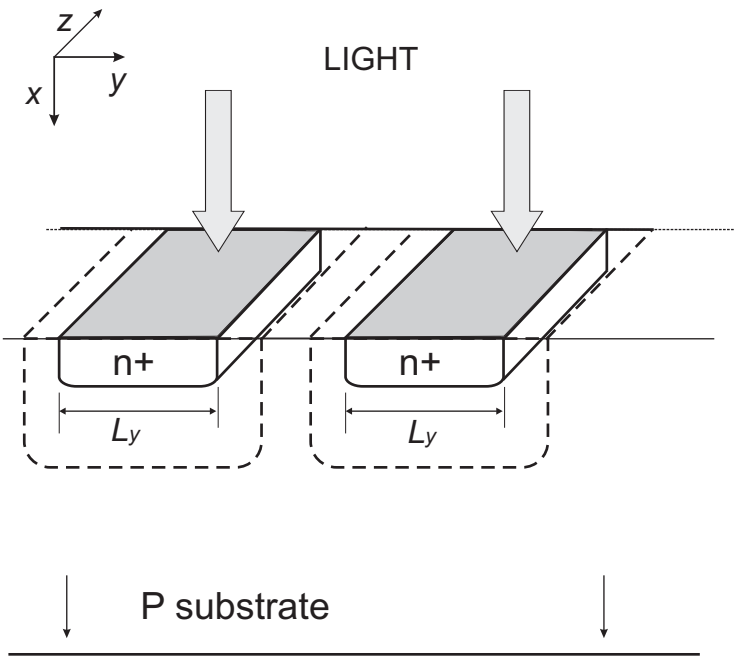


Figure 3.20: Finger n+/p-substrate photodiode in CMOS technology with high-resistance substrate.

of the shallow n^+ is much higher than in the n-well, the diffusion length L_{p1} will be much smaller [3]. The size of the n^+ diffusion layer towards the substrate L_{x1} is also lower. The maximum frequency response is determined mainly by the depth of the n+. In comparison with the n-well region, the holes' diffusion bandwidth is increased while the contribution to the overall current response is decreased because of the larger contribution of the slow substrate current (see Figure 3.20). The depletion region contribution is slightly larger because it is located closer to the diode surface, but its influence in the total current response is not changed significantly.

3.2.4 P+/nwell/p-substrate photodiode with low -resistance substrate in adjoined-well technology

The p+/nwell/p-substrate photodiode consists of two diodes: p+/nwell and nwell/p-substrate. The former photodiode is sort of complement of the previously discussed n+/p-substrate diode. This photodiode structure is illus-

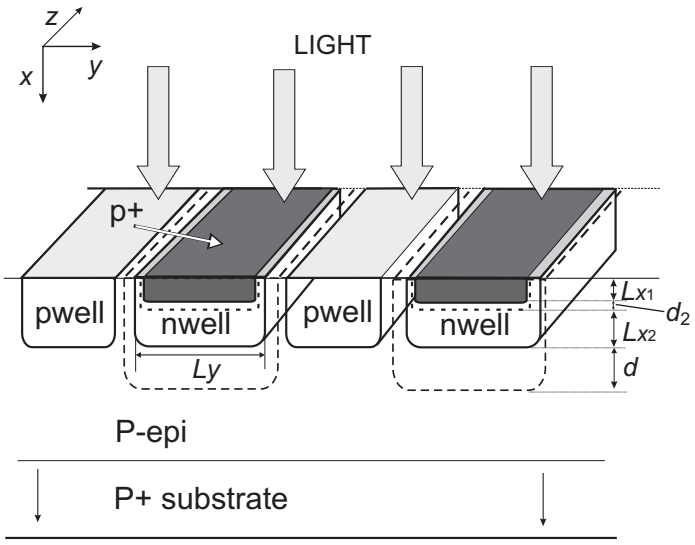


Figure 3.21: Finger p+/nwell/p-substrate photodiode structure in standard CMOS technology with low-resistance substrate and adjoined-wells.

trated in Figure 3.21. There are two vertical junctions (p+/nwell and nwell/p-substrate), and for this reason the diode is often referred to as *double photodiode*. The diffusion current responses are again calculated using a two-dimensional diffusion equation similar to those in 3.8. The main difference in comparison with the nwell/psubstrate and n+/p-substrate photodiode analyzed in previous section is a diffusion response inside nwell region. The boundary conditions for the holes density on every nwell side are shown in Figure 3.22. They are zero since the nwell is enclosed by junctions:

$$p_{\mathrm{n}}|_{x,y@\ \mathrm{boundary}} = 0 \tag{3.22}$$

The calculated current response of the nwell is given in equation (3.23):

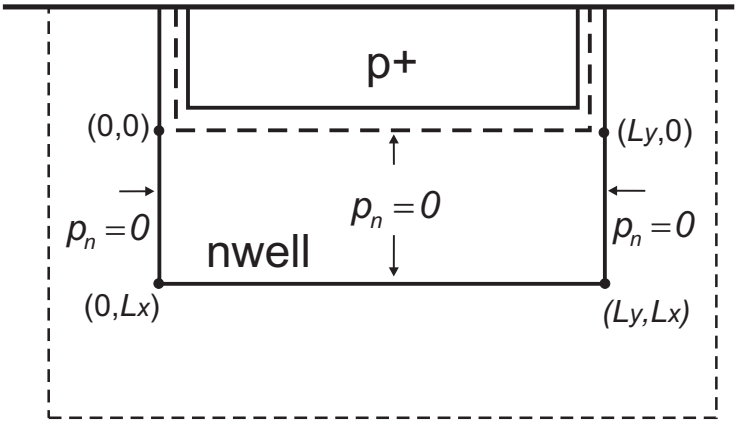


Figure 3.22: The boundary conditions for the hole density inside nwell for the double-photodiode.

$$J_{\mathrm{nwell}_1}(s) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{64e\Phi_0(s)L_{\mathrm{p}}^2[e^{-\alpha(L_{x1}+d_2)} - e^{-\alpha L_{x2}}]}{l\pi^2 L_{\mathrm{e}}} \times \frac{\dfrac{L_y}{L_{\mathrm{e}}(2n-1)^2} + \dfrac{L_{\mathrm{e}}}{L_y(2m-1)^2}}{\dfrac{(2n-1)^2\pi^2 L_{\mathrm{p}}^2}{L_y^2} + \dfrac{(2m-1)^2\pi^2 L_{\mathrm{p}}^2}{L_{\mathrm{e}}^2} + s\tau_{\mathrm{p}} + 1} \tag{3.23}$$

where $L_{\mathrm{e}} = L_{x2} - L_{x1} - d_2$.

For the p+ region, the electron current response is calculated using the nwell response in nwell/p-substrate diode and changing diffusion coefficient and diffusion length as well the depth of the junction: ($D_{\mathrm{p1}} \rightarrow D_{\mathrm{n1}}$, $L_{\mathrm{p1}} \rightarrow L_{\mathrm{n1}}$, and $L_x \rightarrow L_{x1}$). The substrate current response for both diodes is the same due to the same nwell depths and the same doping concentrations.

The total frequency response of the double-photodiode is calculated for two nwell/p+ sizes: first, the nwell width is twice its depth, $L_y = 2$ μm, and second, the nwell width is much larger than its depth, $L_y = 10$ μm. The input wavelength is $\lambda = 850$ nm. The results are presented in Figure 3.23. This picture illustrates that the bandwidths of p+ and nwell currents are mainly determined by the low physical depth of the junctions ($2L_{x1}, 2L_x < L_y$); changing nwell and p+ widths has almost no effect on the cutoff frequency. The bandwidth of the junction-framed nwell current is $f_{\mathrm{3dBnwell}} = 5$ GHz for $L_y = 2$ µm,

and $f_{3\text{dBnwell}} = 4.2$ GHz for $L_y = 10$ µm. That is more than twice the nwell bandwidth of the nwell/p-substrate diode. The distances towards the junctions are lower providing a higher charge gradient and higher net transport in the diffusion process. The current bandwidth of the p+ region is lower than the bandwidth of the nwell current; the calculated value is about 3 GHz for all nwell/p+ widths, (the depth of the p+ is smaller than its width, and it mainly determines the bandwidth). The p+ surface is reflective for the carriers[8]. They are repelled back to the other three p+ sides with the junctions. Thus, these carriers need extra time to start contributing to the overall photocurrent. The calculated intrinsic photodiode bandwidth is 4.6 MHz, and the nwell/p+ widths have no influence on it (see Figure 3.23).

The maximal roll-off value in the photocurrent response is located in the first GHz decade its value is 4.0 dB. This is valid for both diode geometries, as illustrated in Figure 3.23. In the 1-1000 MHz range, the roll-off per decade is 1-2 dB lower than in nwell/p-substrate diode: the value is about 3.5 dB/decade.

The time impulse response of the hole diffusion profile inside nwell is again calculated using the Inverse Laplace transform of equation (3.23). This profile is calculated after 1 ps, 6 ps, 15 ps, 100 ps and illustrated in Figure 3.24. The nwell region is completely surrounded by junctions; for this reason the hole-carrier profile diminish much faster than in the case of the hole profile inside the nwell for the nwell/p-substrate photodiode (see Figure 3.8).

In Table 3.1 and Table 3.2, the bandwidths and the responses of nwell/p-substrate and double-photodiode structures with two different kinds of substrate and two different geometries are summarized. The diode responses are normalized on the response of the nwell/p-substrate photodiode in twin-well technology and with low-resistance substrate.

3.2.5 Finger p+/nwell photodiode

This photodiode structure is obtained by disconnecting the p-substrate in the double-photodiode, as illustrated in Figure 3.25. In this way, it is possible to achieve a high intrinsic bandwidth of the photodiode. The responsivity is however lower than in the other two diode structures because the substrate excess carriers do not contribute the overall photocurrent. For $\lambda = 850$ nm, the calculated responsivity of this photodiode is 18 dB lower (see Figure 3.26).

[8] The surface recombination process is far slower than the 1/MHz and 1/GHz time responses considered in this thesis.

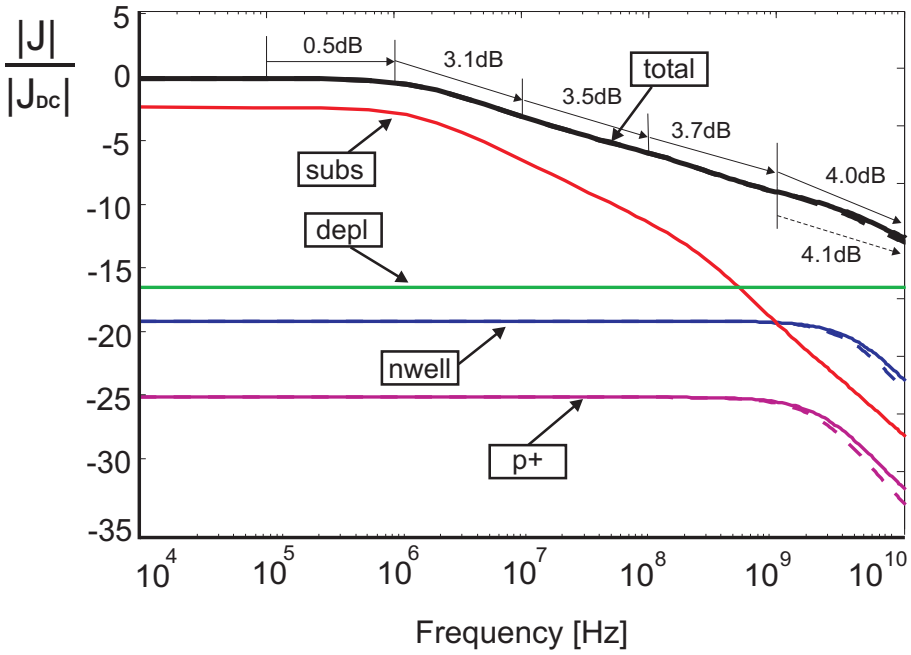


Figure 3.23: The calculated amplitude response of *p+/nwell/p-substrate* photodiode with low-resistance substrate in a adjoined-wells technology: 2 µm (solid lines) and 10 µm nwell size (dashed lines) for $\lambda = 850$ nm.

The charge movement inside the nwell is constant in the lateral direction since there is no lateral current towards the substrate, (the substrate is disconnected). Thus, the nwell width has no influence on the nwell diffusion bandwidth.

The nwell diffusion response is calculated using a one-dimensional diffusion equation, similar to equation 3.10 given in section 3.2.1. The hole density at the top of the nwell (below the depletion region towards p+) is taken to be zero while the gradient of the holes density at the nwell bottom is zero since there is no current flowing. The current response of the nwell is given in equation (3.24):

$$J_{\mathrm{nwell}_2} = \sum_{n=1}^{\infty} \frac{2\ e\Phi L_{\mathrm{p}}^2 (e^{-\alpha(L_{x1}+d_2)} - e^{-\alpha L_{x2}})}{L_{\mathrm{e}} \left(\dfrac{(2n-1)^2 \pi^2 L_{\mathrm{p}}^2}{L_{\mathrm{e}}^2} + s\tau_{\mathrm{p}} + 1 \right)} \tag{3.24}$$

where $L_{\mathrm{e}} = L_{x2} - L_{x1} - d_2$.

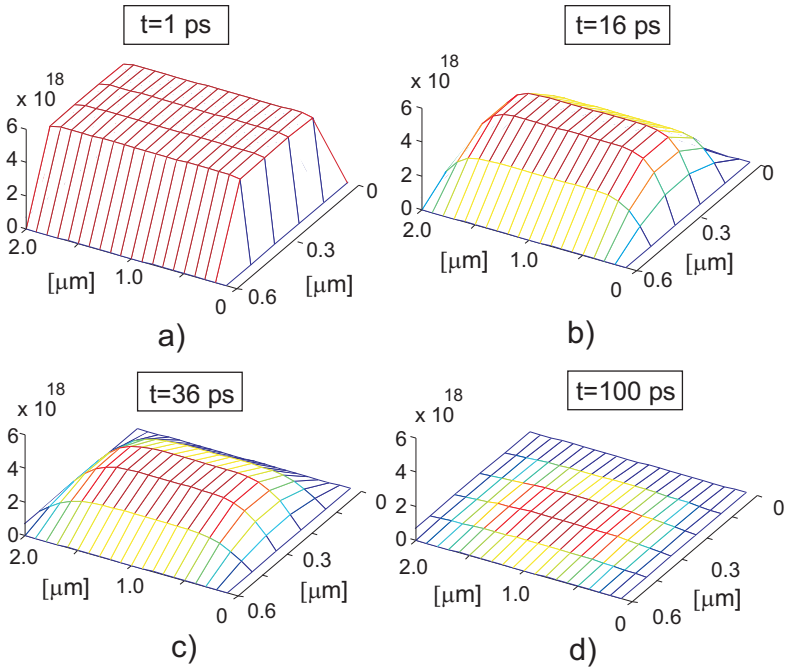


Figure 3.24: The calculated concentration of the excess carriers inside nwell for p+/nwell/p-substrate photodiode (effective depth 0.6 µm), under incident light pulse (10 ps pulse-width). This profile is calculated after 1 ps, 6 ps, 15 ps, 100 ps.

The amount of the nwell current contributing to the overall photocurrent is the same as in the p+/nwell/p-substrate photodiode, but its bandwidth is five times lower: $f_{3\text{dBnwell}}$=1 GHz. The carriers moving towards p-substrate are repelled back and need extra time to diffuse to the top of the nwell where they sink into the junction. This effect causes the decrease in photocurrent bandwidth. The diffusion current response inside p+ is the same as in the double-photodiode. The frequency responses of two finger p+/nwell diodes with 2 µm and 10 µm nwell sizes, for $\lambda = 850$ nm are shown in Figure 3.26.

The overall intrinsic photocurrent bandwidth is about 1 GHz, meaning that this photodiode structure has the highest bandwidth among all photodiodes discussed in the previous sections of this chapter. However, the photocurrent value is small; it is 18 dB lower due to a very low diode responsivity. For proper signal detection, the required optical power has to be more than one order of magnitude higher than typically specified for short-haul optical communication

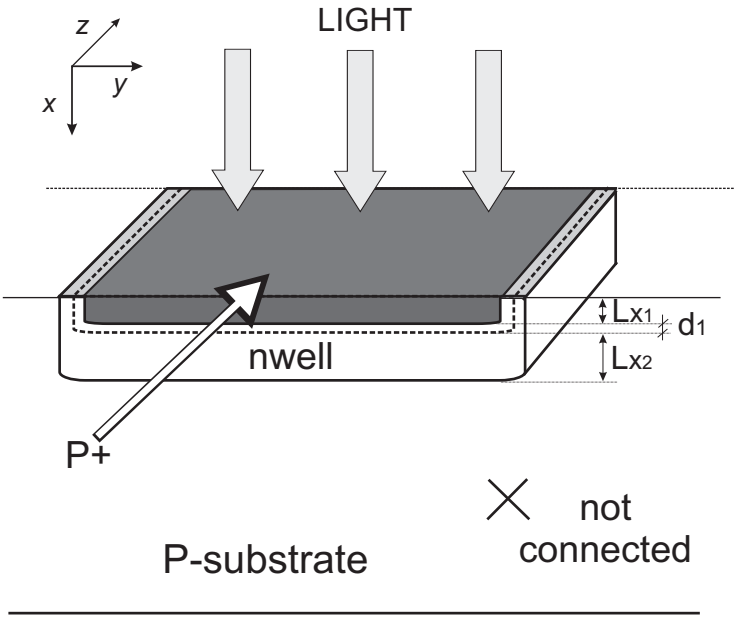


Figure 3.25: Cross-section of p+/nwell photodiode.

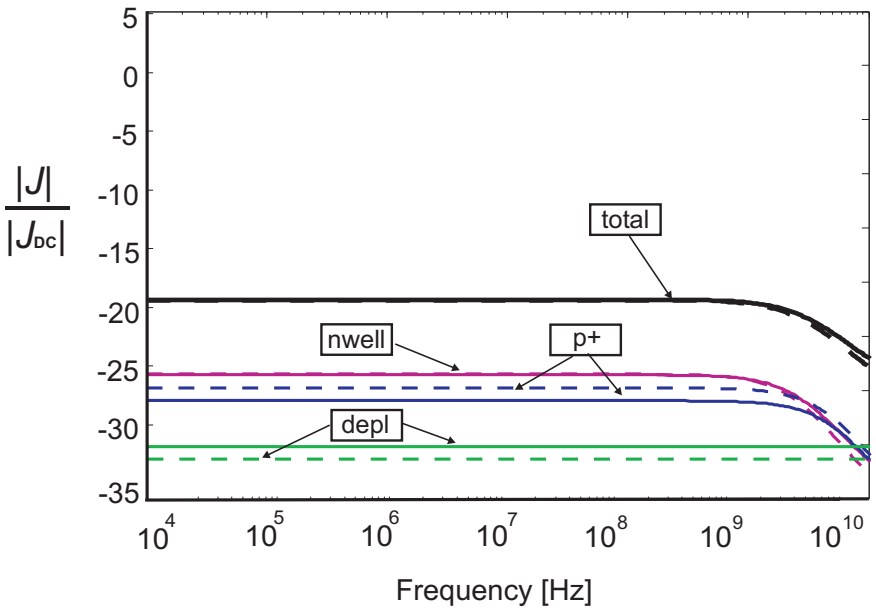


Figure 3.26: The calculated amplitude response of p+/nwell photodiodes with: 2 µm (solid lines) and 10 µm nwell size (dashed lines) for $\lambda = 850$ nm.

Table 3.1: The bandwidth and responsivity of nwell/p-substrate photodiode with various substrates

	$\lambda = 850$nm		$\lambda = 600$nm	
	$L_y = 2\mu$m	$L_y = 10\mu$m	$L_y = 2\mu$m	$L_y = 10\mu$m
nwell BW	930MHz	450MHz	930MHz	450MHz
depl BW ($V_b = 1$ V)	10GHz	10GHz	10GHz	10GHz
high-resistance substrate separate-wells				
subs BW	3.5MHz	3.5MHz	18MHz	18MHz
total BW	4.5MHz	4.5MHz	90MHz	90MHz
maximal roll-of/decade	5	5.4	4	5
normalized responsivity	3dB	3dB	4.3dB	4.3dB
low-resistance substrate separate-wells				
subs BW	5MHz	5MHz	28MHz	28MHz
total BW	8MHz	8MHz	160MHz	150MHz
maximal roll-of/decade	4.5	5.2	4	4.6
normalized responsivity	0dB	0dB	3.5dB	3.5dB
high-resistance substrate adjoined-wells				
subs BW	3.5MHz	3.5MHz	18MHz	18MHz
total BW	3.5MHz	3.5MHz	75MHz	75MHz
maximal roll-of/decade	5.2	5.7	4.7	4.8
normalized responsivity	3dB	3dB	4.3dB	4.3dB
low-resistance substrate adjoined-wells				
subs BW	5MHz	5MHz	28MHz	28MHz
total BW	8MHz	8MHz	160MHz	150MHz
maximal roll-of/decade	4.9	5.7	4.6	5
normalized responsivity	0dB	0dB	3.5dB	3.5dB

[4]. This photodiode structure is comparable with photodiodes designed in Silicon-On-Insulator technology. The analyses in this section showed that for $\lambda = 850$ nm, such photodiodes have bandwidth in the low GHz range, but very low responsivity; thereby, they can not be used in optical communication systems.

Table 3.2: The bandwidth and responsivity of p+/nwell/p-substrate photodiode with various substrates

	$\lambda = 850$nm		$\lambda = 600$nm	
	$L_y = 2\mu$m	$L_y = 10\mu$m	$L_y = 2\mu$m	$L_y = 10\mu$m
p+ BW	3.4GHz	2.9GHz	3.4GHz	2.9GHz
nwell bBW	6GHz	5.2GHz	6GHz	5.2GHz
depl BW ($V_b = 1$ V)	10GHz	10GHz	10GHz	10GHz
p-epi substrate single-well				
subs BW	3.5MHz	3.5MHz	18MHz	18MHz
total BW	5MHz	5MHz	100MHz	100MHz
maximal roll-of/decade	4.7	4.7	3.5	4
normalized responsivity	3dB	3dB	4.3dB	4.3dB
two p-layers substrate single-well				
subs BW	5MHz	5MHz	28MHz	28MHz
total BW	8MHz	8MHz	190MHz	190MHz
maximal roll-of/decade	3.7	3.7	3.5	3.9
normalized responsivity	0dB	0dB	3.5dB	3.5dB
p-epi substrate twin-well				
subs BW	3.5MHz	3.5MHz	18MHz	18MHz
total bBW	4.5MHz	4.5MHz	85MHz	85MHz
maximal roll-of/decade	5.0	5.0	4	4.2
normalized responsivity	3dB	3dB	4.3dB	4.3dB
two p-layers substrate twin-well				
subs BW	5MHz	5MHz	28MHz	28MHz
total BW	8MHz	8MHz	190MHz	190MHz
maximal roll-of/decade	4	4.2	3.9	4.2
normalized responsivity	0dB	0dB	3.5dB	3.5dB

3.2.6 Extrinsic (electrical) photodiode bandwidth

Apart from the intrinsic bandwidth of the "stand-alone" photodiode, the in-circuit photodiode bandwidth is also determined by the extrinsic (electrical) bandwidth. This bandwidth is determined by the diode and interconnect capacitance in combination with the pre-amplifier's input resistance.

The capacitance of high-speed photodiodes depends on the diode area and it is typically in the pF range (using a multimode fiber connection the diode area is

50×50 µm^2). Table 3.3 shows the calculated values of the parasitic capacitances for two nwell widths of nwell/p-substrate and double photodiode in the 0.18 µm CMOS technology: the first nwell width is twice its depth $L_y = 2L_x = 2$ μm, and the second with is much higher than the nwell depth $L_y = 10$ μm.

For the photodiodes in the separate wells technology, the width of the lateral depletion region is much larger than for the diodes in twin-well technology with adjoined wells. The doping concentration of the pwells is about two orders of magnitude larger than in high-resistance substrate. Hence, the calculated depletion region width towards pwells is 7 times smaller. The total diode capacitance is 5-7 times larger for a adjoined-well process in comparison with the separate-well process.

Table 3.3: Parasitic capacitance for different photodiode structures and geometries

	$L_y = 2$ μm	$L_y = 10$ μm
nwell/p-substrate		
separate-wells	0.28 pF	0.27 pF
adjoined-wells	1.6 pF	0.62 pF
p+/nwell/p-substrate		
separate-wells	2.0 pF	1.8 pF
adjoined-wells	3.60 pF	2.20 pF

For all photodiodes discussed in this chapter, there are two general observations about their electrical bandwidth. First, by decreasing the diode nwell width (for a constant diode area), the total junction area of the photodiode increases. As a result, the diode capacitance increases. Second, the implementation of twin-well technology increases diode capacitance too. This implementation particularly changes the nwell/p-substrate diode capacitance.

According to Table 3.3, a nwell/p-substrate photodiode in separate-well technology has the lowest capacitance. The calculated value is about 10 times lower than the capacitance of the adjoined-well double-photodiode. For a reasonably low input impedance of the pre-amplifier, the electrical photodiode bandwidth is about 10 GHz for 0.18 µm CMOS.

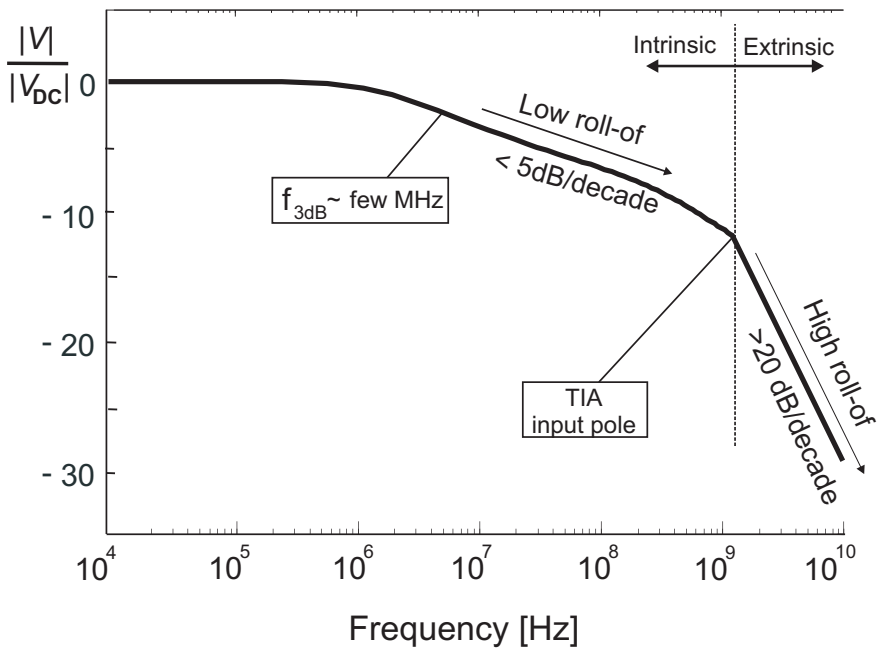


Figure 3.27: Total photodiode frequency response including the *intrinsic* and the *extrinsic* frequency responses.

3.3 Total photodiode bandwidth

The total frequency response of the integrated photodiode with a transimpedance amplifier is the product of the intrinsic and extrinsic photodiode responses. The total diode bandwidth is by approximation the lower between the intrinsic and the extrinsic bandwidth $f_{3\text{dBtotal}} \simeq \min[f_{3\text{dBintrinsic}}, f_{3\text{dBextrinsic}}]$. In order to keep the extrinsic bandwidth in the GHz range, for pF diode capacitance, a subsequent transimpedance amplifier (TIA) should be designed to have a low input impedance. Figure 3.27 shows the total frequency response of nwell/p-substrate photodiode. The intrinsic diode bandwidth clearly dominates the total bandwidth.

Using the p+/nwell photodiode, the total diode bandwidth can be in the low GHz range. However, this diode has 3-4 times (12 dB) lower responsivity in comparison with the other photodiode structures. Thus, the TIA configuration should be chosen according to both the required optical receiver's sensitivity (sets a very high demands) and the amplifier's bandwidth.

For the other three photodiodes structures (nwell/p-substrate, n+ /p-substrate and p+/nwell/p-substrate) using low input resistance TIA's, the calculated electrical bandwidths will be larger than the calculated intrinsic diode bandwidths. Thus, the parasitic capacitance has almost no influence on their total bandwidth.

3.3.1 Noise in photodiodes

The noise generated by a photodiode operating under reverse bias, is a combination of shot noise and Johnson noise. Shot noise is generated by random fluctuations of current flowing through the device. This noise is discovered in tubes in 1918 by Walter Schottky who associated this noise with direct current flow. In fact, it is required that there be dc current flow or there is no shot noise. The dc current is a combination of dark current (I_{r}) and quantum noise (I_{qn}). Quantum noise results from the random generation of electrons by the incident optical radiation. The shot noise is given as [15]:

$$\overline{i_{\mathrm{s}}^2} = (2q(I_{\mathrm{r}} + I_{\mathrm{qn}})BW) \tag{3.25}$$

where $\overline{i_{\mathrm{s}}^2}$ is the shot-noise current, and BW is bandwidth of interest.

The Johnson noise (thermal noise) is due to the load resistance of the photodiode as well as the ambient temperature [8]. The Johnson noise is given by:

$$\overline{i_{\mathrm{j}}^2} = \left(\frac{4kTBW}{R}\right) \tag{3.26}$$

where R presents the resistance at its output.

The total noise current i_{noise} is the root mean square sum of the individual noise current contributions:

$$\overline{i_{\mathrm{noise}}^2} = \overline{i_{\mathrm{s}}^2} + \overline{i_{\mathrm{j}}^2} \tag{3.27}$$

In order to investigate which of the two noise sources is dominant in the optical detection, the ratio between the rms values of the Johnson noise current and shot noise current is taken:

$$\frac{\overline{i_j^2}}{\overline{i_s^2}} = \frac{2kT}{q(I_r + I_{qn})R} \quad (3.28)$$

$$\approx \frac{1}{20(I_r + I_{qn})R} \quad (3.29)$$

The data-rate of interest in this thesis are in the low Gb/s, so the electrical pole at the photodiode output $1/(C_{diode}R)$ should larger than 1 GHz. The capacitance of the CMOS photodiode C_{diode} is related to the photodiode area, which corresponds to the diameter of the multimode fiber $50 \times 50\mu m^2$. Table 3.3 shows that the diode capacitance C_{diode} is in the pF range. As a result, the resistance R at the diode output should be smaller than 160 Ω. Substituting this value for R into equation (3.29), the dc current is obtained, for which the ratio between the rms noise currents is equal to 1:

$$\frac{\overline{i_j^2}}{\overline{i_s^2}} = 1 \approx \frac{1}{20(I_r + I_{qn})160} \quad (3.30)$$

$$\Rightarrow I_r + I_{qn} = 312\mu A \quad (3.31)$$

The measured dark current of the photodiode using Semiconductor Parameter Analyzer HP 4146B is less than 1 nA, for the -1 V bias voltage. Therefore, the effect of the dark current is negligible for the diode shot noise. The dc current is determined by the dc photocurrent (quantum noise). The laser used in experiments presented in this thesis emits the light with the modulation index m smaller than one: $m = 0.2$. As a result, the dc photocurrent is larger than the signal current. In practical applications, the modulation index is higher since the laser is operated around the threshold current not to dissipate the power when the signal is not transmitted.

Equation (3.31) shows that for dc currents smaller than 312 µA, the Johnson noise in the photodiode is larger than the shot noise. The next paragraph describes the practical dc currents that can be expected in the optical detection.

The specified sensitivity of the optical detector in the Gigabit Ethernet standard [4] is -17 dBm average power, which corresponds to the $P_{in} = 40$ µW peak-to-peak power. Chapter 2 described that the maximum photodiode responsivity for $\lambda = 850$ nm is R_{max}=0.63 A/W. Thereby, the maximum ac photocurrent $I_{ph_{AC}}$ is $I_{ph_{AC}} = R_{max}P_{in} = 25.2$ µA. The ratio between ac photocurrent (signal

current) and the dc current (noise current) gives the modulation index of the input signal. If the dc current has the previously calculated value of 312 µA the shot noise current is equal to the Johnson noise current if the modulation index m is:

$$m \approx \frac{I_{\mathrm{ph_{AC}}}}{I_{\mathrm{r}} + I_{\mathrm{qn}}} = 0.08 \tag{3.32}$$

This modulation index is small and it can cause high power consumption since the information is carried out only in a very small portion of light. In practical applications the modulation index m is much higher (> 0.25). As a result, the dc current will be much lower than 312 µA, since the signal amplitude is already defined with Gigabit Ethernet standard. Therefore, the Johnson noise of the photodiode will be larger (and dominant) in comparison to the shot noise.

3.4 Summary and Conclusions

This chapter analyzed the frequency and the time responses of different photodiode structures designed in standard CMOS technology, for $\lambda = 850$ nm. For clear explanation of the physical processes inside photodiode, the available 0.18 µm CMOS technology was chosen to be analyzed in detail, which does not fundamentally influence the photodiode behavior in general. It is a standard 0.18 µm CMOS since that was the available fabrication technology. The photodiodes are first analyzed as stand-alone detectors i.e. without subsequent electronic circuitry. This allows an analysis of the *intrinsic* photodiode behavior. The intrinsic behavior is related to the movement (drift and diffusion) of the generated carriers inside the diode. Based on the frequency analysis, an intrinsic (physical) diode bandwidth is determined. In the second part of this chapter, the diode is investigated as an "in-circuit" element, integrated together with the subsequent electronics. The *electrical* bandwidth of the diode is determined by the diode capacitance and the input impedance of the subsequent amplifier. These two bandwidths determines the total diode bandwidth which is by approximation the lower of the two.

The most often used figure-of-merit for optical detectors is responsivity-bandwidth product. For easier comparison among photodiodes, it is assumed that all photodiode have the same active area, which provides the same maximal responsivity. Thereby, the photodiode performances are compared according to their bandwidth.

The first part of the chapter analyses the intrinsic frequency response of the different photodiode structures in 0.18 µm CMOS technology, for 850 nm wavelength: 1) nwell/p-substrate, 2) n+/p-substrate, 3) p+/nwell/p-substrate and 4) p+/nwell. The wavelength of $\lambda = 850$ nm is now-days in the standard for short-haul fiber communication, for example Gigabit Ethernet standard.

The intrinsic bandwidths of the photodiode structures 1),2) and 3) are almost the same, in the low MHz range. It is mainly determined by the bandwidth of the substrate current which dominate the overall response. The light penetration depth is about 30 µm (90% of the light is absorbed), and most of the excess carriers are generated in the substrate.

For the structure 4), there is no p-substrate contribution and the response is fast (in the GHz range). However, the responsivity of the photodiode is significantly lower (10 times) which makes it unsuitable for practical applications (the sensitivity of the whole optical receiver is reduced for the same amount). In order to increase the bandwidth of the presented diode structures in the GHz range, the proposed method in chapter 4 of this thesis can be used.

The high-resistance and low-resistance p-substrate also determines the photodiode bandwidth. Inside the low-resistance substrate, a larger number of carriers are recombined in comparison with a high-resistance substrate. For this reason, the overall photocurrent is lower on one side, but on the other side the overall response is faster due to the recombination of the carriers that are diffusing later in time. The calculated normalized amplitude of the overall photocurrent is -2 dB with 3 times higher total intrinsic bandwidth: 8 MHz. The temperature influence on the photodiode characteristics will be analyzed in chapter 4.

In the second part of the chapter, the extrinsic (electrical) photodiode bandwidth is discussed. It is determined by both the input capacitance (the diode capacitance and the input capacitance of the TIA) and the input impedance of the TIA. Both have to be minimized for the maximal electrical bandwidth. The diode capacitance is strongly dependent on the diode area (50×50 μm^2) and the diode structure, and it is larger than 0.6 pF. For the maximal electrical bandwidth, the input impedance of the TIA thus, has to be minimized.

For a typical input impedance of the TIA the electrical bandwidth is in the GHz range; therefore it is much larger than the intrinsic photodiode bandwidth for most of the analyzed structures. The minimal parasitic capacitance and thus, the highest electrical bandwidth is achieved with maximal nwell width (using single photodiode).

References

[1] D. Coppée, H. J. Stiens, R. A. Vounckx, M. Kuijk: "Calculation of the current response of the spatially modulated light CMOS detectors", *IEEE Transactions Electron Devices*, vol. 48, No. 9, 2001, pp. 1892-1902.

[2] S. M. Sze: "*Physics of semiconductor devices*", New-York: Wiley-Interscience, 2-nd edition, p. 81, 1981.

[3] G.W.de Jong et al.: "A DC-to-250MHz Current Pre-Amplifier with Integrated Photo-Diodes in Standard CBiMOS, for Optical-Storage Systems", *ISSCC 2002*, s21.8

[4] IEEE 10 Gigabit Ethernet Standard 802.3ae.

[5] S. Radovanović, AnneJohan Annema and Bram Nauta: "Physical and electrical bandwidths of integrated photodiodes in standard CMOS technology", *Electron Device Solid State Circuit 2003*, Hong Kong, pp. 95-98.

[6] S. Radovanović, AnneJohan Annema and Bram Nauta: "On optimal structure and geometry of high-speed integrated photodiodes in a standard CMOS technology", *CLEO 2003*, Taiwan, pp.87, TU4H-9-1.

[7] C. L. Schow, J. D. Schaub, R. Li, and J. C. Campbell, "A 1 Gbit/s monolithically integrated silicon nmos optical receiver", *IEEE J. Select. Topics Quantum Electron.*, vol. 4, pp. 1035-1039, Nov.Dec. 1999.

[8] S. Alexander: "*Optical communication receiver design*", SPIE Optical engineering press, 1997.

[9] Wei Jean Liu, Oscal T.-C. Chen, Li-Kuo Dai and Far-Wen Jih Chung Cheng: "A CMOS Photodiode Model", *2001 IEEE International Workshop on Behavioral Modeling and Simulation*, Santa Rosa, California, October 10-12, 2001.

[10] Y. R. Nosov: "*Switching in semiconductor diodes*", New York: Plenum, 1969, p.14.

CHAPTER 4

Bandwidth improvement

This chapter presents bandwidth improvement of integrated photodiodes in standard CMOS, such that they can be used for Gb/s optical detection. This improvement is achieved by using an inherently robust analog equalizer that compensates (in gain and phase) for the photodiode roll-off in the range from 1 MHz to 1 GHz. The important feature of the equalization method is that complex adaptive algorithms are not required.

4.1 Introduction

Chapter 3 described the frequency and time responses of photodiodes in standard CMOS for $\lambda = 850$ nm. The photodiodes were considered first as standalone devices, which allowed an *intrinsic* diode response calculation. This response was also calculated in general for any nwell width in the diode layout, while the examples of two nwell/n+/p+ sizes are discussed in detail (minimum width and nwell width much larger than the minimum). Further in chapter 3, the diode is considered as an "in-circuit" device, integrated together with the electronic circuitry. This allowed the calculation of an *electrical* (extrinsic) diode response. From the intrinsic and extrinsic diode responses, it is possi-

ble to calculate the intrinsic and extrinsic diode bandwidths. The total diode bandwidth is by approximation minimum between these two.

Chapter 3 showed that diode *intrinsic* frequency characteristics has low roll-off up to the low GHz range. This range is determined by the drift bandwidth $f_{3\text{dBdrift}} = 0.4\ v_\text{s}/W$, where v_s is the saturation velocity of excess carriers and W is the depletion region depth. From the drift bandwidth on, the diode response has high roll-off (>20 dB/decade). For 0.18μm CMOS, the calculated drift bandwidth is about 8 GHz. The electrical diode frequency characteristics has a higher roll-off; from the electrical bandwidth frequency on, the total roll-off is also $\geq$20 dB/decade, due to the electrical poles domination.

This chapter presents an analog equalization method for bandwidth improvement of integrated CMOS photodiodes using the low roll-off feature in the frequency characteristics. The achieved photodiode bandwidth is in the low GHz range, which was limited by the electrical diode bandwidth. The first part of this chapter explains the design of the integrated photodiode and pre-amplifier for the largest bandwidth improvement. It includes both the optimization of the photodiode layout (nwell-width) and the design of transimpedance amplifier (TIA). Once the photodiode and the TIA are optimized, a subsequent analog equalizer is designed.

Important in the design of equalizers is spread both on the characteristic to be compensated and on the equalizer itself. The robustness of the proposed method on spread as well as temperature of the environment is described in the second part of this chapter. The frequency behavior of the pre-amplifier is analyzed using the total *equalized* frequency response. The time domain response of the pre-amplifier was used to investigate the total responsivity while the a bit error rate is measured using digital data analyzer. Section 4.3.1 gives the simulation results and section 4.3.1 the time-domain measurements on the pre-amplifier. Even though a signal-to-noise ratio can be deduced using eye-diagrams, it is more accurate to evaluate the receiver with a bit error rate (BER) measurement. This is analyzed in section 4.2.2.

4.2 Pre-amplifier

For large data-rates (hundreds of Mb/s) and for wavelength of λ = 850 nm, a typical photodiode in standard CMOS causes severe intersymbol interference (ISI) due to its limited bandwidth [1]. In literature, the most common solution

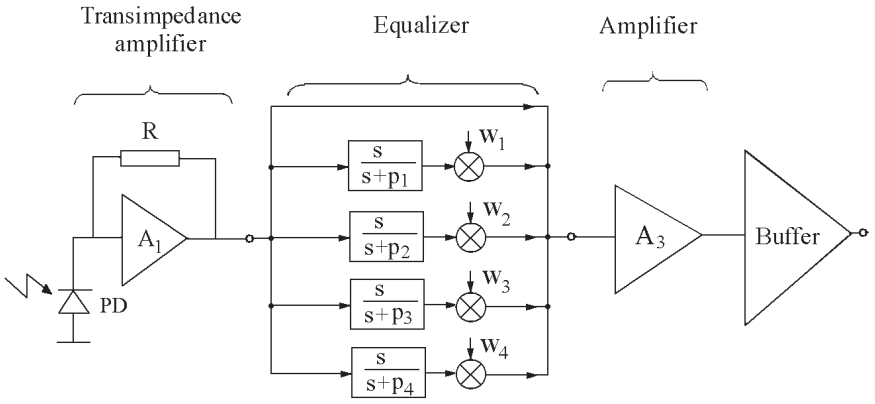


Figure 4.1: Block-diagram of integrated photodiode and preamplifier system using an analog equalizer.

for reducing a ISI is an adaptive equalization in the analog or in the digital domain. Equalization has been widely used in communications applications such as voice-band modems, wireless [2], digital subscriber lines, and ISDN [3], and even at rates close to 500 Mb/s in disk drives [4], [5]. For fiber optics communication, adaptive equalization is typically used for fiber dispersion compensation [6].

This section presents an analog equalizer that compensates (in gain and phase) for the diode photocurrent decay in the range from $f_{3\mathrm{dB}_{\mathrm{diode}}}$ (typically in the MHz range) to 1 GHz. The electrical bandwidth frequency is larger than the compensation frequency in order not to add no extra attenuation to the low roll-off intrinsic compensation[1].

For the maximum equalization range[2], the electrical diode bandwidth should be as large as possible. This bandwidth is maximized by minimizing both the input resistance of the subsequent TIA and the photodiode capacitance. The former one typically increases the power consumption of the pre-amplifier.

The proposed pre-amplifier is presented first as a block-diagram in Figure 4.1. The difference in comparison with the straightforward pre-amplifier

[1]It is certainly possible to use an analog equalizer further in the GHz range where roll-off/decade is >10 dB. However, because of the spread in both the diode and the equalizer, the output signal decreases, which increases bit error rate, as shown in chapter 6. Therefore, some adaptive algorithms are needed. A proposed method in this thesis avoid implementation of such a complex algorithms at cost of speed.

[2]A larger equalization range can provide higher data-rates.

configuration in the optical receiver [7], is that an analog equalizer is placed after the transimpedance amplifier (TIA). In this manner the signal-to-noise ratio (SNR) is maximized [6]. In the following sections, the design of every pre-amplifier stage will be separately discussed.

4.2.1 Photodiode and TIA design

Chapter 3 showed that a photodiode with minimal finger size has the highest intrinsic bandwidth and also the lowest roll-off per decade. Because of the maximal diode capacitance, the electrical diode bandwidth does have demand on input resistance of the TIA. Increasing the finger (nwell) size results in the lower capacitance and higher electrical bandwidth. The total roll-off in the diode characteristics is then 2-3 dB higher up to the low GHz range (see Figure 3.11), which decreases the maximum equalization frequency, as shown in chapter 6.

Subsequent to the photodiode, the transimpedance amplifier (TIA) is typically used current-to-voltage convertor for any optical receiver. The advantage is its high bandwidth, low noise, ease of biasing, and its relative insensitivity to photodiode and parasitic capacitance (having low input resistance). Typically, during the TIA design, the main tradeoffs are in sensitivity (due to noise), speed (bandwidth) and transimpedance gain. The transimpedance gain is typically equal to the feedback resistor for large open-circuit amplifications [8]. If the output voltage signal is small, further amplification is done by a post-amplifier. A large feedback resistance increases the gain, but at the same time, reduces the amplifiers' bandwidth by increasing the input resistance of the preamplifier [8].

In general, there are two transistor configurations for a TIA design: common source (CS) and common gate (CG) [8], [9]; they are shown in Figure 4.2. The implementation of one of the two configurations depends on the TIA performances: required transimpedance, bandwidth, noise, power consumption.

The calculated capacitances of the photodiodes shown in Table 3.3 imply that the input impedance of the TIA, independent on configuration should be low. For the CG-stage, the input impedance is approximated by: $Z_{\mathrm{in}} \approx A_v/R||1/g_{\mathrm{m1}}$, where A_{v} is the voltage gain. The input impedance can be decreased by increasing the amplification of the second amplifier stage. This amplifier is used in [10], as the regulated cascode (RGC) configuration, shown in Figure 4.3. The RGC mechanism enhances the input transonductance by the voltage gain $g_{\mathrm{mb}}R_{\mathrm{b}}$. As a result, the input impedance of the amplifier is very low ($R_{\mathrm{in}} \cong 0$) and the ef-

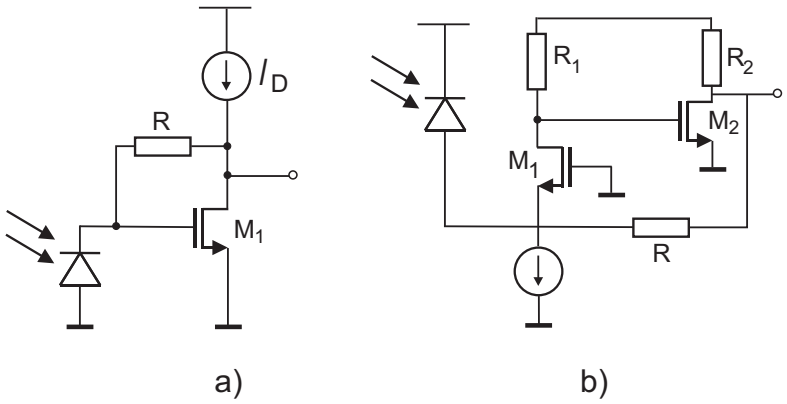


Figure 4.2: Low input impedance transimpedance amplifiers a) common source, b) common gate amplifier.

fect of the large input capacitance is minimized. However, there can be a large stability problem of the amplifier since there are two poles in the loop.

The main issue in the pre-amplifier design was to show the feasibility of the proposed diode bandwidth enhancement by using an analog equalizer. The rest of the electronic circuitry should be design without taking risk of circuit instability. Therefore, the authors have chosen a CS TIA configuration with one amplifier stage. If the transimpedance is not sufficient, a number of voltage amplifiers should be added subsequently to the designed TIA. This amplifier also met the bandwidth requirements (4 GHz) which is shown in the following subsection.

4.2.2 TIA design

Chapter 3 showed that the measured photocurrent from the nwell/p-substrate photodiode for 25 µW (-19 dBm average optical power) input optical signal is 10.2 µA. This corresponds to a diode responsivity of 0.41 A/W. The TIA should be designed such that the signal-to-noise ratio (S/N) is still sufficient for the required bit-error-rate at the certain speed (data-rate).

BER and Inter Symbol Interference

In high-speed data-communication, the achievable data-rates are closely linked to proper bit detection. Today's optical links requires bit error rates (BER)$\leq 10^{-11}$

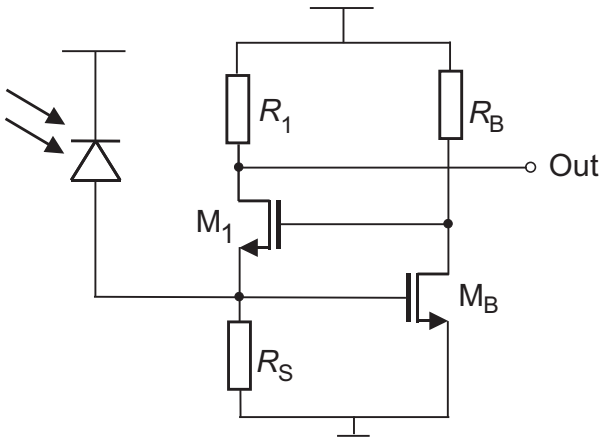


Figure 4.3: Schematic diagram of the RGC input stage.

[11]. The lower the signal-to-noise ratios (SNR) and the higher the intersymbol interference, the higher the bit error rates.

The amount of inter-symbol interference and consequently the theoretical bit error rates of the proposed pre-amplifier are calculated using the procedure described in [1]. We considered a baseband binary data at the transmitter side while a fixed threshold level is used to detect both logic levels (typically half the output value). The time domain current impulse response $J(t)$ is obtained from the inverse Laplace transform of the calculated total frequency response of photocurrent. The rms signal for a bit-period T_b is calculated as [1]:

$$S = \int_0^\infty J(t)[H(t) - H(t - T_\mathrm{b})]dt \tag{4.1}$$

where H denotes Heaviside function. Further, the bit error rate is determined as:

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\frac{S}{\sqrt{ISI_\mathrm{v}^2 + N^2}}\right) \tag{4.2}$$

where ISI_v is statistical variance of inter-symbol interference, S and N are rms signal and noise respectively. As given in [1], the probability of having a digital one n bits before the actual symbol is $P(n) = 1/2$.

Figure 4.4 shows the random bitstream BER of the equalized signal for several signal-to-noise levels. The data-rate is normalized on the electrical band-

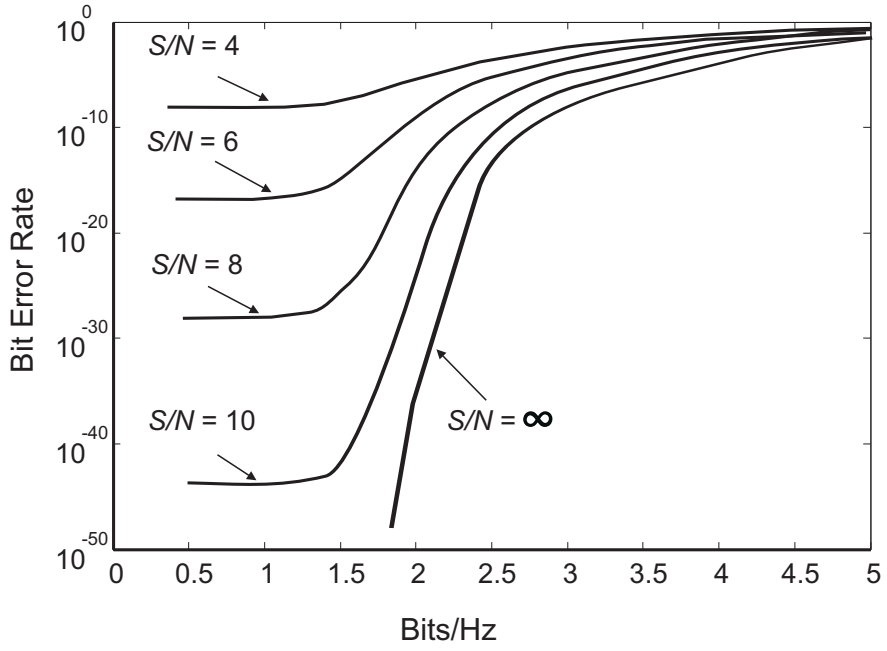


Figure 4.4: Bit error rate vs. normalized data-rate for several signal-to-noise ratios (S/N). Data-rate is normalized on total electrical bandwidth of the system (1.5 GHz).

width of the total system. For 1.5 GHz electrical bandwidth, and specified BER= 10^{-11}, the achievable data-rates are determined by signal-to-noise ratios (see Figure 4.4).

After calculating the theoretical BER for different data-rates and S/N ratios, the next step is to design the TIA. Figure 4.4 showes that BER=10^{-12} can be achieved with data-rate of 3 Gb/s, and a signal-to-noise ratio $S/N = 8$.

The calculated rms value of the square input signal is $S = 1/\sqrt{2} \cdot 10.2$ µA$\approx$ 8 µA. For a $S/N = 8$ and a bandwidth of 1 GHz, the calculated rms value of the input referred noise is $I_N = 28$ pA/$\sqrt{\text{Hz}}$. Using Figure 4.5, the input referred current noise can be approximated as:

$$\frac{I_{\text{N}}^2}{BW} \approx \overline{I_{\text{R}}^2} + \overline{I_{\text{n1}}^2} + 2 \cdot \overline{I_{\text{n2}}^2} \tag{4.3}$$

$$\approx \left(\frac{4kT2g_{\text{m1}}}{3} + \frac{4kT2g_{\text{m2}} \cdot 2}{3} \right) \frac{g_{\text{m1}}}{(g_{\text{m1}}R - 1)^2} + \frac{4kT}{R} \tag{4.4}$$

where g_{m1} and g_{m2} are the transconductances of transistors M_1 and M_2. The

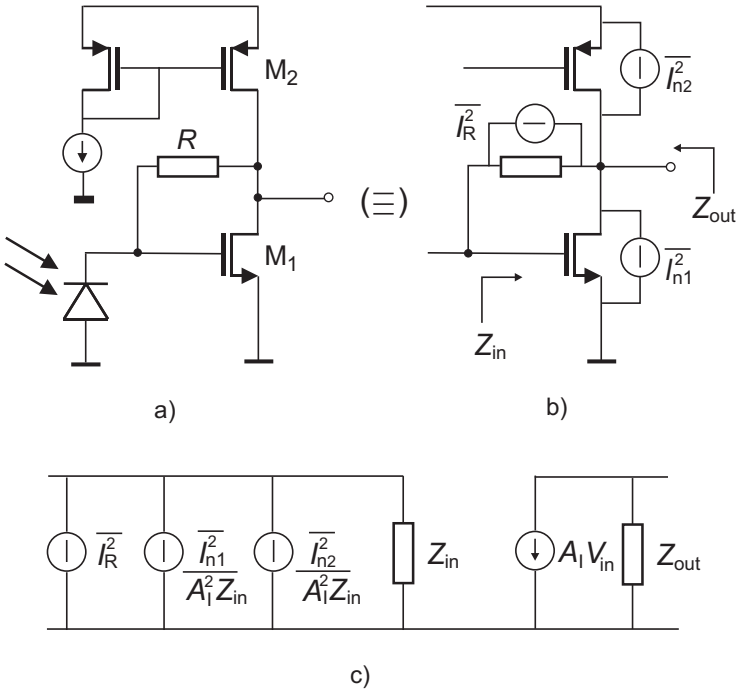


Figure 4.5: The simplified noise model of the CS TIA.

input PMOS transistor in the current mirror is designed to have the same aspect-ratio as the output one. As a result, the transconductances of both transistors will be approximately equal. For the simplicity of the noise calculation, as shown in Figure 4.5, we approximated that the output PMOST is twice as noisy $2\,\overline{I^2_{\rm n2}}$. The input and output impedances are approximated to $Z_{\rm in}{=}Z_{\rm out}{\approx}1/g_{\rm m1}$. The equation (4.4) shows that the input referred noise is directly proportional to the transistors' transconductances and inversely proportional to the feedback resistor R.

For the amplifier bandwidth performance, the input impedance is calculated as:

$$Z_{\rm in} = \frac{R_0 + R + sRR_0C_0}{sC_{\rm in}R_0 + sC_{\rm in}R + s^2C_{\rm in}RR_0C_0 + 1 + sR_0C_0 + g_{\rm m}R_0} \tag{4.5}$$

where R_0 and C_0 are output resistance and capacitance of the transistor M_2, $C_{\rm in}$ is total input capacitance of the amplifier, which is the sum of diode capacitance $C_{\rm diode}$ and input capacitance of the TIA $C_{\rm TIA}$. For very low frequencies (s$\sim$0),

the input impedance is:

$$Z_{\text{in}} = \frac{R_0 + R}{1 + g_{\text{m}} R_0} \approx \frac{1}{g_{\text{m}}} \tag{4.6}$$

This input impedance should be constant in the frequency range of interest (up to equalization frequency). Table 3.3 show the nwell/p-substrate photodiode capacitance for 2 µm nwell width and 50 µm× 50 µm area. This area corresponds to the diameter of the multimode fiber. The capacitance is $C_{\text{diode}} = 1.6$ pF. The required bandwidth is larger than 2 GHz, meaning that the input resistance of the TIA should be lower than 50 Ω ($\approx 1/g_{\text{m1}}$). However, the input capacitance of the TIA C_{TIA} should be added to the total input capacitance, which results in a lower input resistance. The lower resistance is achieved by increasing the g_{m1} value.

The transconductance g_{m1} can be increased by either increasing the aspect ratio W/L of the transistor or by increasing the bias current [8]. The larger the aspect-ratio of the input transistor, the larger the g_{m}, but also the input capacitance of the TIA C_{TIA}. This capacitance can start dominating the total input capacitance C_{in}, which deteriorates the input signal. In best case, the g_{m} should be increased until the ratio $g_{\text{m}}/C_{\text{in}}$ is maximized. The bias current I_{D} is in the low mA range for the required g_{m1} value ($< 40\Omega$). This bias current flows also through PMOS transistor M_2 which has large g_{m2}; this transconductance is comparable to the value of g_{m1} so the transistor M_2 has also significant noise contribution, as shown in equation (4.4).

For the better mirror-performances the length of the transistor M_2 should be larger than minimal. Knowing that the drain current is in the mA range, the aspect. The biasing current flowing through PMOS transistor is in the mA range, and the larger the current the lower the output impedance of the PMOS as shown in Figure 4.6. Thus, the larger the g_{m2}, the smaller its output impedance, and higher the input impedance of the TIA (from equation 4.5); as a result the bandwidth of the TIA decreases.

The larger the g_{m1} and g_{m2} the larger is the input referred noise in the circuit, as shown in equation 4.4. The larger the g_{m} of the transistors M_1 and M_2, the larger the bias current I_{D}. This results in higher power consumption of the circuit. Typical power consumption of the pre-amplifier is in the range 5 mW-20 mW [1], [9].

For the $1/g_m \leq 40\Omega$, the TIA is designed in several steps. If the transistor M_1 is chosen to have the aspect ratio W/L=62/0.18 and overdrive-voltage of V_{gt1}=0.35 V, the calculated value for $g_{\text{m}}\sim$30 mS. The input capacitance of

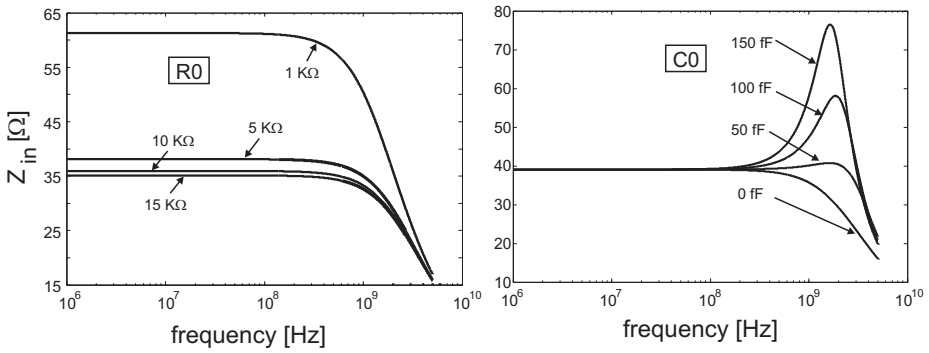


Figure 4.6: The influence of the output resistance and capacitance of the PMOST M_2 showed in Figure 4.5a, on the TIA's input impedance.

the TIA is 150 fF, including the Miller capacitance at the TIA's input. The biasing current is 6 mA. The transistor M_2 is biased such that $V_{\rm gt2} > V_{\rm ds2} = -1.05$ V. With a current of 6 mA, the transistors are sized W_2/L_2=170/0.3 (non-minimal gate length for better mirror-performance). The output resistance of these transistors is approximately R_0=1.7 kΩ and the capacitance is 130 fF. Inserting these values in the equation for the total input impedance of the TIA (4.5), the maximum transresistance value is only $R = 350$ Ω.

One way to increase the transresistance, is to increase the $g_{\rm m1}$ of the input transistor to 35 mS i.e. the aspect ratio is W/L=78/0.18 for the overdrive voltage of V_{gt1}=0.35 V. The input capacitance of the TIA is slightly increased[3]. The bias current is consequently increased to 7 mA. This allowed the transimpedance gain of 0.9 $k\Omega$. However, this transimpedance value was not sufficient, and voltage amplifier should be placed after the TIA. For the best performance, the biasing of the subsequent stage is such that the overdrive voltage of the M_1 transitor in TIA is decreased to V_{gt1}=0.15 V. The aspect ratio of the transistor M_1 is W/L=133/0.18, and the transconductance $g_{\rm m1}$=45 mS. The calculated input resistance of the TIA is 22 Ω, and the bandwidth of the TIA is 2.5 GHz. The calculated power consumption of the TIA is 12.6 mW. The bias current is large enough to overdrive the DC offset photocurrent. The output signal from the VCSELs is usually with modulation index of 0.2 [12]. With photo-responsivity of 0.41 A/W, the calculated DC photocurrent for 25 µW AC optical power is 56 µA, which is less than 1% of the input bias current. Using equation (4.4),

[3]still mainly determined by the photodiode capacitance.

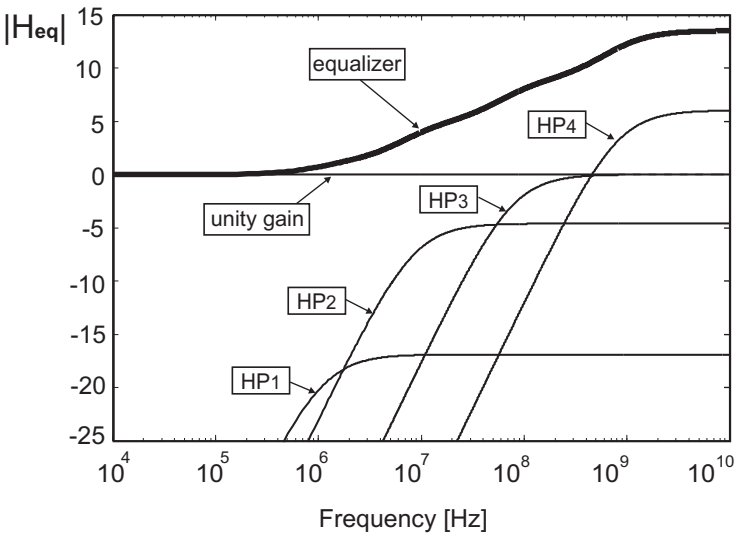


Figure 4.7: The analog equalizer transfer characteristics: unity gain path summarized with 4 high-pass sections.

the calculated input referred noise is 10 pA/$\sqrt{\text{Hz}}$, which is lower than required. However, in the detail noise calculation the noise source of the input PMOS transistor in the current mirror should be included as well as the input impedance of the next stage of the TIA. Nevertheless, the calculated noise value leaves enough margin for the required signal-to-noise ratio $S/N = 8$. One should also notice that the designed TIA is similar to a LNA configuration presented in [13].

4.2.3 Equalizer design

Once the electrical diode bandwidth is optimized, an analog equalizer is designed as the next stage of the pre-amplifier. Chapter 3 described that the total frequency response of the photodiode is the sum of the frequency responses of particular diode regions. This implies that the equalizer should also be designed as the sum of several high-pass filters responses with an overall low *roll-up* ($\sim$ 5 dB/decade). The equalization characteristics in this thesis is inverse of the calculated frequency characteristics of the implemented photodiode. One way to mimic a low roll-up characteristics is summation of the outputs of four parallel first-order high-pass filters (HPF); this is illustrated in Figure 4.1. The equalizer characteristic is illustrated in Figure 4.7.

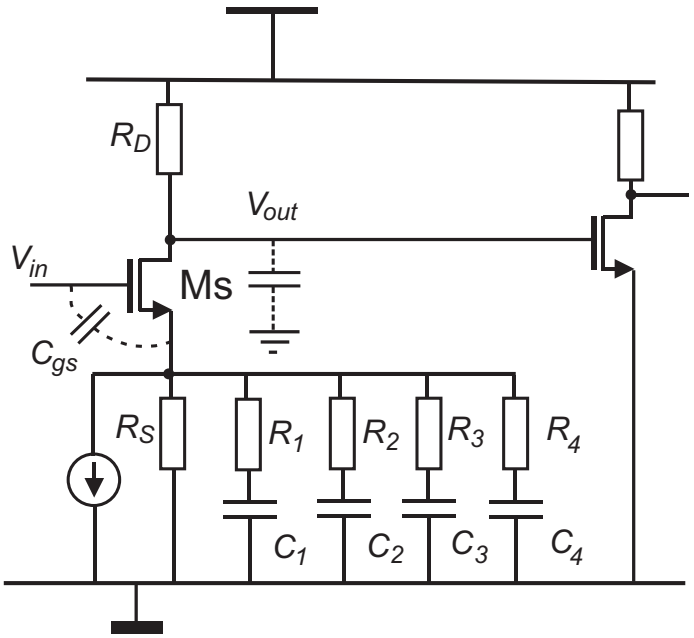


Figure 4.8: The analog equalizer from Figure 4.7 including parasitic capacitances.

One way to realize the analog equalizer is to use a source degeneration (SD) stage with low-pass filter sections in its source [6]; this configuration is shown in Figure 4.8. For unity gain of the amplifier, $R_\mathrm{D} = R_\mathrm{S}$, the transfer function $V_\mathrm{out}/V_\mathrm{in}$ of the equalizer can be written as:

$$\frac{V_\mathrm{out}}{V_\mathrm{in}} \approx -\left(1 + \frac{sR_\mathrm{D}C_1}{1+sR_1C_1} + \frac{sR_\mathrm{D}C_2}{1+sR_2C_2} + \frac{sR_\mathrm{D}C_3}{1+sR_3C_3} + \frac{sR_\mathrm{D}C_4}{1+sR_4C_4}\right) \quad (4.7)$$

The output pole ω_p of the SD stage itself is determined by the output capacitance C_out and the resistance in the drain R_D: $f_\mathrm{p} = 1/R_\mathrm{D}C_\mathrm{out}$. This pole should be high enough not to interfere with the equalization range of the photodiode characteristics, $f_\mathrm{p} > f_{3\mathrm{dB_{elec}}}$.

The lower the R_D, the higher the bandwidth of the SD stage. However, a lower R_D results in lower resistance values $R_{1..4}$ (see equation 4.7) necessary to provide the required roll-up in the characteristics. For fixed $1/R_{1...4}C_{1...4}$ poles, and lower values of $R_{1..4}$ of the equalizer, the capacitance values are higher; this will result in more *area consumption* in the chip, but also a better noise

performance.

According to the diode frequency characteristics shown in Figure 3.14, the total amplitude drop up to 1 GHz is about 15 dB. In order to obtain proper equalization characteristics, the ratio R_{D}/R_4 in equation (4.7) has to be about 4. This is achieved either with large R_{D} or with a low R_4 values. The lower the R_4, the lower the impedance in the source of source-degeneration (SD) stage. On high frequencies where the gate-source capacitance $C_{\mathrm{gs_{Ms}}}$ starts to dominate, the low source impedance lowers the input impedance of the SD. As a result, the bandwidth is decreased. This input impedance of the source-degeneration stage can be expressed as:

$$Z_{\mathrm{in}} = R_{\mathrm{S}}||Z_{\mathrm{eq}} + \frac{1}{sC_{\mathrm{gs}}}[1 + g_{\mathrm{m_{Ms}}}R_{\mathrm{S}}||Z_{\mathrm{eq}}] \tag{4.8}$$

where $Z_{\mathrm{eq}} = Z_1||Z_2||Z_3||Z_4$ is the impedance of the HP filter stages, and $Z_i = R_i + 1/(sC_i)$, for $i = \{1, 2, 3, 4\}$. For rather constant input impedance in the frequency range of interest, C_{gs} should be as small as possible and Z_{eq} should be much larger than R_{S}. Small C_{gs} is related to the small transistor aspect ratio (W/L) which on the other side can limit the g_{m} and the accuracy of the transfer function (4.7).

According to the previous analyzes, for area and power efficiency reasons, the highest zero can be implemented in the circuit using e.g. inductive peaking stage [14], as shown in Figure 4.9. The $g_{\mathrm{m}}C_{\mathrm{gs}}$ combination of the transistor M_{p} together with resistor R, in a certain frequency range behaves as an inductor. The impedance in the drain of transistor M_1 is given as:

$$Z_{\mathrm{d}} = \frac{1}{g_{m_{\mathrm{Mp}}}} \frac{1 + sRC_{\ \mathrm{gs}}}{1 + sC_{\ \mathrm{gs}}/g_{m_{\mathrm{Mp}}}} \tag{4.9}$$

where $R > 1/g_{\mathrm{m}}$. In this manner the circuit with the analog equalizer with one inductive peaking stage and the source degeneration with the first three poles of the equalizer is designed. The overall circuit is shown in Figure 4.9. The simulated frequency response at the pre-amplifier's output, including the photodiode, is shown in Figure 4.10.

With the circuit of Figure 4.9 it is fairly straightforward to compensate the diode characteristic for frequencies where the roll-off is low (for $\leq$5 dB/decade). The equalizer's frequency response is band-limited in order to prevent out of band high-frequency noise from being added to a signal [15].

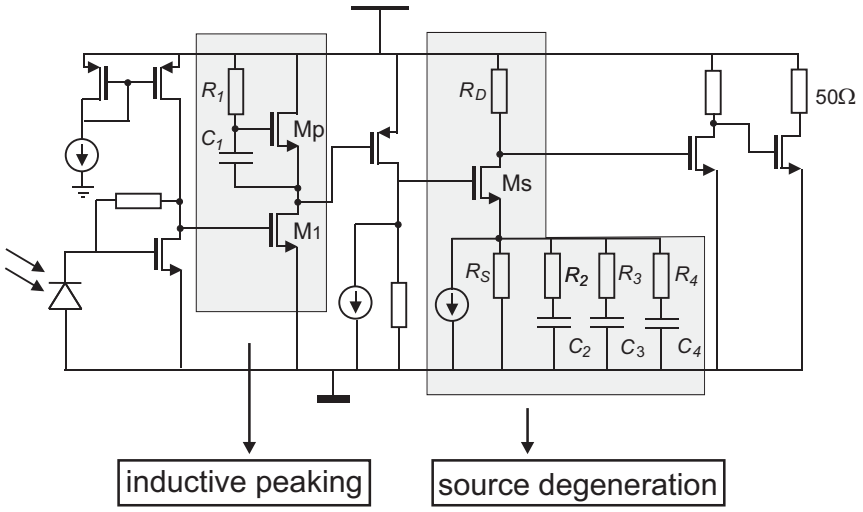


Figure 4.9: The circuit topology of the preamplifier including the analog equalizer.

4.2.4 Robustness on spread and temperature

This chapter presents a major advantage of the proposed pre-amplifier circuit with the analog equalizer: its robustness on spread and temperature. First, the robustness on typical $\pm 20\%$ spread in RC components of the equalizer is analyzed.

A common feature for the calculated characteristics of different photodiode structures shown in Figures 3.12, 3.13, 3.19 and 3.23 is a very low roll-off ($R_{\text{off}} < 6$ dB/decade). The roll-off in these characteristics is compensated by implementing inverse equalization characteristics, as shown in Figure 4.11. Nevertheless, there is a spread in the equalizer components that changes the poles in the equalizer characteristics for $\pm p$, where p is the spread value, typically in the range $\{-0.2, 0.2\}$. As a result, there will be amplitude shifts in the equalized characteristics, which is shown in Figure 4.11.

The maximum amplitude shift A_{shift} (expressed in dB) can be calculated as:

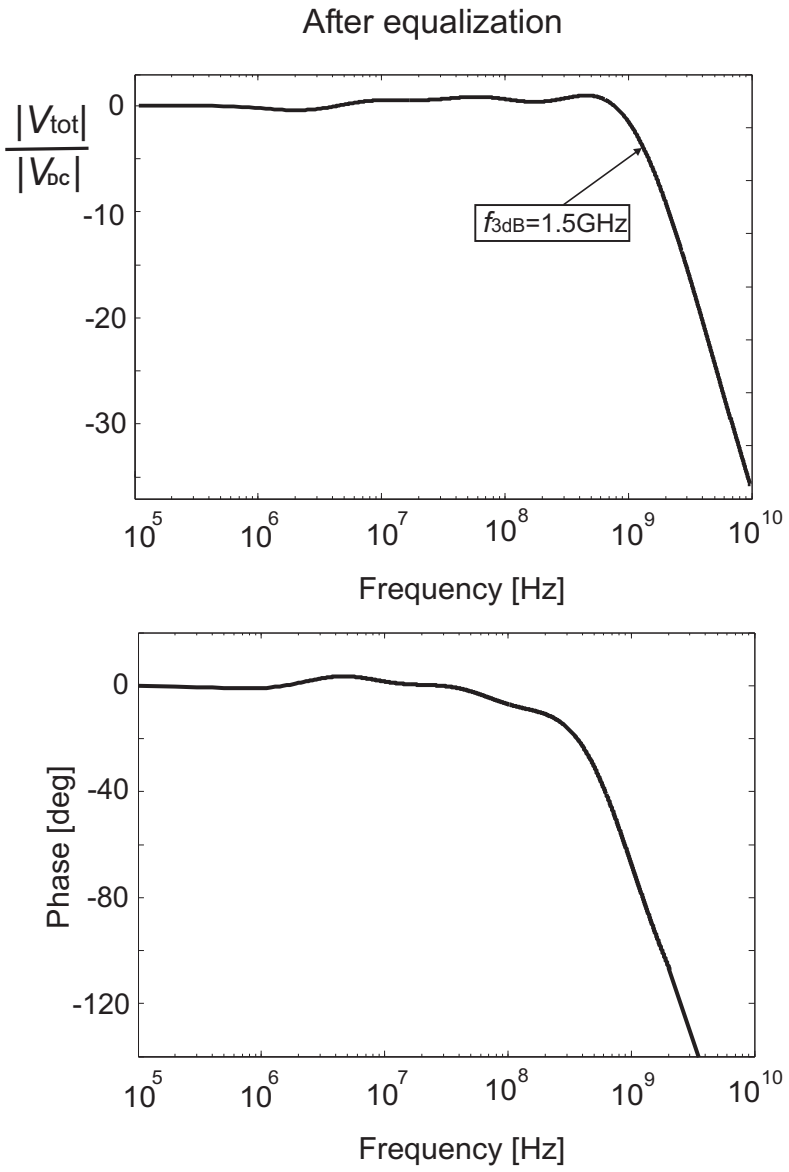


Figure 4.10: The simulated amplitude and phase responses of the photodiode and pre-amplifier after the analog equalizer.

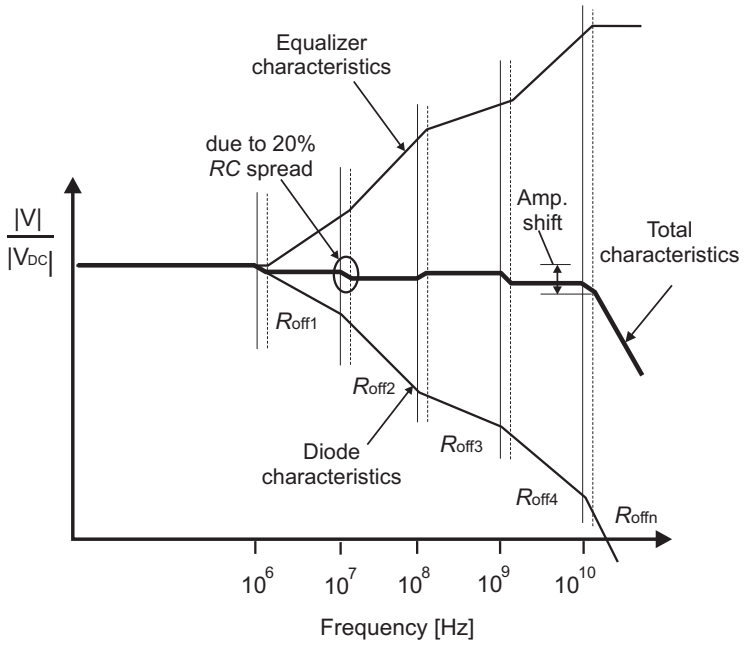


Figure 4.11: The stylistic plot of the amplitude shift in the equalized (total) characteristics due to the RC spread in the equalizer. The amplitude shift is roll-off (R_{off}) dependent.

$$\begin{aligned} A_{\text{shift}} &= \max\left(\sum_{i=1}^{n}(R_{\text{off}_i} - R_{\text{off}_{i-1}})\log_{10}(1+p)\right) \\ &- \min\left(\sum_{i=1}^{n}(R_{\text{off}_i} - R_{\text{off}_{i-1}})\log_{10}(1+p)\right) \end{aligned} \tag{4.10}$$

According to the calculated characteristics of photodiodes analyzed in chapter 3, the maximum amplitude shift up to 1 GHz range is lower than ± 0.5 dB for $\pm 20\%$ RC spread in the equalizer. The small amplitude shift slightly modifies the output signal as well as the total bit-error rate in the system, as shown in equations (4.1) and (4.2).

In optical communication systems and optical detection in general, there is always specified value for BER [11], [16], [17]. Adapting the modified BER value back to the specified one, results in the change of the data-rate. Using equations (4.1) and (4.2), the BER is calculated for a total amplitude shift of 0.5 dB (the

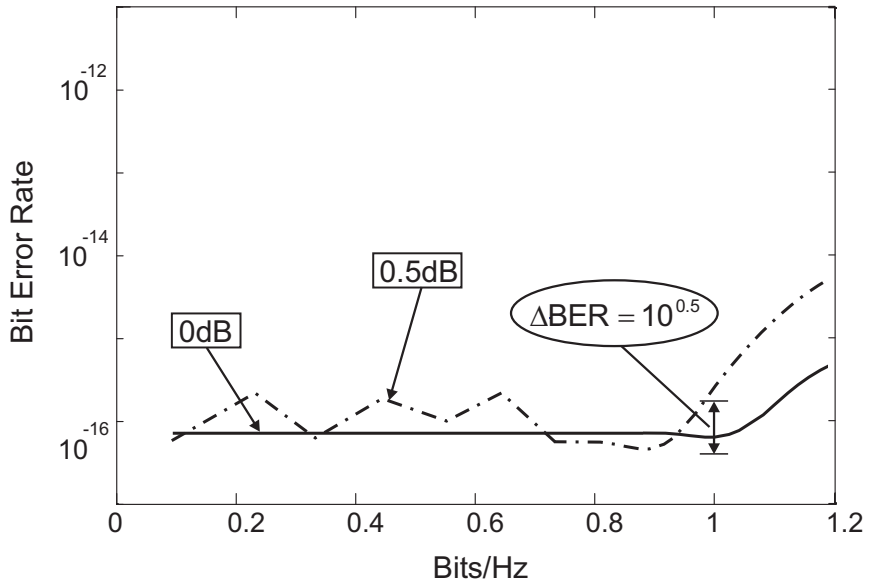


Figure 4.12: A change in BER due to the amplitude shift in the equalized photodiode characteristics. The data-rate is normalized with the upper equalization frequency.

amplitude characteristics has the roll-off of 6 dB/decade). The result is shown in Figure 4.12. The BER is increased only $\Delta\text{BER}\sim 10^{0.5}$. Adopting this BER value to its original one (without RC spread), the maximum data-rate decreases 6 %. Therefore, the proposed pre-amplifier is very robust on the spread, due to the low roll-off in the diode characteristics.

Robustness on temperature

The frequency characteristics of the photodiodes analyzed in chapter 3, is calculated again following the same procedure with the included change of the temperature in environment. The temperature range is taken to be $T = 250$–350 K. The temperature dependent terms in photocurrent equations (3.8), (3.15) and (3.23) are diffusion lengths L_{n} and L_{p}. The diffusion lengths are defined as $L_{\text{n,p}} = \sqrt{D_{\text{n,p}}\tau_{\text{n,p}}}$ [18], where $D_{\text{n,p}}$ is the diffusion coefficient of the n, p minority carriers respectively and $\tau_{\text{n,p}}$ is the minority carrier lifetime. The diffusion coefficients are presented with equations (4.11), (4.12). These coefficients are product of the term that linearly depend on temperature and the mobility of the minority carriers. These mobilities are also temperature dependent as shown in

equations (4.13), (4.14) [18]. The nominator of the second terms in both equations decreases dominantly with temperature (with $T^{-2.33}$). As a result, the overall mobility of the minority carriers drops with temperature which is shown in Figure 4.13. The same figure presents that the higher the doping concentration ($N_{\rm a}$, $N_{\rm d}$), the lower the decrease of the mobilities with temperature. The doping concentrations in CMOS photodiodes are such that the mobility dependence counteracts the linear temperature dependence in diffusion coefficients.

Figure 4.13 shows that the change of the diffusion coefficients with temperature is maximal for the lowest doped substrate. For the nwell and p+ regions, the diffusion coefficient does not change more than 20%. The corresponding diffusion lengths $L_{\rm n,p}$ having a square-root dependence on the diffusion coefficients, will have even smaller temperature dependence. Inserting the values of the diffusion lengths in the photocurrent equations (3.8), (3.15) and (3.23), the deterministic gain error is only $0.2-0.3$ dB/decade. This implies that proposed pre-amplifier is temperature robust too.

$$D_{\rm n}(N_{\rm a},T) = \frac{kT}{q}\mu_{\rm n}(N_{\rm a},T) \tag{4.11}$$

$$D_{\rm p}(N_{\rm d},T) = \frac{kT}{q}\mu_{\rm p}(N_{\rm d},T) \tag{4.12}$$

$$\mu_{\rm n}(N_{\rm a},T) = 88t_{\rm n}^{-0.57} + \frac{7.4\cdot 10^{8}T^{-2.33}}{1+\dfrac{N_{\rm a}}{1.26\cdot 10^{17}t_{\rm n}^{2.4}}0.88t_{\rm n}^{-0.146}} \tag{4.13}$$

$$\mu_{\rm p}(N_{\rm d},T) = 54.3t_{\rm n}^{-0.57} + \frac{1.36\cdot 10^{8}T^{-2.33}}{1+\dfrac{N_{\rm d}}{2.35\cdot 10^{17}t_{\rm n}^{2.4}}0.88t_{\rm n}^{-0.146}} \tag{4.14}$$

where k is the Boltzman constant, $N_{\rm a}$, $N_{\rm d}$ are the doping concentrations of the p and n-region in the photodiode, and $t_n = T/300$.

4.3 Results

The integrated photodiode and pre-amplifier were fabricated in a standard 0.18 µm CMOS technology. The technology has five metal layers and one polysilicon layer available. The chip micrograph of the integrated nwell/p-substrate photodiode and the pre-amplifier is shown in Figure 4.15. The total circuit area is 145×305 µm^2. The overall area including bondpads is 0.7×0.4 mm^2.

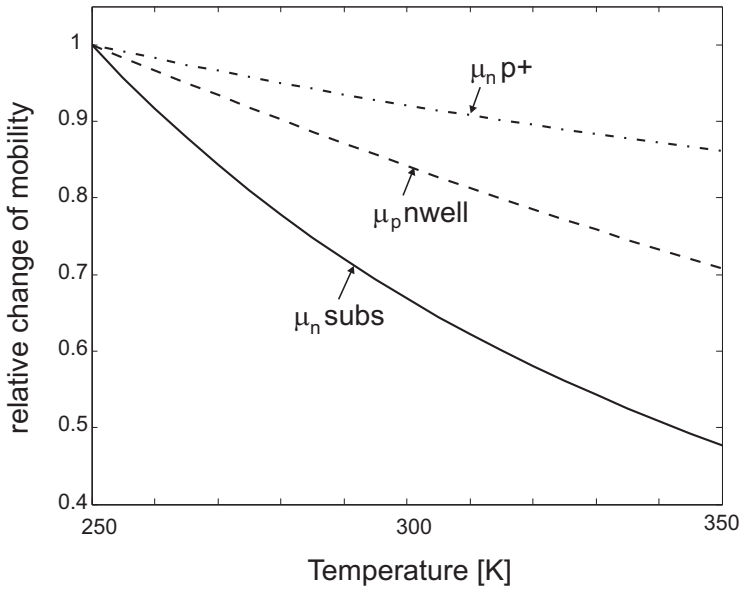


Figure 4.13: Normalized change of the mobility of minority carriers inside diode regions with temperature. The values are normalized with their maximum values.

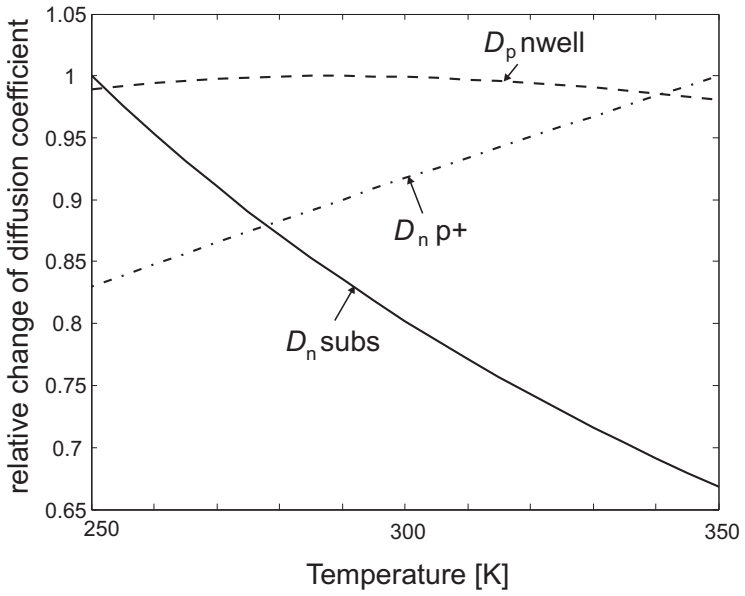


Figure 4.14: Normalized change of the diffusion coefficient of minority carriers with temperature. The values are normalized with their maximum values.

A minimal nwell-width finger photodiode (2 μm) is used as a photodetector (illustrated in Figure 3.2). The size of the photodiode is $50 \times 50 \mu m^2$ corresponding to the diameter of the multimode fiber used to couple the light from laser to the chip. The calculated diode capacitance is 1.6 pF. The input resistance of the TIA, realized as a common source stage with resistive feedback, is 33 Ω.

An 850-nm VCSEL is used as a light source for testing the system performance. The 50 Ω output buffer (shown in Figure 4.9) is AC coupled to the input of the digital oscilloscope with the external coupling capacitance of $C_{\text{ext}} = 0.47$ µF. The bias voltage is 1.8 V.

Figure 4.16 shows the eye diagram of the integrated pre-amplifier *without* and *with* the analog equalizer. For the system without equalizer, the calculated maximal speed for BER$<10^{-11}$ is 10 Mb/s. However, an eye-diagram for 50 Mb/s input with BER$= 10^{-7}$ is measured since that is a minimal speed of the used digital data analyzer. The input light power is 25μW peak-to-peak, which is -19 dBm average optical power. For comparison, the required optical signal power for Gigabit Fiber Ethernet systems [11] is about -17 dBm, meaning that the used optical power is sufficient.

It is important to notice that this is the AC optical power of the signal. Typically, the index modulation of of the lasers are between 0.2-0.5. The VCSEL LV1001, $\lambda = 850$ nm, produced by OEPIC company [12], that is used in our experiment has index of modulation about 0.2. Thereby, the biasing point of the laser is such that the DC optical power is four times larger than AC power. This DC optical power causes a DC photocurrent that behaves as offset current.

The eye-diagram shown in Figure 4.16b presents the achieved 3 Gb/s using the analog equalizer in the pre-amplifier. The input light power (AC) is again 25 µW peak-to-peak and the achieved BER$<10^{-11}$. This is an order of magnitude larger BER than calculated (10^{-12}). This difference is due to the noise in the circuit supply, but also due to the limiting amplifier. The BER measurements are described in section 4.3.2. The complete receiver consumes approximately 34 mW + 16 mW for the 50 Ω output buffer realized as a common source stage (see Figure 4.9).

4.3.1 Robustness of the pre-amplifier: simulations and measurements

Section 4.2.4 showed that the big advantage of proposed pre-amplifier circuit with an analog equalizer is its robustness on spread and temperature. In order to

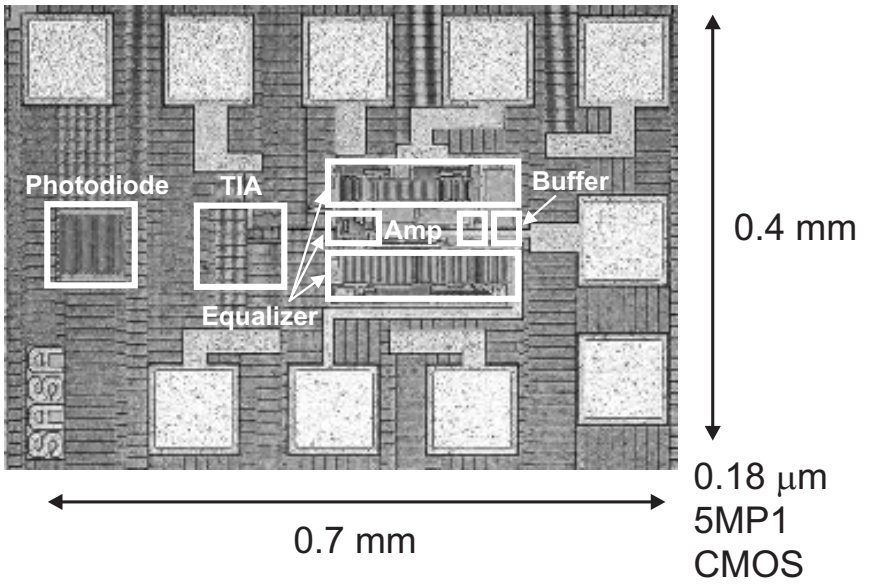


Figure 4.15: Chip micrograph of the integrated nwell/p-substrate photodiode and pre-amplifier with an analog equalizer in a fully standard CMOS.

confirm this, a number of simulations and measurements are presented further in the following sections section.

A. Robustness on spread - *simulations*

The effect of typical ±20% spread in the RC components of the equalizer on the total transfer characteristics, is presented first with asymptotic approximation shown in Figure 4.17. Due to the low roll-off in the photodiode characteristics, the total amplitude shift is less than ±0.5 dB. This is also confirmed with a simulated symbol response (in time) of the pre-amplifier. The spread included is ±20% and the obtained result is presented in Figure 4.18. This figure illustrates that the robustness against the spread is high i.e. the equalization is spread insensitive, for the expected spread during fabrication in standard CMOS.

The same simulations were done including ±50% component spread in the equalizer. First, it is assumed that there is no spread in the equalizer components $RC_{\text{filter}_{\text{ideal}}}$. Then the spread is included in the equalizer components $RC_{\text{filter with spread}}$. The simulation results are presented in Figure 4.19. This figure shows that for RC spread of ±30% the output signal amplitude is changed

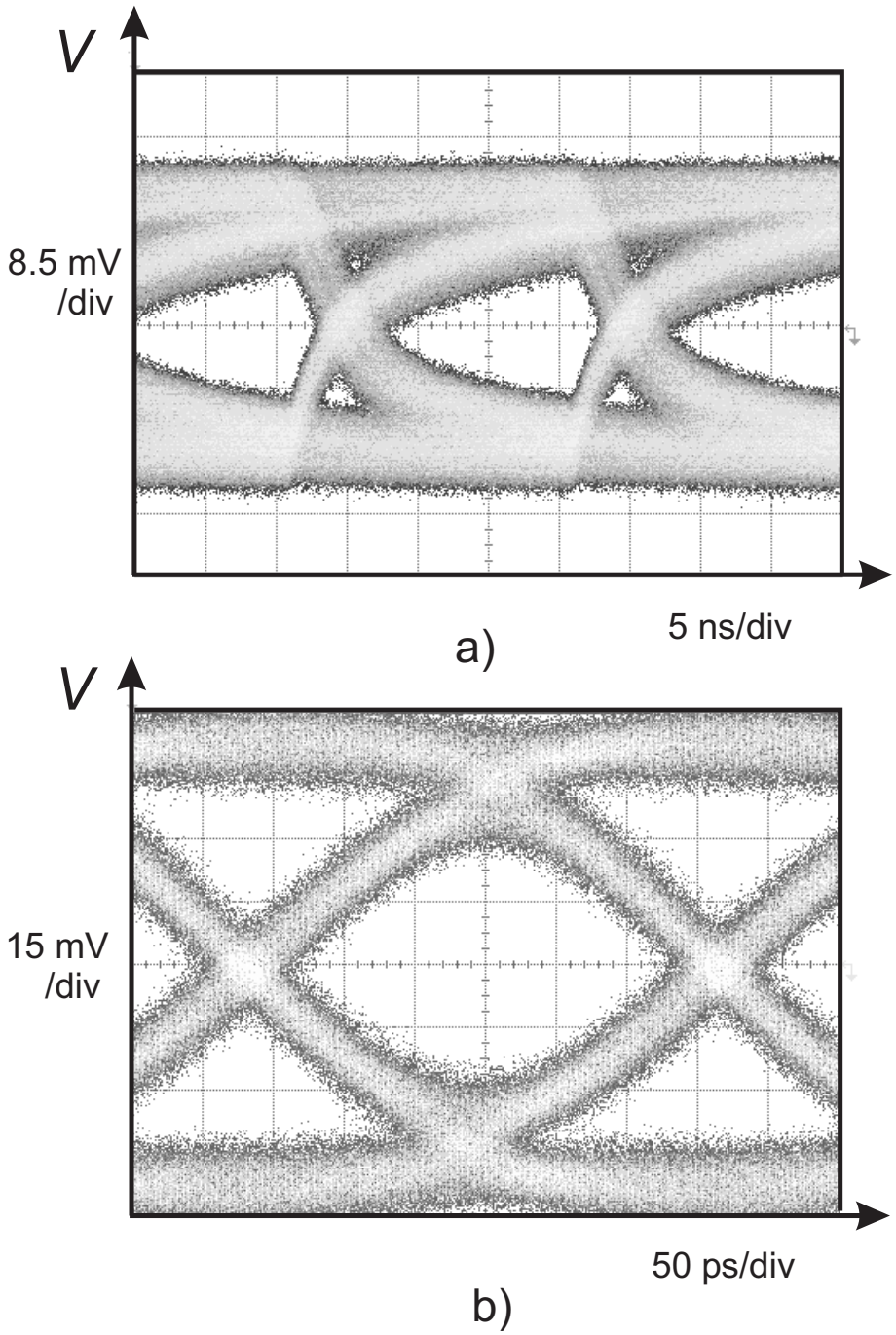


Figure 4.16: Eye-diagram of the nwell/p-substrate CMOS photodiodes a) without equalizer 50 Mb/s, BER=10^{-7}, b) with an analog equalizer 3Gb/s, BER=10^{-11}.

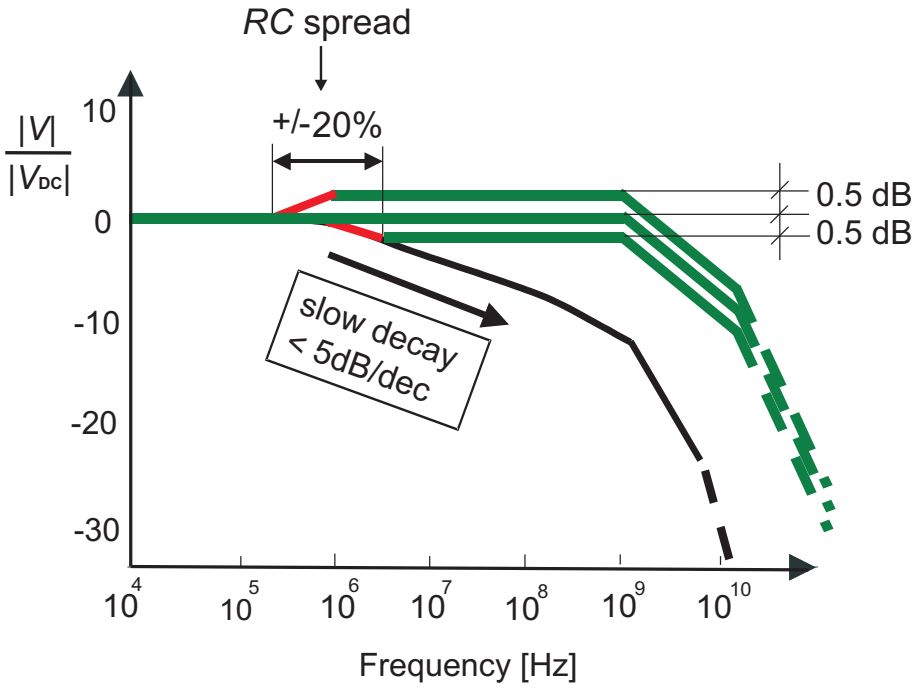


Figure 4.17: The asymptotic approximation of the ±20% *RC* spread-effect in the equalizer on the total pre-amplifier transfer characteristics. The amplitude shift is less than ±0.5 dB.

only 10%. However, the maximum output signal is very dependent on the biasing R_{bias} of the circuit shown in Figure 4.7, which is further sensitive on the spread. For the expected ±20% spread in biasing resistances, the maximum output signal can change up to 20% (see Figure 4.19).

B. Robustness on spread - *measurements*

The robustness of the pre-amplifier circuit on spread is confirmed with the measurements on the circuit shown in Figure 4.15. During the chip-layout design, all *RC* filter components are placed as a number of smaller components (fingers) connected using the highest metal layer. Removing some of this metal-layer connections, the *RC* values can be changed. In the experiment, we changed the values for ±20% after the fabrication using the Focused Ion Beam (FIB). The eye-diagrams are measured at the output of the pre-amplifier for 3 Gb/s data-rate, and the results are shown in Figure 4.20. In order to compare this measured

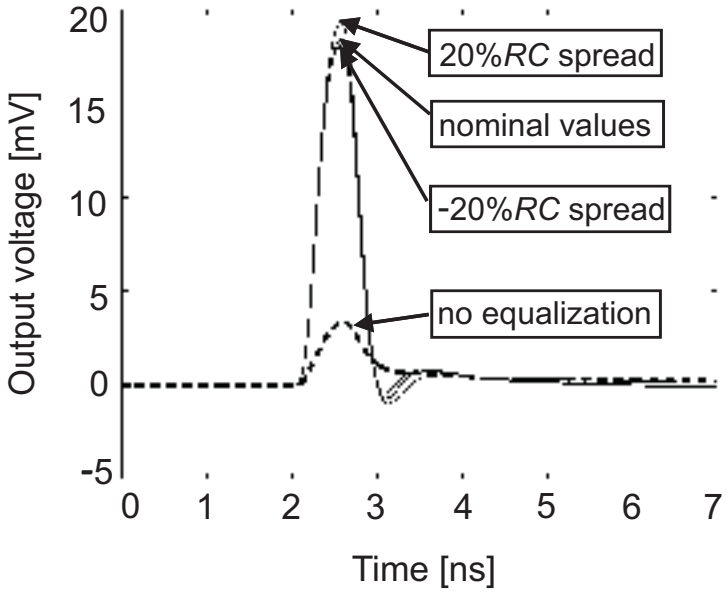


Figure 4.18: The symbol response of the pre-amplifier including the $\pm 20\% RC$ spread in the equalizer. The amplitude is change for ± 0.45 dB.

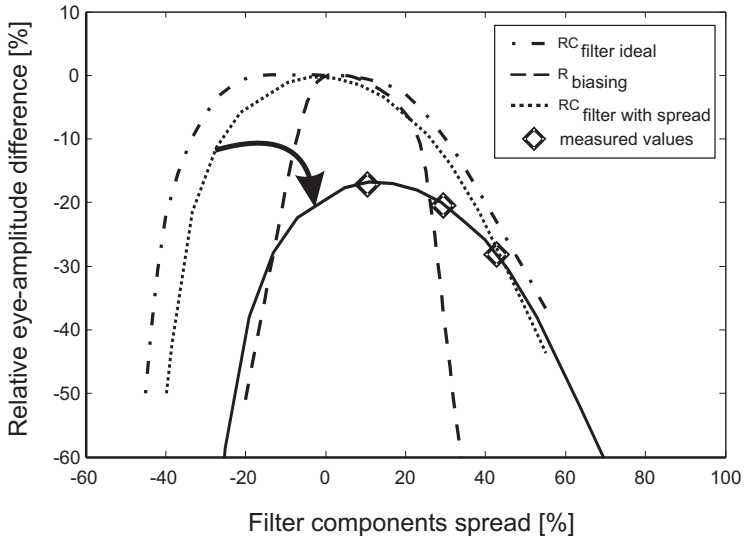


Figure 4.19: Simulated relative eye-amplitude change [%] at the equalizer's output vs. spread in RC filters as well as spread in R biasing of the circuit.

results with the simulation, the simulated curve with the included spread in the equalizer $RC_{\text{filter with spread}}$, is shifted down the biasing curve R_{biasing}. This is illustrated in Figure 4.19. The R_{biasing} curve includes the spread influence of the source degeneration (SD) resistances R_D, R_S and the resistance of the PMOS amplifier in front of the SD stage (see Figure 4.9). The measured results show that the circuit is not optimized for the maximum output signal. The values of the biasing resistors should be improved (changed for about -25%, or using even different biasing) for the better circuit performance.

According to Figure 4.20, the eye-diagram amplitude including the modified RC values in the equalizer, is changed less than 6%. Therefore, these measurements confirm a high robustness of the pre-amplifier on spread in the equalizer.

C. Pre-amplifier with various photodiode structures

In order to measure the impact of photodiode spread, the optical detector circuit was also implemented using a different photodiode, with the same pre-amplifier circuit. This section presents the pre-amplifier integrated with p+/nwell/p-substrate photodiode (double photodiode). The same pre-amplifier and analog equalizer used with nwell/p-substrate diode (shown in Figure 4.9) are used here. The filter parameters in the equalizer are thus, not optimized for the double photodiode characteristics shown in Figure 3.23. In this manner, the feasibility of the implementation of the pre-amplifier with other photodiode structures is tested. The result is presented in Figure 4.21.

The achieved data-rate is 2.5 Gb/s with 38.5 µW optical power. With the expense of approximately 2 dB higher input optical power, but without optimizing the equalizer, a very high data-rate is achieved. By optimizing the HF filter parameters, the simulated data-rate of the system is also 3 Gb/s for the previously used optical power of 25 µW.

4.3.2 BER measurements

In the proposed pre-amplifier, an approximate value of the signal-to-noise ratio can be deduced from Figure 4.16. However, having some intersymbol interference, the receiver is evaluated with a BER measurement.

An 850 nm VCSEL LV1001 from OEPIC company [12] is used as a light source for measuring the BER of the integrated pre-amplifier and photodiode showed in Figure 4.15. The light was coupled from the laser to the photodiode using multimode fiber with 50 µm core-diameter. The fiber length is 1 m.

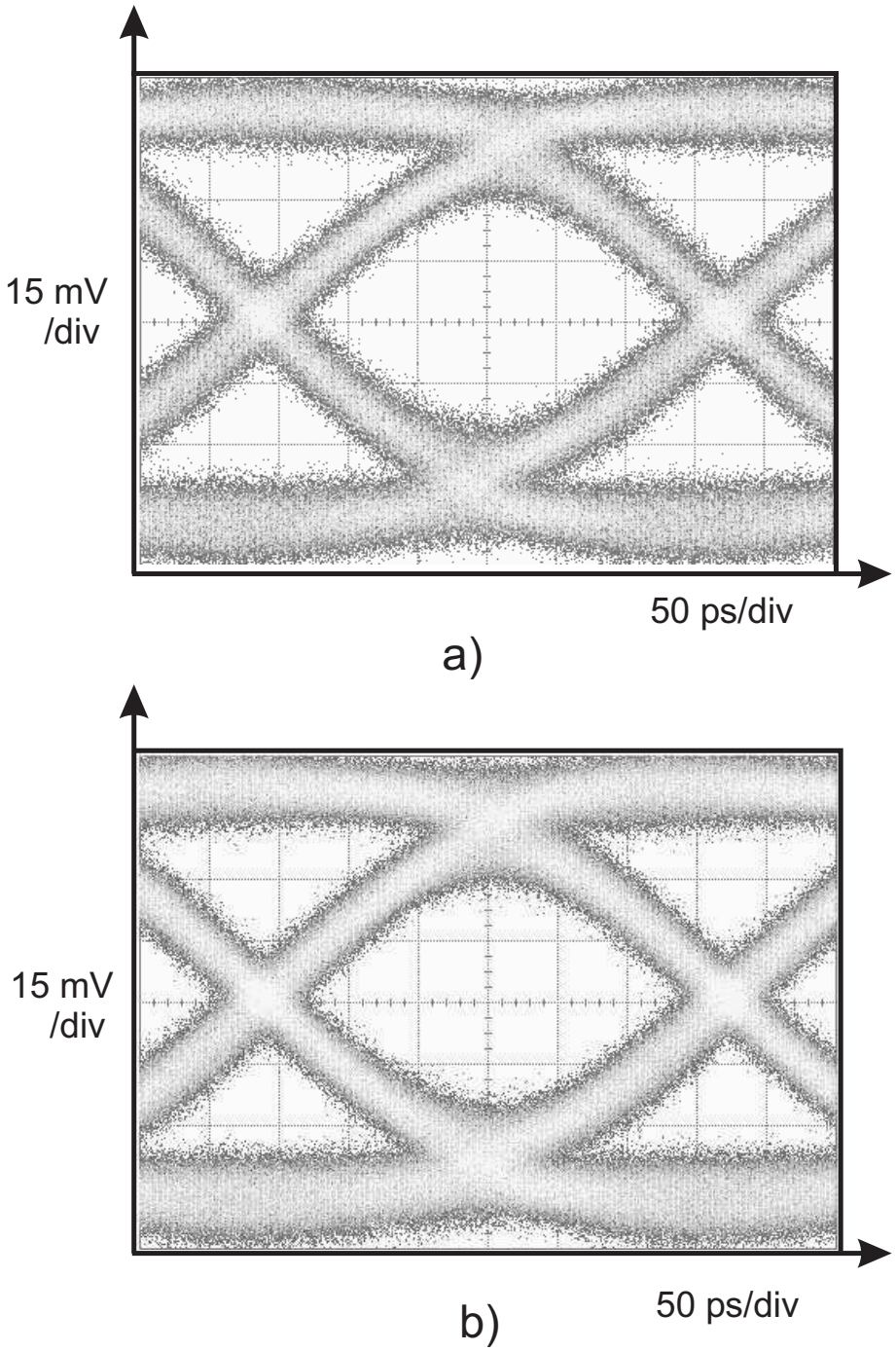


Figure 4.20: Eye diagrams on 3 Gb/s data-rate of the pre-amplifiers including a) $+20\%RC$ spread and b) $-20\%RC$ spread in the equalizer. BER is 10^{-10}.

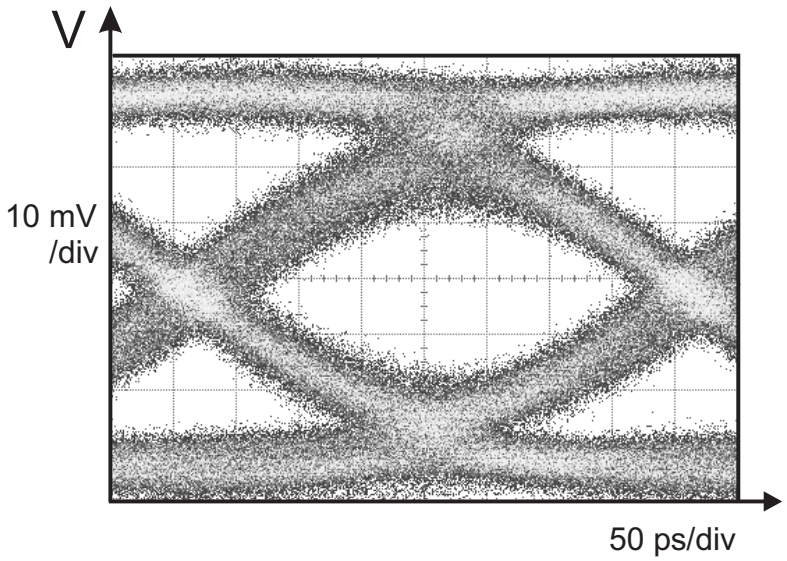


Figure 4.21: 2.5 Gb/s eye-diagram of the p+/nwell/p-substrate CMOS photodiodes with same analog equalizer used for nwell/p-substrate diode, BER=10^{-11}.

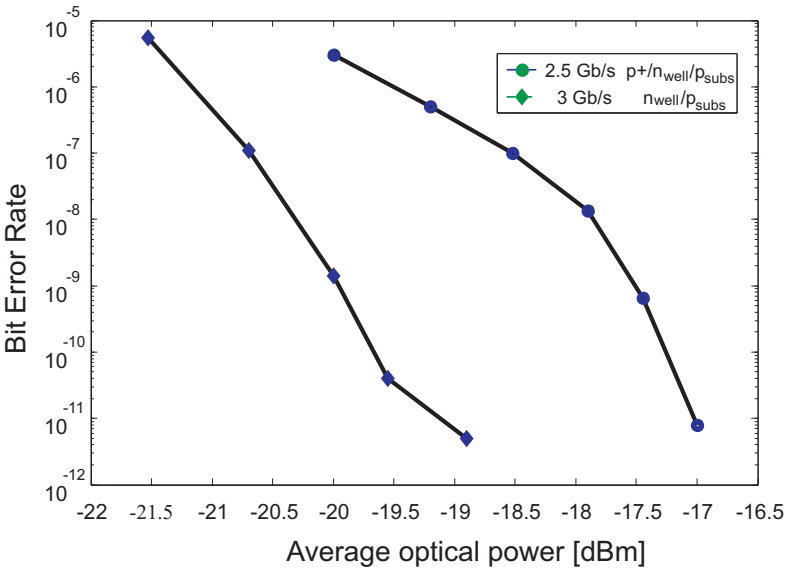


Figure 4.22: Bit error rate vs. input optical power

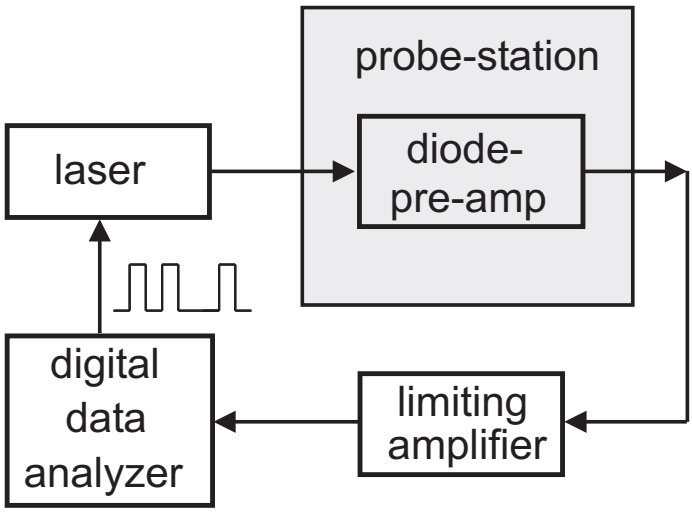


Figure 4.23: Block schematic of the measurements set-up for BER measurements.

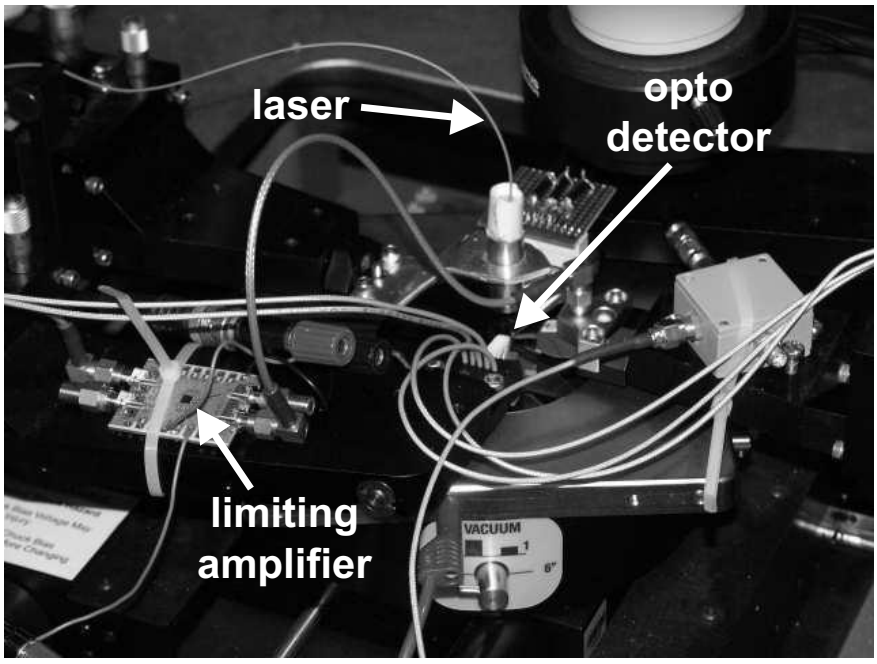


Figure 4.24: The photo of the measurements set-up for BER measurements.

The AC optical power at the fiber's output is deduced by measuring the DC optical power around the operating point of the laser. The bias is changed for the values of the modulation voltages. The optical power is measured using HP 8153A Lightwave Multimeter. The shape of the output signal and the AC optical power were also measured using the reference photoreceiver PT1003 from OEPIC company, which consists of a PIN photodetector, integrated with an InGaPHBT transimpedance amplifier (TIA). The maximum operating data-rate of this reference photoreceiver is 10 Gb/s.

For easier handling, the chip was glued on the larger surface, in our case glass. The on-chip measurements were done using probe-station. The output voltage is measurements using GSG probe ACP40 (for up to 40 GHz). The insertion loss of the coaxial cable, used to measured the output signal, is calibrated up to 4 GHz. The chip is supplied with DC voltage using Eye-pass probe [19] which should provide a stable-supply in the frequency range of interest (the specified frequency range is up to 20 GHz). The DC *current* supply was set using coaxial cables and GS pico-probes.

The multimode laser was mounted on a micro-manipulator of the probe-station perpendicularly to the chip, as shown in Figure 4.24. This manipulator allows better alignment between the fiber and the photodiode.

The laser was modulated with the pseudorandom bitstream of 2^{31}-1 from the MP1632C digital data analyzer (generator/analyzer platform). Since the swing of the signal at the pre-amplifier's output was not large enough for the proper BER measurements, a limiting amplifier was placed after the pre-amplifier as shown in Figure 4.23. The limiting amplifier is an L1001 fabricated by OEPIC company. The input and output resistance of the amplifier are 50Ω and the operating frequency between 100 kHz up to 5 GHz.

Figure 4.22 shows the measured BER versus average light input power for 3 Gb/s. The sensitivity at the BER of 10^{-11} is around -19 dBm, which is 2 dBm better than the one defined in the Gigabit Ethernet standard for short-haul optical communications [11].

4.4 Conclusions and Summary

The proposed optical detector architecture with an analog equalizer, increases the bandwidth of the state of the art CMOS detector [1] by 4.5 times for $\lambda = 850$ nm, without reducing the photo-responsivity. A 3 Gb/s data-rate is

achieved with 25 µW peak-to-peak light input power and BER$<10^{-11}$. With respect to conventional CMOS detectors [20] almost three orders of magnitude higher data-rates are achieved. This data-rate is over a factor 4 higher than that of state-of-the-art fully integrated CMOS detectors for $\lambda = 850$ nm.

The high-speed optical detector with an analog equalizer is very robust. Firstly, it was shown that due to the low roll-off of the photodiode characteristics, the robustness against spread is high. Secondly, the change in temperature results in a very small change in the diffusion coefficient as well as the lifetime of the minority carriers. Therefore, robustness on the temperature is high too. Thirdly, it was shown that the equalization is robust for photodiode spread by using different CMOS photodiode structures and layouts. The proposed optical detector can be implemented for the whole wavelength sensitivity range $\lambda = 400 - 850$ nm, (discussed in detail in chapter 6). For lower wavelengths in this range, the equalization is required over a smaller frequency range, since the roll-off starts at higher frequencies. The proposed design is also implementable in any CMOS technology which is discussed in chapter 6 of this thesis.

References

[1] D. Coppée, H. J. Stiens, R. A. Vounckx, M. Kuijk: "Calculation of the current response of the spatially modulated light CMOS detectors", *IEEE Transactions Electron Devices*, vol. 48, No. 9, 2001, pp. 1892-1902.

[2] H. Lin, R.C. Chang, H. Chih-Hao, L. Hongchin: "A flexible design of a decision feedback equalizer and a novel CCK technique for wireless LAN systems Circuits and Systems", 2003. *ISCAS '03*, Proceedings of the 2003 International Symposium on , Volume: 2, 25-28 May 2003, pp.II153-II156.

[3] D. Inami, Y. Kuraishi, S. Fushimi, Y. Takahashi, Y. Nukada, S. Kameyama, A. Shiratori: "An adaptive line equalizer LSI for ISDN subscriber loops", *IEEE Journal of Solid-State Circuits*, Volume: 23, Issue:3, June 1988, pp.657-663.

[4] W.L. Abbott, H.C. Nguyen, B.N. Kuo, K.M. Ovens, Y. Wong, J. Casasanta: "A digital chip with adaptive equalizer for PRML detection in hard-disk drives", *Solid-State Circuits Conference,* 1994. Digest of Technical Papers. 41st ISSCC., 1994 IEEE International , 16-18 Feb. 1994, pp.284-285.

[5] A. Bishop, I. Chan, S. Aronson, P. Moran, K. Hen, R. Cheng, L.K. Fitzpatrick, J. Stander, R. Chik, K. Kshonze, M. Aliahmad, J. Ngai, H. He, E. daVeiga, P. Bolte, C. Krasuk, B. Cerqua, R. Brown, P. Ziperovich, K. Fisher: "A 300 Mb/s BiCMOS disk drive channel with adaptive analog equalizer", *Solid-State Circuits Conference*, 1999. Digest of Technical Papers. ISSCC. 1999 IEEE International , 15-17 Feb. 1999, pp.46-49.

[6] K. Azadet, E. Haratsch, H. Kim, F. Saibi, J. Saunders, M. Shaffer, L. Song, and M. Yu: "Equalization and FEC techniques for optical transceivers", *Journal of Solid-State Circuits*, vol. 37, March 2002, pp. 317-327.

[7] W. Etten and J.vd Plaats: "*Fundamentals of Optical Fiber Communications*", Prentice-Hall, 1991.

[8] B. Razavi: "*Design of Analog CMOS Integrated Circuit*", McGraw-Hill Higher Education, 2001.

[9] S. Alexander: "*Optical communication receiver design*", SPIE Optical engineering press, 1997.

[10] S. Park, H.J. Yoo: "1.25 Gb/s regulated cascode CMOS transimpedance amplifier for gigabit ethernet applications", *IEEE Journal of Solid State Circuits*, vol. 39, January 2004, pp. 112-121.

[11] IEEE 10 Gigabit Ethernet Standard 802.3ae.

[12] http://www.oepic.com/hm021206/Products.asp

[13] F. Bruccoleri: "*Wide-band low-noise amplifier techniques in CMOS*", PhD thesis, ISBN: 90-365-1964-0, 2003.

[14] C. H. Lu, W. Z. Chen: "Bandwidth enhancement techniques for transimpedance amplifier in CMOS technology", *ESSCIRC 2001*, 18-20 September 2001, Villach, Austria, pp.192-195.

[15] P.Amini and O.Shoaei: "A low-power gigabit Ethernet analog equalizer", *ISCAS 2001*, pp 176-179.

[16] Deluxe Global Media Service Ltd. CD and DVD technology - CD physical specification. www at url: http://www.disctronics.co.uk/technology/cdbasis/cd_{specs}, May 2004.

[17] Philips semiconductors, TZA1045 photodiode and pre-amplifier IC for CD and DVD applications, data sheet, November 2002.

[18] M. Shur: "*Physics of semiconductor devices*", Prentice-Hall, New Yersey, 1997.

[19] http://www.cascademicrotech.com/index.cfm/fuseaction/pg.view/pID/124

[20] S. Radovanović, AnneJohan Annema and Bram Nauta: "Physical and electrical bandwidths of integrated photodiodes in standard CMOS technology", 16-19. December, *EDSSC2003*, HongKong, pp. 95-98.

[21] Tanja van der Meer: "*Design of a PDIC for CD/DVD-system in standard CMOS technology*", Master Thesis, June 2004.

CHAPTER 5

High-speed photodiodes for $\lambda = 400$ nm

This chapter presents both time domain and frequency domain analyzes of monosilicon and polysilicon photodiodes in a standard 0.18 μm CMOS technology, for $\lambda = 400$ nm.

For monosilicon diodes, the maximum calculated intrinsic bandwidth is up to 6 GHz. The photodiodes designed in twin-well technology have smaller bandwidth (maximum up to 3 GHz) because of the limited size of the vertical depletion region (fast drift current contributes less to the overall photocurrent). Measurements on p+/nwell/p-substrate photodiode designed in 0.18 μm CMOS, showed that the total diode bandwidth is 1.7 GHz, which was limited by the electrical diode bandwidth.

This chapter presents also a lateral polysilicon photodiode that has an intrinsic bandwidth far in the GHz range. The electrical bandwidth is also high due to a very small parasitic capacitance (<0.1 pF). For $\lambda = 400$ nm, the achieved quantum efficiency is only 2.5% due to a very small light sensitive diode volume. The diode active area is limited by a narrow depletion region and small depth is limited by the technology.

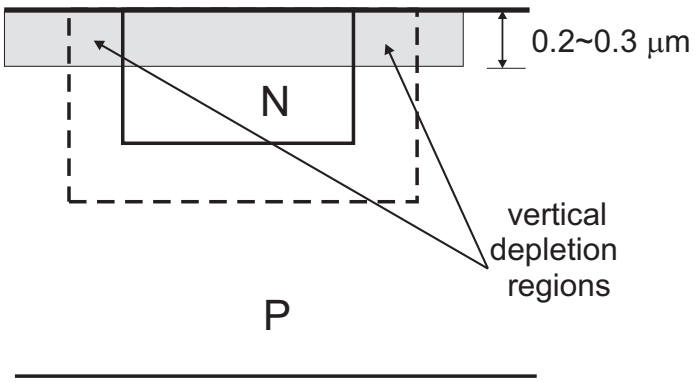


Figure 5.1: Light absorbtion in silicon photodiode for $\lambda = 400$ nm.

5.1 Introduction

Chapter 2, section 2.9, showed that the lower the wavelength of the input signal, the higher the light absorption coefficient α. For the boundaries of the CMOS sensitivity range $\lambda \in [400,850]$ nm, the difference in light penetration depth is almost 70 times. At the lower boundary ($\lambda = 400$ nm), the 1/e-absorption depth is only about 0.2 µm. In both previous and modern CMOS processes (up to 0.13 µm technology), this depth is certainly less than or equal to the shallowest junction available (n+ or p+). Hence, the whole amount of light is absorbed very close to the diode surface. As a result, the overall photocurrent is determined by the (fast) diffusion inside n+/p+/nwell regions and the drift photocurrent generated in the vertical depletion regions as shown in Figure 5.1.

The responsivity of a CMOS photodiode and consequently the photocurrent, is low for $\lambda = 400$ nm as shown in chapter 2, Figure 2.11. The energy of the incoming photons is $h\nu$ and for the same input optical power P_{in}, the number of photons $P_{\text{in}}/h\nu$ is minimal. Therefore, the photocurrent is also minimal. Moreover, the surface recombination process is now very important for $\lambda = 400$ nm. Due to the low light-penetration depth, the surface recombination of the carriers becomes significant i.e. it is 30% larger [1] than for wavelengths with larger penetration depths (depths larger than 600 nm). The maximum photodiode responsivity is about 0.23 A/W.

Due to the very low penetration depth of the light, the lateral photodiode structure becomes important in the overall photodiode response [2], [3]. There are two main advantages of using lateral structures:

• for the diodes with large intrinsic (depletion) region most of the carriers are generated in this region. As a result, the total photodiode bandwidth can be in tents of GHz range. For the diodes without a significant intrinsic layer, carriers are generated in the n-region and the p-region close to the diode surface. Depending on the depth of the n and p-regions, the diffusion of these carriers can be fast, which results in higher diode bandwidth; this will be shown in the following part of this chapter.

• a surface recombination is less dominant for the carriers generated incidently in the vertical depletion region. The lateral photodiodes in standard CMOS technology with maximal vertical junctions can be designed by making the nwell not adjunct to the pwell; this provides large depletion region between nwell(n+) and p-epi layer. However, the available fabrication process for a photodiode was twin-well CMOS with one complementary mask for nwell and pwell. Thereby, the designed diodes have narrow vertical depletion regions since the doping concentration of the pwell is typically much higher than the one in p-epi. The total drift current contribution is smaller, which according to analysis in chapter 3, results in a lower total diode bandwidth. Nevertheless, this bandwidth is still in 100ths of MHz up to a few GHz range. Both the frequency and the time response of the twin-well photodiodes are investigated in sections 5.2, 5.3 and 5.5. For comparison, a separate-well nwell/p-substrate photodiode is analyzed in section 5.4.

5.2 Finger nwell/p-substrate diode in *adjoined-well* technology

The nwell/p-substrate diode in *twin-well* technology is illustrated in Figure 5.2. In this chapter we present the photodiode frequency response and time response on a Dirac light pulse for $\lambda = 400$ nm. Both responses are calculated following the procedure explained in chapter 3. The absorption depth of light is lower than the junction depths so the substrate does not contribute to the overall photocurrent. The overall photocurrent is determined mainly by a nwell diffusion current plus a small drift current generated in vertical depletion regions between wells. The diffusion current generated inside the pwell (with excess electrons) is considered as a part of the nwell current with the amplitude that corresponds to the ratio between the nwell and pwell size.

A few ps after the incident light pulse, most of the excess carriers are gener-

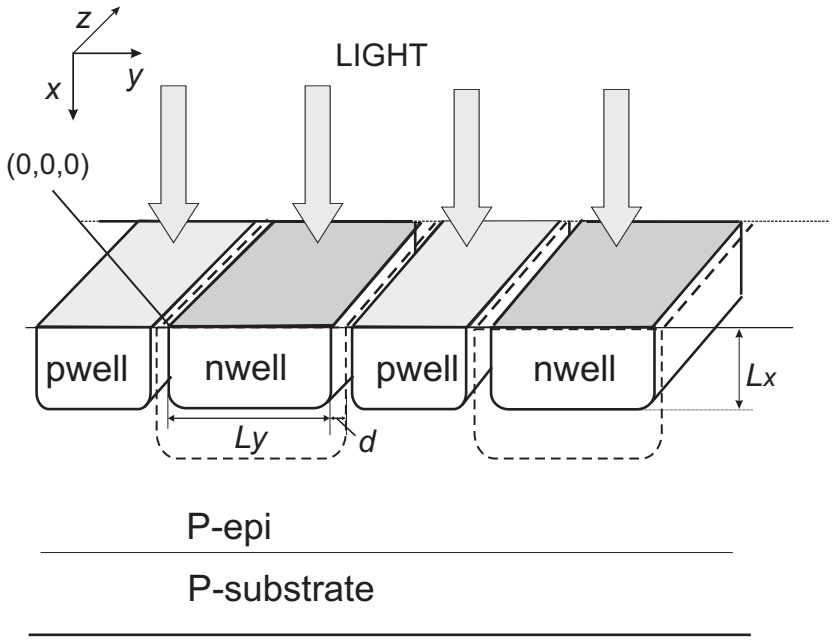


Figure 5.2: Finger nwell/p photodiode structure with *low resistance* substrate and adjoined-wells in standard CMOS technology.

ated in the upper side of the nwells i.e. in the area close to the photodiode surface. At that time, the largest gradient of hole density is towards the *nwell-side junctions*. Further in time, a large number of holes diffuses towards the nwell bottom-junction due to the excess carrier concentration gradient. However, the diffusion process, as discussed in chapter 3 for λ=850 nm, takes now longer until the first carriers need to reach the nwell-bottom junction. Even then, the hole gradient towards the nwell-bottom is lower than the hole gradient toward the nwell-side, as shown in Figure 5.3. Therefore, the lateral dimension of the nwell mainly determines the diffusion speed; it is the most important dimension in the diffusion process inside nwell for $\lambda = 400$ nm.

The total photocurrent is the sum of the diffusion current presented with equation 3.8, and the drift current given in equation 3.17 in chapter 3. The nwell/p-substrate photodiode frequency response is illustrated in Figure 5.4. The response is normalized with its DC photocurrent on 400 nm. Figure 5.4 shows the importance of the nwell width on the diode intrinsic bandwidth. This

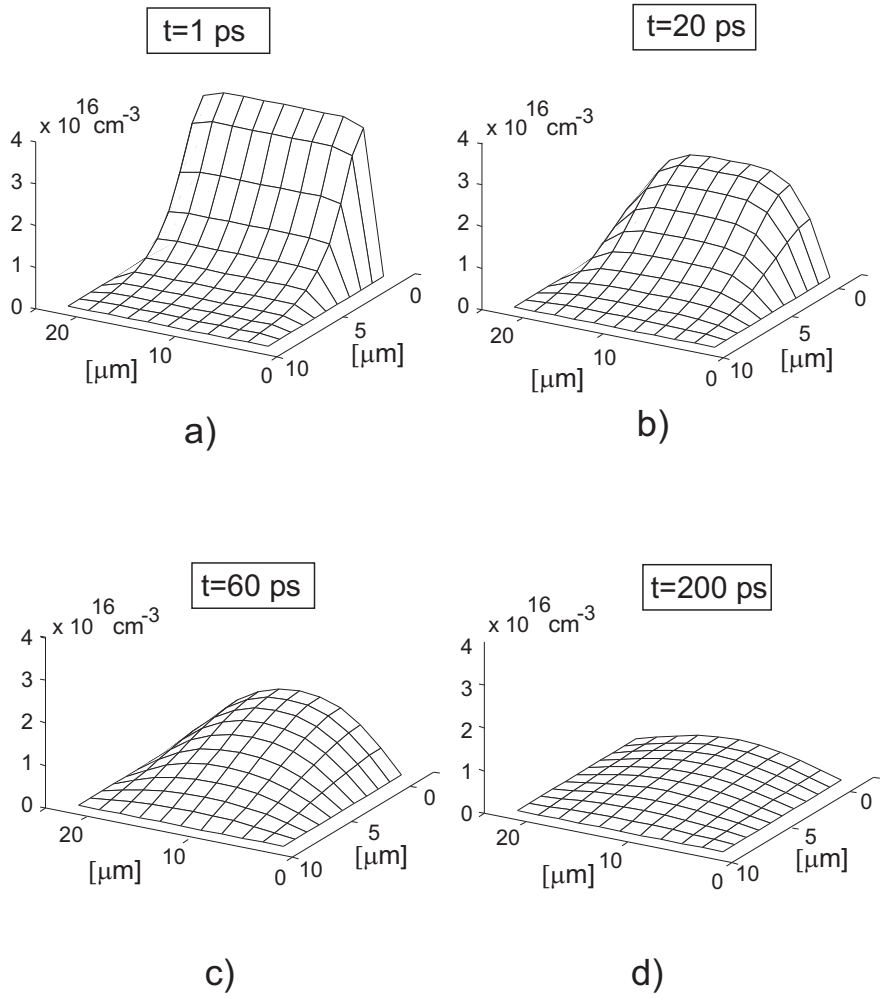


Figure 5.3: The calculated hole diffusion profile inside the nwell for a minimum nwell width, 2 µm, under incident light pulse ($\lambda = 400$ nm, 10 ps pulse-width). This profile is calculated after 1 ps, 20 ps, 60 ps, 200 ps.

bandwidth is 700 MHz when nwell width is minimum [1] (2µm) and 370 MHz when nwell width is much larger then its depth, 10 µm. Therefore, *for maximum intrinsic bandwidth, the diode nwell size should be minimal.*

5.3 Finger n+/nwell/p-substrate diode

In standard CMOS technology, it is possible to place a shallow n layer (n+), at the top of the nwell region as shown in Figure 5.5. This section presents the large influence of n+ layer inside the nwell on the total photocurrent response for $\lambda = 400$ nm.

The depth of the n+ layer in standard 0.18 µm CMOS technology is larger than the 1/e-absorbtion depth of $\lambda = 400$ nm. The frequency response of the n+/nwell diffusion current is calculated using the system of *two* two-dimensional diffusion equations, similar to those given in equation 3.10. Nevertheless, the electron doping concentration should be changed for the hole doping concentration and corresponding vertical positions of the p-substrate layers are interchanged with the **vertical** positions of the nwell/n+ regions.

In order to solve the obtained system of equations, two boundary conditions between the two n-layers are used, plus the one boundary at the n+ top and the boundary at nwell bottom, as shown in Figure 5.6. These conditions are related to both the current density and the minority carrier concentration. Due to the continuity of currents, the current densities are equal in a plane between the two layers:

$$-qD_{\mathrm{p1}}\frac{\partial p_{\mathrm{n+}}(x,s)}{\partial x}|_{x=L_{\mathrm{n+}}} = -qD_{\mathrm{p2}}\frac{\partial p_{\mathrm{nwell}}(x,s)}{\partial x}|_{x=L_{\mathrm{n+}}} \tag{5.1}$$

The second boundary condition is related to the continuity of the concentration of the minority carriers:

$$p_{\mathrm{n+}}(L_{\mathrm{n+}},s) = p_{\mathrm{nwell}}(L_{\mathrm{n+}},s) \tag{5.2}$$

where $L_{\mathrm{n+}}$ is the depth of the n+ region, and s denotes frequency domain. The photodiode surface is again assumed to be reflective i.e. the *hole gradient* at the diode surface is taken to be zero, since the recombination process is far slower than the MHz or GHz frequencies of interest in this thesis. The other boundary condition for the *hole concentration* at the nwell bottom is zero.

[1] In standard CMOS minimum nwell width is typically twice its depth.

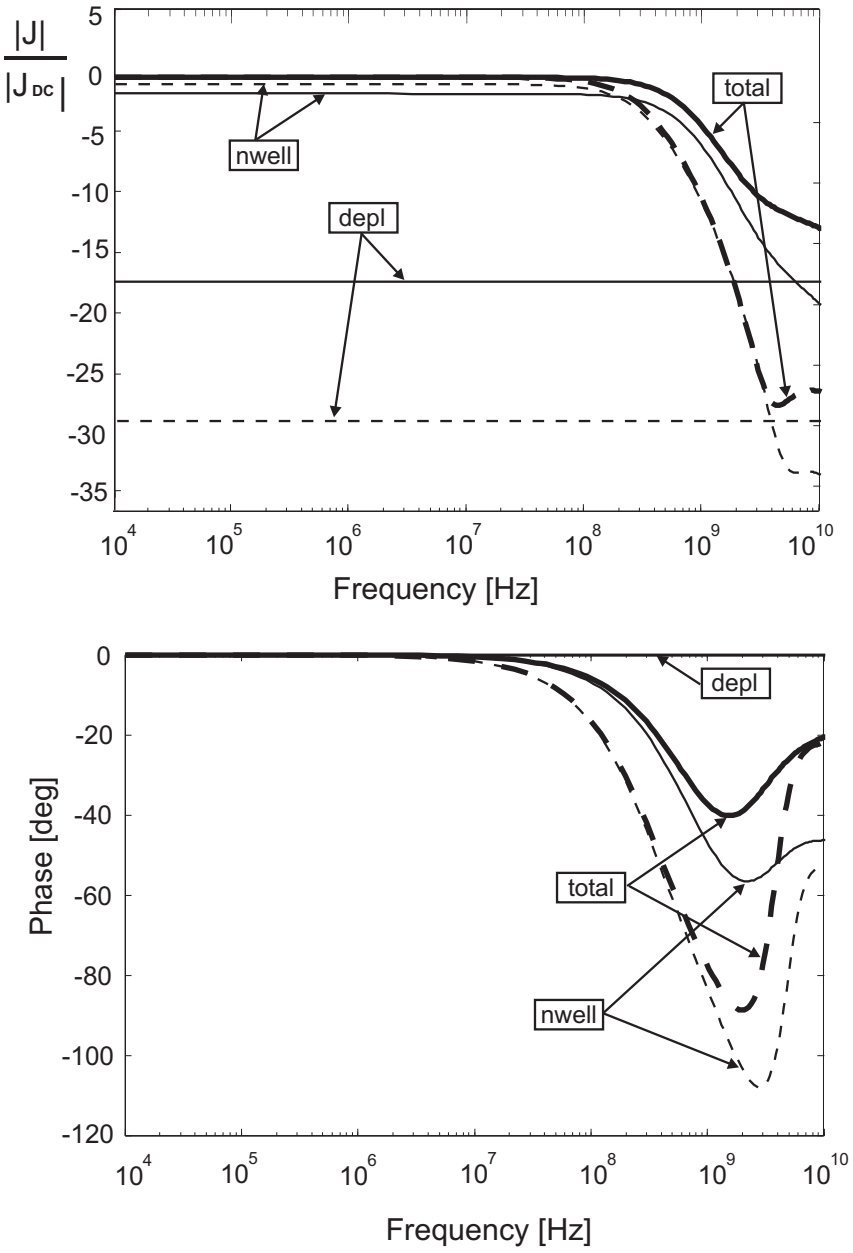


Figure 5.4: The calculated total photocurrent response of nwell/p-substrate photodiodes in a *twin-well* technology for the minimum nwell width (solid-line) and nwell width much larger than its depth 10µm (dashed-line) for $\lambda = 400$ nm.

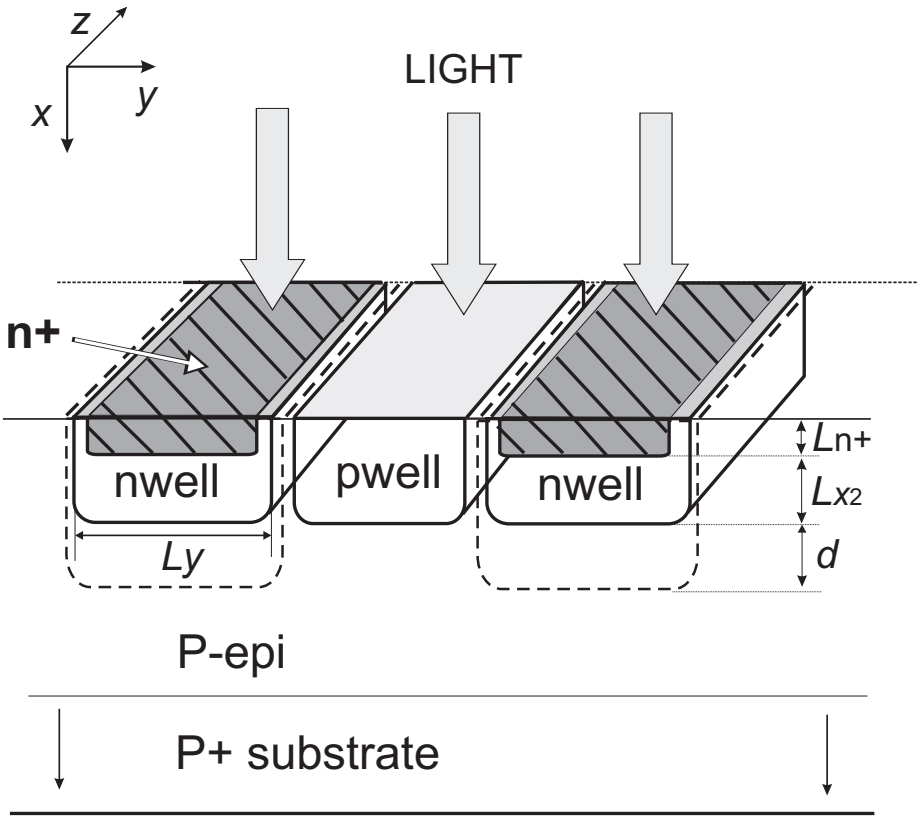


Figure 5.5: Finger nwell/p-substrate photodiode with n+ layer at the top of the n-well region.

The total n+/nwell p-substrate frequency response for $\lambda = 400$ nm is illustrated in Figure 5.7. A maximum bandwidth of this photodiode is about 300 MHz for 2 µm nwell width and 170 MHz only for 10 µm nwell width. These bandwidth are twice as low as the corresponding bandwidths of the photodiode without n+ layer because the diffusion constant D_{p1} is twice as low for the n+ region than D_{p2} for the nwell region due to the higher majority carrier concentration. Thus, a highly doped n+ region at the top of the nwell (nwell/p-substrate diode) *decreases* maximum intrinsic bandwidth for more than *two times*.

Electrical bandwidth

For the photodiode area of 50×50 $\mu\mathrm{m}^2$, corresponding to the core-diameter of the multimode fiber, the calculated photodiode capacitance from the process technology parameters is given in Table 3.3. It ranges from 0.6-1.6 pF for 10 µm

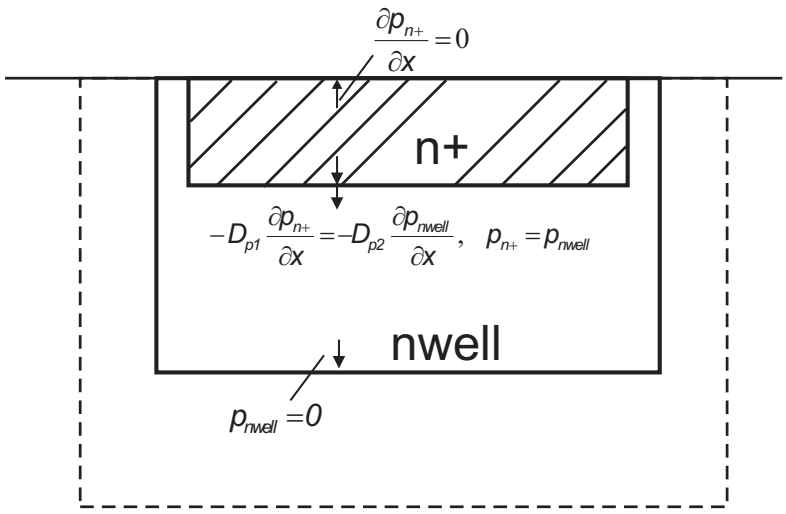


Figure 5.6: The boundary conditions for the calculations of nwell/n+ diffusion curent.

and 2 µm nwell size, respectively. For the input resistance of the subsequent transimpedance amplifier lower than 130 Ω, the extrinsic photodiode bandwidth is larger than the intrinsic diode bandwidth. For larger transimpedances, the nwell size i.e. photodiode capacitance does influence the total photodiode bandwidth. The lower the nwell width, the lower the electrical bandwidth.

5.3.1 Time domain measurements

The calculated total photodiode bandwidth is confirmed with measurements. The diode is fabricated in standard 0.18 µm CMOS with minimum nwell-width 2 µm. The diode layout is presented in Figure 5.8. On the transmitter side, a picosecond blue-light laser with $\lambda = 400$ nm was used. The light was focused into a multimode fiber using a system of lenses, as shown in Figure 5.9. The pulse width of the picosecond laser is 200 ps and the power is 1 mW. The output of the n+/nwell p-substrate photodiode is presented in Figure 5.10. This output voltage is measured using RF pico-probe and coaxial cable which was terminated with 50Ω of the oscilloscope.

Using Figure 5.10, the photodiode bandwidth is calculated using rising edge τ_r and falling edge τ_f, with a simplified formula [2]:

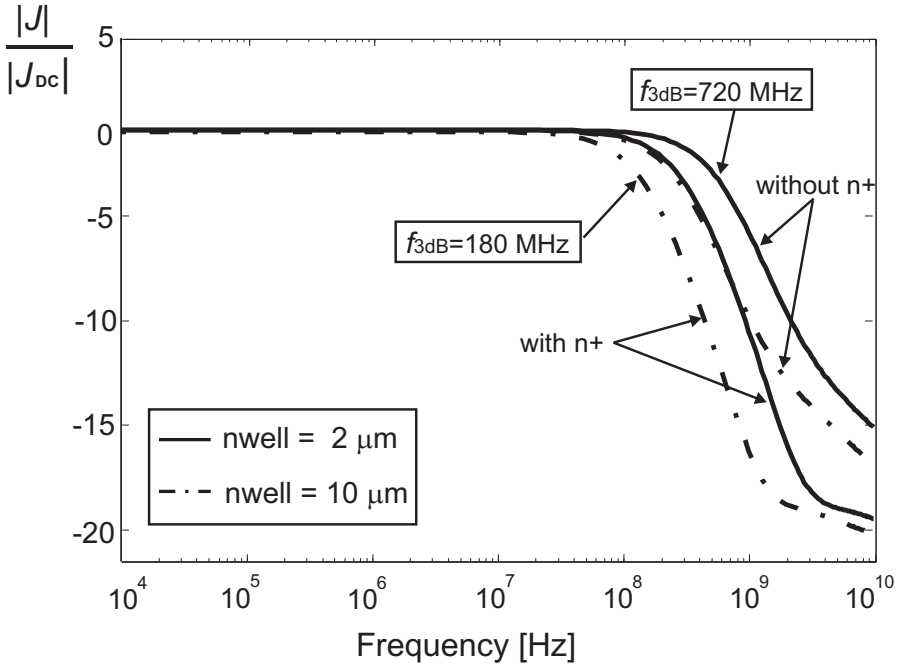


Figure 5.7: The calculated intrinsic photocurrent response of nwell/p-substrate photodiode in a *twin-well* technology *with* n+ layer at the top (solid line) and *without* n+ layer (dashed-dot line) for $\lambda = 400$ nm. This current is calculated for two nwell widths: minimum nwell width that is typically twice its depth in standard CMOS (2 µm), and nwell width much larger than its depth (10 µm).

$$f_{3\mathrm{dB}} = \frac{ln(9)}{\pi(\tau_\mathrm{r} + \tau_\mathrm{f})} \tag{5.3}$$

The obtained bandwidth value is 280 MHz. Figure 5.7 shows that the calculated bandwidth is about 300 MHz; hence, the measured photodiode bandwidth closely corresponds to the calculated bandwidth.

5.4 Finger n+/p-substrate photodiode in *separate-well* technology

The frequency analysis of n+/p-substrate photodiode is similar to the previously analyzed photodiode. However, the main difference is in the size of the vertical

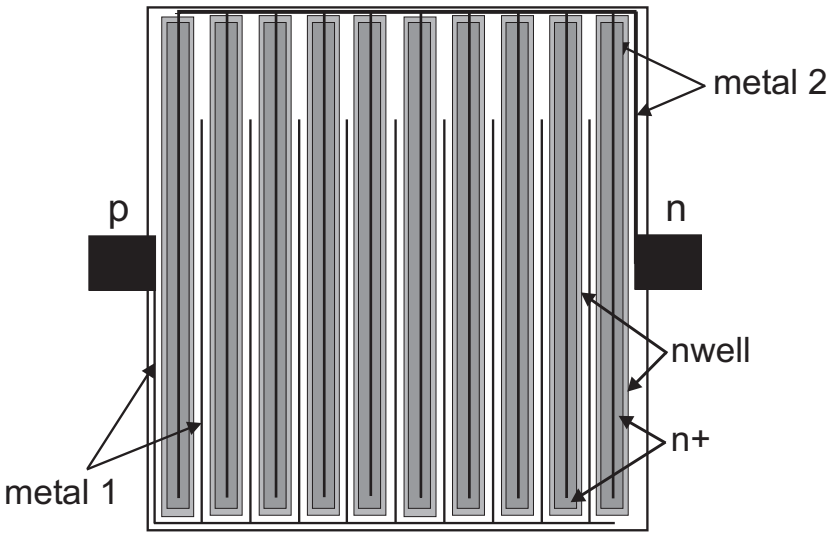


Figure 5.8: A nwell/n+/p-substrate photodiode with 2 µm nwell width.

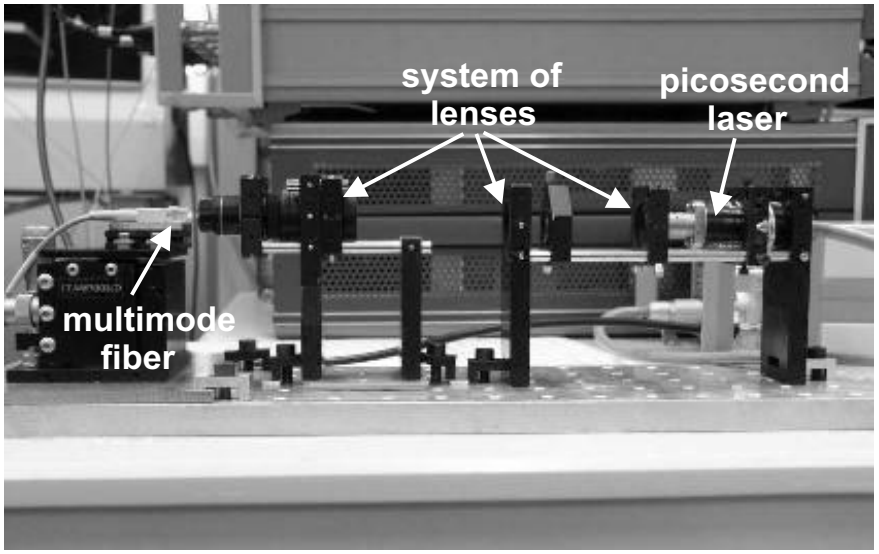


Figure 5.9: Focusing the light from the picosecond blue laser into the multimode fiber by using system of lenses.

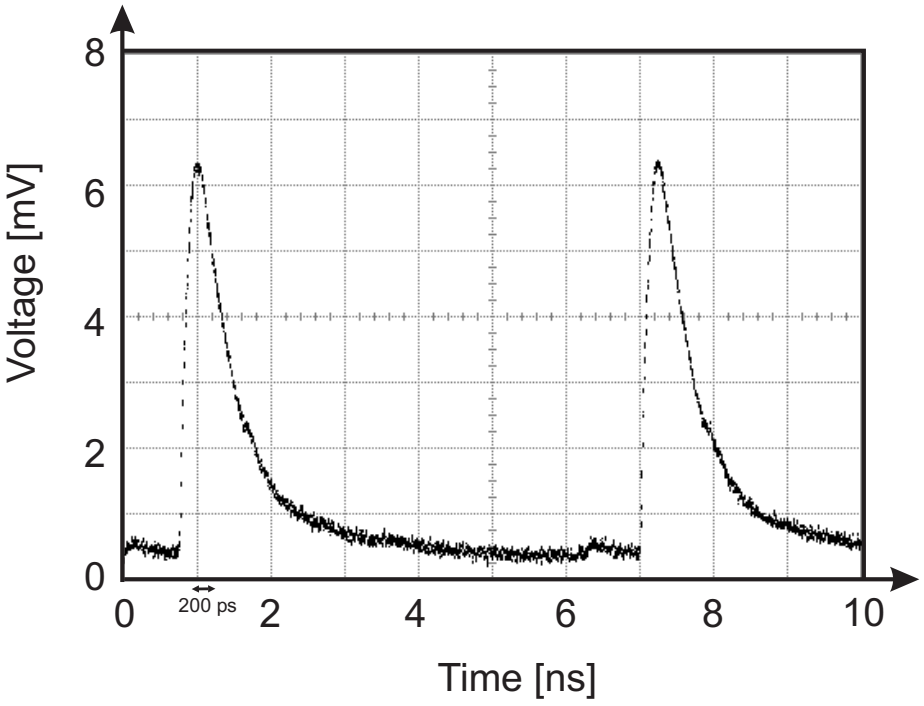


Figure 5.10: Transient response of the nwell/n+ p-substrate photodiode on 200 ps input light pulse($\lambda = 400$ nm) with a 50 Ω load resistor.

depletion regions between n+ fingers and p-epi region; here it is larger, which results in a faster total intrinsic response. The calculated intrinsic bandwidth for this photodiode is 6 GHz. In [2], a 9-finger n+/p-substrate photodiode is presented for high-speed data rate. The photodiode was designed in standard 1-µm CMOS technology. The doping concentration of the epitaxial layer in CMOS technology that they are using is very low ($< 10^{15}$ cm^{-3}). On the other hand, the doping concentration of the shallow n+ region is very high (10^{20} cm^{-3}) resulting in the large depletion region. The measured bandwidth for 400 nm wavelength was 470 MHz and it is limited by the electrical photodiode bandwidth (the resistance subsequent to the photodiode is 1 kΩ).

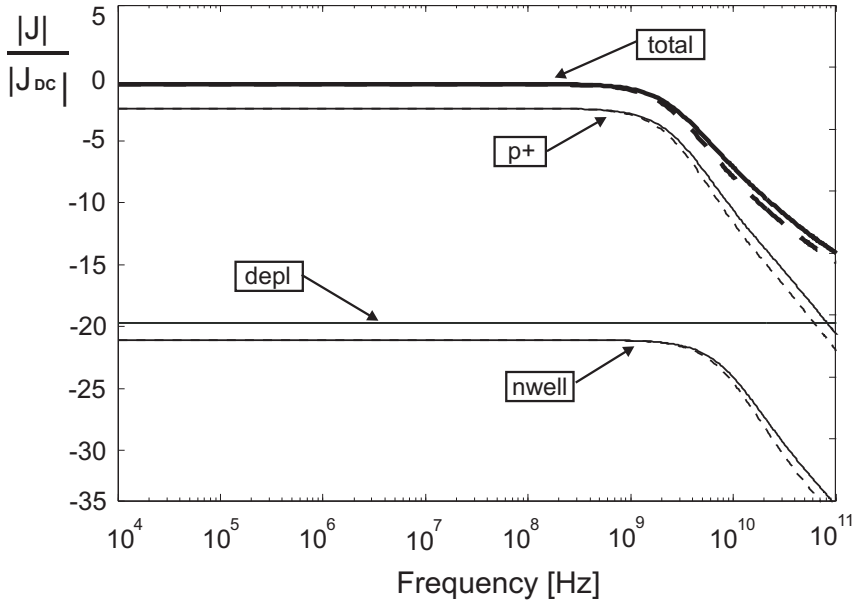


Figure 5.11: The calculated total photocurrent response of p+/nwell/p-substrate photodiode in a *adjoined-well* technology with 2 µm nwell size (solid-line) and 10 µm nwell size (dashed-line) for $\lambda = 400$ nm.

5.5 Finger p+/nwell/p-substrate in *adjoined-well* technology

The double photodiode structure, p+/nwell/p-substrate, was already analyzed for λ=850 nm, in chapter 3, section 3.2.4. This section investigates the frequency response of this photodiode on a Dirac light pulse for $\lambda = 400$ nm, using the same analyzes from chapter 3. First, the calculated intrinsic response is shown in Figure 5.11.

Most of the carriers are again generated close to the photodiode surface inside p+ region and vertical depletion regions. The depth of the p+ region is low (<0.3 µm) so most of the carriers are generated close to p+ bottom junctions. Thereby, the diffusion process is fast and the calculated bandwidth using equation (3.8) is 2 GHz. The total calculated bandwidth including the drift current in vertical junctions is 2.8 GHz.

Electrical bandwidth

For a photodiode area of 50×50 μm^2, the calculated photodiode capacitance from the process technology parameters is given in Table 3.3. It ranges from 2.2 pF-3.6 pF for 10 µm and 2 µm p+/nwell size, respectively. For the input resistance of the subsequent transimpedance amplifier (TIA) lower than 15 Ω, the extrinsic photodiode bandwidth is larger than intrinsic diode bandwidth. This low input resistance value can be realized typically with the cost of larger power consumption and also area consumption [4]. For the minimal input capacitance the photodiode capacitance should also be minimal; this is achieved by designing a single p+/nwell/p-substrate photodiode (without fingers).

5.5.1 Time domain measurements

The response of the single p+/nwell/p-substrate diode is measured in the time domain. The top view of this photodiode is shown in Figure 5.12. A picosecond blue-light laser with $\lambda = 400$ nm was used as a transmitter. The pulse width of the picosecond laser is 200 ps and the maximum optical power is 1 mW. The output signal is measured using RF pico-probe and the coaxial cable terminated with 50 Ω load of oscilloscope. The bondpads for n-contact and p-contact and reversed in comparison with previous diode. Therefore, the output voltage presented in Figure 5.13 is also inverted.

According to the time response, the approximated bandwidth of the single p+/nwell/p-substrate photodiode using equation (5.3) is 1.4 GHz. The calculated electrical diode bandwidth (first order) including diode capacitance and the output impedance (R_{out}=50 Ω, C_{in}=1.7 pF) is 1.7 GHz. Thereby, there is 20% difference between the measurements and simulation. This difference is even lower if the bondpad capacitances (∼100 fF) are also included in the bandwidth calculation.

In real applications, typically there is a transimpedance amplifier subsequent to a photodiode. Designing the TIA with low input impedance (<25Ω) the total bandwidth of the photodiode is equal to the intrinsic diode bandwidth of 2.8 GHz. Therefore, this photodiode can be used for Gb/s data-rate optical detection, for input wavelength $\lambda = 400$ nm.

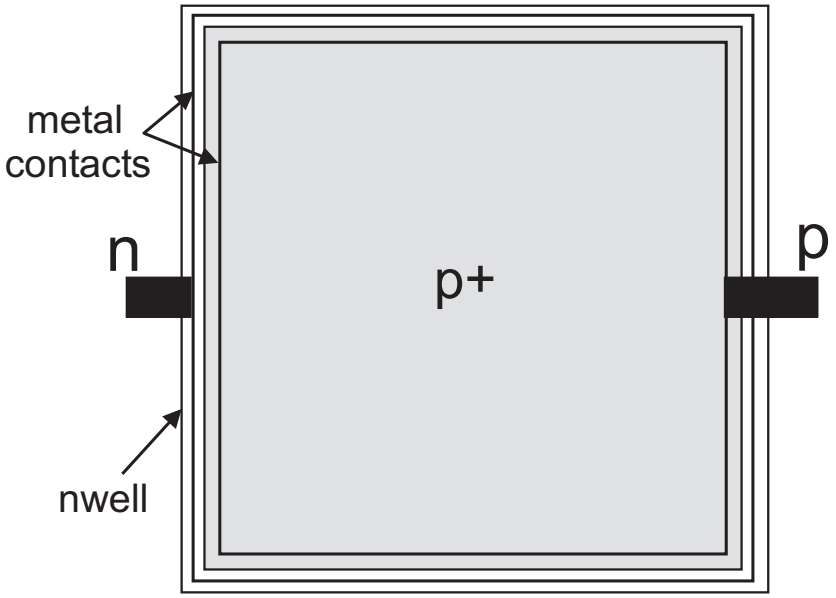


Figure 5.12: Top view of the single p+/nwell/p-substrate photodiode.

5.6 p+/nwell photodiode

Another type of photodiode is p+/nwell diode, presented in chapter 3. For $\lambda = 400$ nm, the excess carriers are generated close to the photodiode surface i.e. close to the p+/nwell junction. As a result, the diode substrate will not contribute to the overall photocurrent. By disconnecting the substrate, the photodiode capacitance decreases. Therefore, the electrical photodiode bandwidth is larger with no responsivity penalty.

The width of p+/nwell fingers does not influence the intrinsic photodiode bandwidth because of the shallow depth of the p+ region, as was described in chapter 3, section 5.5. On the other side, the larger the size of the p+/nwell fingers the lower the diode capacitance. Thereby the maximum intrinsic and extrinsic bandwidth are obtained using one nwell finger i.e. using a single p+/nwell photodiode.

For $\lambda = 400$ nm, the light penetration depth (< 0.2 µm) is typically lower than nwell depth. Thereby, cancelling p-substrate will not influence the overall photodiode responsivity. The calculated frequency response is similar to that illustrated in Figure 5.13.

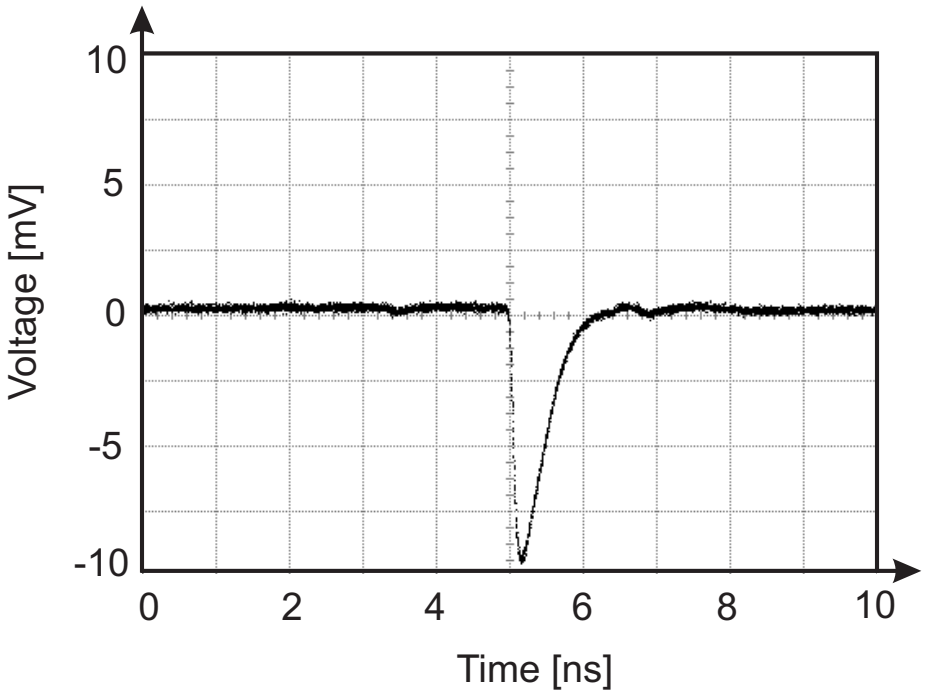


Figure 5.13: Transient response of p+/nwell/p-substrate photodiode on 200 ps input light pulse($\lambda = 400$ nm) using a 50 Ω output resistor.

The calculated capacitance of the single p+/nwell photodiode with the dimensions of 50 µm × 50 µm is 1.7 pF for 0.18 µm CMOS. For the input resistance of the subsequent TIA lower than 30 Ω the electrical diode bandwidth is larger than the intrinsic diode bandwidth. The total photodiode bandwidth is thus equal to the intrinsic bandwidth.

5.7 Polysilicon photodiode

In nowadays MOS transistor processes, there is a polycrystalline silicon (polysilicon) layer above the silicon-oxide exploited as a gate terminal for both NMOST and PMOST. The doping concentration of that polysilicon layer is high ($1 \cdot 10^{20}\text{cm}^{-3}$) with the doping charge corresponding to a type of the MOS transistor. Using this two opposite types of polysilicon we made a p-n junction, creating thus lateral polysilicon diode (Figure 5.14.).

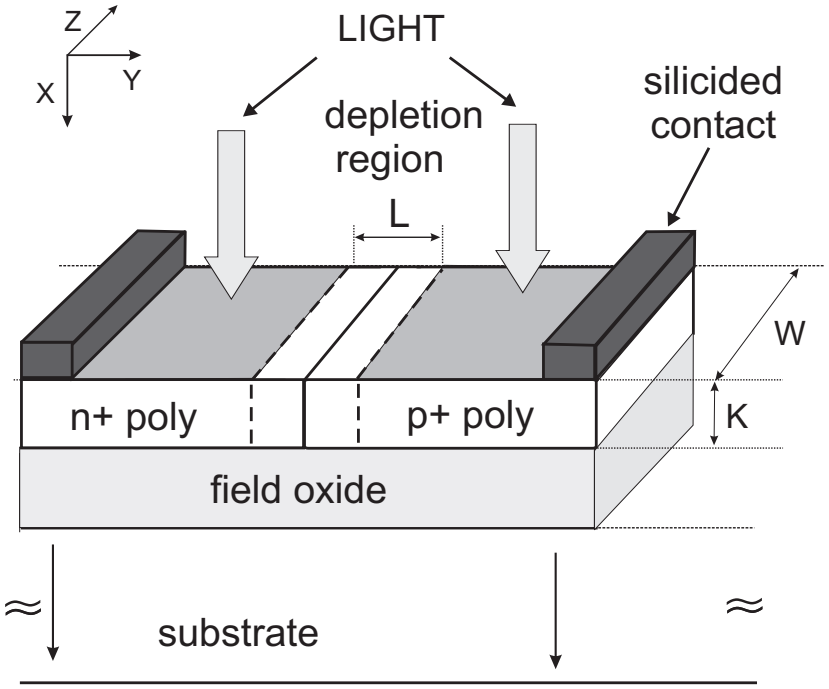


Figure 5.14: Lateral polysilicon photodiode in CMOS technology.

The paramount advantage of polysilicon photodiode in comparison with the monosilicon one is that there are no slow-diffusive carriers coming from the substrate. This is valid for all wavelengths of interest.

The total diode response is the sum of two responses: fast drift current inside depletion region and fast diffusion current inside n+ and p+. The latter is due to a high doping concentration inside n+ and p+. The lifetime of the excess carriers is low (ps range [5]), the diffusion lengths $L_{\mathrm{n,p}}$ are short (300 nm) [5]) and only those carriers generated close enough to the junctions are collected as a photocurrent. The rest of the excess carriers are recombined.

The main differences of *polysilicon* in comparison with *monosilicon* photodiode concerning a photo-responsivity and a data-rate are:

• the absorption coefficient α is four times the absorption coefficient of monosilicon photodiodes [5], [6], (for the same material depth, the responsivity of polysilicon diode is higher for all wavelengths $\lambda \in [400, 850]$ nm)

• the electron mobility μ_n is approximately four times lower than the mobil-

ity of monosilicon photodiode [7], which can limit the bandwidth of the polysilicon photodiodes.

The depth of the poly layer in standard CMOS technology is typically less than 600 nm. Therefore, the polysilicon layer is sensitive mainly for lower wavelengths ($\lambda < 600$ nm).

The first lateral pn junction in polysilicon found in literature was designed and investigated by J. Manoliu in 1972 [8]. The dopant concentrations on both sides are very high, about $2 - 5 \times 10^{15}\text{cm}^{-2}$. Compared to the p-n junctions in a single-crystal Si, polysilicon diodes carry much higher current densities. For many years after, the knowledge on polysilicon diodes' behavior was maturing and in 1994 [7], a thorough theoretical and numerical analyzes on this diode is presented while the obtained results showed good agreement with the measurements. High leakage current in a polydiode was explained with the field enhanced effect, where a large number of carriers typically trapped in the grain boundaries, are released due to a high electric field. The analysis on this effects is out of scope of this thesis, and the reader should refer to [7].

The lateral polysilicon photodiode is often implemented for Electrostatic discharge (ESD) protection [9] due to high Zener-effect on relatively low voltages (2-3 V). According to literature, the PIN polysilicon photodiode is first introduced as a high-speed photodetector in 1994 [10]. The time-response measurements using the high power pulse-laser with $\lambda = 514$ nm, showed that the cut-off frequency of their polysilicon photodiode was 5 GHz. The doping concentrations of n+ and p+ poly regions were very high $3.3 \cdot 10^{20}\text{cm}^{-3}$.

In 1997, the PIN polysilicon resonant-cavity photodiode with silicon-oxide Bragg reflectors was introduced with a speed in the GHz range [11]. Doping concentrations of n+ poly is $2 \cdot 10^{20}$ and p+ poly $4 \cdot 10^{19}\text{cm}^{-3}$. The absorption thickness of the polysilicon was 500 nm which is 8 times higher in comparison with the previously reported poly-diodes. This also implies that this photodiode is suitable for wavelengths in the range between $\lambda \in [400, 600]$ nm. For the upper boundary, the maximal amount of the absorbed light is about 50%. In [11], a quantum efficiency of 40% is reported for input wavelength light $\lambda = 640$ nm. The responsivity measurements as well as the high frequency measurements for $\lambda = 400$ nm were not presented in the paper.

In this chapter we present lateral polysilicon photodiode in standard CMOS technology. The main difference in comparison with the diode in [10] and [11] is in following:

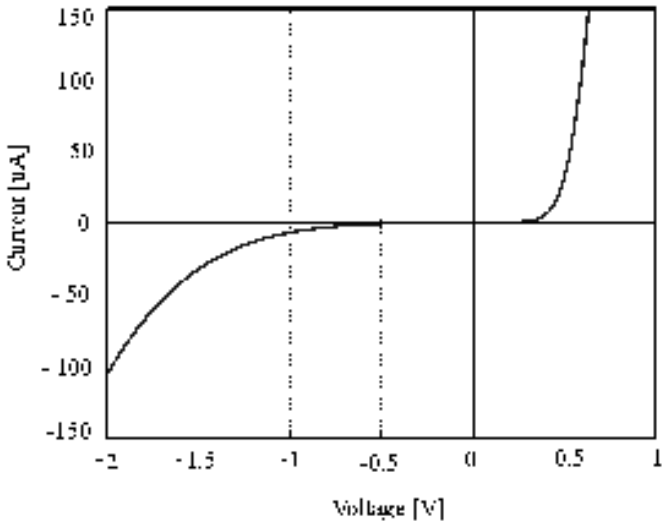


Figure 5.15: Measured DC current (without light) of "jagged" polysilicon photodiode in standard CMOS technology. The lateral diode dimensions are 45×45 μm.

- there is no intrinsic (low doped or undoped poly) layer between n+ and p+ regions; as a result, a light sensitive area is smaller.
- a depth of the poly-diode is 0.2 µm i.e. smaller than reported ones;
- designed in standard CMOS technology, this poly-diode can be easily integrated with the rest of the electronic circuitry. This is very suitable for low-cost, high-speed optical detector design. Moreover, an array of detectors can be easily designed which increases the overall data-rate for the cost of minimal chip area (simple embedding). This is valid for all silicon photodiodes.

Figure 5.15 shows the diode I-V characteristic. The large leakage current is due to grain-boundary trap-assisted band-to-band tunnelling and field-enhanced emission [7].

During the chip processing the masks for the $n+$ and $p+$ layers shown in Figure 5.14 are not perfectly aligned. This might influence the size of effective width of polydiode depletion region. However, measurements on number of devices on the same wafer showed that the effects of misalignment as seen as

spread in sensitivity were not observed. It is also important that the measured chips are exposed on the different reticles.

The carrier lifetime in polysilicon diode depends on recombination rates of holes and electrons and it is proportional to the concentration of recombination centers [8]. It is also inversely proportional to the grain size of polysilicon. In 0.18 μm CMOS technology, the grain size is about 50-60 nm [12], which causes the carrier lifetime to be very short, about $\tau_{\mathrm{n,p}} = 50$ ps [5]. Since in this case the diffusion speed of carriers is mainly determined by their lifetime, the diffusion bandwidth will be far in the GHz range.

5.7.1 Pulse response of poly photodiode

The major speed limitation in all monosilicon photodiodes lies in the very slow diffusion of excess carriers generated deep into the substrate when using long-wavelength light. This section analyzes the intrinsic processes in the poly photodiode including the drift and the diffusion of carriers inside the depletion region as well as the diffusion of carriers outside this region. The latter is not negligible for narrow poly photodiodes without an intrinsic region.

The current response of polysilicon detector is mainly determined by the minority carrier lifetimes τ_n and τ_p, saturation drift velocities v_s and diffusion of minority carriers inside depletion region; the last one is important if the width of the depletion region is larger than excess carrier diffusion lengths [10]. If $n(x,t)$ is the excess electron concentration and $p(x,t)$ is the excess hole concentration, the transport of these carriers *inside* the junction can be described with drift-diffusion equations as [13],[10]:

$$\begin{aligned}
\frac{\partial n(x,t)}{\partial\, t} &= D_{\mathrm{n}}\frac{\partial^2 n(x,t)}{\partial x^2} \pm v_{\mathrm{n}}\frac{\partial n(x,t)}{\partial x} - \frac{n(x,t)}{\tau_{\mathrm{n}}} + g(t,x) \\
\frac{\partial p(x,t)}{\partial\, t} &= D_{\mathrm{p}}\frac{\partial^2 p(x,t)}{\partial x^2} \mp v_{\mathrm{p}}\frac{\partial p(x,t)}{\partial x} - \frac{p(x,t)}{\tau_{\mathrm{p}}} + g(t,x)
\end{aligned} \tag{5.4}$$

where τ_{n} and τ_{p} are the excess carrier lifetimes, D_{n}, D_{n1}, D_{p} and D_{p1} are the diffusion coefficient of the electrons and holes outside and inside depletion region, respectively, $g(x,t)$ is the volume generation rate due to a light input, and v_{n} and v_{p} are the hole and electron drift velocities. In general, these velocities depend on the electric field. Since the photodiode is reversely biased, there is strong electric field inside the depletion region so drift velocities are maintained

at their saturation values.

When the input light pulse is incident on the device, the generation rate $g(x,t)$ is:

$$g(x,t) = \Phi(1-R)[H(x) - H(x-L)]\frac{(1-e^{-\alpha K})}{K}\delta(t) \tag{5.5}$$

where Φ is the incident light flux, R is reflectivity of the surface, K is the depth of the polysilicon layer, l is the width of the polysilicon layer, α is absorption coefficient and H and δ are Heaviside and Dirac pulses, respectively.

One way to solve drift-differential equations is first to simplify them by two substitutions. The substitution $n(x,t) = \exp(-t/\tau_n)N(x,t)$ is placed into thedrift-diffusion equation (5.4), where τ_n is the electron recombination lifetime. This reduces this equation to:

$$\frac{\partial N(x,t)}{\partial t} = D_n\frac{\partial^2 N(x,t)}{\partial x^2} \pm v_{\mathrm{n}}\frac{\partial N(x,t)}{\partial x} + g(t,x) \tag{5.6}$$

Then, substituting $\zeta = x \pm v_{\mathrm{n}}\, t$ and $\theta = t$ into equation 5.5, the following partial differential equation is obtained

$$\frac{\partial N(\xi,\theta)}{\partial \theta} = D_{\mathrm{n}}\frac{\partial^2 N(\xi,\theta)}{\partial \zeta^2} + g(\zeta,\theta) \tag{5.7}$$

The above equation is a well-know equation of thermal conduction [13] and the final solution (after restoring the variables) is:

$$\begin{aligned} n(x,t) = {} & \Phi(1-R)e^{-\frac{t}{\tau_n}}\frac{(1-e^{-\alpha K})}{K}H(t) \\ & \times\frac{1}{2}\left[\mathrm{erf}\left(\frac{L-x\mp v_{\mathrm{n}}t}{2\sqrt{D_n t}}\right) + \mathrm{erf}\left(\frac{x\pm v_{\mathrm{n}}t}{2\sqrt{D_{\mathrm{n}}t}}\right)\right] \end{aligned} \tag{5.8}$$

A similar analytic expression follows for holes, by simply replacing $\pm v_{\mathrm{n}}$ with $\mp v_{\mathrm{p}}$ and D_{n} with D_{p}.

The associated photocurrent $i_1(t)$ can be obtained by volume integration [10] of the conduction current density which consists of the photo-generated carriers moving over the graded depletion region, and dividing the result by the depletion region width L :

$$i_1(t) = \frac{qW}{h\nu}(1-R)\frac{(1-e^{-\alpha K})}{K}\Phi H(t) \times \sum_{j=n,p} e^{-\frac{t}{\tau_j}} [E_1(t, v_j, D_j) + E_2(t, v_j, D_j)] \tag{5.9}$$

where W is the width of the poly photodiode (see Figure 5.14). The functions $E_1(t, v_j, D_j)$ and $E_2(t, v_j, D_j)$ are defined in terms of error functions and exponential functions, respectively:

$$\begin{aligned} E_1(t, v_j, D_j) = \quad & -(D_j + v_j^2 t)\mathrm{erf}\left(\frac{v_j t}{2\sqrt{D_j t}}\right) \\ & -\frac{1}{2}\mathrm{erf}\left(\frac{L - v_j t}{2\sqrt{D_j t}}\right)[v_j^2 t - v_j L + D_j] \\ & +\frac{1}{2}\mathrm{erf}\left(\frac{L + v_j t}{2\sqrt{D_j t}}\right)[v_j^2 t + v_j L + D_j] \end{aligned}$$

$$\begin{aligned} E_2(t, v_j, D_j) = \quad & \frac{v_j\sqrt{D_j t}}{\pi}\left[\exp(-\frac{(L - v_j t)^2}{4D_j t}) \right. \\ & \left. + \exp(-\frac{(L + v_j t)^2}{4D_j t}) - 2\exp(-\frac{v_j t^2}{4D_j t})\right] \end{aligned} \tag{5.10}$$

For the case where the diffusion inside the junction is negligible and excess carrier lifetime is longer than the carrier transit time (as is the case for CMOS poly-diodes without an intrinsic layer), the impulse response of the polysilicon diode can be simplified to:

$$i_1(t) = qW\Phi(1-R)\frac{(1-e^{-\alpha K})}{K}\delta(t)\sum_{j=n,p} v_j(L - v_j t)H(L - v_j t) \tag{5.11}$$

where L is the length of the poly photodiode (see Figure 5.14).

For polydiodes with a large intrinsic layer, the recombination lifetime is much shorter than the transit time and the impulse response is given in [10]. In general, the time response of the CMOS poly photodiode on any input light signal can be calculated by means of convolution between the signal and the impulse response [13]. In this manner, the photocurrent due to Gaussian irradiation (in time) is obtained by convolving the input with $i_1(t)$. Here we assumed that the time duration of the input Gaussian pulse is 100 ps, and the width of de-

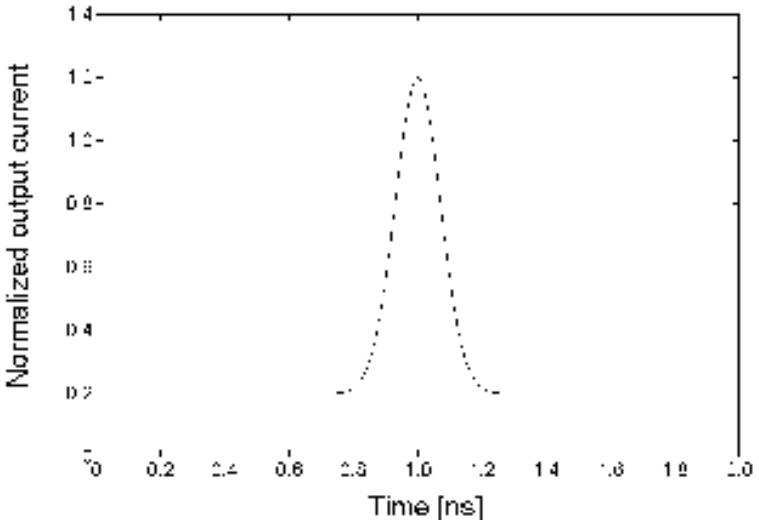


Figure 5.16: Simulated drift current time response on input Gaussian light pulse with 100 ps time duration (FWHM). Diffusion inside depletion region is neglected.

pletion region L is 100 nm, which is realistic in standard CMOS processes due to the high doping concentrations of n-poly and p-poly regions. The result is presented in Figure 5.16. This figure shows that the time response of depletion region of poly diode, with neglected diffusion, is as fast as the input Gaussian pulse. Therefore, there is no speed limitation far in a GHz range.

Because of the narrow depletion region, the diffusion length of the excess carriers is larger than the depletion region width and there are almost no carriers recombined in this region. Moreover, the excess carrier profile $n, p(x,t)$ is almost constant and the simplified formula for the drift frequency $f = 0.4v_s/L$ can be used. The calculation presented in this section are however useful for the lateral photodiodes with larger depletion widths (> 1 µm).

If the recombination process dominates the response of the polysilicon diode, one can take the recombination time much shorter than the transit time. Thereby, the impulse current response of the lateral polysilicon diode can be expressed as [13]:

$$i(t) = \frac{q}{h\nu L}(1-R)[1-\exp(-\alpha K)]\theta(t)\sum_{j=\mathrm{n,p}} v_j \exp(\frac{t}{\tau_j}) \qquad (5.12)$$

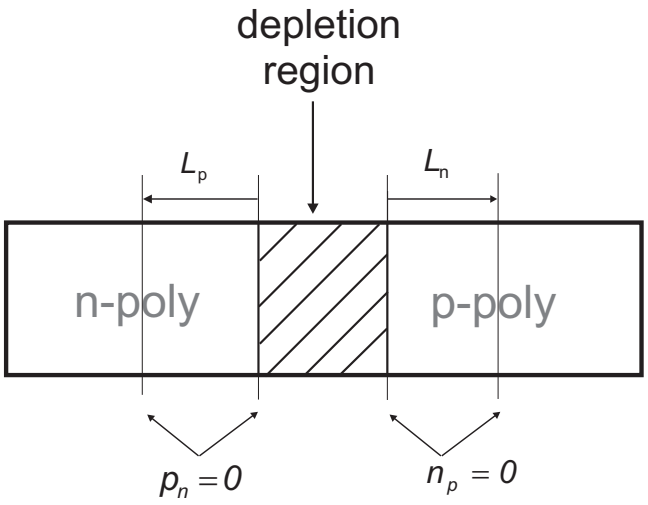


Figure 5.17: The boundary conditions for the diffusion current inside polysilicon diode.

5.7.2 Diffusion current outside the depletion region

If the polysilicon photodiode is realized using two highly (inversely) doped regions without an intrinsic layer in between, the width of the depletion region is very small and the diffusion current outside of this region will also contribute the overall photocurrent. This diffusion current is calculated in the n-region and p-region using the procedure similar to those explained in chapter 3. Here, the one-dimensional lateral diffusion equation is solved. Starting from the diffusion equations, the carrier profile is calculated using the boundary conditions shown in Figure 5.17:

- the excess carriers concentration on the edge of depletion region is zero.
- the excess carriers concentration on the diffusion distance L_j, j=n, p is zero.

From the calculated carrier profile, the diffusion current can be further calculated at the border of the depletion region:

$$i_2(t) = 4qWL\Phi\frac{1-e^{-\alpha K}}{K}\sum_m \sum_{j=n,p} \frac{L_j}{\tau_j} e^{-[(1+(2m-1)^2\pi^2]\frac{t}{\tau_j}} \tag{5.13}$$

If the photodiode consists of N n-p fingers, the total photocurrent i_{tot} is directly proportional to the number of fingers, $i_{\text{tot}} = N \cdot i_2(t)$. Due to the exponential

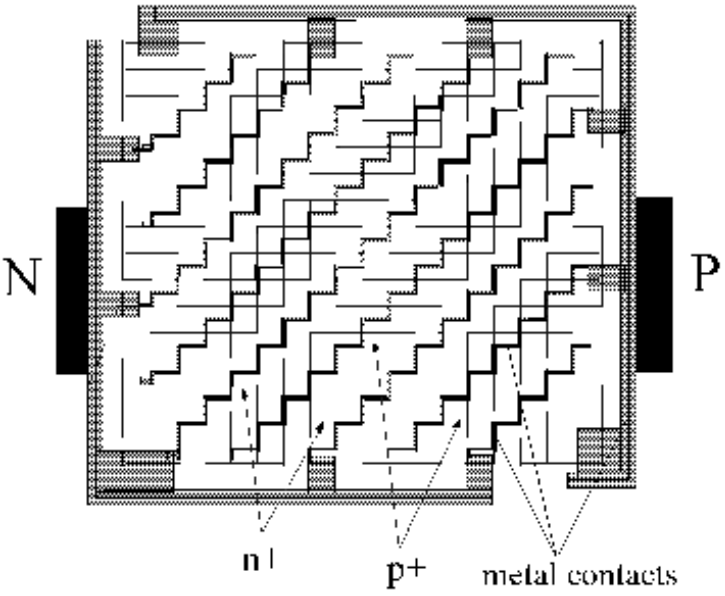


Figure 5.18: "Jagged" poly photodiode with an order of magnitude larger light sensitive area in comparison with a single poly photodiode. The lateral diode dimensions are 45×45 μm.

term in the equation 5.13, the speed of the diffusive response is mainly determined by the lifetime of the excess carriers. Noting that this lifetime is short (50 ps), the response speed of diffusion current component will be in hundreds of ps range. The overall photocurrent is obtained as the sum of drift and diffusion currents.

5.8 Frequency characterization of the polysilicon photodiode

The light sensitive part of a poly photodiode consist only of a small depletion region area plus the area outside this region proportional to a diffusion length of holes and electrons. This diffusion length is very small in comparison with that in a monosilicon diode. The depth of the polysilicon in standard CMOS technology (K in Figure 5.14) is only about 0.2μm and it also contributes to the poor responsivity of poly photodiode on vertical incident light. According to this, a single polydiode would have very small active area and very low quantum

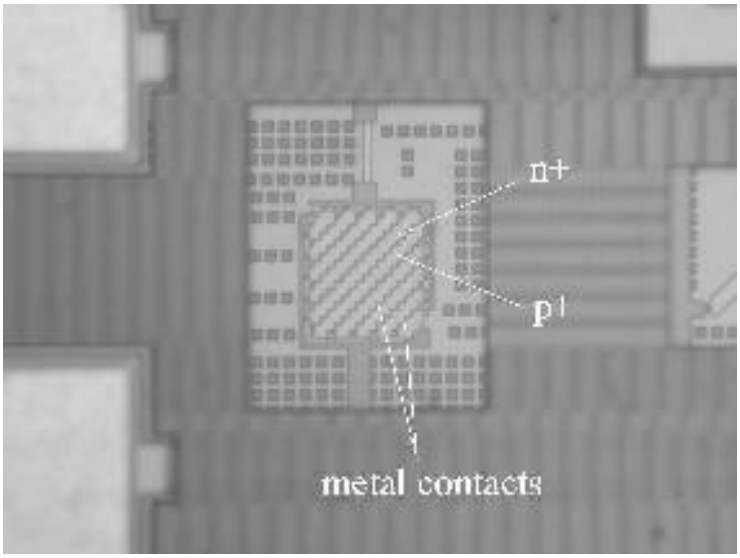


Figure 5.19: Layout of "jagged" poly photodiode designed to increase the overall light sensitive area.

efficiency (<1 %). In order to increase an active photodiode area, a "jagged" polysilicon diode consisting of a number of polydiodes connected in parallel was designed (Figure 5.18.). The calculated overall active area is about 13 times larger than that in the single polydiode. This implies that the expected output signal is 22 dB larger than in a single polydiode. However, there are rounding effects at the many corners in the poly p-n structure that decreases this value.

The poly-diode is designed in standard 0.18 µm CMOS technology, and the diode-layout is shown in Figure 5.19. The measurements of the photocurrent showed that the photocurrent is 17 dB larger than the measured photocurrent of the single polysilicon photodiode. As expected, the depletion area increased less than proportional.

The frequency response of the photocurrent is measured using Agilent E4404E Spectrum Analyzer. The response of the polysilicon photodiode is measured from 1 MHz up to 6 GHz. For frequencies up to 1 GHz, the signal from the photodiode was amplified using a Minicircuits ZFL 1000LN 0.1-1000 MHz amplifier. For frequencies above, we used 0.5-26.5 GHz Agilent 83017A Microwave system amplifier.

The transmitter part consist of the 850 nm 10 Gb/s VCSEL and its driver

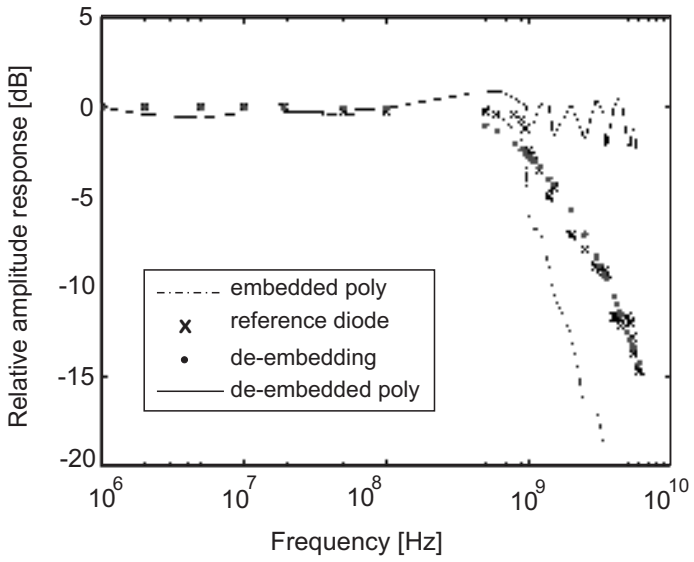


Figure 5.20: Frequency response of de-embedded polysilicon photodiode.

amplifier. An HP 8665B frequency synthesizer was used as a modulating signal source up to 6 GHz. The signal was coupled into the photodiode using the multimode fiber with 50 µm core diameter.

The same setup is for calibration purposes used to measure a reference photodiode (Tektronix SA-42) response, which has according to specifications, 7 GHz-3 dB frequency. The response of the reference diode in the setup is presented in Figure 5.20.

The polysilicon photodiode frequency characteristic was de-embedded out of the parasitic bondpad capacitance and parasitic capacitance of the picoprobes. The de-embedded frequency characteristic is presented in Figure 5.20. The characteristics is almost flat up to 6 GHz, meaning that the measured bandwidth of the polydiode is even larger. The high intrinsic (physical) bandwidth is due to a short excess carrier lifetime (about 50 ps [5]). The calculated capacitance of the poly-diode using the 0.18 µm technology parameters is small ~ 0.2 pF, which can result in the large electrical bandwidth (depending on the input resistance).

5.8.1 Noise in polysilicon photodiodes

Large leakage current in polysilicon photodiodes for rather low values of reverse voltages (20µA for 1.5 V) can cause large noise of the photodiode. For this reason, the following section presents a leakage current in a polysilicon photodiode.

5.8.2 Dark leakage current in the polysilicon diode

Figure 5.15 shows the diode reverse I-V characteristic. The leakage current is large as a result of the grain-boundary trap-assisted band-to-band tunnelling and field-enhanced emission rate [8], [14]. Also, since the doping concentration of both diode regions is high the width of the depletion region is very small, even though the junction behaves as a graded one. The reverse current is given with [7]:

$$J_{\mathrm{r}} = qN_{\mathrm{t}}kT\pi\frac{\sigma v_{\mathrm{th}}n_{\mathrm{i}}}{2}\frac{W_{\mathrm{d}}(V_{\mathrm{R}})}{L_{\mathrm{g}}}\exp\left[\left(\frac{\alpha}{E_0}\right)^{n}(V_{\mathrm{R}}+V_{\mathrm{b}})^{\frac{2n}{3}}\right] \tag{5.14}$$

where

$$\alpha = \left(\frac{9}{32}\frac{qa}{\epsilon}\right)^{\frac{1}{3}} \tag{5.15}$$

with a [cm^{-4}] is the dopant concentration gradient and V_{R} and V_{b} are applied and built-in potentials voltages respectively. E_0 is a threshold electric field in the depletion region from which the emission amplification becomes significant (depends on the temperature and on the material nature), v_{th} is thermal velocity often given as $v_{\mathrm{th}} = \sqrt{3kT/m_e, h}$ where m_e, h is the mass of the electron or hole, n_i is intrinsic carrier concentration, σ is an effective capture cross-section [cm^2], $W_{\mathrm{d}}(V_{\mathrm{R}})$ is depletion region width [cm], L_{g} is the grain size in polysilicon [cm], N_{t} is the grain boundary trap density [cm^{-2}eV^{-1}] and n is the exponential argument which generally varies between 0 and 1.5.

The above equation includes field enhancement of the emission rates of traps in the depletion region [15]. The value of E_0 depends also on the junction area. In our case we took the approximated value of $E_0 = 2{\cdot}10^5$ V/cm. H.C. de Graaf et. al. showed in [16] that the trap energy distribution N_{t} is U-shaped with the broad minimum around mid-gap. For most purposes it can be approximated by a homogeneous distribution with $N_t = 3 - 5 \cdot 10^5$ cm^{-2}eV^{-1}. The capture-cross section for polysilicon is about $\sigma = 10^{-15}$cm^2, and the thermal velocity is about $1.2 \cdot 10^5$ m/sec.

According to both the calculations and the measurements of the reverse diode characteristic, one can conclude that if the reverse voltage value is higher than 0.7 V, the leakage current is higher than 500 nA. The high leakage current multiplied by the GHz bandwidth results in a high shot-noise value. This value together with additional Johnson noise, decreases the sensitivity of the polysilicon photodiode. For higher voltages (>1.5 V), the value of the leakage current can be even higher than the magnitude of the photocurrent, and this poly-diodes can be used only in high optical-power applications like detection of pulsed light signals and for trigger applications.

5.9 Time domain measurements

The characterization of the polysilicon photodiode is also performed in the time domain. Firstly, a picosecond laser with $\lambda = 650$ nm was used as a transmitter. The pulse width of the picosecond laser is 200 ps and the peak optical power is 1 mW (0 dBm). This rather large optical power was necessary due to the low quantum efficiency of poly-diode, which will be shown in section 5.10 of this chapter. The light was coupled from the laser to the poly-diode using multimode fiber with 50 µm core-diameter. The poly-diode was not packaged, and "on-chip" measurements were done using RF probes. The diode DC biasing of V_R=-0.5 V, was provided using a bias-tee. The larger (negative) voltages causes large leakage currents ($> 0.5\mu$A), as shown in Figure 5.15. This leakage currents can be even larger than the photocurrent, and thus strongly limit the bit-detection which is shown in the following part of the chapter.

The alignment of the fiber on the photodiode was done using micro-manipulators of the probe-station. By shining the light from the pulse-laser, the RF signal from the poly photodiode shown in Figure 5.18 is measured first without the external amplifier. However, the signal from the photodiode was too weak to be seen at the digital oscilloscope with the minimal sensitivity of 1 mV/div. For this reason, the external amplifier with 750 Ω transimpedance was placed after the photodiode, and the result is shown in Figure 5.21. The maximum measured output voltage is 1.2 mV, meaning that the maximum photocurrent is 1.6 µA. Since the maximum input optical power is 1 mW, the poly-diode responsivity as well as its quantum efficiency is clearly very low. The exact numbers are given in section 5.10.

The pulse width in Figure 5.21 is about 1 ns which is larger than calculated

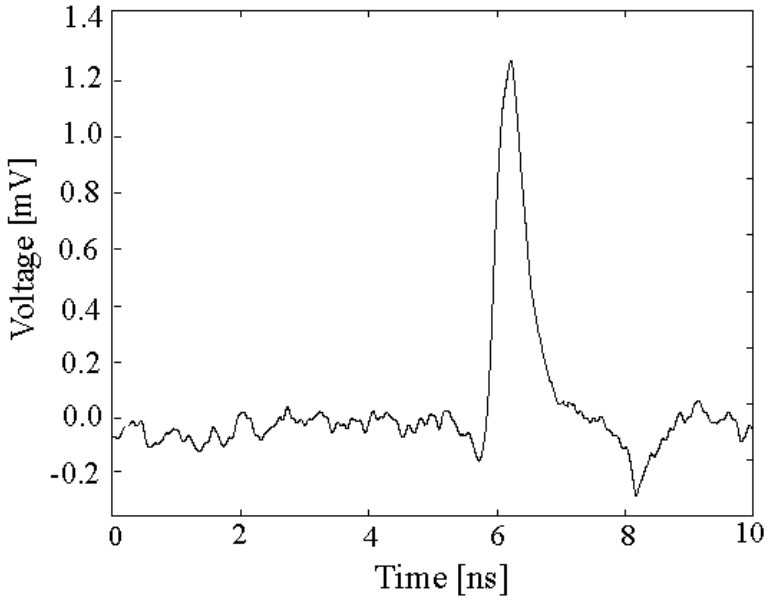


Figure 5.21: Transient response of poly photodiode on 200 ps input light pulse (transimpedance 750 Ω, $\lambda = 650$ nm).

in previous sections of this chapter. This can be explain with a fact that poly-diode is *embedded* inside the resistances, capacitances and inductances of the bondpads, connectors, and series resistances of the diode itself. In order to *de-embedd* [10] the poly-diode we used the Tektronix SA-42 photodetector with 7 GHz-3 dB performance. The same TIA and coaxial cables are used for the measurements. The measured time response of this photodetector on 650 nm, is presented in Figure 5.22.

Having a reference diode response within the measurements set-up (embedded), and calculating the diode response according to its manufacture specifications (de-embedded), it is possible to calibrate out the measurements set-up (de-embedding). The time-domain convolution between the de-embedded diode response R_{deemb} and the de-embedding R_{d} results in the embedded diode response R_{emb}, presented with Fredholm integral [17]:

$$R_{\mathrm{emb}}(t) = \int_{-\infty}^{\infty} R_{\mathrm{deemb}}(t-x) R_{\mathrm{d}}(x) dx \tag{5.16}$$

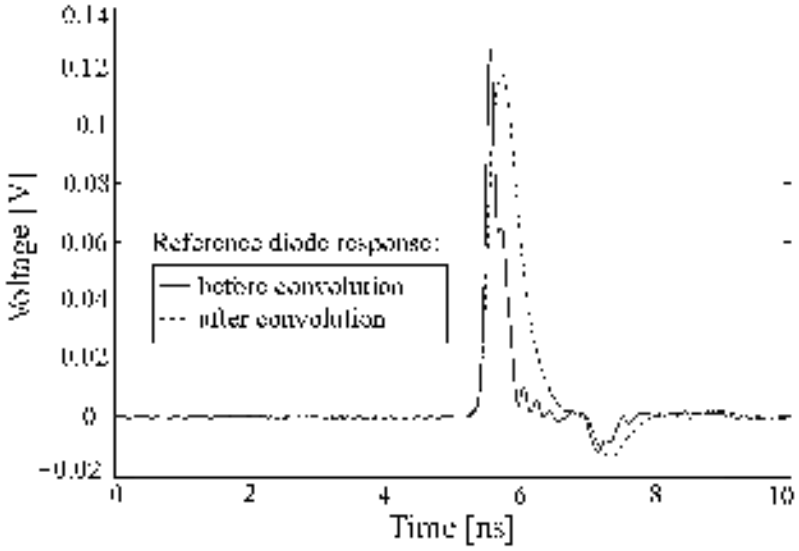


Figure 5.22: Transient response of the reference photodiode (7 GHz-3 dB, transimpedance 750 ohm, $\lambda = 650$ nm) and its convolution with the *difference* between embedded and de-embedded poly photodiode)

In order to do the de-embedding i.e. calibrating out the measurement equipment, the equation above has to be solved for $R_{\mathrm{d}}(x)$. This is a complex deconvolution problem that can be only solved numerically [17]. The problem should be discretize in order to obtain a structured (Toeplitz) matrix. This matrix can speed up the computations with the FFT algorithm. Nevertheless, one still need to add regularization. To circumvent this complex solving algorithms, we took the de-embedding R_d as a sum of complex exponential functions, which were adapted such that the equation (5.16) is satisfied. The parasitic capacitances of the bondpads were also included.

Further, the convolution between a reference diode signal and the de-embedding R_d is calculated and the result is shown in Figure 5.22. The estimated speed of the de-embedded polysilicon photodiode is at least as fast as a reference diode which has 7 GHz cutoff frequency.

Secondly, a picosecond laser with $\lambda = 400$ nm was used as a transmitter and the output signal from the polysilicon photodiode is presented in Figure 5.23. The signal shape is similar to that shown in Figure 5.22 with four times larger signal amplitude. This complies with the earlier (theoretical) findings reported

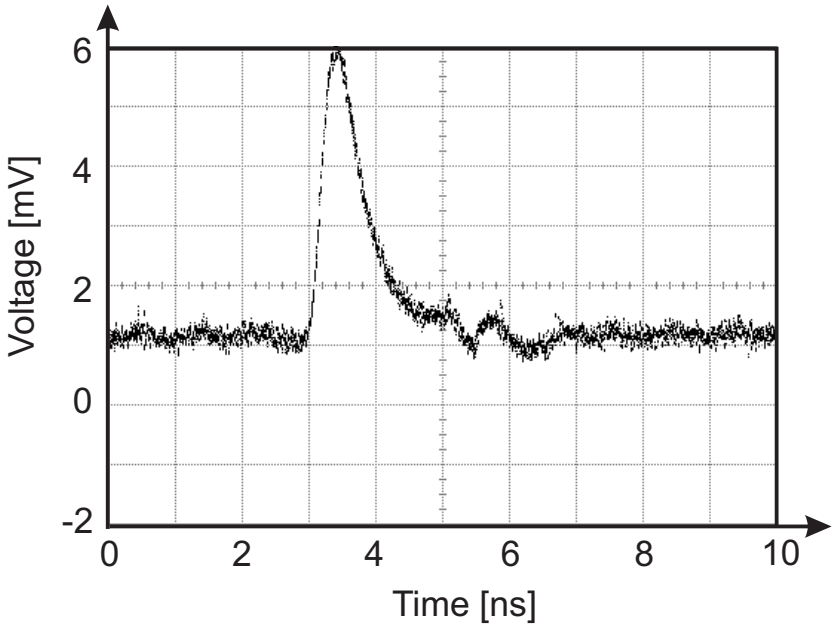


Figure 5.23: Transient response of embedded poly photodiode on 200 ps input light pulse (transimpedance 750 Ω, $\lambda = 400$ nm).

in chapter 2, Figure 2.9, since the absorbtion coefficient of light in polysilicon is four time larger.

5.10 Quantum efficiency and sensitivity

An important feature of polysilicon is that the light absorption depth is four times larger than in monosilicon. Therefore, for the same depth of the polysilicon and silicon material, the expected quantum efficiency (QE) is larger for the polysilicon [5]. Previous chapters showed that the photocurrent of the polysilicon photodiode is 1.6 µA for 1 mW input optical power. The calculated responsivity of the poly photodiode is thus only 1.6 mA/W. The metal coverage area of the polydiode shown in Figure 5.19 is 15 %, meaning that the optical power absorbed by the active area of the poly-diode is 8.5 mW. However, the responsivity of the poly-diode is still very low. Using equation (2.14) from section 2.5.2, the maximum calculated responsivity (η=1) for $\lambda = 650$ nm is 0.52 A/W. By dividing the calculated poly-diode responsivity and the maximum re-

sponsivity, the calculated quantum efficiency is only $\eta = 0.3\%$. For blue-light $\lambda = 400$ nm, the calculated photocurrent is 8 µA for 1 mW optical power; this results in the responsivity of 8 mA/W. The maximum calculated responsivity for $\lambda = 400$ nm is 0.32 A/W. Therefore, the calculated quantum efficiency is thus $\eta = 2.5\%$.

Using simplified formula for the maximum achievable quantum efficiency for both wavelengths $\eta_{\max} \sim 1 - e^{-\alpha K}$, the calculated values are 21% and 97% respectively. The active (light sensitive) detector area A_{eff} is can be approximately determined using the following equation:

$$\eta_{\text{meas}} = \eta_{\max} \frac{A_{\text{tot}}}{A_{\text{eff}}} \tag{5.17}$$

where A_{tot} is a total photodiode area. For the simplicity of calculations, the bottom reflection of light is neglected as well as the reflection on the air/semiconductor interface. The calculated value of the active poly-diode area is less than 2% implying very thin depletion regions as well as the small diffusion area outside it[2].

BER and S/N ratio

For 25 µW peak-to-peak input optical power (-19 dBm average optical power) which is specified in the Gigabit Ethernet standard, the calculated photocurrent of the polydiode for 650 nm and 1.6 mA/W responsivity is 40 nA. For $V_R = -0.5$ V reverse bias of the polydiode, the measured leakage current is 180 nA. In this subsection, we present the data-rate and the bit-error-rate analyzes using the procedure explained in chapter 3, with the assumption that the subsequent TIA is noise-free. The signal is already calculated (40 nA) and in order to calculate the noise, we use equation (3.25). To achieve S/N=8 i.e. a noise current of 5 nA, the bandwidth of the polydiode for -19 dBm input optical power is limited to only ∼400 MHz. With the included noise in the subsequent TIA, the calculated bandwidth is decreased. Moreover, for larger bias voltages (V_{R} >-1 V), the leakage current of the poly-diode increases dramatically as shown in Figure 5.15. This large leakage can limit the poly-diode bandwidth in the low MHz range.

The improvement of the quantum efficiency in poly-diodes is typically done using two methods. First, it is increased by building the light reflectors i.e.

[2]Multiplying the maximum quantum efficiency $\eta_{\max}$ with the calculated active poly-diode area $21\% \cdot 2\%$ the poly-diode quantum efficiency for $\lambda = 650$ nm is about 0.4%.

design a resonant-cavity photodiode [11]. This is however not available in standard CMOS technology. The second method is to design a PIN poly photodiode [10], which includes non-doped polysilicon layer, which is also not available in standard CMOS technology.

5.11 Conclusion

For the lower boundary of the CMOS wavelength-sensitivity range $\lambda = 400$ nm, the bandwidth of the photodiodes in 0.18 µm CMOS technology is in the range of 170 MHz up to 6 GHz.

The maximum intrinsic bandwidth of 6 GHz, is achieved with nwell/p-substrate photodiode designed in separate-well technology because of the maximum depletion region area. Using the adjoined-well technology, the maximum calculated intrinsic bandwidth is achieved using single p+/nwell photodiode and it is up to 3 GHz. This diode structure is obtained by disconnecting the diode substrate. The influence of the nwell/p+ width (number of fingers) is negligible on the intrinsic bandwidth, because the small p+ depth is the minimal (and most important) distance for the diffusion process. The number of nwell/p+ fingers however, does influence the overall bandwidth: the higher the number of fingers i.e. the lower the nwell/p+ width, the higher the photodiode capacitance. Consequently, the electrical bandwidth decreases.

For the nwell/p-substrate photodiode in the adjoined-well technology, the nwell width is very important in the diode intrinsic bandwidth. The highest calculated bandwidth is 700 MHz achieved with a minimal nwell-width (2 µm). Larger nwell widths decrease the diode intrinsic bandwidth by almost a factor two. In addition, the n+ layer, which may exist at the top of the nwell, decreases the diode bandwidth further by a factor two. This is because the high majority carrier concentration inside n+ decreases the minority carrier diffusion constant and thus, the bandwidth. The maximum bandwidth of nwell/p-substrate photodiode is obtained with the minimum width of the nwell region and without the n+ layer at the nwell-top.

In the second half of this chapter a lateral polysilicon photodiode in standard 0.18 µm CMOS technology is analyzed. The analytical calculations, and the measurements in the frequency and the time domain showed that polysilicon photodiode has very large bandwidth $f_{3\mathrm{dB}} > 6$ GHz. Due to the small excess carrier lifetime, the slow diffusion limitation on the intrinsic (physical)

polydiode bandwidth is negligible. The electrical bandwidth limitation is also minimal: the small diode parasitic capacitance is proportional to the low depth of the polysilicon layer. The big advantage of polydiode is that the parasitic capacitance towards the substrate is also very low because of the thick field oxide layer in comparison with the conventional thin gate oxide.

The disadvantage of the polydiode in standard CMOS technology is the low quantum efficiency ($\leq$2.5 %). This is because of the very small light sensitive area: the *width* of the depletion region is small because of the high doping concentrations of the n-region and p-region. The *depth* of the poly-diode is limited by the technology. The "out of junctions" active diode area is also small due to the small diffusion lengths of the excess carriers. These diffusion lengths are determined by the short carrier lifetime (50 ps). However, the quantum efficiency can be improved by building light reflectors i.e. to design a resonant-cavity photodiode. Another method is to design of PIN poly photodiode with large intrinsic (light sensitive) area. Nevertheless, these methods are *not included* in standard CMOS technology.

References

[1] Wei Jean Liu, Oscal T.-C. Chen, Li-Kuo Dai and Far-Wen Jih Chung Cheng: "A CMOS Photodiode Model", *2001 IEEE International Workshop on Behavioral Modeling and Simulation*, Santa Rosa, California, October 10-12, 2001.

[2] H. Zimmermann, H. Dietrich A. Ghazi, P. Seegebrecht: "Fast CMOSIntegrated Finger Photodiodes for a Wide Spectral Range", *ESSDERC 2002*, pp. 435-438, 24-25 September, Italy

[3] S. M. Sze: "*Physics of semiconductor devices*", NewYork: WileyInterscience, 2-nd edition, p. 81, 1981.

[4] B. Razavi: "*Design of Analog CMOS Integrated Circuits*", McGraw-Hill, 2001.

[5] Kamins T.: "*Polycrystalline silicon for integrated circuits and displays*",, Boston : Kluwer Academic Publisherse, 2-nd edition, p. 240, 1998.

[6] McKelvey, J. P.:"*Solid-State and Semiconductor Physics*", New York: Harper& Row, pp. 340, 1966.

[7] A. Aziz, O. Bonnaud, H. Lhermite and F. Raoult: "Lateral polysilicon pn diode: Current-voltage characteristics simulation between 200K and 400K using a numerical approach", *IEEE Trans. on Electr. Devices*, vol. 41, pp. 204-211, 1994.

[8] Manoliu and T. Kamis, "P-N junction in polycrystalline-silicon films", *Solid-State Electronics*, Vol. 15, pp. 1103-1106, 1972.

[9] Ker, M. D, Chen T.Y., Wang T.H. and Wu C.Y: "On-chip ESD protection design by using polysilicon diodes in CMOS process", T*rans. Soild-State Circuits*, vol. 36, no. 4, pp.676-686, April 2001.

[10] Kim, D. M., Lee, J. W., Dousluoglu, T., Solanki, R. and Qian, F: "High-speed lateral polysilicon photodiodes", *Semiconductor Sci. Technology*, vol. 9, pp. 1276-1278, 1994.

[11] Diaz, D.C., Scho, C.L., Qi, J, Campbell, J.C.: "High-speed Polysilicon Resonant-Cavity Photodiode with $S_iSO_2 - S_i$ Bragg reflector", *Photonics Technology Letters*, vol. 9, no.6, pp. 806-808, June, 1997.

[12] Plummer J., Deal M., Griffin P.: "*Silicon VLSI technology; fundametals, practice and modelling*", Prentice Hall, p. 560, 2000.

[13] Lee, J.W., Kim., D. M.: "Analytic time domain characterization of *p-i-n* photodiodes: effects of drift, diffusion, recombination, and absorption", *J. Appl. Phys.*, vol. 6, pp. 2950-2958, March 1992.

[14] M. Dutoit and F. Sollberger: "Lateral Polysilicon p-n Diodes", S*olid-State Science and Technology*, vol.125, No.10, pp1648-1651, October 1978.

[15] D. W. Greve, P. Potyraj and A. Guzman: "Field-enhanced emission and capture in polysilicon pn junctions", *Solid-State Electronics*, vol. 28, pp. 1255-1261, 1985.

[16] H.C. de Graaf, M. Huybers: "Grain-boundary states and the caharcteristics of lateral polysilicon pn junctions", *Solid-State Electronics*, vol. 25, pp. 67-71, 1982.

[17] P. C. Hansen: "*Deconvolution and regularization with Toeplitz matrices*", Numerical Algorithms, vol. 29, 2002, pp. 323-378.

CHAPTER 6

High-speed integrated photodiodes for 400 nm<λ<850 nm and for various CMOS technologies

This chapter analyzes the frequency behavior of photodiode in standard CMOS in general for the whole wavelength-sensitivity range 400 nm<λ<850 nm. The maximum junction depths are taken between 4 μm and 0.05 μm, which correspond to older and the new-coming CMOS technologies. Independent on CMOS technology, for all wavelengths for which 99% of the light is absorbed in the area above the maximum junction depth, the lower the wavelength, the smaller the total photocurrent and the lower the diode bandwidth. This bandwidth is however still in the hundreds of MHz range.

For the wavelengths with the 1/e-absorption depth much larger than the deepest junction depth, the substrate current dominates the total response. The larger the wavelength, the lower the photodiode bandwidth. Minimum photodiode bandwidth is limited in the low MHz range. The bandwidth of the photodiodes can be improved using the analog equalization method, given in chapter 4. This chapter analyzes the starting and the maximal equalization frequencies depending on CMOS technology and input wavelength.

6.1 Introduction

Chapters 3 and 5 showed that the photodiode intrinsic bandwidth increases more than two orders of magnitude when the input wavelength ranges from 850 nm to 400 nm. For 850 nm, the substrate current dominates the total response of the photodiode designed in standard CMOS technology. The bandwidth of the substrate current is limited in the low MHz range and so is the total diode bandwidth. Chapter 3 described that the photodiode layout (nwell width) does not influence the total diode bandwidth for $\lambda = 850$ nm. On the other hand, chapter 5 described that for $\lambda = 400$ nm, the nwell width is very important for the total bandwidth. The smaller the width, the larger the bandwidth. The substrate has no influence on the photocurrent, since there are no carriers generated in this diode region.

This chapter analyzes the frequency response of CMOS photodiode in general for the whole wavelength sensitivity range 400nm<λ<850 nm. For this purpose, a device-layer/p-substrate photodiode is introduced in Figure 6.1. Device layers are those that make an active semiconductor device (diodes in this case), like nwell, n+ and p+ layers and this layer is assumed as the top layer of CMOS photodiode. The influence of different photodiode regions on the total photocurrent will be also analyzed in detail. The depth of the device-layer is in the range from 4 µm to 0.05 µm.

In the second part of the chapter, using the calculated frequency behavior of the photodiode, we will calculate the maximum equalization frequency i.e. maximum bandwidth improvement of a CMOS photodiode by using an analog equalization method explained in chapter 4.

Equation 3.1, can be rewritten as a function of CMOS technology-scale i.e. function of the corresponding depth of the device-layer L_x, as well as input wavelength λ. In general, all photocurrent components defined with equation (3.1) are also dependent on the dimensions of the corresponding diode regions. However, to stress the importance of the technology and wavelength on the total diode response, these two parameters are written in equation (6.1). For the case where the device-layer depth is lower than the light absorption depth $1/\alpha$, all diode layers contribute to the overall photocurrent:

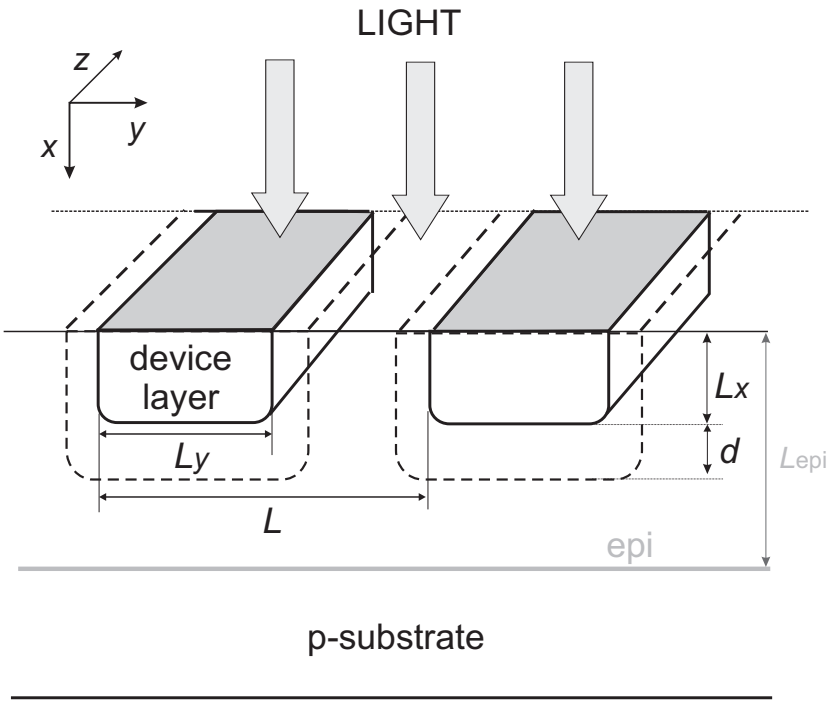


Figure 6.1: A general photodiode in standard CMOS technology.

$$
\begin{aligned}
I_{\text{tot}}(L_x,\lambda,s) &= I_{\text{drift}}(L_x,\lambda)\frac{1}{1+\dfrac{s}{s_{\text{drift}}(L_x,\lambda)}} \\
&+ \sum_k I_{\text{diff}_k}(L_x,\lambda)\frac{1}{\sqrt{1+\dfrac{s}{s_{\text{diff}_k}(L_x,\lambda)}}}
\end{aligned}
\tag{6.1}
$$

where $I_{\text{diff}_k}(L_x,\lambda)$ and $I_{\text{drift}}(L_x,\lambda)$ are the amplitudes of the diffusion and the drift regions of photodiode and s_{diff_k} and s_{drift} are the poles of the diffusion currents and drift current, respectively. Chapter 3 described that diffusion processes in general have a low roll-of (∼10 dB/decade). For this reason, the drift currents are approximated here as square-root pole functions. This also provides easier understanding of the total photodiode behavior.

For device-layer depth larger than the light absorption depth $1/\alpha$, the total photocurrent consists of the diffusion current in the device-layer and the drift

current in the vertical and lateral depletion regions. The diffusion current is presented with equation (3.8), and the drift current with equation (3.17). The following sections of this chapter present the photocurrent components from equation (6.1) as well as the total photocurrent of the device-layer/p-substrate photodiode as a function of CMOS technology and input wavelength. The amplitude of the photocurrent components will be normalized with the amplitude of the total photocurrent I_{tot}. The input wavelength takes any value from the range 400 nm<λ<850 nm.

6.1.1 Device layer - photocurrent amplitude

The first photocurrent component from equation (6.1) that is analyzed, is the device-layer diffusion current. This section investigates in particular *maximum amplitude* of this diffusion current, as a function of the CMOS technology and wavelength. The quantum efficiency is assumed to be maximal (η=100 %).

The amplitude of the device-layer photocurrent is calculated following the procedure given in section 3.2.1. That section analyzed the nwell diffusion current as a function of frequency. Replacing the hole diffusion length L_p in equation (3.8) with general diffusion length L_{gen}, the maximum amplitude of the diffusion current in the device-layer (nwell, n+ or p+) is obtained when $s\tau_p \ll 1$:

$$\begin{aligned} I_{\text{diff}_{DL}}(L_x,\lambda) &= 32\Phi_0 \frac{eL_{\text{gen}}^2\alpha}{l\pi^2} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{(2n-1)\pi e^{-\alpha L_x} + (-1)^{\frac{(2n-1)-1}{2}}\alpha L_x}{4\alpha^2 L_x^2 + (2n-1)^2\pi^2} \\ &\times \frac{\frac{2L_{x1}}{L_y}\frac{1}{2n-1} + \frac{L_y}{2L_x}\frac{2n-1}{(2m-1)^2}}{\frac{(2n-1)^2\pi^2 L_{\text{gen}}^2}{4L_x^2} + \frac{(2m-1)^2\pi^2 L_{\text{gen}}^2}{L_y^2} + 1} \end{aligned} \tag{6.2}$$

where L_x and L_y are the corresponding width and length of the device-layer (shown in Figure 6.1) and l is the distance between subsequent device-layers. The diffusion length L_{gen} mainly depends on a doping concentration of the layer [1].

The calculated device-layer current $I_{\text{diff}_{DL}}(L_x,\lambda)$ is illustrated in Figure 6.2. For all wavelengths for which the 1/e-absorption depth is larger than the device-layer depth L_x, the *larger* the wavelength, the lower the photocurrent of the device-layer. On the other side, for the wavelengths for which the light is almost

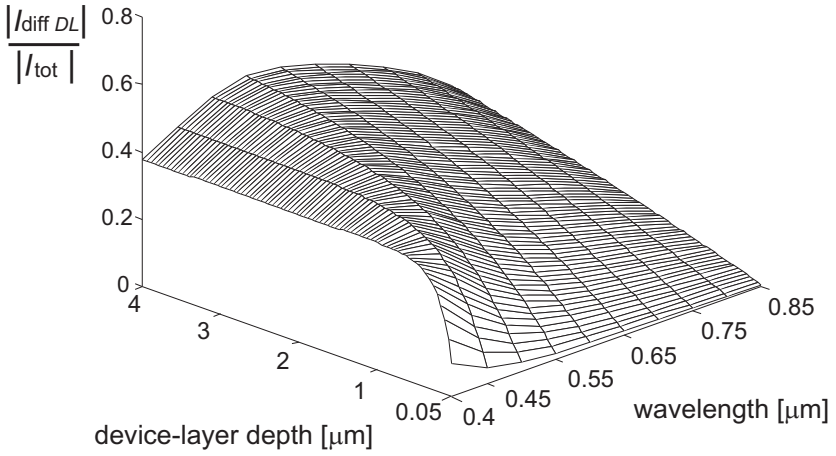


Figure 6.2: The normalized photocurrent of device-layer region versus layer depth (technology dependent) and input wavelength, for device-layer/p--substrate photodiode. The device-layer photocurrent is normalized with a total diode photocurrent.

completely absorbed in the device-layer (99%), the photocurrent corresponds to the diode responsivity shown in Figure 2.11.

Apart from the maximum photocurrent, the other parameter of interest in the device-layer response is its bandwidth $1/(2\pi s_{\mathrm{diff_{DL}}})$.

6.1.2 Device layer - photocurrent bandwidth

The device-layer current is the diffusion current of minority carriers. The bandwidth of the diffusion current depends on the wavelength and CMOS technology. For easier understanding of this dependence, first a light absorption[1] inside the device-layer is illustrated in Figure 6.3. For all wavelengths with the 1/e-absorption depth *larger* than device-layer depth L_x ($L_x \ll 1/\alpha_1, 1/\alpha_2$), the light absorption inside device layer is almost constant with a distance:

$$\exp(-\alpha_1 x) \approx \exp(-\alpha_2 x) \approx Const. \tag{6.3}$$

[1]Light is absorbed exponentially inside the photodiode as explained in chapter 2.

Therefore, the excess carrier concentration gradient towards the junctions can be taken to be independent on input wavelength, [2]. Chapter 3 explained that device-layer bandwidth is proportional to the excess carrier concentration gradient. This gradient is constant, and the bandwidth is wavelength-independent; it can be expressed using equation (3.8) as:

$$f_{3\mathrm{dB}} \approx \frac{\pi D_{\mathrm{gen}}}{2}\left(\left(\frac{1}{2L_x}\right)^2+\left(\frac{1}{L_y}\right)^2+\left(\frac{1}{L_{\mathrm{gen}}}\right)^2\right) \tag{6.4}$$

where D_{gen} is the general diffusion constant of the excess carriers.

Figure 6.3 shows that for wavelengths with the 1/e-absorption depth *smaller* than the device-layer depth (L_x), the total photocurrent is generated completely in the layer. For very small wavelengths λ_4, λ_5, the light is completely absorbed inside the layer. The lower the wavelength, the further the distance between the deepest generated carriers and the junction. The excess carrier concentration gradient drops which results in the smaller bandwidth, as shown in chapter 3. Nevertheless, for λ_4, λ_5, the change in excess concentration gradient with the wavelength is still small, since the deepest carriers are generated away from the bottom junction. The bandwidth dependence on the input wavelength can be approximated from equation (3.8), with $(\lambda/\lambda_{1/e})^{1/3}$, where $\lambda_{1/e}$ is the wavelength with 1/e-absorption depth equal to the layer depth L_x:

$$f_{3\mathrm{dB}} \approx \left(\frac{\lambda}{\lambda_{1/e}}\right)^{\frac{1}{3}}\frac{\pi D_{\mathrm{gen}}}{2}\left(\left(\frac{1}{2L_x}\right)^2+\left(\frac{1}{L_y}\right)^2+\left(\frac{1}{L_{\mathrm{gen}}}\right)^2\right) \tag{6.5}$$

For very small device-layer depths ($< 0.2\mu$m), the doping concentration of the layer is usually higher. As a result, the diffusion constant of the minority carriers is smaller [1]. The device-layer bandwidth, presented in equation (6.5) is proportional to the ratio D_{gen}/L_x and D_{gen}/L_y. This bandwidth increases with the downscaling of the device-layer: the diffusion constant D_{gen} decreases "slowly" in comparison with the downscaling of L_x and L_y.

Chapter 3 described that the maximal bandwidth is achieved with the minimum device-layer width. Typically, the minimum device-layer width in standard CMOS is twice the layer depth. This is therefore an assumption taken in this section for the calculation of the maximum device-layer bandwidth as the function of technology. The bandwidth is calculated using equation (6.5) and the result is illustrated in Figure 6.4. For larger layer-depths the device-layer bandwidth decreases, as shown in equations 6.4 and 6.5.

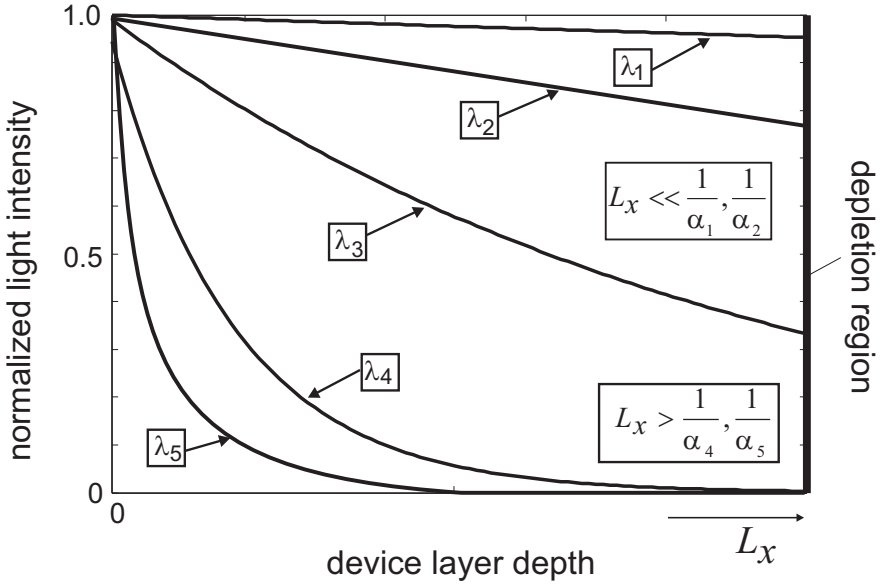


Figure 6.3: Normalized light intensity inside the device-layer region for different wavelengths: $\lambda_5 < \lambda_4 < \lambda_3 < \lambda_2 < \lambda_1$.

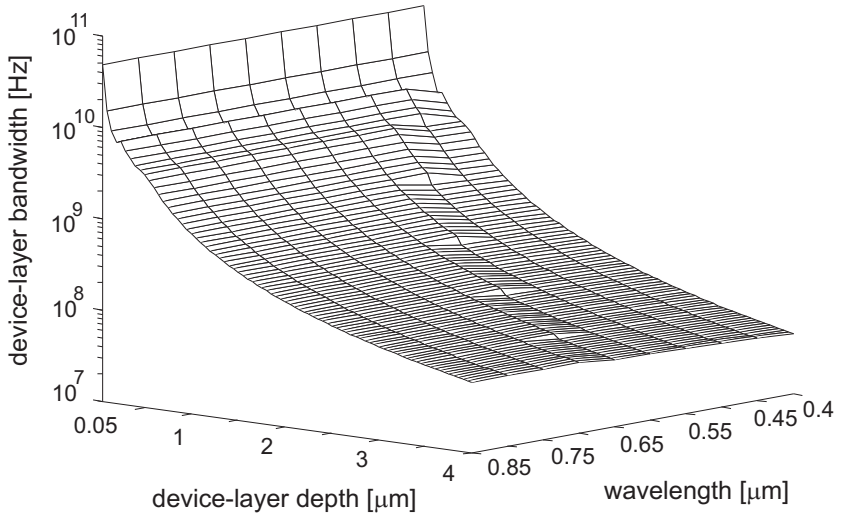


Figure 6.4: Bandwidth of device-layer diffusion current versus layer depth (technology dependent) and input wavelength. It is assumed that the device-layer width is twice its depth $L_y = 2L_x$.

Up to this point, it can be concluded that for wavelengths with the 1/e absorbtion depth larger than device-layer depth, the bandwidth and the photocurrent of the device-layer are opposite in nature. The lower the layer depth, the higher the bandwidth, but the lower the photocurrent in the device-layer. This results in the smaller contribution to the overall photocurrent. For wavelengths for which the light is completely absorbed in the device-layer, the bandwidth and the photocurrent are proportional in nature. The lower the wavelength, the lower the photocurrent (due to the lower responsivity, see chapter 2), and the lower the device-layer bandwidth (since the majority of the carriers are generated further from the bottom junction).

The influence of the other diode regions, the substrate and the depletion regions, on the total photocurrent is analyzed in the following sections of this chapter. The substrate photocurrent and drift photocurrent will be again calculated for CMOS technology in general, and for the whole wavelength sensitivity range.

6.1.3 Substrate current-photocurrent amplitude

The largest part of the photodiode in standard CMOS technology is its substrate. Chapter 3 described the substrate current response for $\lambda = 850$ nm. Light penetration depth for this wavelength is about 30 µm. Therefore, for all recent and future CMOS technologies (feature-size <0.5 µm) in general, substrate current has the largest contribution on the overall photocurrent. Chapter 3 analyzed two types of p-substrate: high-resistance and low-resistance substrates. It was shown that the calculated photodiode bandwidth is larger for low-resistance substrate: typically slow-diffusion carriers generated far away (20 µm) from the junctions need µs time to reach these junctions. However, inside low-resistance substrate they are recombined and do not contribute to the total photocurrent. On the other side the photocurrent is lower because of the recombined carriers.

Chapter 5 described a photodiode behavior on $\lambda = 400$ nm. Most of the carriers are generated close to a photodiode surface i.e. close to diode junctions. These carriers diffuse faster towards junctions and the bandwidth is in hundreds of MHz range. The influence of the substrate current on the overall photocurrent is negligible.

This section presents substrate current response of the photodiode in CMOS technology for the whole wavelength sensitivity range. Firstly, amplitude of the

substrate photocurrent $I_{\mathrm{diff_{subs}}}(L_x, \lambda)$ from equation (6.1) is calculated for *high-resistance* substrate, using the simplified equation (3.16), from section 3.2.1:

$$I_{\mathrm{diff_{subs}}}(L_x, \lambda) = e\alpha L_{\mathrm{n}} e^{-\alpha L_x} \frac{1}{\alpha L_{\mathrm{n}}} \tag{6.6}$$

The result is shown in Figure 6.5. Secondly, the photocurrent is calculated for *low-resistance* substrate using the procedure explained in chapter 3, section 3.2.1. For the easier comparison with the first substrate type, the result is also illustrated in Figure 6.5.

For all wavelengths having the absorbtion depth smaller or equal to the depth of the epi-layer L_{epi} (see Figure 6.1) in low resistance-substrate, the substrate photocurrent is comparable for both substrate types due to the similar "active" diode layers. The depth of the epi-layer is assumed to be three times larger than device-layer depth, which is the case for the available 0.18 µm CMOS technology. For wavelengths with absorbtion depths larger than the depth of epi-layer, the photocurrent for low-resistance substrate decreases: the larger the wavelength, the larger the number of the recombined carriers.

6.1.4 Substrate current-photocurrent bandwidth

Chapter 3 showed that the substrate current bandwidth does not depend on the device-layer depth i.e. it is independent on CMOS technology. In order to explain this, we can use Figure 6.6. Changing the depth of the device-layer for constant input wavelength has the same effect as changing the intensity of the input signal i.e. changing the photon flux from Φ_1 to Φ_2:

$$\Phi_1 e^{-\alpha x_2} = \Phi_1 e^{-\alpha \Delta x} e^{-\alpha x_1} = \Phi_2 e^{-\alpha x_1} \tag{6.7}$$

In [1] and [3] it was shown that photodiode bandwidth does not dependent on the intensity of the input light signal. Therefore, the *substrate current bandwidth is independent on CMOS technology.* It only depends on the input wavelength. Using the equation (3.15) in chapter 3, and substituting the corresponding diffusion lengths for electrons in high-resistive and low-resistive substrates L_{n1} and L_{n2}, the bandwidth of the substrate current is calculated and presented in Figure 6.7.

For the wavelengths having the absorbtion depth smaller or equal to the depth of the epi-layer L_{epi}, the diffusion lengths for both substrate types are comparable. Thereby, the bandwidth for both substrate types, $s_{\mathrm{diff}high}(L_x, \lambda)$

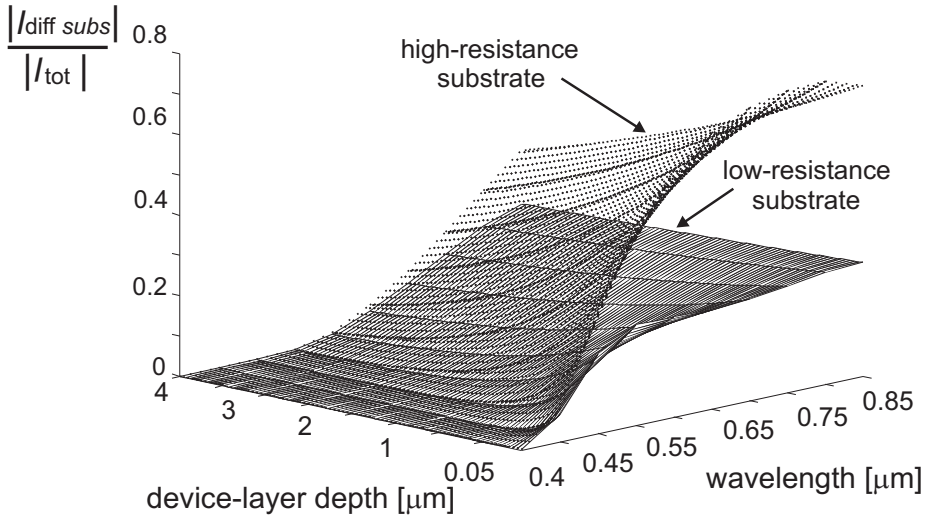


Figure 6.5: Normalized photocurrent of the substrate versus device-layer depth (technology dependent) and input wavelength, for photodiode in standard CMOS. The depth of the L_{epi} is assumed to be three times larger than the device-layer depth. The substrate current is normalized with the total diode current.

and $s_{\text{diff}low}(L_x, \lambda)$ are also comparable. For all other wavelengths, the bandwidth of the low-resistive substrate current is higher.

6.1.5 Depletion region current

Independent on CMOS technology, the depth of the lateral depletion region towards substrate and the width of the vertical depletion regions is rather constant since it is mainly determined by the concentration of the lower-doped region, in this case the substrate. However, the absorbed amount of light changes in both depletion regions with the technology and wavelength. The depth of the vertical depletion region decreases with new technologies (lower device-layer depth), and the lateral depletion region is located closer to the diode surface.

The amplitude of depletion region photocurrent $I_{\text{drift}}(L_x, \lambda)$ (from equation (6.1)) as a function of technology and wavelength, can be calculated as:

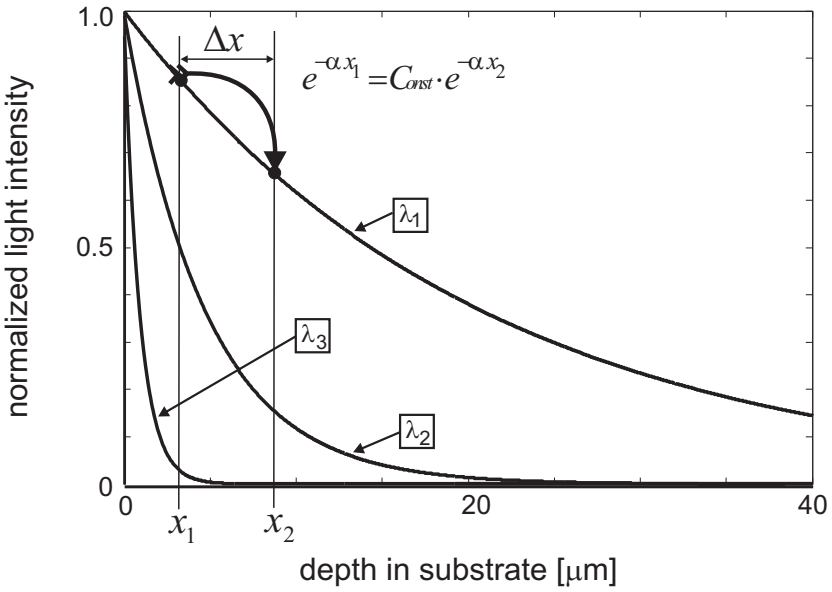


Figure 6.6: Normalized light intensity inside the substrate for different wavelengths: $\lambda_3 < \lambda_2 < \lambda_1$.

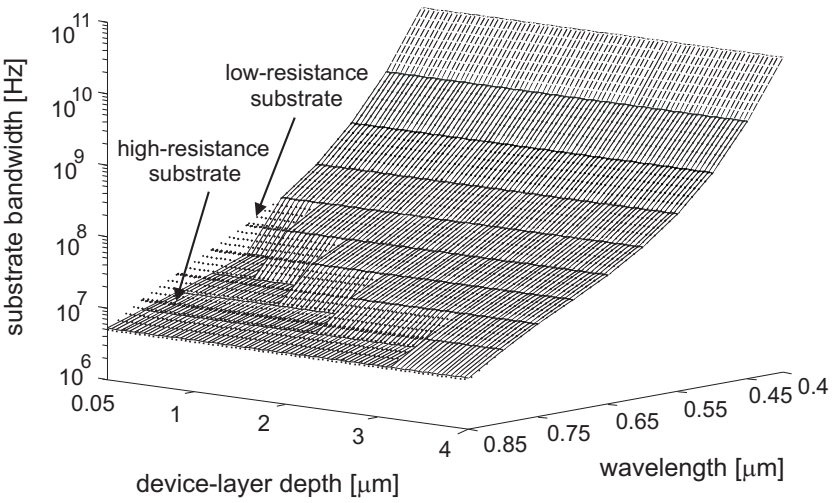


Figure 6.7: A bandwidth of substrate diffusion current versus device-layer depth (technology dependent) and input wavelength.

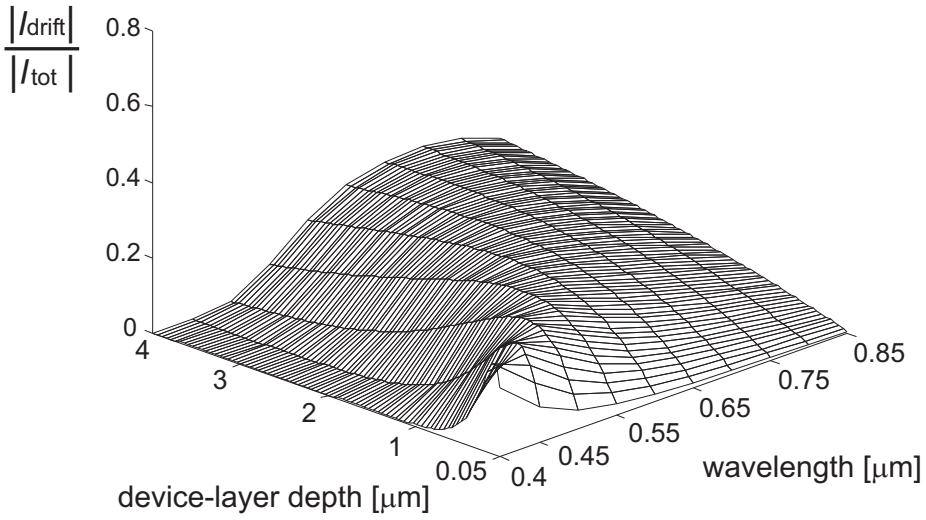


Figure 6.8: Normalized photocurrent of the depletion region versus device-layer depth (technology dependent) and input wavelength. A depletion region current is normalized with the total photocurrent.

$$\begin{aligned} I_{\text{drift}}(L_x,\lambda) &= \Phi e(e^{-\alpha L_x} - e^{-\alpha(L_x+d)})\frac{A_{\text{total}}}{A_{\text{eff}_{\text{lat}}}} \\ &+ \Phi e(1-e^{-\alpha L_x})\frac{A_{\text{total}}}{A_{\text{eff}_{\text{ver}}}} \end{aligned} \tag{6.8}$$

The result is illustrated in Figure 6.8.

Chapter 2, described that light is absorbed exponentially inside a photodiode. Thereby, the lower the wavelength as well as the lower the device-layer depth, the larger the value of the drift photocurrent inside depletion region. This is illustrated in Figure 6.8.

6.1.6 Depletion region - photocurrent bandwidth

The bandwidth of the depletion region current (drift current) is typically an order of magnitude higher than the bandwidth of diffusion currents in the photodiode [2]. For this reason, chapters 3, 4 and 5, assumed that the drift current is independent on frequency. In this chapter, for better accuracy the depletion

region current is approximated as an one-pole function in the frequency domain (see equation 6.1) with the bandwidth:

$$\frac{1}{2\pi} s_{\text{drift}}(L_x, \lambda) \approx \frac{0.4}{\tau_{\text{t}}} \tag{6.9}$$

since the depletion region behavior is similar to that in capacitor. In the equation (6.9), τ_{t} represents the average carrier transit-time.

For a long wavelength light, λ_1, λ_2 (see Figure 6.3), there is a uniform carrier generation inside the depletion region. The average transit time is $t_{\text{t}} = W/2v$, where v is the velocity of the slower carriers. For a large electric fields, velocity of both types of carriers are saturated, [1] around the same value: $v_{\text{h}} = v_{\text{e}} = v_{\text{s}}$.

For short wavelength light λ_4, λ_5 shown in Figure 6.3, the electron-hole pairs are generated near the upper, device-layer side of depletion region. In this case, it is the holes which must travel across the depletion region. A transit time is related to the velocity of holes v_h, and it is twice the average transient time due to the twice as large distance: $t_{\text{t}} = W/v_{\text{h}}$. Thereby, for the case when carriers are generated closer to the upper-side of the depletion region, the bandwidth is half the bandwidth for the uniform carrier generation. In standard CMOS technology, the drift bandwidth is typically 5-10 GHz.

6.1.7 Total photocurrent

Previous sections analyzed the photocurrent components of photodiode in CMOS technology for the whole wavelength sensitivity range. The sum of all components gives the total photocurrent, as shown in equation 6.1. Maximum photocurrent for any wavelength is almost independent on CMOS structure i.e. independent on CMOS technology [4]. This is explained with the fact that the quantum efficiency is determined mainly by absorption coefficient α and the diffusion length of the minority carriers [1].

For device-layer depths larger than absorption depth of light, $\alpha L_x > 1$, the total photocurrent is mainly determined by device-layer diffusion plus the drift current from the vertical depletion region. For twin-well technology, the influence of this vertical depletion region is almost negligible. For the input wavelengths 650 nm$< \lambda <$ 850 nm, independent on the CMOS technology, the total photocurrent is determined by the diffusion currents and the drift current.

Photodiode bandwidth is however dependent on both the technology and the wavelength. This bandwidth is calculated numerically as a -3 dB fre-

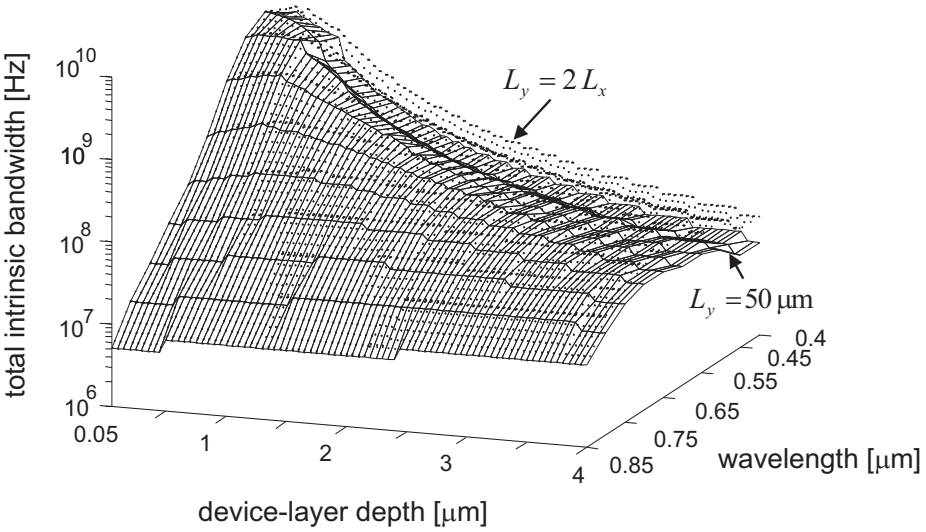


Figure 6.9: The total intrinsic bandwidth of general device-layer/p-substrate photodiode in standard CMOS versus device-layer depth and input wavelength. This bandwidth determines the starting equalization frequency.

quency of the total photocurrent given in equation (6.1). The bandwidth is calculated for both the minimum and the maximum device-layer widths, $L_y = 2L_x$ and $L_y \gg L_x = 50$ µm (corresponding to the core of the multimode fiber), respectively and the result is presented in Figure 6.9.

The lower the device-layer depth, the larger the excess carrier concentration gradient towards the bottom junction. As a result, the diode bandwidth is larger [2]. The drift current of the vertical depletion regions between the diode fingers (see Figure 6.1), is important only for a minimal device-layer width; for the maximal layer-width, the lateral depletion area in comparison with the device-layer area is almost negligible. The maximal calculated bandwidth is lower than 200 MHz and mainly determined by diffusion processes. The minimum bandwidth is in the low MHz range, which is typically the case for the deep absorption depths of light (> 15 µm). Then, the slow diffusive substrate carriers dominate the overall response.

6.2 Bandwidth improvement-analog equalization

6.2.1 Equalization range

Chapter 4 showed that using an analog equalizer, the bandwidth of integrated photodiodes in standard CMOS can typically be improved up to the low GHz range. A slow "roll-off" in the intrinsic diode response is compensated by a slow "roll-up" obtained as a sum of high-pass filter responses. Maximum equalization frequency is determined by the electrical diode bandwidth. The *equalized* diode characteristics was robust on typical 20% RC spread in the equalizer, due to the very slow roll-off (< 5 dB/decade).

This section presents the starting and the maximum equalization frequency of CMOS photodiodes including the $\pm 20\%$ RC spread in the equalizer, for the whole wavelength sensitivity range. In general, frequency characteristics and equalization range of CMOS photodiode strongly depend on the technology and input wavelength.

A *starting equalization frequency* is located below the intrinsic photodiode bandwidth. Namely, the starting compensation frequency corresponds to the roll-off start in photodiode characteristics. The intrinsic diode bandwidth for the whole wavelength sensitivity range is illustrated in Figure 6.9. The lower the intrinsic bandwidth, the larger the equalization range for the same maximum equalization frequency.

6.2.2 RC spread and maximum equalization range

The roll-off in photodiode frequency characteristics is technology and wavelength dependent which can be concluded by solving equation (6.1) for any technology-wavelength case. Chapter 4 showed that due to a spread in the equalizer, there is an amplitude shift in the equalized characteristics which changes the output signal as well as the total bit-error rate (BER) in the system. Using equations (4.1) and (4.2), the BER is calculated for the total amplitude shift of 0 dB, 0.4 dB, 0.8 dB and 1.2 dB, which corresponds to the roll-offs of 0 dB/decade, 5 dB/decade, 10 dB/decade and 15 dB/decade. The result is shown in Figure 6.10. For very low roll-off (~ 5 dB/decade), BER is increased only $\Delta\text{BER}\sim 10^{0.4}$. Adopting this BER value to its original one (without RC spread), the maximum data-rate decreases 5 %. However, for a 0.8 dB amplitude shift, the BER is increased $10^1 \div 10^2$. Changing this BER value to its original one (without amplitude shift) the data-rate decreases 15 %. For a

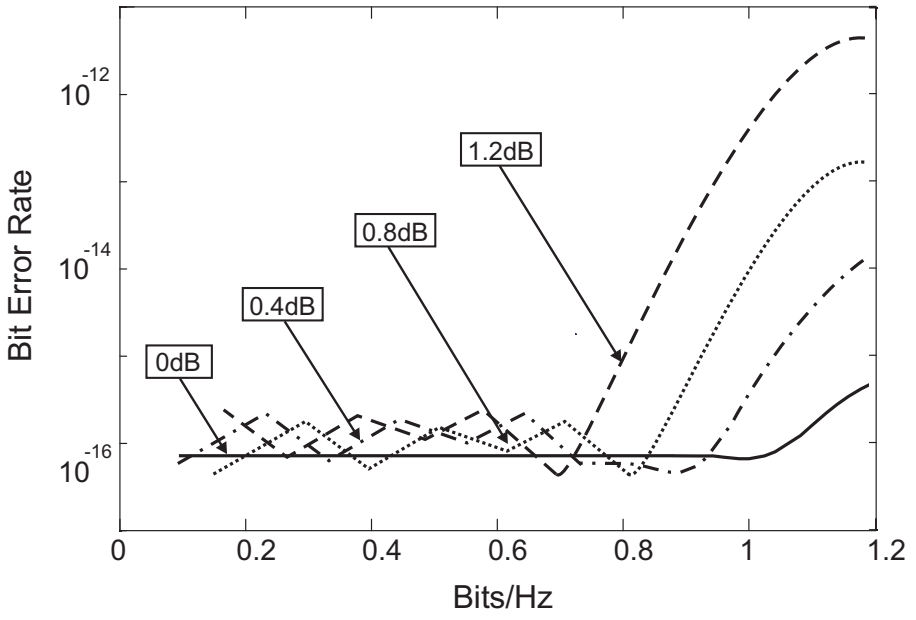


Figure 6.10: A change in BER versus amplitude shift in the equalized photodiode characteristics. The data-rate is normalized with the upper equalization frequency. Depending on the roll-off in the diode transfer characteristics i.e. depending on a CMOS technology and input wavelength, there are different amplitude shifts for a typical ±20% spread in RC components of the equalizer.

1.2 dB amplitude shift, the data-rate decreases about 30 %. We can conclude that for larger roll-off values, there are larger amplitude shifts in the equalized characteristics and the BER increases.

The lower the tolerant amplitude shift, the lower the lower the maximum equalization frequency $f_{\max}$, as shown in Figure 6.11.

Maximum equalization frequency is calculated first for a minimum device-layer width, which is taken to be twice its depth. As a first step, the total photocurrent characteristics are calculated using equation (6.1) for the whole wavelength sensitivity range; the device-layer depth of device-layer/p-substrate diode is changed from 0.05 µm ÷ 4 µm. From this characteristics, the corresponding roll-offs are deduced. With respect to the these roll-offs, and including the ±20 % RC spread, the maximum amplitude shift is calculated using equation (4.10). This amplitude shift determines the maximal equalization frequency

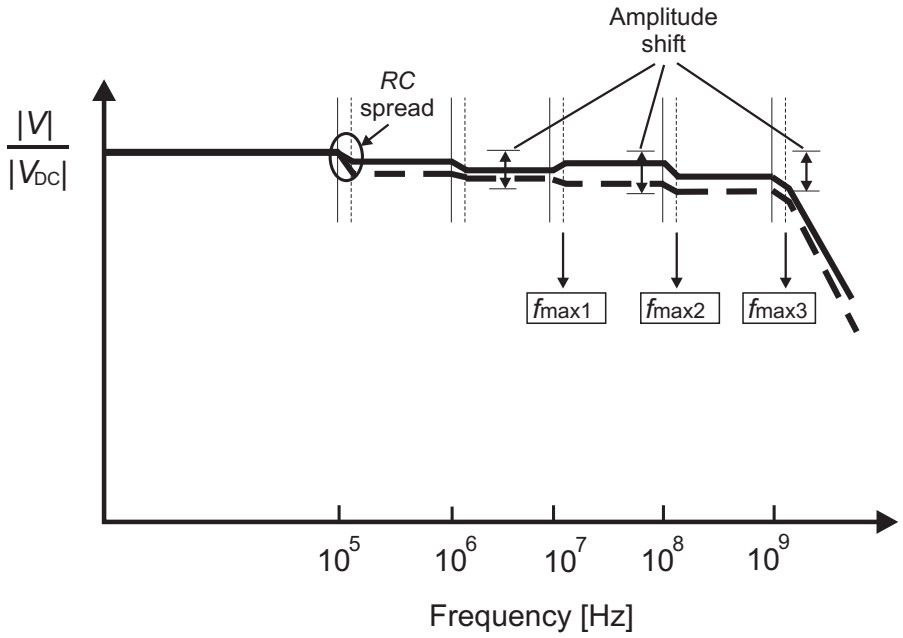


Figure 6.11: An amplitude shift in equalized characteristics due to RC spread in equalizer. A maximum tolerant amplitude shift determines the maximum equalization frequency ($f_{\max 1}$, $f_{\max 2}$, $f_{\max 3}$).

as shown in Figure 6.11. The maximum equalization frequencies for two amplitude shifts in the equalized characteristics: 0.4 dB and 0.8 dB, are presented in Figure 6.12 a) and b). The maximum equalization frequencies below 1 GHz are not presented, since we are interested mainly in solutions for Gb/s optical detection.

Further, the maximum equalization frequency is calculated for the maximum device-layer width i.e. for the single-photodiode. The result is shown in Figure 6.13 a) and b). The drift bandwidth f_{drift} determines the maximum possible equalization frequency.

For wavelengths with the absorbtion depth smaller than device-layer depth $\alpha L_x > 1$, the maximum equalization frequency has the lowest value among all analyzed wavelengths and technologies. The photodiode characteristic is determined mainly by diffusion inside the device-layer. Chapter 5 showed that roll-off in the diffusion characteristic approximately corresponds to the one-pole func-

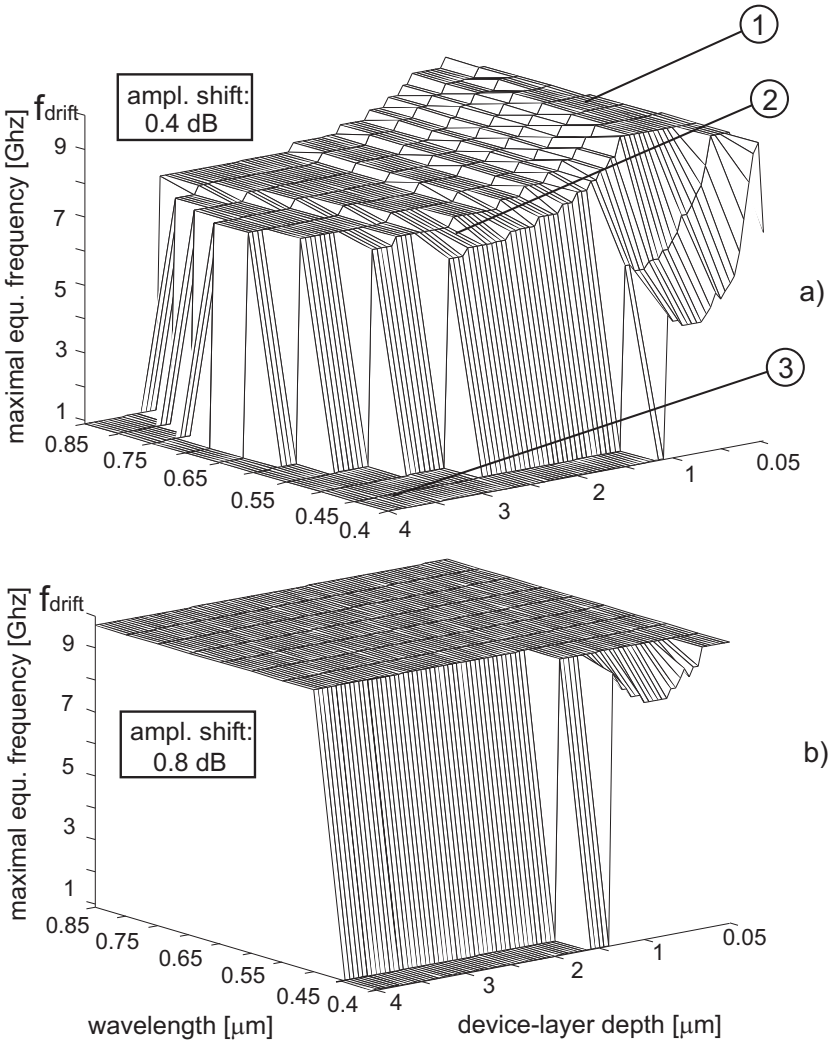


Figure 6.12: A maximum equalization frequency for a device-layer/p-substrate (high-resistance substrate) photodiode in standard CMOS, having the width of the device-layer twice its depth ($L_y = 2L_x$). Two permissible amplitude shifts are assumed in the equalized photodiode characteristics: a) 0.4 dB and b) 0.8 dB; amplitude shifts are due to a spread in RC components of the equalizer (typically $\pm 20\%$). 1. maximum possible equalization frequency, 2. maximum equalization frequency is lower because the amplitude shift is larger than 0.4 dB, 3. maximum equalization frequency is below 1 GHz (values not shown); for low wavelengths the maximum equalization frequency is situated close to the device-layer diffusion bandwidth.

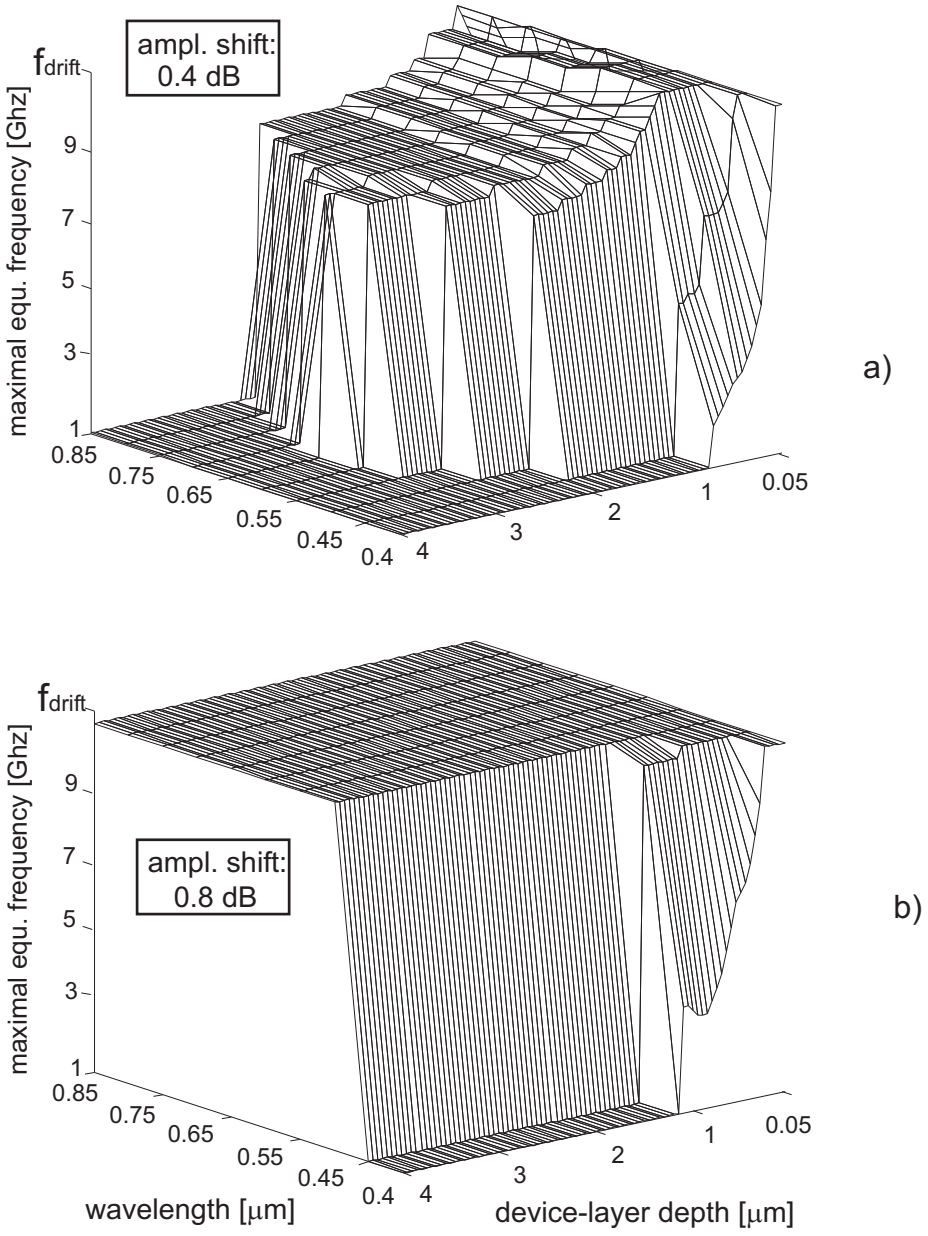


Figure 6.13: A maximum equalization frequency for a photodiode in standard CMOS, having a maximal width (L_y = 50 µm) of the device-layer, for two permissible amplitude shifts in the equalized characteristics: a) 0.4 dB and b) 0.8 dB.

tion (20 dB/decade) for the minimum device-layer width, and to the triple-pole function for maximum device-layer width. The equalization of these curves is very sensitive on RC spread because of the large amplitude shift. Therefore, the maximum equalization frequency is situated close to the -3 dB bandwidth frequency of photodiode. The bandwidth improvement is *minimum*.

For wavelengths with the 1/e absorbtion depth larger than device-layer depth $\alpha L_x < 1$, all diffusion and drift current components determines the roll-of in the photodiode characteristics, as shown previously in this chapter. Chapter 3 described that for these wavelengths, the roll-off in the diode characteristics is low (< 5 dB) up to the drift-bandwidth frequency f_{drift} shown in Figures 6.12 and 6.13 (typically up to 8 GHz). The equalization frequency is then *maximum*. These maximum frequencies are illustrated in Figures 6.12, and 6.13, for the minimum and the maximum device-layer widths, respectively.

For a minimal device-layer width, the maximum equalization frequency is equal or higher than for the maximum device-layer width (single photodiode). This is because of the lower roll-off in its frequency characteristics which allows higher equalization frequency (for typical ±20% spread).

According to the analyzes in this section, the maximum equalization frequency is determined by intrinsic processes inside photodiode: for $\alpha L_x > 1$, it is mainly determined by device-layer diffusion, and for $\alpha L_x < 1$ by the drift bandwidth.

Electrical diode bandwidth

The maximum equalization frequency analyzed in the previous section is determined by intrinsic processes inside photodiode. In order to preserve this maximum equalization frequency, the electrical diode bandwidth should be larger.

Chapter 3 showed that a photodiode in twin-well technology with adjoined-wells has a much higher capacitance than in separate-well technology. For constant ratio between the width and the depth of the device-layer $L_y/L_x = K$, the total diode capacitance is independent on CMOS technology:

$$C_{\text{diode}} = \epsilon_0 \epsilon_r \frac{KL_x}{d} \frac{ZY}{KL_x} \tag{6.10}$$

$$= \epsilon_0 \epsilon_r \frac{YZ}{d} \tag{6.11}$$

where Y and Z are the photodiode dimensions in y and z-directions, and with the assumption that the substrate doping concentration is almost independent

on technology (d=const). However, for larger values of K, the photodiode capacitance decreases. As a result, the electrical diode bandwidth increases. A minimum capacitance is achieved using single device-layer diode i.e. taking the maximum value of K.

6.3 Summary and Conclusions

The intrinsic frequency characteristics of CMOS photodiode is dependent on both the *input wavelength* and the *technology*. Consequently, the diode intrinsic bandwidth strongly depends on these two parameters. The maximum diode bandwidth without equalization is achieved using:

• *minimal wavelength*, $\lambda = 400$ nm. The bandwidth is 8 GHz i.e. approximately equal to the diffusion photodiode bandwidth.

• *technology* for which the device-layer depth corresponds to the absorption depth of light, $1/\alpha_{400}$.

For other input wavelengths and technologies, it is possible to achieve bandwidth improvement using the analog equalization method, explained in chapter 4. This bandwidth improvement is however wavelength, technology and device-layer width dependent.

For the device-layer depths larger than 1/e absorbtion depth of light $\alpha L_x > 1$, the intrinsic diode response is determined by device-layer diffusion and the drift current of the vertical depletion region. This drift current is important only for a minimal device-layer width. For a maximal layer-width, the vertical depletion area in comparison with the device-layer area is almost negligible. The roll-off for a minimal layer-width is about 15 dB/decade, while for a maximal device layer width (50 µm) it corresponds to a triple-pole characteristics (∼60 dB/decade). The larger the roll-off, the larger the amplitude shift in the equalized characteristics due to the ±20% RC spread in the equalizer. This amplitude shifts lowers the bit-error-rate in the system. The change in BER for 15 dB/decade roll-off is $\Delta BER = 10^3$, which reflects in the data-rate decrease of 25%. The maximum equalization frequency is situated close to the diffusion bandwidth frequency and the bandwidth improvement is *minimal*.

For the device-layer depths smaller than absorbtion depth of light αL_x <1, the roll-off in the intrinsic diode characteristics is < 10 dB/decade. The calculated amplitude shift for the 5 dB/decade roll-off is 0.4 dB. This shift increases BER for $\Delta BER = 10^{0.4}$, while for 10 dB/decade, the roll-off is 0.8 dB and

$\Delta BER = 10^{1.5}$, assuming 20 % RC spread. The effective data-rates are decreased for 5% and 15% respectively. We can conclude that for low roll-off of <5 dB/decade, there is very small change in the BER and the data-rate, so the system is very robust on the spread in the equalizer components. The maximum equalization frequency is determined by the drift bandwidth i.e. up to a 8-10 GHz range. For this reason, the bandwidth improvement of CMOS photodiode is *maximal*.

For the *minimal device-layer width*, the maximum equalization frequency is higher than for the maximum device-layer width (single photodiode) due to the lower roll-off in the frequency characteristics. This low roll-off allows higher equalization frequency for typical $\pm 20\%$ RC spread in the equalizer components. The maximum equalization frequency is determined by intrinsic processes inside photodiode. For low wavelengths it is determined by device-layer diffusion, and for large wavelengths by the drift bandwidth.

According to analyzes in this section, the maximum bandwidth improvement can also be achieved using single photodiode. This significantly simplifies layout design and maximize light-sensitive area due to the minimal metal contact area.

References

[1] S. M. Sze: "*Physics of semiconductor devices*", New-York: Wiley-Interscience, 2-nd edition, 1981, p. 81.

[2] D. Coppée, H. J. Stiens, R. A. Vounckx, M. Kuijk: "Calculation of the current response of the spatially modulated light CMOS detectors", *IEEE Transactions Electron Devices*, vol. 48, No. 9, 2001, pp. 1892-1902.

[3] S. Alexander: "*Optical communication receiver design*", SPIE Optical engineering press, 1997.

[4] I. Brouk, and Y. Nemirovsky : "Dimensional effect in CMOS photodiodes", *Solid State Electronics*, vol. 46, 2002, pp. 19-28.

CHAPTER 7

Conclusions

7.1 Summary and Conclusions

7.1.1 Chapter 1: Motivation

The motivation of the project *fiber-to-the-chip* is presented in chapter 1. In the future high-speed communication system (> 20 Gb/s), the optical interconnect may become important since straightforward electrical connections suffers from poor impedance matching (results in distorted signals), crosstalk and significant Electro-Magnetic noise which degrades the system performances.

The main goal of the project was to design high-speed (Gb/s range) and low-cost optical receiver for short-haul communication. Fast receivers were already available, but their cost was two orders of magnitude too high since they are designed in expensive technologies like InP or InGaAs. For low cost and simple integration with electronic circuitry, an inexpensive CMOS technology was viable candidate for the design; it is nowadays a straightforward technology for all digital systems and processors. However, a disadvantage of the CMOS photo-receiver is limited speed of the photodetector. It is typically more than two orders of magnitude too low (10 Mb/s).

7.1.2 Chapter 2: Introduction

Apart from the EM interference, a number of optical interconnection advantages in comparison with the straightforward wired connection are summarized in chapter 2. Further in the chapter, an overview of an optical communication system consisting of transmitter, optical fiber and receiver is presented. On the transmission side, typically used light sources are LEDs and lasers. In the last couple of years, widely used laser are VCSELs, which can emit light from a relatively small area (5-25 microns) on the surface of the chip, directly above the active region. In this way, a compact size and high reliability are achieved.

Single-mode fibers (small core size) and multimode fibers (large core size) are used as transmission media in the system. The latter has higher dispersion that limits the data-rate and it is mainly used for short-haul communication. Nevertheless, for short-haul communication, a multimode fiber has some advantages: light is launched into a multimode fiber with more ease. The higher numerical aperture and the larger core size make fiber connections straightforward. Moreover, during fiber splicing, core-to-core alignment becomes less critical. Another advantage is that multimode fibers permit the use of light-emitting diodes (LEDs).

On the receiver side, the requirements of the ideal photodetector are also presented. In addition, main parameters in the photodetection process are discussed: responsivity, quantum efficiency, absorption depth and absorption coefficient.

7.1.3 Chapter 3: High-speed photodiodes for λ=850 nm

A typical figure-of-merit for photodetector is its bandwidth-responsivity product. For λ=850 nm, the bandwidth of the photodiodes in standard CMOS is in the low MHz range, which is the main limiting factor for the Gb/s optical detection. A common feature of the frequency characteristics of the analyzed photodiodes is a low roll-of (<5 dB/decade) in the frequency range up to a few GHz. This stems from the fact that total photocurrent is the sum of the diffusion currents having low roll-offs (10 dB/decade) and the drift current which has by approximation "flat" frequency response up to the 10 GHz range. The low roll-off feature is used for the bandwidth improvement of CMOS photodiodes by using analog equalization explained in chapter 4.

7.1.4 Chapter 4: Bandwidth improvement

Using an analog equalizer that compensates (in both gain and phase) the low roll-off of the photodiode frequency decay, the bandwidth of the integrated photodiode in standard CMOS was increased more than two orders of magnitude. In this way it was possible to achieve the Gb/s optical detection in standard CMOS technology. The important feature of the proposed solution is that complex adaptive algorithms are not required. A 3 Gb/s bit-rate is achieved with BER$< 10^{-11}$ and average input optical power of P_{in}=25 μW (-19 dBm). This is over half an order of magnitude speed-increase in comparison with a state-of-the-art photodetectors. The proposed pre-amplifier with the analog equalizer is spread and temperature robust: due to the low roll-of, shift of $\pm$20% in the RC-network of the equalizer does not influence the output amplitude more than $\pm$0.4 dB. This changes the BER for $\Delta BER = 10^{0.4}$ i.e. the bit-rate is changed 5% only. Spread robustness is confirmed using both the simulations and the measurements. An on-chip RC network is designed as a number of RC fingers connected with the highest metal layer. Using Focused Ion Beam it was possible to choose RC values by removing the highest metal layer. The measurements showed that with $\pm$20% change in the RC values, the output amplitude is changed no more than 6%. The measured BER$=< 10^{-10}$ with average input optical power of 25 μW.

7.1.5 Chapter 5: High-speed photodiodes for λ=400 nm

This chapter presented time and frequency behaviors of photodiodes in standard 0.18 μm CMOS for λ=400 nm. Due to the low light-penetration depth (<0.2 μm), most of the excess carriers are generated close to the diode surface. The total photocurrent is mainly determined by the diffusion current of the top photodiode layer(s). The diffusion bandwidth is high (up to 5 GHz) since the distance between the generated carriers and the junctions is small ($< 1\mu$m). This allows high minority concentration-gradient and thus fast diffusion process. The total diode bandwidth is about 5 GHz too. This diodes can be used for the low Gb/s optical detection without using analog equalizer. The photodiode frequency characteristics for 400 nm wavelength have high roll-off: 15 dB/decade for the minimal nwell width (2 μm) and 60 dB/decade for the maximum nwell width (50 μm). This large roll-off enables significant bandwidth improvement Using analog equalizer, which is shown in chapter 6 of this thesis.

In this chapter we presented also a lateral polysilicon photodiode that has a

frequency bandwidth far in the GHz range: the measured bandwidth of the poly photodiode was 6 GHz, which figure was limited by the measurement equipment. However, the quantum efficiency of poly-diodes is very low (< 8 %) due to the very small diode active area: the depletion region is very small due to the lack of intrinsic layer and the diode depth is limited by the technology.

7.1.6 Chapter 6: High-speed photodiodes for 400 nm<λ<850 nm

This chapter analyzes frequency response of photodiode in CMOS technology in general, for the whole wavelength sensitivity range 400 nm<λ<850 nm. The maximum calculated diode bandwidth without equalization is 8 GHz (for the moderate diode responsivity 0.4 A/W), and it is achieved with the minimum wavelength $\lambda = 400$ nm, and minimum device-layer width.

For other input wavelengths and technologies, it was possible to achieve bandwidth improvement using the analog equalization method from chapter 4. This bandwidth improvement was also wavelength, technology and device-layer width dependent. For the wavelengths having absorbtion depth smaller than device-layer *depth*, the intrinsic diode response is mainly determined by device-layer diffusion. The roll-off for the minimal layer *width* corresponds approximately to one-pole frequency characteristic (20 dB/decade). For larger layer-widths, the roll-off increases up to triple-pole characteristic (60 dB/decade). After equalization and including the *RC* spread, there are amplitude shifts in the equalized characteristics larger than 1.5 dB. This causes the change in BER for more than $10^2 \div 10^3$, i.e. the decrease in effective bit-rate is more than 30%. Thus, the possible bandwidth improvement is *minimal*.

For wavelengths having the absorbtion depth larger than device-layer depth (αL_x<1), the roll-off in the intrinsic diode characteristic is relatively low (<10 dB/decade). For 5 dB/decade roll-off and ±20% *RC* spread, the change in bit-rate is only 5 %. The equalization method explained in chapter 4 is then very robust on *RC* spread in the equalizer and the bandwidth improvement is large; maximum bandwidth improvement i.e. maximum equalization frequency is determined by the drift bandwidth frequency (up to $8 \sim 10$ GHz).

In addition to the maximum bandwidth improvement, the electrical diode bandwidth should be maximal i.e. it should be typically larger than drift diode bandwidth. In this way, the electrical bandwidth will not limit the maximum equalization frequency which is then determined by intrinsic processes inside

photodiode: for low wavelengths it is determined by device-layer diffusion bandwidth, and for large wavelengths by the drift bandwidth. For the maximum electrical bandwidth, a photodiode should be designed to have minimal capacitance e. a single photodiode should be used.

Papers and conference contributions

∗ **Papers in 2002:**

[1]. Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"Monolithically Integrated Photo-Diodes in Standard CMOS Technology for high speed optical communication: General Consideration and Analysis"*, ProRISC conference, Veldhoven, The Netherlands, November 2002, Best poster award.

∗ **Papers in 2003:**

[1] Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"High-speed lateral polysilicon photodiode in standard CMOS"*, ESSDERC 2003, Estoril, Portugal, 16-18 September 2003, pp.521-524.

[2] Saša Radovanović, Anne-Johan Annema and Bram Nauta : *"Physical and electrical bandwidths of integrated photodiodes in standard CMOS technology"*, The First IEEE Conference on Electron Devices and Solid-State Circuits EDSSC2003, Honkong, December 16-18, 2003, pp.95-98.

[3] Saša Radovanović, Anne-Johan Annema and Bram Nauta:*" On optimal structure and geometry of high-speed integrated photodiodes in a standard CMOS technology"*, The 5th Pacific Rim conference on Lasers and Electro-optics,2003, Taipei, Taiwan, p.87.

[4] Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"Analysis of the high-speed polysilicon photodetector in Standard CMOS Technology"*, ProRISC 2003, Veldhoven, Netherlands.

[5] Bjorn de Boer, Saša Radovanović, Anne Johan Annema, Bram Nauta: *" Graphical Analysis of the Non-Linear PLL"*, ProRISC 2003, Veldhoven, Netherlands.

∗ **Papers in 2004:**

[1] Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"3 Gb/s monolithically integrated photodiode and pre-amplifier in standard 0.18um CMOS"*, ISSCC February 2004, San Francisco, USA, pp. 472-473.

[2] Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"3 Gb/s integrated photodiodes and pre-amplifier in a fully standard CMOS technology"*, to be submitted to the Journal on Solid-State Circuit.

[3] Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"Integrated Photodiodes in Standard CMOS Technology for CD and DVD Applications"*, MIEL conference, Niš, Serbia-Montenegro, May 2004, pp.246-249.

[4] Saša Radovanović, Anne-Johan Annema and Bram Nauta: *"Integrated photodiode and pre-amplifier for 3 Gb/s bit-rate in a standard CMOS technology"*, accepted for the conference ProRISC 2004, Veldhoven, Netherlands.

[5] Saša Radovanović, Anne-Johan Annema and Bram Nauta: "*Bandwidth of Integrated Photodiodes in Standard CMOS for CD/DVD Applications*", accepted for the Journal of Microelectronics Reliability, Elsevier, 2004.

Acknowledgements

Writing down this part of the thesis should inevitably remind me that the dedicated part of my life towards PhD degree should come to an end very soon. But I cannot notice that clearly, partly because I am still busy with the writing and arranging the administrative things for the defense, and partly because that this was really something that I wanted and that it was so much fun to me that I let the good feeling continue. I am certainly glad that my PhD project was finished successfully, and there are many people that have contributed to this success. I would like to mention them in this dedicated part of the book.

On the first place it is my promotor Bram Nauta, who surf my way through big waves in the ocean of IC design. Then, my supervisor Anne Johan Annema who was there as a lighthouse to watch that I have chosen the safe path towards the successful ending of the project. Thank you guys for making sure I learned how to act scientifically and gave me all the resources, help, and attention that I needed. Your efforts to explain things clearly and simply, made the IC design fun for me.

A precarious, trilling and finally joyed part of my work was certainly done in the laboratory. For his comprehensive help with the set-up and his patient assistance with measurements equipment I would like to thank Henk de Vries. For making sure that software runs properly on my computer, I need to thank Frederik Reenders and Cor Bakker. The necessary Cadence support was always generously offered by Gerard Wienk and I thank him too.

Inspiring conversation were possible with Erik Klumperink, who help me in understanding many things in IC design, and he shared many insights from his experience. I also thank him for the grand dinner we had on the whirling restaurant in San Francisco.

For plesureous everyday atmosphere at work, first I would like to thank to my officemate Arnoud for providing a stimulating and fun environment in which to learn and develop. His tremendous help with the Dutch language especially at the beginning (and ever after) was invaluable for me. I am very grateful to my other colleague Vincent who helped me a lot with cracking the LaTex. I thank to Mustafa for many nice discussions we had and for a pleasant time in Eindhoven. I would also like to acknowledge David and Svetislav for their advices on my difficulties.

I owe a great deal to my master students Bjorn, Tanja and Chen-Chen.

My overwhelming thanks goes to the people from Semiconductor group. Above all I would like to thank prof. Jurriaan Schmitz for his valuable advices about photodiode design as well as for helping me in writing the paper for ESSDERC. I learned a lot of Silvaco (I used it for simulation of semiconductor optical detectors) from working with Sander and Cora. I would also like to acknowledge Alexey who help me with always tricky questions. I thank to Jay for his great companionship in Estoril. For a great cooperation on the TeraHertz project I would like to show my gratitude to Jaap Flokstra, Arne, Arun, Gideon and all other people.

One big thank you to the scientists from Vrije Universiteit Brussels working in the group of prof. Martin Kuijk. Their precious work in the area of photodetector design in standard CMOS technology gave me a powerful inspiration for the step forward that I made in this very challenging field.

I would also like to acknowledge people from Philips Natlab for helping me with chip(s) finishing. I am specially grateful to Henk ter Meer, for his time and effort to help me.

I am grateful to the members of my promotion committee J. Schmitz, W. van Etten, M. Kuijk, A. Driessen and I. Radovanović, for taking time to read this dissertation and to evaluate my research.

A number of people in my home town did not loose sight of me. They repeatedly encouraged me by e-mail or asking me directly. I thank specially to my friend Aca who assisted many times when it was up to sorting out things of mind.

For the great after-work time and sporting atmosphere, especially in my first

two years of study, I am grateful to many friends: Petko, Zoki, Boris, Dule, Raša, Žika, Marko, Vojkan, Ljuba and the members of Arruba recreational team.

My overwhelming thanks goes to dear friends Stanča i Vladana, and their charming daughter Anja. We spent a lot of the time together, sharing the experience, knowledge and fun. I am very grateful to Vlada for providing stimulation, inspiration and motivation during the study.

There are also a number of people who do not work with me but have been friends for all this years and I would also like to thank them for their support: Goran, Bilja, Boris, Jelena, Saša, Nataša, Rossano, Helene, Ruud, Dejana, Wilko, Katerina, and many others that I didn't mentioned and who should also be here. I am also grateful to my friends Duško for his encourage and support and Džoni with whom I had a great time especially in my first days in The Netherlands.

My dearest parents. I would certainly not writing this thesis without their tremendous support, encouragement and care. With their help, I have made it through all the steps necessary to reach this point in life. They did their best so I can achieve my best in life; I am grateful to them for molding my personality.

My twin brother Igor, my best friend, my everyday support, my "added-power" in whole life and work. I am proud having him as a brother! I am exceedingly grateful to Igor for being there all along. I am also enormously grateful to his wife Jelena, who always support me when I needed. Thank you Jelenče, you are great.

My two adorable girls. Vesna and Sara. Choosing the wife to share all good and bad in life, I am every day more convinced that I got a precious and dearest friend for a life also. Hard days, nights, necessitate perseverance and "wind beneath my wings" was and is my wife Vesna. The paramount help and support she gave me can be reflected in only a few words of her: "Thank God that you are finishing". My baby-daughter Sara for sparked the fire of happiness in my life. She actually change the course of my PhD studies to the best point, she shined the brightest light in my life, she ... Heavens, she is constantly interrupting me in writing this acknowledgement. It seems that she doesn't have much interest in my writing. Her pants are full again. Yes, yes, I'll be in a second with you sweety...

Summary

This thesis describes high-speed photodiodes in standard CMOS technology which allow monolithic integration of optical receivers for short-haul communication. The electronics for (multiple users) long-haul communication is very expensive (InP, GaAs), but the usage is justified by the large number of users. For short distance communication, the number of users is small, and more cost-effective solutions were needed. An integrated optical receiver in standard CMOS technology is a viable candidate. As a result, the optical communication can be enabled directly to CMOS chips. Moreover, there is a possibility to implement matrices of optical detectors that operate on many parallel optical channels. Using parallel channels, the overall data-rate is significantly increased. An important advantage of monolythically integrated optical receiver is that ground-bounce issues and bond-wire inductance-capacitance problems are largely eliminated.

The first part of the thesis presents the advantages of optical communication in comparison with the copper-wired communication. The main disadvantage of the latter one is that at high data-rates major problems occur: poor impedance matching results in distorted signals, signal losses due to conduction and radiation losses, significant Electro-Magnetic interference noise which degrades the system performances. Using optical communication system these problems are circumvented. The main focus of the thesis is given at the receiver side; the objective was to design a low-cost Gb/s optical detector that can be easily

integrated with the rest of electronic circuitry.

As specified in the short-haul optical communication standard, the transmitted light has $\lambda = 850$ nm wavelength. However, on this wavelength the speed of the photodiodes in standard CMOS is in the low MHz range. Therefore, it is the speed limiting component in the overall Gb/s optical receiver design. This low speed is intrinsic to CMOS photodiodes, due to the occurrence of slow diffusive carriers. Designing p+/nwell photodiode (ignoring the p-substrate), the photodiode bandwidth is in the low GHz range, but the responsivity is very low: 18 dBm smaller than specified in communication standard, which makes them non-applicable for short-haul communication. On the other side, photodiodes that meet the specified responsivity have bandwidth that is two orders of magnitude too low.

A solution for increasing the bandwidth for more than two orders of magnitude is introduced in chapter 4. Using an analog equalizer, which equalizes in gain and phase the photodiode intrinsic characteristics, the bandwidth is increased more than two orders of magnitude. The presented circuit enables usage of the slow photodiodes for data-rates in the low GHz range, without reducing responsivity. The analytical analyses and a number of measurements are presented illustrating the correct operation and robustness of the proposed method. Although the presented design is for $\lambda = 850$ nm, the aproach is suitable for all wavelengths in the CMOS senistivity range from $\lambda = 400 - 850$ nm.

Samenvatting

Dit proefschrift beschrijft het ontwerp van hoge snelheids fotodiodes in een standaard CMOS technologie. Het doel is monolithische integratie van optische ontvangers voor hoge snelheids communicatie over korte afstanden.

Voor lange afstands optische communicatie waar InP en GaAs technologie wordt gebruikt worden de hoge kosten van de zender en ontvanger gerechtvaardigd door het grote aantal gebruikers.

Voor korte afstands optische communicatie is het aantal gebruikers veel kleiner, en is er behoefte aan een goedkopere oplossing. Een geintegreerde optische ontvanger, gerealiseerd in een standaard CMOS proces is hier een haalbaar alternatief.

Behalve de voordelen van prijstechnische aard biedt een optische ontvanger in CMOS nog andere voordelen: omdat een matrix aan optische ontvangers op een IC gerealiseerd kan worden, wordt grootschalig parallellisme, teneinde de communicatiesnelheid nog verder te verhogen, mogelijk. Een ander voordeel van een geïntegreerde optische ontvanger is dat de ontvanger in de directe nabijheid van de verdere bewerkingselektronica geplaatst kan worden, en er zo veel parasitaire problemen voorkomen worden.

In het eerste deel van dit proefschrift wordt korte-afstands optische communicatie vergeleken met klassieke elektrische communicatie. Het blijkt dat optische communicatie, vooral bij hogere datasnelheden, veel voordelen biedt. Minder verlies, eenvoudigere detectie door het grotendeels uitblijven van vervorming, en

minder problemen met overspraak veroorzaakt door EM straling behoren tot de belangrijkste. De nadruk van het eerste deel ligt op de optische ontvanger; het doel was het ontwerp van een prijsgunstige Gb/s optische detector die eenvoudig met de rest van de signaalverwerkende elektronica geïntegreerd kon worden.

De standaarden voor korte-afstands optische communicatie specificeren dat het gebruikte licht een golflengte heeft van $\lambda = 850$ nm. Dit is voor een CMOS fotodiode ongunstig; gangbare ontwerpen halen een bandbreedte van slechts enkele MHz vanwege de langzame diffusie van mobiele ladingsdragers in het CMOS. De beperkte snelheid van de fotodiode is hiermee een van de belangrijkste knelpunten bij de realisatie van een Gb/s optisch communicatiesysteem in CMOS.

De bandbreedte van een CMOS fotodiode kan vergroot worden tot het lage GHz gebied door gebruik te maken van een p+/nwell fotodiode (waarbij het p-substraat genegeerd wordt), maar dit gaat ten koste van de gevoeligheid. Er zijn ontwerpen bekend die weliswaar de benodigde snelheid halen, maar met 18 dB minder gevoeligheid dan vereist. Dit maakt ze als ontvanger voor korte-afstands communicatie onbruikbaar. Aan de andere kant bestaan er fotodiodes die wel de vereiste gevoeligheid bezitten, maar een bandbreedte die twee ordes van grootte te klein is.

In hoofdstuk vier presenteren we een oplossing die tegelijkertijd aan de bandbreedte en aan de gevoeligheidseis voldoet. Met behulp van een analoge equalizer die de amplitude en fasekarakteristiek van de fotodiode compenseert, blijken we de bandbreedte van de fotodiode met meer dan twee ordes van grootte te kunnen vergroten, zonder dat dit ten kostte gaat van de gevoeligheid. Hiermee is de gepresenteerde schakeling geschikt voor gebruik tot snelheden van een paar GHz.

Aan de hand van zowel een analytische beschouwing als meetresultaten wordt gedemonstreerd dat de gepresenteerde oplossing naar behoren functioneert en ook robuust is. Hoewel het huidige prototype werkt op $\lambda = 850$ nm, is de methode geschikt voor alle golflengten waar CMOS gevoelig is, van $\lambda = 400 - 850$ nm.

Biography

Saša Radovanović was born on August 11, 1974, in Peć, Serbia. He received the Graduate engineer degree in 1997 and M.Sc. degree in 2000 at the University of Niš, Serbia. In January 2000, he joined the IC-design group of the MESA+ Reasearch Institute, University of Twente, The Netherlands, where he received his PhD degree in December 2004. He is currently working with National Semiconductor BV, Delft, The Netherlands. His research interest include design of high-frequency transmitters and receivers for optical storage and optical communication systems.