Tesi di Laurea Magistrale
Corso di Laurea Magistrale in Ingegneria Elettronica/Electronic Engineering
Devices and Technologies for Integrated Electronics and Optoelectronics

Silicon-on-germanium waveguide photodetectors: a multiphysics computer-aided design approach

Matteo Giovanni Carmelo Alasio

* * * * * *

Supervisors
Dr. Alberto Tibaldi, Relatore
Prof. Michele Goano, Co-relatore
Prof. Francesco Bertazzi, Co-relatore
Dr. Marco Vallone, Co-relatore

Politecnico di Torino
July 17, 2020

Contents

Summary

The past few years have witnessed a staggering increase of information traffic, growing incessantly with the complexity and quality of the services provided to users around the world. A remarkable example can be found in the social media history: starting from Facebook (quite focused on text interactions), people swept to Instagram (strongly image-oriented), and at present one of the most fashionables media is TikTok (based on video sharing). The paradigm enabling this evolution has been "centralized computation", most of the data processing being processed in data centres, where most of traffic is localized, and then delivered to the end users. In this context, the role of short-range interconnects is becoming comparable or even greater than that of telecommunications, leading to the requirement of fast, low-power optoelectronic devices.

One of the main limitations of this architecture is the interface between the optical domain, pertinent to data communications, and the electrical domain, which is focused on data processing. Until an all-optical computer will be technologically feasible, one of the most promising solutions to attack this bottleneck is silicon photonics (SiPh), which promises a synergistical, low-cost and fully CMOS-compatible integration of optical and electronics systems. Exploring new device concepts or even just optimizing consolidated geometries requires extensive and expensive trial-and-error prototyping campaigns, each prototype requiring manufacturing a wafer from masks to epitaxy.

Aiming to overcome the state of the art SiPh receiver subsystems, the scope of this thesis is developing a computer-aided design framework for waveguide photodetectors. Also from a simulation perspective, dealing with the optical and electrical domains is delicate, as it requires a joint treatment of Maxwell's equations and of a carrier transport model, resulting in a self-consistent multiphysics picture: in the first instance, the spatially-resolved photogenerated carrier distribution is evaluated from absorbed photon density through a full-wave

electromagnetic simulation (optical model); then, this is used as a source term in an electrical simulation obtained from the solution of the electron and hole continuity equations with drift-diffusion constitutive relations, coupled to the quasistatic Poisson's equation (electrical model).

The thesis is organized as follows:

- The first chapter reviews the state of the art of waveguide photodetectors, introducing the figures of merit necessary to assess the electro-optical performance (average photogeneration, responsivity, modulation response).

- The second chapter describes the details of the simulators used in this work. The electromagnetic section is based on Synopsys RSoft FullWave: a 3D finite-difference time-domain (FDTD) solver of Maxwell's equations. The electrical simulator is Synopsys Sentaurus Device, which solves the Poisson-drift-diffusion system with a finite-box method based on the Scharfetter-Gummel discretization of the drift-diffusion relations. The coupling strategy behind the multiphysics approach is described in detail, with emphasis on possible model simplifications allowing to reduce the computational burden (low carrier generation rates).

- The third chapter is focused on two silicon-germanium waveguide photodetectors operating in both O (1.31 µm) and C (1.55 µm) bands: one based on mode-evolution, one on butt coupling. After describing their operation by means of approximate semi-analytical models, the two devices are optimized and compared on the basis of multiphysics simulations. The results of these simulations campaigns are supporting prototyping iterations carried out by an industrial partner (Cisco Systems).

Chapter 1

Photodetectors: a general introduction

Very fast communication systems are spreading all around the world and the role of fast optoelectronic devices is essential in order to guarantee the ultra-wide bandwidth high transmission rate needed. Photodetectors convert light into electrical signals, so the challenge consists in optimisation of these devices, leading to even higher transmission rate. Moreover, the combination with conventional Silicon integrated circuit (SIC) is fundamental for the development of Silicon Photonics (SP), where the integration of different materials allows high performance integrated systems.

This first chapter wants to be an overview of the main photodetectors implementation and an introduction to the main problem and figures of merit discussed in this thesis.

1.1 Photodetectors

The multiphysics approach proposed in chapter 3 that takes into account the solution of both optical and electrical problem (Maxwell equation and semiconductor drift-diffusion model) has multiple applications, but it is particularly suitable for the simulation of optoelectronic devices. The focus is waveguided photodetectors, one of the most important implementations of photodetectors where light is guided with a dielectric waveguide to the detecting region.

In order to better place in context the important role of photodetectors, we start from the basic building blocks of a optical communication system, an ideal scheme that describes the role of the different components involved in the communication chain (see figure 1.1).

An optical communication system consists of three fundamental blocks:

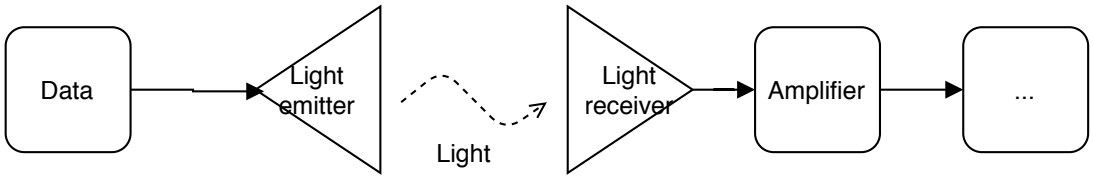


Figure 1.1: Block scheme of an ideal optoelectronic transmitter-receiver communication system

1. the transmitter: digital data coming from the transmitter system (like for example a microprocessor or a microcontroller) are converted into analog signal that is used to modulate a light source (for example a laser),

2. the channel (or link): it is usually based on optical waveguide (for example optical fibers),

3. the receiver: it converts the modulated light into a current that contains the information transmitted.

In the receiver block, photodetectors are the fundamental components as they convert the light received from the transmitter into a current (called *photocurrent*), which can be managed by high speed electronic systems in successive chain blocks (such as amplifiers, decoders, ...).

The conversion process of light into a current takes advantage of a generation-recombination mechanism of carriers inside the semiconductor. The interaction of light (electromagnetic field) with the semiconductor generates electron-hole pairs that contribute to the overall current. In order to make an effective photodetector, the photogenerated electron-hole pairs must be extracted from the photodetector. This is usually done by means of an external electric field applied on the device, which is expected to generate an output current. A current is present even if no light is illuminating the device, in which case it is referred to as *dark current*. For example, in detectors based on *pn* junctions, dark current is related to the inverse saturation current of the *pn* junction.

In order to describe a photodetector we have to define suitable figures of merit that allow us to compare different configurations. The first is the responsivity, usually indicated with $\mathcal{R}$, measured in A/W. This quantity links the input optical power to the output current of the detector taking into account the dependence of the detector on the optical filed (such as light wavelength) and the intrinsic parameters of the device.

For low input optical power this relation is approximately linear, however, increasing the input optical power the electron-hole pair generation mechanism

may saturate, screening the capability of the detector to convert light into a current. This effect is due to the high number of photogenerated carriers that do not allow to incoming light to interact with the semiconductor, generating new electron-hole pairs. The photodetector is dependent on the wavelength λ of the optical input field. In fact, the photogeneration process is dependent on the absorption profile of the material, so the responsivity is strongly influenced by the input optical wavelength. As a result, the responsivity can be written as a function of the wavelength $\mathcal{R} = \mathcal{R}(\lambda)$.

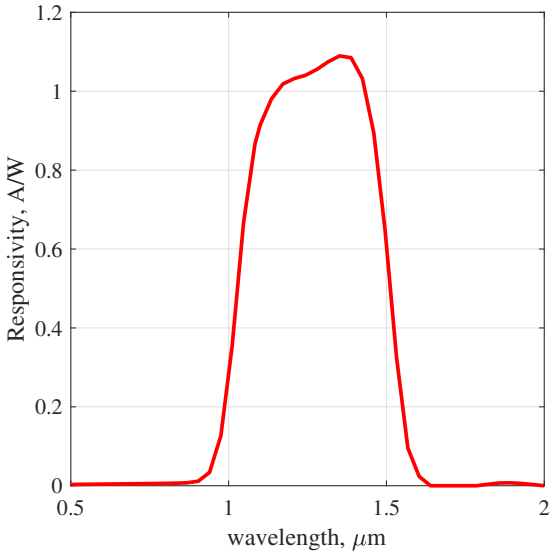


Figure 1.2: Example of the behaviour of the responsivity of a pin photodetector with respect to the input optical wavelength

Device capacitance and parasitic resistances are responsible for the responsivity dependence, indicating that not all of the photogenerated carriers have been collected. Responsivity can be considered memory-less only if the input optical signal is slowly varying in time, i.e., the input signal spectrum is significant only below the device cutoff frequency. If the variation in the modulated signal is too fast, the output current is not able to follow instantaneously the input optical field, due to the delay mechanism in carrier transport introduced by device capacitance and parasitic resistances. So, a frequency-dependent responsivity $\mathcal{R}(\omega)$ relates the photocurrent component at ω to the input optical power. With this in mind, the main parameters we have to take into account are the responsivity with respect to the input optical power and the frequency behaviour of the device with respect to the modulation of the input optical power, which is called *frequency response* of the device.

In order to better understand the photodetector operation we can introduce

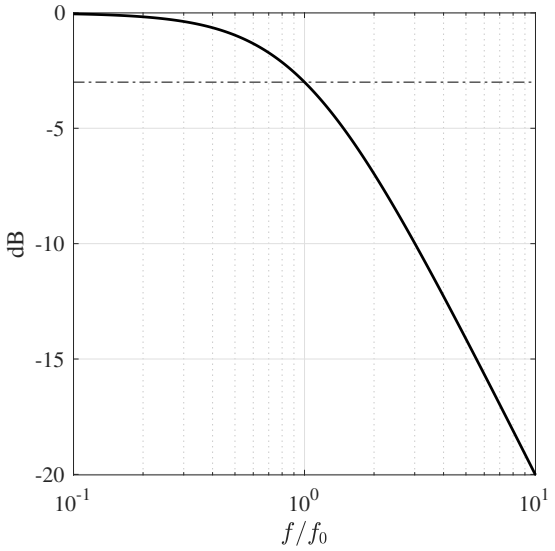


Figure 1.3: Frequency response of a general photodetector, with frequency axis normalized with respect to the cutoff frequency

a circuit model. A simple ideal electric model of the photodetector consists of a current generator with a very high impedance, with parasitic resistance and capacitance in parallel with the generator. In figure 1.4 it is reported the circuit schematic of the circuit model of a photodetector.

A photodetector can be implemented in different ways, either by changing the detector material, or by changing the photodetector structure. Different materials have been exploited, but in this thesis the focus is on the Germanium detector. Several types of semiconductor-based detector configurations exist, namely

1. *bulk* photodetector (such as photoresistors, photoconductors),

2. *junction-based* photodetector (such as *pn* or *pin* photodiode),

3. *avalanche* photodetectors (which exploit internal gain by means of a controlled avalanche process),

4. *phototransistors* (quite similar to bipolar junction transistors, where the amplified base current is photogenerated).

Each of these implementations has advantages and disadvantages in terms of responsivity, frequency response and costs. For example bulk photodetector are very simple, low cost devices, but their response in terms of responsivity and frequency is far from ideal.

This thesis focuses on junction-based devices, and in particular to waveguide photodetectors with a configuration provided by the literature [2], where the photogeneration process takes place in a reverse-bias *pin* junction and the light reaches the detector by means of a optical waveguide. The idea behind

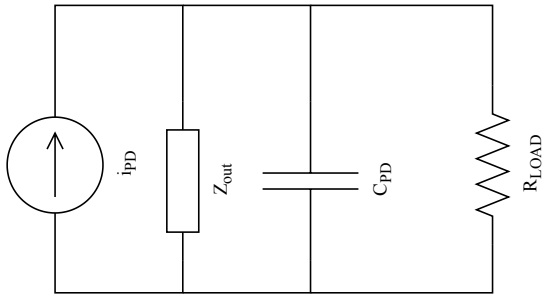


Figure 1.4: Circuit model of a generic photodetector, reproduced from [6]

the operation of a *pin* device is that photocarriers are removed by the junction reverse electric field, thus increasing the diode reverse saturation current (the dark current). A first simple implementation of a junction-based photodetector is given by a *pn* photodiode, but photons can be absorbed in the depleted region (and adjacent diffusion region) that leads to a poor frequency response, limited by the transit time and by the lifetime of photogenerated carriers. On the other hand, a more sophisticated and effective solutions are in pin photodiodes. Here, electrons holes generation occurs in a large intrinsic region placed between two high-doping layers where an almost constant electric field can be applied from an external source. The width of the intrinsic layer can be made large enough with respect to the absorption length (define in eq. (1.2)) in order to make the related photocurrent contribution dominant over the photocurrent originating from the diffusion regions.

The very possibility of carriers generation is related to the material absorption profile and, in particular, to the absorption threshold, i.e. the minimum energy that photons must have in order to be absorbed[6]. Absorbed photons must satisfy the following condition:

$$E_{ph} = \hbar\omega \geq E_g \longrightarrow\Longrightarrow \lambda[\mu\text{m}] \leq \frac{1.24}{E_g[\text{eV}]}. \tag{1.1}$$

Both direct and indirect bandgap semiconductors can be exploited in photodetectors; a direct bandgap semiconductor have the minimum of the conduction band and the maximum of the valence band align with the same momentum, while indirect bandgap semiconductors have the minimum of the conduction band and

the maximum of the valence band at different momentum values. Direct bandgap materials typically have higher absorption, which results into higher speed[9].

Along a line parallel to the propagation direction (for example the z axis), the light is absorbed according to the formula

$$P_{in}(z) = P_{in}(0)\mathrm{e}^{(-\alpha z)} = P_{in}(0)\mathrm{e}^{(-z/L_\alpha)}, \tag{1.2}$$

where $L_\alpha = 1/\alpha$ indicates the absorption length. The dimension of the detector must be consistent with the absorption length, i.e. the detecting region must be sufficiently long. In particular, in one-dimensional case, the length d of the detector should be larger than L_α,. Unfortunately a large d corresponds to higher transit time, reducing the cutoff frequency, so this is a major limitation to the detector speed[6]. In junction-based detectors, however, the absorption region that is depleted acts as a capacitor, whose capacitance can be approximated as a parallel plate capacitor. So the absorption region capacitance C is inversely proportional to d limiting the detector speed, due to RC cutoff. If in the one dimensional case d, the length of the detector, is too big, the transit-time higher, limiting the speed of the device and decreasing the cutoff frequency, while the RC limited speed increases, thus leading to a design trade-off.

1.2 Photodetector comparisons

The goal of this thesis is to analyse with CAD tools and the multiphysics approach two types of waveguide photodetector. In such context it is important to define some quantities which will allow us to compare different structures and configurations. From the electrical standpoint, photodetectors are one-ports with an optical input port[6]. Assume that the input optical power around wavelength λ is $p_{in}(t)$ and that the output current is i_{PD} (including the photocurrent i_L and the dark current i_d); the photodetector is generally characterized by the following constitutive relation:

$$i_{PD}(t) = f\left(p_{in}(t), v_{PD}(t); \frac{\mathrm{d}}{\mathrm{d}t}, \lambda\right), \tag{1.3}$$

where v_{PD} is the detector applied bias and λ is the wavelength of the optical input field[6]. The time derivative stands for the memoryless behaviour of photodetectos, but as we explained at the beginning of the chapter, if the input optical signal is slowly varying with respect to the cutoff frequency of the device, this dependence is not present. The output current, can be rewritten as:

$$i_{PD} = i_L + i_d, \tag{1.4}$$

where the dark current i_d is the current in the absence of optical power, and the photocurrent i_L is the current contribution due to incident light[6]. The dark current can be defined as:

$$i_d = f\left(0, v_{PD}(t); \frac{\mathrm{d}}{\mathrm{d}t}, \lambda\right), \tag{1.5}$$

while the photocurrent can be defined as

$$i_L = f\left(p_{in}(t), v_{PD}(t); \frac{d}{dt}, \lambda\right) - i_d. \tag{1.6}$$

In general, the relation between the photodetector current and the optical power is nonlinear and with memory. But, for a slowly varying $p_{in}(t)$ and low input optical power, it is possible to approximate the relation to a memoryless relation, linearly approximated as:

$$i_{PD}(t) = i_L + i_d \approx \mathcal{R}\left(\lambda, v_{PD}\right) p_{in}(t) + i_d\left(v_{PD}\right). \tag{1.7}$$

In many photodetectors both the dark current and the responsivity are, in fact, virtually independent from v_{PD}, and the dark current is small and in these conditions it is possible to neglect it

$$i_{PD}(t) = \mathcal{R}(\lambda)p_{in}(t) + I_d \approx \mathcal{R}(\lambda)p_{in}(t). \tag{1.8}$$

The linear dependence for the photocurrent i_L typically holds for input optical powers $P_{in} \ll P_{in,\,\mathrm{sat}}$, where $P_{in,\,\mathrm{sat}}$ is the saturation optical power at which the photocurrent saturates at $I_{L,\,\mathrm{sat}}$.

1.2.1 Responsivity

As mentioned (and discussed in the chapter 2), the detector photocurrent (and therefore the responsivity) can be in principle derived by integrating the optical generation rate G_o. The generation rate describes the number of photogenerated carriers per unit volume and unit time, and can be directly obtained from the absorbed optical power. If we integrate the optical generation rate G_o over the device active volume[6]:

$$I_L = q \int_V G_o\left(\underline{r}, P_{in}\right) \mathrm{d}\underline{r}, \tag{1.9}$$

the responsively can be directly evaluated and With this definition it is possible to define two types of responsivity: from I_L the device responsivity can in turn be obtained as:

$$\mathcal{R} = \frac{I_L}{P_{in}} \tag{1.10}$$

or

$$\mathcal{R}_{\text{differential}} = \frac{\mathrm{d}I_L}{\mathrm{d}P_{in}}. \tag{1.11}$$

The two definitions coincide if the current-power characteristic is linear.

In the simulation that will be considered in chapter 3, the generation rate is evaluated with a numerical solver that implements FDTD method to solve the optical (Maxwell) problem. Nevertheless, a simple derivation of G_o can be useful to derive an ideal, best-case limit for the detector parameter. As a first step, we directly relate G_o to the optical power as follows

$$\frac{\mathrm{d}\tilde{P}_{in}(x)}{\mathrm{d}x} = -\alpha\widetilde{P}_{in}(x) \rightarrow \frac{\text{Energy absorbed}}{t \cdot V} = -\frac{\Delta\tilde{P}_{in}}{\Delta x} = \alpha\widetilde{P}_{in}, \tag{1.12}$$

where we have differentiated the power equation with respect to x and defined $\widetilde{P}_{in} = P_{in}/A$ as the optical power density ($\mathrm{W/m^2}$), A being the detection area[6]. Dividing by the photon energy $E_{ph} = \hbar\omega$, we obtain:

$$\begin{aligned}\frac{(\text{ Energy lost})/(t \cdot V)}{\text{Photon energy } \hbar\omega} = \frac{\alpha\widetilde{P}_{in}}{\hbar\omega} = \frac{\text{Number of photons absorbed}}{t \cdot V} = \\ = \frac{\text{Number of e-h pairs generated}}{t \cdot V} = G_o,\end{aligned} \tag{1.13}$$

hence

$$G_O = \frac{\alpha \widetilde{P}_{in}}{\hbar\omega}, \tag{1.14}$$

where G_o is the optical generation rate associated to the external photon flux, i.e. the number of electron-hole pairs generated per unit time and volume. Since the optical power density exponentially decreases with x, so the same behaviour is followed by the optical generation rate:

$$G_o(x) = \frac{\alpha \widetilde{P}_{in}(x)}{\hbar\omega} = \frac{\alpha \widetilde{P}_{in}(0)}{\hbar\omega} \mathrm{e}^{-x/L_\alpha} = G_o(0)\mathrm{e}^{-x/L_\alpha}. \tag{1.15}$$

For simplicity, let us now assume that all the incoming optical power is absorbed, and all of the generated electron hole pairs are collected as a current in the external circuit. We have:

$$\frac{I_L}{q} = A\int_0^\infty G_o(x)\mathrm{d}x = A\int_0^\infty \frac{\alpha \widetilde{P}_{in}(x)}{\hbar\omega}\mathrm{d}x = -\frac{A}{\hbar\omega}\int_0^\infty \frac{\mathrm{d}\widetilde{P}_{in}(x)}{\mathrm{d}x}\mathrm{d}x \approx \frac{P_{in}(0)}{\hbar\omega}, \tag{1.16}$$

therefore we have

$$\frac{I_L}{q} = \frac{P_{in}(0)}{\hbar\omega}, \tag{1.17}$$

where $P_{in}(0)$ is the incident power. From this simplified model, it follows that the photocurrent indeed linearly depends on $P_{in}(0)$ through the responsivity $\mathcal{R}$:

$$I_L = \frac{q}{\hbar\omega}P_{in}(0) = \mathcal{R}P_{in}(0). \tag{1.18}$$

Using power and current densities, we similarly have $J_L = \mathcal{R}\widetilde{P}_{in}(0)$.

The above analysis is based on the assumption that each incident photon generates an electron in the external circuit, and leads to an ideal, best-case value for the responsivity[6]. In these conditions it is possible to get

$$\mathcal{R} = \frac{q}{\hbar\omega} = \frac{q}{E_{ph}}, \tag{1.19}$$

which holds when all of the incident photons are absorbed and converted into the external short-circuit current.

In the best-case conditions described above, the responsivity is a function of the photon energy, and it has a maximum $\mathcal{R}_{max}$ that can be computed. The photon energies below the absorption threshold, that is given by the energy gap of the material, gives a responsivity equals to zero; just above the threshold a sharp increase of α is present (since electron hole pairs start to be generated), and the $\mathcal{R}$ have its maximum for $E_{ph} \approx E_g$, i.e.:

$$\mathcal{R}_{max} \approx \frac{q}{E_g} = \frac{1}{E_g[\mathrm{eV}]} \approx \frac{\lambda[\mu\mathrm{m}]}{1.24}. \tag{1.20}$$

Finally, for $E_{ph} > E_g$, the responsivity ideally decreases with increasing E_{ph}

$$\mathcal{R}(E_{ph}) \approx \mathcal{R}_{max} \frac{E_g}{E_{ph}}, \tag{1.21}$$

according to this behaviour, for energies close to the threshold, the responsivity approximately follows the absorption coefficient, while for higher energies it decreases like the inverse of the photon energy. This behaviour can be seen in figure 1.2, in which a MgCdTe sample is used as detector.

Very large maximum responsivity values are achieved in far infrared detectors, due to the inverse dependence of $\mathcal{R}_{max}$ on the energy gap. Responsivity for long wavelength infrared detectors (used in communications [18]) have its maximum an order of magnitude of 1 A/W. Additional detector figures of merit are the internal quantum efficiency η_Q and the external (or device) quantum efficiency η_x. The internal quantum efficiency is defined as:

$$\eta_Q = \frac{\text{generated pairs}}{\text{photons reaching the active region}}, \tag{1.22}$$

where typically $\eta_Q \approx 1$. On the other hand, η_x, that is defined as the *external quantum efficiency*, is directly related to the responsivity, indeed

$$\eta_x = \frac{\text{collected pairs}}{\text{incident photons}} = \frac{I_L/q}{P_{in}/\hbar\omega} = \frac{\hbar\omega}{q}\mathcal{R} < \eta_Q. \tag{1.23}$$

In general, $\eta_x \leq 1$ in the absence of gain. If we assume ideal operation $\eta_x = 1 \equiv \eta_Q$.

But the number of electrons flowing in the external circuit in real device can be substantially lower than the number of incident photons, so the responsivity can be much smaller than the maximum value evaluated with the ideal method (1.20). In fact, the incident light from the source to the detector region has to pass trough a number of steps before being converted into a current. Some of the steps are

1. Part of the power from the source is reflected at the waveguide-photodetector interface due to possible dielectric mismatch;

2. A fraction of the power is absorbed by the waveguide

3. A fraction of the power is absorbed in regions outside the detector region, so it does not contribute to the overall current;

4. A fraction of the optical power is absorbed by metals, that usually heat up;

5. A Fraction of the power is transmitted through the PD without being absorbed.

So, only part of the power is absorbed in the detector region, contributing to the output current of the device. Devices must be optimised so that only a small fraction of the input optical signal is lost during the path from the source to the detecting region, or transmitted by the detector.

1.2.2 Photodetector electrical bandwidth

The responsivity concept can be readily extended to describe the detector frequency response, i.e. the dependence of the output current with respect to the modulation frequency of the input optical signal[6]. Let us assume that the device operates in linear condition, i.e. with low input optical power. Let us recall the equation (1.3), where for simplicity the dependence with respect to the input optical wavelength has been removed (since we are considering the modulation frequency of a specific optical wavelength, so there is not a loss in generality since we are considering just a single wavelength)

$$i_{PD}(t) = f\left(p_{in}(t), v_{PD}(t), \frac{\mathrm{d}}{\mathrm{d}t}\right) \tag{1.24}$$

and we perform a kind of small signal analysis, separating the DC and signal components[1].

$$P_{in} = P_{in,0} + \hat{p}_{in}(t), \quad V_{PD} = V_{PD,0} + \hat{v}_{PD}(t), \quad I_{PD} = I_{PD,0} + \hat{\imath}_{PD}(t). \tag{1.25}$$

If we now assume sinusoidal modulation of the input optical signal, it is possible to associate phasors to the signal components. Again, this is not a loss in generality, since the complete modulated signal can be recovered as a sum of harmonic components:

$$\hat{p}_{in}(t) = \mathcal{R}\left(\hat{P}_{in}\mathrm{e}^{\mathrm{j}\omega t}\right), \hat{v}_{PD}(t) = \mathcal{R}\left(\hat{V}_{PD}\mathrm{e}^{\mathrm{j}\omega t}\right), \hat{\imath}_{PD}(t) = \mathcal{R}\left(\hat{I}_{PD}\mathrm{e}^{\mathrm{j}\omega_m t}\right), \tag{1.26}$$

where $\omega = 2\pi/f$ is the light angular modulation frequency. Linearising around a DC working point we obtain:

$$I_{PD,0} + \hat{\imath}_{PD}(t) = \underbrace{f\left(P_{in,0}, V_{PD,0}, 0\right)}_{I_{PD,0}} + \left.\frac{\partial f(\mathrm{d}/\mathrm{d}t)}{\partial p_{in}}\right|_0 \hat{p}_{in}(t) + \left.\frac{\partial f(\mathrm{d}/\mathrm{d}t)}{\partial v_{PD}}\right|_0 \hat{v}_{PD}(t). \tag{1.27}$$

[1] the subscript 0 and the upper letter denote the DC working point

We are interested in the second and third terms, that are respectively the small-signal photocurrent $\hat{\imath}_L$ and dark current $\hat{\imath}_d$.

With the just introduced phasor notation we can express the small-signal detector current $\hat{\imath}_{PD}$ as[6]:

$$\hat{\imath}_{PD}(t) = \hat{\imath}_L(t) + \hat{\imath}_d(t) = \mathcal{R}\left(\mathcal{R}(\omega)\hat{P}_{in}\mathrm{e}^{\mathrm{j}\omega t}\right) + \mathcal{R}\left(Y_{PD}(\omega)\hat{V}_{PD}\mathrm{e}^{\mathrm{j}\omega t}\right), \tag{1.28}$$

where $\mathcal{R}(\omega)$ is the complex small-signal responsivity, and $Y_{PD}(\omega)$ is the detector small-signal admittance. Therefore, the phasor related to $\hat{\imath}_{PD}(t)$ is given by:

$$\hat{I}_{PD}(\omega) = Y_{PD}(\omega)\hat{V}_{PD}(\omega) + \hat{I}_L(\omega), \tag{1.29}$$

where the signal photocurrent phasor $\hat{I}_L(\omega)$ is linearly related to the signal optical power phasor as:

$$\hat{I}_L(\omega) = \mathcal{R}(\omega)\hat{P}_{in}(\omega) \tag{1.30}$$

The complex responsivity $\mathcal{R}(\omega)$ that describe the detector small-signal frequency response is typically a low-pass filter of the modulation frequency. So only the frequency components below the cutoff frequency are significant. The bandwidth is defined as the frequency $f_{3\,\mathrm{dB}}$ at which the responsivity drops by 3 dB with respect to the DC value, in particular we have:

$$20\log_{10}\left|\frac{\Re\left(\omega_{3\,\mathrm{dB}}\right)}{\Re(0)}\right| = -3 \rightarrow \Re\left(f_{3\,\mathrm{dB}}\right) = \frac{1}{\sqrt{2}}\Re(0). \tag{1.31}$$

The cutoff frequency that has been described refers to the short-circuit photocurrent, and it is therefore independent from the detector loading but only on transit time, high-frequency cutoff, ... All these effects are typically considered in the intrinsic cutoff frequency.

Finally, the overall detector response is also affected by the load impedance and by parasitic (extrinsic) elements, such as parasitic resistance. The main load-related cutoff mechanism is the *RC* cutoff, caused by the combined effect of the device capacitances with the load resistance. A quantitative evaluation of the total cutoff frequency can be based on the simplified equivalent circuit of the photodetector in Fig. 1.4. In the frequency domain, the photodetector can be modeled by the current-voltage phasor relation:

$$I_{PD}(\omega) = \left[Y^i_{PD}(\omega) + Y^x_{PD}(\omega)\right] V_{PD}(\omega) + I_L(\omega), \tag{1.32}$$

where $I_L = \mathcal{R}(\omega)P_{\mathrm{in}}$ is the short-circuit photocurrent component at ω, Y^i_{PD} is the detector intrinsic admittance, and Y^x_{PD} is the detector admittance. The load

impedance Z_L, that may describe the input impedance (or resistance for simplicity) of amplifiers, for example. So, capacitive and resistive load, joined with the intrinsic cutoff frequency, influence the bandwidth of the detector. Luckyly, the RC cutoff may be handled at a circuit level. Assuming $Z_L = R_L$ and a total detector capacitance C_{PD}, the current on the load $I_{R_L} = -I_{PD}$ is given by

$$I_{R_L}(\omega) = -\frac{I_L(\omega)}{1 + j\omega R_L C_{PD}} \rightarrow |I_{R_L}(\omega)| = \frac{|I_L(\omega)|}{\sqrt{1 + \omega^2 R_L^2 C_{PD}^2}}, \tag{1.33}$$

Therefore, even if $I_L(\omega) = \mathcal{R}P_{in}$, with $\mathcal{R}$ not depending on the frequency, the responsivity of the loaded detector has the expression:

$$|\mathcal{R}_l(\omega)| = \frac{\mathfrak{R}}{\sqrt{1 + \omega^2 R_L^2 C_{PD}^2}}, \tag{1.34}$$

which is frequency-dependent[6].

In junction-based detectors the DC current is small (dark current) and can be considered bias-independent (or very weakly dependent), the detector large-signal model can be defined as a capacitive admittance with two current generators modeling the photocurrent i_L that is linearly dependent on the optical power[6]. The voltage-dependent photocurrent or a nonlinear detector input admittance, (as in *pn* photodiodes), or even other effects, can be implemented directly at a circuit level.

Chapter 2

Multiphysics approach: models for carriers transport and FDTD method

The modelling of semiconductor devices started in the middle of the past century with the increasing semiconductor devices downscaling and performance optimisation. One of the first mathematicians was Van Roosbroeck, that studied semiconductor devices and he formulated the so-called fundamental semiconductor device equations.

The semiconductor device equations are a system of nonlinear partial differential equations that involves electrostatic potential and carrier density distributions along the spatial domain, as reported in the following.

The first general approach to solve the system of equations was an approximate analytical solution, which allows the definition of design rules and gives an intuitive understanding of the problem, how carriers move, etc. But the miniaturisation process became stronger and stronger, and so more accurate solution were needed. Therefore, the fully analytic approach became obsolete while numerical simulation has become more and more relevant over the years, allowed by the fast development of very high speed computers with large memory.

The first to suggest the numerical approach was Gummel for the bipolar transistor. De Mari applied the fully computational approach to *pn*-junction diodes.

But standard discretisation methods (like finite difference method) suited for the analysis of small semiconductor devices, since instabilities are present and as enormous amount of computational resources are needed in order to obtain acceptable results. Scharfetter and Gummel developed a more sophisticated discretization method, which is still in use nowadays, solving the instability problem [11].

2.1 Modeling of semiconductor devices

A semiconductor is a material whose electrical properties fall between insulators and conductors. A possible representation of a semiconductor can be done by means of a band structure that describes the range of energy levels that electrons may have inside the semiconductor, as well as the ranges of energy levels that they may not have within it [17]. A simplified representation includes two energy bands, namely the valence band and the conduction band, separated by the energy gap E_g.

Since the semiconductor devices of interest in this work operate at room temperature (around $T = 300\,\mathrm{K}$) or above, it is possible that some electrons have enough thermal energy to be promoted from the valence to the conduction band, leaving behind positive charges called holes.

Electrons and holes in the semiconductor can be described by means of densities (or concentration), usually expressed in cm^{-3} (n and p respectively). Electrons and holes can interact with different entity, such as an external applied electric field (or an applied voltage to the semiconductor device), photons, as well as with other particles (like phonons) [6].

In the following, the electronic devices that are considered work in low energy condition, i.e. most of electrons and holes are situated in the minimum of the conduction band and holes in the maximum of the valence band, so the effective mass approximation can be used. This situation occurs when the semiconductor is in a condition near its thermodynamic equilibrium position (simply called equilibrium). The effective mass approximation can be obtained from the complete description of the band structure of the semiconductor considered, with the fundamental period of the reciprocal space (of the periodic lattice of the semiconductor) along the *irreducible wedges* (called *first Brillouin zone*[9]). With this description, an electron energy-momentum relation is found that can be used to completely describe the electronic band structure of the semiconductor. Possible methods to evaluate the band structure of a semiconductor can be found in [3].

So, in low energy condition, almost all electrons and holes fall in the minimum of the conduction band E_c and the maximum of the valence band E_v respectively, and these two points (with their neighboring points) can be approximated with a parabolic function, leading to the effective mass approximation [9].

Since semiconductors have a band structure, it is reasonable to describe the states that can be occupied in the two bands with the number of states per unit volume for electron in the conduction band, defined as N_c, and the number of states per unit volume for holes in the valence band, defined as N_v, both functions

of the energy. They are also called density of states (DOS) and in a bulk (also called 3D) semiconductor they have the following expressions

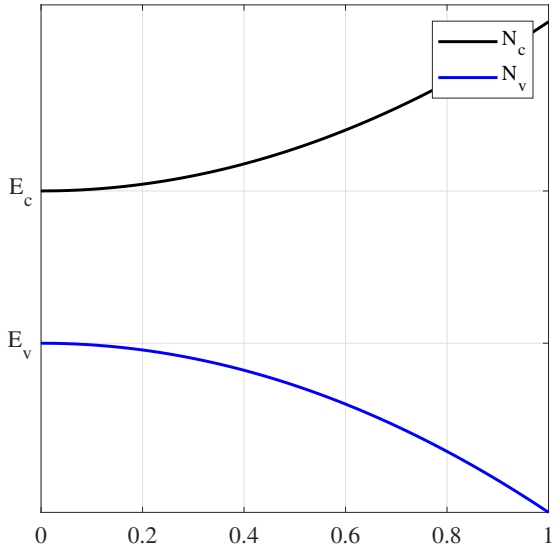


Figure 2.1: Qualitative plot of the dependence of the density of states with respect to the energy

$$N_c(E) = \frac{4\pi}{h^3}\left(2m^*_{n,D}\right)^{3/2}\sqrt{E - E_c}, \tag{2.1}$$

$$N_v(E) = \frac{4\pi}{h^3}\left(2m^*_{h,D}\right)^{3/2}\sqrt{E_v - E}. \tag{2.2}$$

Near the edge of the valence band, holes may belong to different bands very close one to the other (light holes and heavy holes), so the valence band DOS typically is larger than the one of the conduction band[16].

Electrons and holes follow at equilibrium the Fermi-Dirac distribution that describes the electrons and holes equilibrium occupation as:

$$f_n(E) = \frac{1}{1 + \mathrm{e}^{\frac{E - E_F}{k_B T}}} \tag{2.3}$$

$$f_h(E) = \frac{1}{1 + \mathrm{e}^{\frac{E_F - E}{k_B T}}}, \tag{2.4}$$

where the Fermi level E_F is constant in the whole system.

In Fig. 2.2 the Fermi-Dirac $f_n(E)$ has been plotted at different temperature. At $T = 0\,\mathrm{K}$ no electrons have energy above the Fermi level E_F, i.e. no electrons have enough energy to belong to the conduction band. On the contrary, increasing temperature, the probability to find electrons with energy above the Fermi

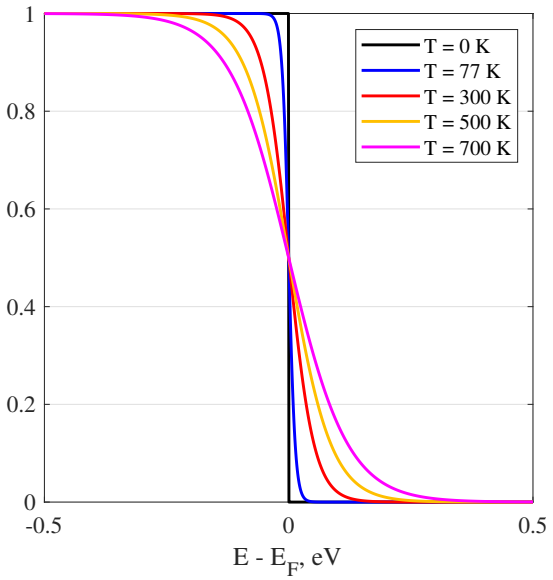


Figure 2.2: Fermi-Dirac Statistic for electrons with different temperature as a function of the difference between the energy of the particle and the Fermi level E_F

level (i.e. in the conduction band) increases. The same argument, but with opposite sign can be applied to holes Fermi-Dirac statistic, since $f_h(E) = 1 - f_n(E)$. The Fermi-Dirac distribution are essential because it can be interpreted as the availability of state, or better, it can be seen as the probability that a state with a certain energy is available. This is crucial when dealing with optoelectronic devices, since when the electromagnetic wave interact with the semiconductor, new electron-hole pairs are Created. But, to be possible, some "space"in the conduction band and in the valence band must be present.

Integrating the product between the density of states and the statistical distributions over all energies, we have:

$$n = \int_{E_c}^{\infty} N_c(E) f_n(E) \mathrm{d}E, \tag{2.5}$$

$$p = \int_{-\infty}^{E_v} N_v(E) f_h(E) \mathrm{d}E, \tag{2.6}$$

since N_c has a physical meaning only from E_c to $\approx \infty$[1], only this energy range is considered for the conduction band, while for N_v only from $\approx -\infty$ to E_v has a physical meaning, so only this energy range is considered for the valence band.

[1] N_c represents a density of state, so only real values has a physical meaning, and since the semiconductor has a forbidden gap, only hor energy outside the forbidden gap the integral is different from zero. Same for N_v

The integral 2.5 and 2.6 can not be evaluated analytically, so a numerical solution must be used. In other words equations 2.5 and 2.6 are the product of the all possible states and the probability that each state is occupied, leading to the actual concentration of electrons and holes in the semiconductor.

In many devices, the semiconductor is non degenerate, i.e. if the Fermi level lies within the energy gap, so the Fermi-Dirac distribution can be approximated with the Boltzmann approximation:

$$f_n(E) \underset{E \gg E_F}{\approx} \mathrm{e}^{\frac{E_F - E}{k_B T}}, \tag{2.7}$$

$$f_h(E) \underset{E \ll E_F}{\approx} \mathrm{e}^{\frac{E - E_F}{k_B T}}. \tag{2.8}$$

In the degenerate case the Fermi level can fall into the conduction or valence bands, and the previous condition is violated. In such case, the full Fermi-Dirac statistics has to be used. Substituting the Boltzmann approximation in equation 2.5 and 2.6 the integral can be evaluated

$$n = \int_{E_c}^{\infty} N_c(E) f_n(E) \mathrm{d}E = N_c \mathrm{e}^{\frac{E_F - E_c}{k_B T}} \tag{2.9}$$

$$p = \int_{-\infty}^{E_v} N_v(E) f_h(E) \mathrm{d}E = N_v \mathrm{e}^{\frac{E_v - E_F}{k_B T}} \tag{2.10}$$

where we define the effective densities of states as:

$$N_c = 2\frac{\left(2\pi m^*_{n,D} k_B T\right)^{3/2}}{h^3}, \tag{2.11}$$

$$N_v = 2\frac{\left(2\pi m^*_{h,D} k_B T\right)^{3/2}}{h^3}. \tag{2.12}$$

Doping is the process that substitutes atoms of the semiconductor (in the case of Si, which belongs to the IV group of the period table) with atoms of neighboring groups of the period table, such as boron (III group) or arsenic (V group)[6].

A semiconductor can be doped with a donors, whose density is usually denoted by N_D. Donors are elements capable of providing an additional electron when substituting an atom of the native semiconductor lattice. The additional electron is weakly bounded to the donor atom, so it can be easily ionized and enter the conduction band, contributing to conduction. In this case, the semiconductor is called n-type. Semiconductors can also be doped with acceptors, whose concentration is usually denoted by N_A. They have opposite behaviour

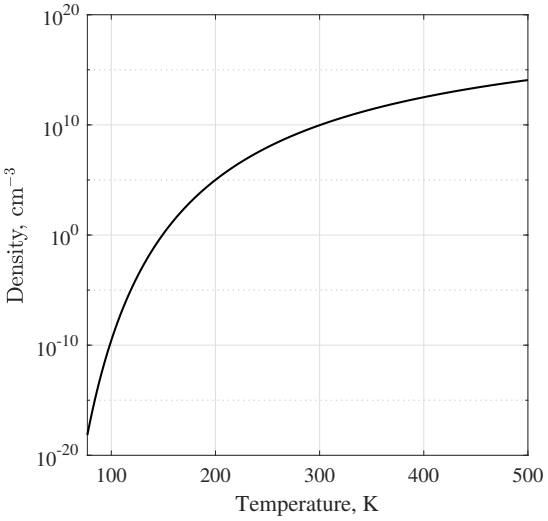


Figure 2.3: Intrinsic carrier density for Silicon as a function of the temperature

with respect to donors and can therefore attract an electron from the valence band leaving behind a hole. The semiconductor is called p-type.

In an *intrinsic* (or undoped) semiconductor at equilibrium, electron and holes have the same concentrations, since the overall semiconductor is neutral. So $p = n = n_i$, which implies that

$$n_i = N_c \mathrm{e}^{\frac{E_{Fi}-E_c}{k_B T}} = p_i = N_v \mathrm{e}^{\frac{E_v - E_{Fi}}{k_B T}}. \tag{2.13}$$

where $E_{F,i}$ is the Fermi level for an intrinsic semiconductor, and it can be evaluated as

$$E_{Fi} = k_B T \log \sqrt{\frac{N_c}{N_v}} + \frac{E_c + E_v}{2}. \tag{2.14}$$

Moreover, the intrinsic concentration can be directly evaluated only with the effective densities of states (2.11), the temperature T and the energy gap, i.e. all quantity of the material considered:

$$n_i p_i = n_i^2 = N_c N_v \mathrm{e}^{-\frac{E_g}{k_B T}}. \tag{2.15}$$

The intrinsic density grows exponentially with T. If the intrinsic concentration is of the order of the doping, the doping becomes ineffective, leading to one of the main limitations in semiconductor operation at high temperature.

In figure 2.3 an example of this effect is shown with a Silicon sample (data from [8]).

The *mass action law* shows that, in equilibrium conditions, with and without doping, the product of the concentrations n and p does not depend on the position

of the Fermi level

$$np = n_i^2 \tag{2.16}$$

and substituting the equation 2.15 in the equation it is clear that the product does not depends on the Fermi level E_F.

In this thesis only room temperature cases are of interests, i.e. dopants are in the so so called *saturation region of dopants*, where dopants are fully ionized.

From the expressions of the electron and hole densities, the Fermi level can be evaluated. In n-type semiconductors, the Fermi level increases with respect to E_{Fi}, becoming closer to the conduction band edge. On the other hand, for p-type semiconductors the Fermi level decreases and becomes closer to the valence band edge. For very high doping (with respect to the intrinsic concentration and near the value of the effective densities of states), donors and acceptors can no longer be assumed to be fully ionized (or electrically activated), but their ionization is related to the position of the Fermi level[6].

2.2 Semiconductor transport

After briefly describing the behaviour of the semiconductor at thermodynamic equilibrium, it is important to describe the effects of applied electric field to a semiconductor. The main differences with respect tot the equilibrium case is that at equilibrium the average carrier velocity is zero, while, when an electric field is applied, average carrier velocity increase. This means that, although carriers have zero ensemble average velocity, the root mean square (r.m.s.) carrier velocity (also called the thermal velocity) is extremely high. Quasi-Fermi levels are introduced to describe the behaviour of electrons and holes concentrations, substituting the single Fermi level with two separate quasi-Fermi levels E_{Fn} and E_{Fh} as follows:

$$f_n\left(E, E_{Fn}\right) = \frac{1}{1 + \mathrm{e}^{\frac{E - E_{Fn}}{k_B T}}} \underset{E \gg E_{Fn}}{,} \tag{2.17}$$

$$f_h\left(E, E_{Fh}\right) = \frac{1}{1 + \mathrm{e}^{\frac{E_{Fh} - E}{k_B T}}} \underset{E \ll E_{Fh}}{\approx} . \tag{2.18}$$

Applying again the Boltzmann approximation we get

$$f_n\left(E, E_{Fn}\right) \approx \mathrm{e}^{\frac{E_{Fn} - E}{k_B T}}, \tag{2.19}$$

$$f_h\left(E, E_{Fh}\right) \approx \mathrm{e}^{\frac{E - E_{Fh}}{k_B T}}. \tag{2.20}$$

Substituting the Boltzmann approximation in the carrier densities it becomes:

$$n = N_c \mathrm{e}^{\frac{E_{Fn} - E_c}{k_B T}}, \quad p = N_v \mathrm{e}^{\frac{E_v - E_{Fh}}{k_B T}}, \tag{2.21}$$

while the mass action law 2.16 is changed, introducing a difference in the two quasi-Fermi levels:

$$np = n_i^2 \mathrm{e}^{\frac{E_{Fn} - E_{Fh}}{k_B T}}. \tag{2.22}$$

In particular, if $np > n_i^2$ (for $E_{Fn} > E_{Fh}$) carriers injection occurs, while if $np < n_i^2$ (for $E_{Fn} < E_{Fh}$) carriers depletion occurs.

In the degenerate case, the Boltzmann approximation is not valid the expression of the charge density uses the Fermi-Dirac integrals. So the density of states can be rewritten substituting 2.17 and 2.18 in 2.5 and 2.6 respectively, obtaining:

$$n = \frac{2}{\sqrt{\pi}} N_c \mathcal{F}_{1/2}\left(\frac{E_{Fn} - E_c}{k_B T}\right), \tag{2.23}$$

$$p = \frac{2}{\sqrt{\pi}} N_v \mathcal{F}_{1/2}\left(\frac{E_v - E_{Fh}}{k_B T}\right). \tag{2.24}$$

Let us notice that the equilibrium case with the new quantity definitions is just a particular case, i.e. at equilibrium E_F is constant along the whole define, this means that if $E_{F,n} = E_{F,h}$, the semiconductor is at equilibrium.

In the presence of an applied field, the ensemble average velocity assumes a value proportional to the electric field:

$$\underline{v}_{n,av} = -\mu_n \underline{\mathcal{E}}, \quad \underline{v}_{h,av} = \mu_h \underline{\mathcal{E}}. \tag{2.25}$$

where μ_n and μ_h are the electron and hole mobilities, measured in $\mathrm{cm}^2/\mathrm{Vs}$.

If the applied electric field is low with respect to the saturation electric field value, the mobility is called *low-field mobility* and it does not depends on the interaction with lattice vibrations (phonons), impurities etc. Typically, it decreases with increasing doping and increasing temperature. For very large electric fields (with values depending on the semiconductor) the average velocity saturates:

$$v_{n,av} \to v_{n,\,\mathrm{sat}}\,, \quad v_{h,av} \to v_{h,\,\mathrm{sat}} \tag{2.26}$$

where the saturation velocities have magnitude around $10^7 \mathrm{cm/s}$.

The motion of electrons and holes due to the presence of an electric field is called the drift motion and gives rise to the drift current density:

$$\begin{aligned} \underline{J}_{n,\mathrm{dr}} &= -qn\underline{v}_{n,av} = qn\mu_n \underline{\mathcal{E}}, \\ \underline{J}_{h,\mathrm{dr}} &= qp\underline{v}_{h,av} = qp\mu_h \underline{\mathcal{E}}. \end{aligned} \tag{2.27}$$

Scattering mechanism is the main reason why a saturation velocity is present. However, for extremely small time (ps) or space (less than few microns) scales, electrons and holes are free to move in the crystal without any scattering events. In such conditions, the average carrier velocity can be much higher, in the presence of strong electric fields, than the static saturation velocity. The free motion of carriers is called *ballistic motion*, while the increase of the average velocity is calle *velocity overshoot*.

When a carriers concentration gradient is present, carriers tends to eliminate the gradient, leading to a diffusion current, that can be define as

$$\underline{J}_{n,\mathrm{d}} = qD_n \nabla n, \tag{2.28}$$

$$\underline{J}_{h,\mathrm{d}} = -qD_h \nabla p. \tag{2.29}$$

where D_n and D_h are the electron and hole diffusivities, respectively. At equilibrium the diffusivities and mobilities follow the Einstein relation $D_{n,p} = (k_B T/q)\,\mu_{n,p}$.

2.3 Generation and recombination

Carriers generation and recombination processes are described by generation rates $G_{n,p}$ i.e. the number of electrons (n) or holes (p) generated per unit time and volume, and $R_{n,p}$ for the number of (n) or holes (p) recombining per unit time and volume. It is possible to define the electron and hole net recombination rates:

$$U_n = R_n - G_n, \quad U_h = R_h - G_h. \tag{2.30}$$

In DC stationary conditions, $U_n = U_h$, while in time-varying conditions the instantaneous net recombination rates of electrons and holes can be different one from the other.

Generation recombination processes can be phonon-assisted, thermal assisted, photon-assisted (optical generation-recombination, the one of our interest), or assisted by other electrons or holes. Moreover, recombination and generation can occur through interband transitions (direct mechanisms), or assisted by intermediate trap levels in the forbidden band (due to impurities of the crystal, by doping, ...) (indirect mechanism)[6]. In direct-bandgap semiconductors (as Ge), direct optical generation rate mechanism typically is the dominant one.

2.3.1 Trap-assisted (Shockley-Read-Hall) recombination

The trap-assisted generation recombination mechanism is called *Shockley-Read-Hall (SRH)* generation recombination. In a semiconductor consider a trap density of N_t, and each of these traps introduce in the forbidden gap, an energy level E_t. Transitions due to thermal processes from the valence to the conduction bands are easier if assisted by a trap level, since two successive transitions with lower energy differences are much more probable than one single transition with energy bigger than the energy gap.

In stationary conditions [6], the net trap-assisted recombination rate can be expressed as:

$$U_{SRH} = \frac{np - n_i^2}{\tau_{h0}^{SRH}\,(n + n_1) + \tau_{n0}^{SRH}\,(p + p_1)}, \tag{2.31}$$

where:

$$\tau_{h0}^{SRH} = \frac{1}{r_{ch}^{SRH} N_t}, \quad \tau_{n0}^{SRH} = \frac{1}{r_{cn}^{SRH} N_t}. \tag{2.32}$$

The parameters r_{ch}^{SRH} and r_{cn}^{SRH} are the trap capture coefficients for electrons and holes,the coefficient g is an dimensionless parameter, called the trap degeneracy factor, E_{Fi} is the intrinsic Fermi level, close to midgap, while:

$$p_1 = n_i g \mathrm{e}^{\frac{E_{Fi}-E_t}{k_B T_0}}, \quad n_1 = n_i \frac{1}{g} \mathrm{e}^{-\frac{E_{Fi}-E_t}{k_B T_0}}. \tag{2.33}$$

2.3.2 Auger recombination (or impact ionization)

The electron or hole assisted recombination is called *Auger* recombination, and the related rate is proportional to $p^2 n$ and $p n^2$, implying proportionality not only with respect to the colliding populations (electrons and holes) but also to the population of the energy suppliers. The inverse process of the Auger recombination is the generation by *impact ionization*. Due to this dependence, the Auger recombination is important (and is indeed an unwanted competitor of the radiative recombination) in high-injection devices. In high-field conditions (i.e. for fields of the order of 100kV/cm), electrons and holes gather enough energy from the electric field between two successive scattering events (i.e. collisions with phonons, impurities or - less important - other carriers) to be able to interact with another electron and promote it to the conduction band[6].

Each electron or hole is therefore able to generate, over a certain length, a number of electron-hole pairs, that undergo in turn the same process (energy increase, scattering and e-h pair generation). The resulting chain can lead to diverging current, i.e. to avalanche breakdown in the semiconductor.

The phenomenon can be described by the following carrier generation model:

$$\nabla \cdot \underline{J}_n = -qG_n - qG_h, \quad \nabla \cdot \underline{J}_h = qG_n + qG_h, \tag{2.34}$$

where:

$$G_\alpha = \frac{1}{q} \alpha_\alpha(\mathcal{E}) J_\alpha, \quad \alpha = n, h. \tag{2.35}$$

The impact ionization coefficients α_n and α_h or α and β,with dimension of cm^{-1}, show a strong increase with the electric field.

2.3.3 Electromagnetic wave-semiconductor interaction

At a microscopic level, an EM wave with frequency f (or free-space wavelength $\lambda = c/f$) is interpreted as a collection of photons of energy $E_{ph} = hf = \hbar\omega$[6].

Photons can be considered zero-mass particles, whose speed is c_0 (i.e. the speed of light in free space) and c_0/n_r in materials, where n_r is the refracting

index of the material r. A photon has ad momentum $\underline{p} = \hbar\underline{k}$ with $\underline{k}$ as the propagating wave wavevector.

From the same microscopic point of view, a semiconductor can be seen as container electrons and holes that may interact with photons. The interaction can be seen as a scattering process (or a collision), in fact, charged particles in motion are subject to the electric field and to magnetic field of electromagnetic wave. We take into account that a semiconductor is not a dielectric or a conductor, so its response is different with respect to other type of materials since band-to-band processes allows transitions of carriers from one band to the other. In fact, the transition that may occur with the interaction of the carriers with the electromagnetic field is due to absorption or emission of a photon from the semiconductor itself . The relation between the photon energy and the the wavelength λ of the electromagnetic field is: [6]

$$E_{ph} = hf = \frac{hc_0}{\lambda} = \frac{1.24}{\lambda|_{\mu\mathrm{m}}}\mathrm{eV}. \tag{2.36}$$

It is clear that for a wavelength of the order of 1 µm the order of magnitude of the photon energy is of 1 eV. This is important, and we anticipate it here, because the interaction between the electromagnetic wave and the semiconductor is depended on the energy gap of the semiconductor. It is possible to divide the interaction in three different cases with respect to the value of E_{ph} and the energy gap E_g:

- $E_{ph} < E_g$, as in radiofrequency (RF), microwaves, or far infraRed: the interaction is weak and no band-to-band processes are involved. This kind of interaction is typical of a dielectric material [6];
- $E_{ph} \approx E_g$ and $E_{ph} > E_g$, as in Near InfraRed (NIR), visible light and Ultra Violet (UV). Light interaction is strong since band-to-band processes (and the corresponding generation - recombination of electron and hole pairs) are present. When a generation event happen, a photon has been absorbed, when a recombination event happen, a photon has been emitted;
- $E_{ph} \gg E_g$, as for X rays: this is an high-energy ionizing interactions between photons and the semiconductor, i.e. a photon may cause the generation of a high-energy electron-hole pair, that may lead to avalanche breakdown of the semiconductor[16].

The interaction between the electromagnetic field and the semiconductor involves at least a photon and an electron-hole pair, and the interaction can be

one of the following:

- Photon absorption, corresponding to a generation of an electron hole pair: a valence band electron receive the energy from the interaction, and it is promoted to the conduction band leaving a hole in the valence band. Since the process involves an exchange of energy, the electromagnetic wave decreases its amplitude and power.

- Photon stimulated emission, i.e. an electron hole pair recombination: a photon interact with the semiconductor and it emits another photon with the same frequency and wavevector as the first one. The new photon is coherent with the previous one, i.e. it increases the amplitude of the field and the power. A gain process can be defined to describe this event.

- Photon spontaneous emission, that involves the recombination of an electron hole pairs: it is different from the previous case since the emission is not correlated with another photon: the electron hole pair recombines to provide the photon energy. The emitted photon is incoherent, i.e. the process does not amplify an already existing wave, but the excitation of a new electromagnetic field is possible

Absorption and stimulated emission are practically the same process with time reversal, i.e. in the first one the photon is "converted"into an electron hole pair, while in the second case an electron hole pair is "converted "into a photon.

Quantum mechanics is the basis for the interaction of the electromagnetic wave and the semiconductor, through the so-called *perturbation theory*[13]. In the perturbation theory, the interaction must satisfy rules. The first coincides with classical collision, i.e. the total energy and momentum must be conserved during the interaction; the second consists of the so called *selection rules*; The selection rule states that some interactions are not allowed, even if they satisfy energy and momentum conservation.

2.3.4 From EM wave to optical generation rates

We have introduced the basic concepts of the interaction of light (the electromagnetic field) with a semiconductor, but we still have to define the quantities involved. The common way of introducing the electromagnetic interaction is through a generation-recombination rats, that describe the total number of electron-hole pairs generated or recombined for unit volume and unit time.

The electromagnetic wave is described inside the material with the dielectric constant. In this case, a complex dielectric constant can be defined $\epsilon(hf) = \epsilon' - \mathrm{j}\epsilon''$ and, equivalently, the complex propagation constant $k = \omega\sqrt{\epsilon\mu_0} = \bar{\alpha} + \mathrm{j}\beta$. The absorption coefficient is positive in general but can be positive or negative, depending if stimulated emission is present, corresponding to gain $\bar{g}$. A net gain can be defined as $g = \bar{g} - \alpha$. These points of view are just for the electromagnetic wave, but for the semiconductor a radiative recombination rate R_o and a radiative generation rate G_o or the net radiative recombination rate $U_o = R_o - G_o$ are defined.

The goal is to related the scattering processes with the parameters of the electromagnetic wave[12] (absorption, gain, net gain). Let us assume that the electromagnetic wave in the medium propagates as a single frequency plane wave. This is not a loss in generality, since the complete spectrum of the field can be reconstructed from the superposition of a series of single frequency plane waves. We have for the electric field

$$E = E_0 \mathrm{e}^{-\gamma z} = \underline{E}_0[\mathrm{e}^{-\bar{\alpha}z}\mathrm{e}^{-\mathrm{j}\beta z}].$$

And the corresponding optical power as

$$P_{op} = P_{op}(0)\mathrm{e}^{-2\bar{\alpha}} \equiv P_{op}(0)\mathrm{e}^{-\alpha z}, \tag{2.37}$$

where $\bar{\alpha}$ is the field attenuation, and $\alpha = 2\bar{\alpha}$ is the absorption coefficient. We describe now the electromagnetic wave as a bunch of photons, travelling in the semiconductor, whose photon density ρ_{ph} must satisfy, in steady state, to the continuity equation:

$$\left.\frac{\mathrm{d}\rho_{ph}}{\mathrm{d}t}\right|_{em,abs} \equiv W_{em} - W_{abs} = \frac{\mathrm{d}\Phi_{ph}}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\rho_{ph}\frac{c_0}{n_r}\right), \tag{2.38}$$

where Φ_{ph} is the photon flux,$W_{em}$and W_{abs} are respectively the scattering rate for the emission and the absorption. Expressing the scattering rates one has:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\rho_{ph}\frac{c_0}{n_r}\right) = \rho_{ph} w N_{cv}\left[f_n f_h - (1-f_n)(1-f_h)\right] + \frac{1}{V} w N_{cv} f_n f_h, \tag{2.39}$$

where f_n and f_h are the Fermi integrals (2.17) and (2.18), describing the occupation probabilities, N_{cv} is called the *joint density of states*, and it takes care of describing both the density of states of the valence and the conduction band. The first term is including the stimulated emission and absorption rates that is proportional to ρ_{ph}, while the second term is associated to spontaneous emission.

The first term lead to gain or loss of radiation, the second term contribute only to the process that generates photons propagating in random directions (i.e. spontaneous emission). The product of the Fermi integrals arise from the fact that an initial(final) electron hole pair is present, as well as the photon.

Absorption (or gain) decreases (or increases) exponentially the photon density in the propagation direction (z):

$$\rho_{ph}(z) = \rho_{ph}(0)\mathrm{e}^{-\alpha z}\mathrm{e}^{\bar{g}z}, \tag{2.40}$$

Substituting in the previous equation we can identify two important quantities:

$$\begin{aligned} \alpha &= \tfrac{n_r}{c_0} w N_{cv} \left(1 - f_n\right)\left(1 - f_h\right) \\ \bar{g} &= \tfrac{n_r}{c_0} w N_{cv} f_n f_h \end{aligned}$$

Finally, the net gain $g = \bar{g} - \alpha$ can be expressed as:

$$g = \frac{n_r}{c_0} w N_{cv} f_n f_h - \frac{n_r}{c_0} w N_{cv} \left(1 - f_n\right)\left(1 - f_h\right) = \frac{n_r}{c_0} w N_{cv} \left(f_n + f_h - 1\right). \tag{2.41}$$

2.4 Mathematical modelling of semiconductors

The analysis done in the previous sections provided us an overview of the behaviour of semiconductors either at equilibrium and out of equilibrium. It can be summarised with the following statements

- If an electric field is applied, carriers experience a drift motion whose velocity depends on material quantities such as electron hole mobility;
- If a gradient concentration of carriers is present, a diffusion motion of excess carriers tries to balance the carriers difference;
- Generation recombination mechanism are fundamentals for carriers transport and semiconductor interaction with light. In fact, a generation rate is the linking point between the electromagnetic world and the semiconductor world.

The last steps needed to close the mathematical model are the link between the electrostatic potential and the band diagram, so that carriers are completely defined, and the so called *continuity equation*, that describes the microscopic current behaviour.

A possible choice for the state variables of the mathematical model are the electrons and holes concentration (n and p respectively) and the electrostatic potential (or simply, the potential) φ. All the other quantities and parameters, suh as the mobility, the relative dielectric constant, ..., are implicit to the model, i.e. they are inside the equations that have been introduced before and often not directly expressed in the model.

The link between the carriers concentrations and the potential is the *Poisson equation* that in quasi-stationary condition describes its solution gives the relation of the potential and the carriers concentrations. In fact, Poisson's equation is expressed as [10]

$$\nabla^2\phi = -\frac{q}{\epsilon}\rho, \tag{2.42}$$

whit ρ defined as

$$\rho = \left(+\mathrm{q}N_D^+\right) + \left(-\mathrm{q}N_A^-\right) + (+\mathrm{q}p) + (-\mathrm{q}n). \tag{2.43}$$

N_D^+ describe the ionisated donors concentration (in full ionisaiton condition considered equal to N_D) and N_A^- describe the ionisated acceptors concentration (in full ionisaiton condition considered equal to N_A). It is important to remark

that n and p describe carriers that are free to move inside the semiconductor, while N_D^+ and N_A^- are impurities fixed charges.

Now that the links between charges and potential is made, we have to link the potential to the energy level, so that the carrier concentration formulas with Fermi-Dirac (or Boltzmann) statistic can be used. From the energy point of view the potential is linked to the energy by $E = -q\varphi$, but only energy differences (as for potential) are of physical interest. So we have to chose which energy level is more suited for our work and we set the free energy level $U_0 = -q\varphi$, as done in TCAD Sentaurus [15].

The other fundamental relation is the so called *continuity equation*, that describe the conservation of charges in motion. It can be written as [5]

$$\frac{\partial q}{\partial t} + \nabla \cdot \underline{J} + U_q = 0, \tag{2.44}$$

that for electrons can be written as

$$\frac{\partial n}{\partial t} - \frac{1}{\mathrm{q}} \nabla \cdot \underline{J}_n + U_n = 0, \tag{2.45}$$

and for holes

$$\frac{\partial p}{\partial t} + \frac{1}{\mathrm{q}} \nabla \cdot \underline{J}_h + U_h = 0. \tag{2.46}$$

But the current is described by means of a drift term and a diffusion term, hence

$$\begin{aligned} \underline{J}_n &= \underline{J}_{n,\mathrm{t}} + \underline{J}_{n,\mathrm{d}} = -\mathrm{q} n \mu_n \nabla \phi + \mathrm{q} D_n \nabla n \\ \underline{J}_h &= \underline{J}_{h,\mathrm{t}} + \underline{J}_{h,\mathrm{d}} = -\mathrm{q} p \mu_h \nabla \phi - \mathrm{q} D_h \nabla p. \end{aligned} \tag{2.47}$$

Finally, recalling that $\nabla\varphi = -\mathcal{E}$ and dividing the contribution of electron and holes, the mathematical model can be summarized for a 3D semiconductor in time Waring condition as [5]

$$\begin{aligned} \frac{\partial n}{\partial t} &= -\nabla \cdot (-n\mu_n \underline{\mathcal{E}} - D_n \nabla n) - U_n \\ \frac{\partial p}{\partial t} &= \nabla \cdot (-p\mu_h \underline{\mathcal{E}} + D_h \nabla p) - U_h \\ \nabla \cdot \underline{\mathcal{E}} &= \frac{\mathrm{q}}{\epsilon} \left(N_D^+ - N_A^- + p - n \right). \end{aligned} \tag{2.48}$$

2.5 Heterostructures

Semiconductor devices with different materials are fundamental for high speed application. In fact, using different materials the relative dielectric constant profile of a device can be engineered to achieve high performances, carrier and radiation confinements and different absorption profile. Moreover, the possibility of integrating different materials on Silicon substrate, allow different possibility, such as the compatibility with Silicon integrated technology.

Different crystals with different lattice constants can be grown on top of each other by epitaxial techniques[2]., but only under certain condition the final structure can be used to improve the electric performance. In fact, when a layer is grown on top of another layer with a different lattice constant, interface defects called *misfit dislocations* are present. These defects are electrons or holes traps. They capture the electron and reducing the electronic property of the structure, such as the overall conduction of the structure. Therefore, the resulting structure is unsuited to the development of an electron device.

However, if the lattice constant difference between the substrate and the epitaxial layer is low, an almost ideal crystal can be grown. The resulting structure is called a *heterostructure*, with unique electronic property defined by the two material used. In photodetectos, different materials are essential since they allow the absoption of the desired wavelength, and the compatibility with Silicon is fundamental for integration and high speed system. In this work, the material chosen as detector is from literature [2], but semiconductor alloys of three or more semiconductor (called compounds [6]) can be used since the alloy of different semiconductor make possible engineering the lattice constant and the energy gap. With this solution, match of the substrate lattice constant and the optimisation of the energy gap is possible. An example of compound semiconductor in infrared detector is MgCdTe (usually called MERCATEL).

The material discontinuity of the heterostructure leads to important electronic and optical properties, such as confinement of carriers since they experience potential well due to the discontinuity of valence and conduction bands, and confinement of radiation since different materials have different refractive index.

When heterostructures have a slight mismatch in the lattice constant in the

[2]In an epitaxial process, the substrate wafer acts as a seed for the new layer. Epitaxial processes are differentiated from the melt-growth processes since in that the epitaxial layer can be grown at a temperature substantially below the melting point, typically 30% to 50% lower. The common techniques for epitaxial growth are chemical vapor deposition (CVD) and molecular beam epitaxy (MBE)[4]

two layers tensile or compressive strain of one layer to the other is present. In this case, it is called *pseudomorphic* structure or strained heterostructures. Strain often allows for an improvement of the material transport or optical properties.

Electronic and optoelectronic devices require to be grown on a suitable substrate. In practice, the only semiconductor substrates readily available are those which can be grown into monocrystal ingots through Czochralsky techniques[3], such as Si and few others. The use of graded buffer layers made up of a mixture of the two materials allows to exploit mismatched substrates, since it distributes the lattice mismatch over a larger thickness, removing strain and defects far away from the interface.

The band structure of a semiconductor depends on the lattice constant and the crystal structture [9]. Heterojunctions introduce significant variation of the band structure of the device, and they allow quantized structures (such as quantum well or quantum dot) that also introduce variation in the density of states, with important consequences in terms of optical properties. Moreover, strain allows other degrees of freedom, like controlling the degeneracy between heavy and light hole subbands and changing the effective mass value for carriers (often increasing their velocity). As already recalled, lattice-matched or pseudomorphic junctions between different semiconductors allow for photon confinement (through the difference in refractive indices), carrier confinement (through potential wells in conduction or valence bands) and quantized structures such as superlattices, quantum wells, quantum dots, quantum wires[6]. The affinity (Anderson) rule describe with a certain accuracy the bands discontinuities. According to this rule the conduction band discontinuity is the affinity difference, the valence band discontinuity is the difference in ionizations energies.

Depending on the material parameters, several band alignments are possible: the band of the second semiconductor can be bigger than the one for the first semiconductor, and the electron affinity can be very different, leading to different situation. These differences are organised in types.

The quantum behaviour of carriers in narrow potential wells between wide-gap and narrow-gap semiconductors can be analyzed applying the Schrödinger equation to electrons or holes. Solution of the Schrödinger equation enables to evaluate the energy levels and subbands, given the well potential profile. Depending of the type of confinement induced by the different layers used, the

[3]The Czochralski method allows the growth of monocrystals of extreme purity of a semiconductor. This process is mainly used in the growth of silicon blocks, which are obtained with the form of cylindrical loaves [4]

reduced dimensionality structures can be divided in:

1. Confinement in one direction (quantum well), particles are confined along one direction by a potential well but are free to move along the other two;
2. Confinement in two directions (quantum wire): particles are confined along two direction but they are free to move along the remaining one
3. Confinement in three directions (quantum dot): particles are entirely confined and cannot move.

In this thesis the focus is on Si and Ge, an indirect bandgap alloy important for electronic applications (heterojunction bipolar transistors) but also for detectors and electroabsorption modulators. From table 2.1, it is possible to evaluate the lattice difference $\Delta a = 5.658\,\mathring{A} - 5.4318\,\mathring{A} = 0.2162\,\mathring{A}$, where the Ge lattice constant is approximately 4 % bigger with respect to the Si lattice constant.

Parameter	Ge Value	Si Value
Crystal structure	Diamond	Diamond
Dielectric constant	16.2	11.7
Effective electron masses m_l	$1.6m_0$	$0.98m_0$
Effective electron masses m_t	$0.08m_0$	$0.19m_0$
Effective hole masses m_h	$0.433m_0$	$0.49m_0$
Effective hole masses m_{lp}	$0.043m_0$	$0.19m_0$
Electron affinity	4.0 eV	4.05 eV
Lattice constant	5.658 Å	5.4318 Å

Table 2.1: Germanium parameters at 300 K, taken from [7] and [8]

2.6 FDTD

This section focuses on the study of simulation software programs for microdevices such as Synopsys TCAD Sentaurus (RSoft FullWave tool). This software program is used to solve Maxwell's equations in two or three spacial domanis by means of the so-called *Finite-Difference Time-Domain* method *FDTD*, providing 3D CAD modelling and a rich library of properties of semiconductors.

The main challenge consists in integrate different kind of solver on the same device, but it is possible with Synopsys TCAD Sentaurus. It allows the user to evaluate the optical field in the domain of interest, then it converts the evaluated field distribution taking care of the different material parameters in to a generation rate. The first step consists in studying how FDTD methods can be applied to simulate electromagnetic phenomena in micro optoelectronic devices.

2.6.1 Framework

This section is dedicated to the study of FDTD method used to simulate computational electrodynamics problems by finding approximate solutions to the associated system of differential equations. This method produces approximate solutions for a wide range of problems in electromagnetic theory.

We recall that Maxwell equations in a linear and isotropic medium are given by [1]

$$\begin{aligned} \frac{\partial \mathcal{D}}{\partial t} &= \nabla \times \mathcal{H} - J \quad \text{(Faraday's Law)}, \\ \frac{\partial \mathcal{B}}{\partial t} &= -\nabla \times \mathcal{E} \quad \text{(Ampère's Law)}, \\ \nabla \cdot \mathcal{B} &= 0 \quad \text{(Gauss's Law for magnetism)}, \\ \nabla \cdot \mathcal{D} &= \rho \quad \text{(Gauss's Law for electric fields)}. \end{aligned} \tag{2.49}$$

In order to simply the problem, we can assume that the constitutive relations are the following:

$$\begin{aligned} \mathcal{D} &= \epsilon_0 \mathcal{E} \\ \mathcal{H} &= \tfrac{1}{\mu_0} \mathcal{B} \end{aligned} \tag{2.50}$$

The FDTD method allows to compute numerical solutions of Maxwell's equation by means of a finite difference approximation for both space and time derivatives. The method falls in the category of "resonance region"techniques[19]. This means that the characteristic dimensions of the domain of interest and the wavelength have the same order of magnitude.

The FDTD method can be applied to a wide range of problems, however, in some cases, it can become computationally too expensive. For example, if the domain is very small compared to a wavelength, quasi-static approximations can provide more efficient solutions.

The basic step in the construction of a finite difference scheme consists into the approximation of the first derivative of the unknown function. To fix ideas, let us perform a simple computation. Consider the Taylor series expansions of a function $f(\cdot)$ around the point x_0, and consider offsets given by $\pm\delta/2$, in particular we write:

$$\begin{aligned} f\left(x_0+\tfrac{\delta}{2}\right) &= f(x_0)+\tfrac{\delta}{2}f'(x_0)+\tfrac{1}{2}\left(\tfrac{\delta}{2}\right)^2 f''(x_0)+\tfrac{1}{6}\left(\tfrac{\delta}{2}\right)^3 f'''(x_0)+o(\delta^3), \\ f\left(x_0-\tfrac{\delta}{2}\right) &= f(x_0)-\tfrac{\delta}{2}f'(x_0)+\tfrac{1}{2}\left(\tfrac{\delta}{2}\right)^2 f''(x_0)-\tfrac{1}{6}\left(\tfrac{\delta}{2}\right)^3 f'''(x_0)+o(\delta^3), \end{aligned} \tag{2.51}$$

where the primes indicate differentiation. Subtracting the second equation from the first one, yields:

$$f\left(x+\frac{\delta}{2}\right)-f\left(x-\frac{\delta}{2}\right)=\delta f'(x_0)+\frac{2}{3!}\left(\frac{\delta}{2}\right)^3 f'''(x_0)+o\left(\delta^3\right). \tag{2.52}$$

Dividing by δ, we have

$$\frac{f\left(x+\frac{\delta}{2}\right)-f\left(x-\frac{\delta}{2}\right)}{\delta}=f'(x_0)+O\left(\delta^2\right), \tag{2.53}$$

thus, we have obtained an approximate expression for the derivative of $f(\cdot)$ at x_0. Notice that such approximation is δ^2 -accurate.

All in all, we have derived the so called central-difference approximation:

$$f'(x_0)\approx\frac{f\left(x+\frac{\delta}{2}\right)-f\left(x-\frac{\delta}{2}\right)}{\delta}. \tag{2.54}$$

2.6.2 Yee's Algorithm

The FDTD algorithm was first proposed by K. Yee. The algorithm employs second-order centred differences. The fundamental steps in the construction of this numerical scheme are the following (reported from the Yee original article [19]):

1. discretise space and time for the electric and magnetic fields;
2. replace all the derivatives in Ampere's and Faraday's laws with finite differences;

3. use the current values of the electric fields to compute the magnetic fields one time-step into the future;

4. use the values of the magnetic fields obtained in previous step to compute the electric field one time-step into the future;

5. iterate the previous two steps until the fields have been obtained over the desired time interval.

The complete method with different applications can be found in [14] and [20]

2.6.3 Stability of the method

In order to obtain accurate simulations, the size of the grid cells must be sufficiently small in relation to the minimum wavelength we consider. As proposed by K. Yee in his paper[19], in order to ensure computational stability in three dimensions, we set [20]

$$c\Delta t < \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}, \tag{2.55}$$

where c is the speed of light. For simplicity we can choose $\Delta x = \Delta y = \Delta z$. To ensure that the method produces reliable outputs, we add the following condition on Δx :

$$\Delta x < \frac{\lambda_{min}}{10}. \tag{2.56}$$

2.6.4 One-dimensional example

Let us consider a simple one-dimensional problem which allows to better understand the method. In particular, consider a plane wave travelling in the z direction, assuming "free space" as propagation medium. In the one-dimensional case, we can focus the attention on the terms E_x and H_y, as the other terms vanish. We are left with

$$\begin{aligned} \frac{\partial E_x}{\partial t} &= -\frac{1}{\epsilon_0}\frac{\partial H_y}{\partial z}, \\ \frac{\partial H_y}{\partial t} &= -\frac{1}{\mu_0}\frac{\partial E_x}{\partial z}, \end{aligned} \tag{2.57}$$

and this system of equations admits a plane wave traveling in the z direction as a solution. As we are considering a centred difference scheme, we have to evaluate the functions E_x and H_y at the midpoints, i.e. at points that are

shifted in space by half a cell (in this case, $\Delta x/2$) and in time by half a time step ($\Delta t/2$). More specifically, we have:

$$\begin{aligned} \frac{E_x^{n+\frac{1}{2}}(k) - E_x^{n-\frac{1}{2}}(k)}{\Delta t} &= -\frac{1}{\epsilon_0}\frac{H_y^n\left(k+\frac{1}{2}\right) - H_y^n\left(k-\frac{1}{2}\right)}{\Delta z}, \\ \frac{H_y^{n+1}\left(k+\frac{1}{2}\right) - H_y^n\left(k+\frac{1}{2}\right)}{\Delta t} &= -\frac{1}{\mu_0}\frac{E_x^{n+\frac{1}{2}}(k+1) - E_x^{n+\frac{1}{2}}(k)}{\Delta z}, \end{aligned} \tag{2.58}$$

where n is the discretised time index and k is the discretised space index. The explicit FDTD equations can be derived from equations, resulting in

$$\begin{aligned} E_x^{n+\frac{1}{2}}(k) &= E_x^{n-\frac{1}{2}}(k) + \tfrac{\Delta t}{\epsilon_0 \Delta z}\left(H_y^n\left(k-\tfrac{1}{2}\right) - H_y^n\left(k+\tfrac{1}{2}\right)\right) \\ H_y^{n+1}\left(k+\tfrac{1}{2}\right) &= H_y^n\left(k+\tfrac{1}{2}\right) + \tfrac{\Delta t}{\mu_0 \Delta z}\left(E_x^{n+\frac{1}{2}}(k) - E_x^{n+\frac{1}{2}}(k+1)\right) \end{aligned} \tag{2.59}$$

Notice that these equations have to be complemented with suitable discrete boundary conditions when the problem requires it.

The discrete equations can be implemented in a computer code directly, however, in order to avoid computational issues due to distinct amplitudes of E and H, we introduce the new quantity $\tilde{E} = \sqrt{\frac{\epsilon_0}{\mu_0}}E$. Substituting we obtain

$$\begin{aligned} E_x^{n+\frac{1}{2}}(k) &= E_x^{n-\frac{1}{2}}(k) + \tfrac{1}{\sqrt{\mu_0\epsilon_0}}\tfrac{\Delta t}{\Delta z}\left(H_y^n\left(k-\tfrac{1}{2}\right) - H_y^n\left(k+\tfrac{1}{2}\right)\right) \\ H_y^{n+1}\left(k+\tfrac{1}{2}\right) &= H_y^n\left(k+\tfrac{1}{2}\right) + \tfrac{1}{\sqrt{\mu_0\epsilon_0}}\tfrac{\Delta t}{\Delta z}\left(E_x^{n+\frac{1}{2}}(k) - E_x^{n+\frac{1}{2}}(k+1)\right) \end{aligned} \tag{2.60}$$

2.6.5 Remarks on the method

The FDTD method has been very well studied and, when applicable, it is accurate and robust. Systematic errors and, consequently, imprecision in the results are inevitable due to the space-time discretisation, however all the errors can be controlled and made small with a suitable choice of the mesh grid. Indeed, a finer mesh (space-time discretisation) leads to more accurate results, but this determines increased computational costs.

This method can be applied to obtain (approximate) solutions for broadband systems since the time dependence is considered from the very beginning. It is not necessary to carry out repeated tests at various frequencies. Moreover, the electric and magnetic fields are computed step by step allowing us to observe their evolution in time and, eventually, convergence to a stationary state. As a remark, spatial inhomogeneity of the domain can be included in the model with minor adjustments to the method.

The stability condition is satisfied provided that the grid cells are sufficiently small. Consequently, in order to obtain acceptable computational costs, we must consider very small spatial domains. This means that the FDTD method is not adequate for the study of diffraction caused by large objects. The simulation domain is necessarily bounded, and therefore boundary conditions play a significant role and have to be chosen carefully depending on the problem at hand. We also highlight the following properties of the method.

- The Yee's algoritms gives approximate solutions for both electric and magnetic fields in time and space using coupled Maxwell's curl equation rather than solving for electric field alone (or magnetic field alone) with a wave equation. This choice makes the method more robust than others.
- Yee's algorithm centers the electric field components and the magnetic field components in a three-dimentional space so that every the electric field component is surrounded by four magnetic field components and viceversa.
- the method preserves the orthogonality of the fields;
- the location of electric field and magnetic field components in the grid and the central difference operations on these components implicitly enforce each of the two Gauss's Laws. Thus, the Yee mesh is divergence-free with respect to its electric and magnetic field.
- the time-stepping process is fully explicit, allowing us to avoid problems with coupled equations and matrix inversion.
- The resulting time-stepping algorithm is non dispersive, that is, numerical wave modes propagating in the mesh do not spuriously decay due to a non physical artifact of the time-stepping algorithm.

Chapter 3

Silicon-germanium waveguide photodetectors

In this chapter we discuss the properties of two different waveguide photodetectors and the analysis of their configurations with a multiphysics approach, where the simulations have been performed with Synopsys TCAD *Sentaurus* (electrical problem) and with RSoft FullWave (optical problem). The junction based *pin* photodetectors is at the basis of our study. We will consider designs from the literature, see [2]. The devices considered are made up of a Silicon (Si) substrate with a Germanium (Ge) detecting region, with one contact placed on top of Ge and the other on parallel to the detector.

The main difference between the two configurations lies in the waveguide-detector coupling, impacting on the absorption of light and therefore on the overall device performance. In fact, the waveguide is used to guide the optical signal without (relevant) losses up to the detecting region, where it is converted into a current. A silicon dioxide (SiO_2) cladding is surrounding the device. In the first of the two configurations considered the waveguide brings the light to the bottom of the Ge detector region, and it is called *butt coupling* photodetector. In the other configuration, *i.e.,* the *mode evolution* photodetector, the waveguide is alongside the Ge detector region, placed at the same height of the substrate, without being connected to it; here, light is detected through an evanescent mode coupling between the waveguide and the detector. The reference structures are shown in Figs. 3.2 and 3.1. The z axis is chosen as the direction of the propagation of light from the source to the detector, the x axis is used as the lateral width axis and the y axis is related to the height axis. The z axis has the origin at the beginning of the Ge detector region while the x axis has the origin in the middle of the Ge detector region (so the detector is approximately

symmetric). The y axis has the origin at the interface between the Si substrate and the Ge detector.

Starting from the literature geometries, design optimizations aimed at optimizing the responsivity and the frequency response are presented. The literature case is called the *reference* case, both for the evanescent coupling and the butt coupling. Apart from the waveguide, the geometry is the same for both reference configuration. The main section is the Germanium detector region, whose width, length and height are 1.5 µm, 12 µm and 0.8 µm, respectively. The Ge is doped near the top contact, with a p-type thin layer 50 nm thick, with a Gaussian doping profile [4] whose top value of $10^{19}\,\mathrm{cm}^{-3}$ is placed at the interface with the top contact. The Si substrate height is 0.22 µm and it exceeds of 0.5 µm the length of the detector. It is n-doped uniformly with a concentration of $10^{20}\,\mathrm{cm}^{-3}$. The electric contacts are placed one on the top of the Ge detector, with a small W contact on top of which a Cu layer is present, while the side contact is made up of 19 pillars of W, equally spaced, with another Cu layer on top. In table 3.1 detailed dimensions are reported.

Material	Width (μm)	Height (μm)	Length (μm)	Doping (cm^{-3})
Intrinsic Germanium	1.5	0.8	12	0
Germanium Doped Region	1	0.05	11.5	1e19
Silicon	3	0.22	12	1e20
Side contact	0.8	1.19	12	
Top Contact	1	0.25	12	
Waveguide	0.4	0.22		

Table 3.1: Device geometry parameters used in the simulations

Focusing on the Si waveguide, it has the same height of the Si substrate, and its width is 0.4µm; along the detector, the waveguide is tapered. In all the simulations, the source field launched in the Si waveguide is its fundamental mode.

The goal of our analysis is to find a balanced solution which allows the detectors to be used both at at a wavelength of 1.31 µm and 1.55 µm, without limiting too much the device performance.

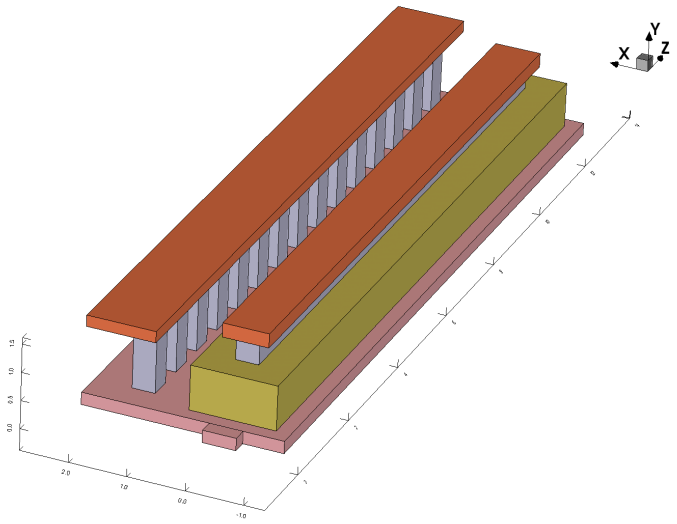


Figure 3.1: Butt coupled photodetector, ideal structure, side view. The waveguide is facing towards the substrate of the detector, on top of which a Ge region is present.

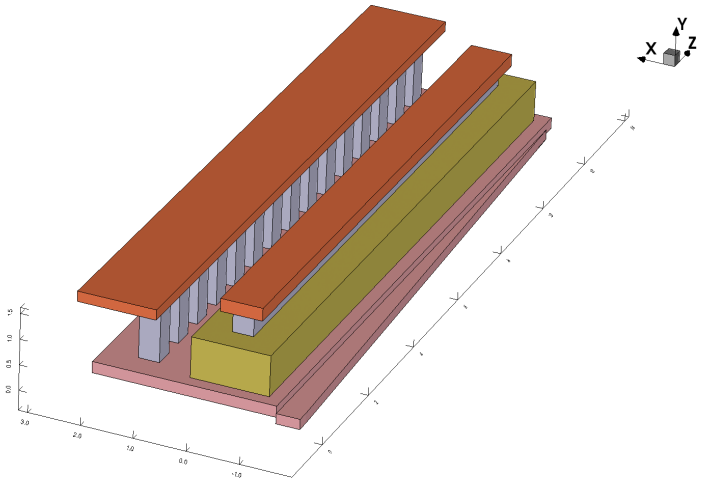


Figure 3.2: Mode evolution photodetector, ideal structure, top view. The waveguide is parallel to the detector, so an evanescent coupling between the detector and the waveguide is possible

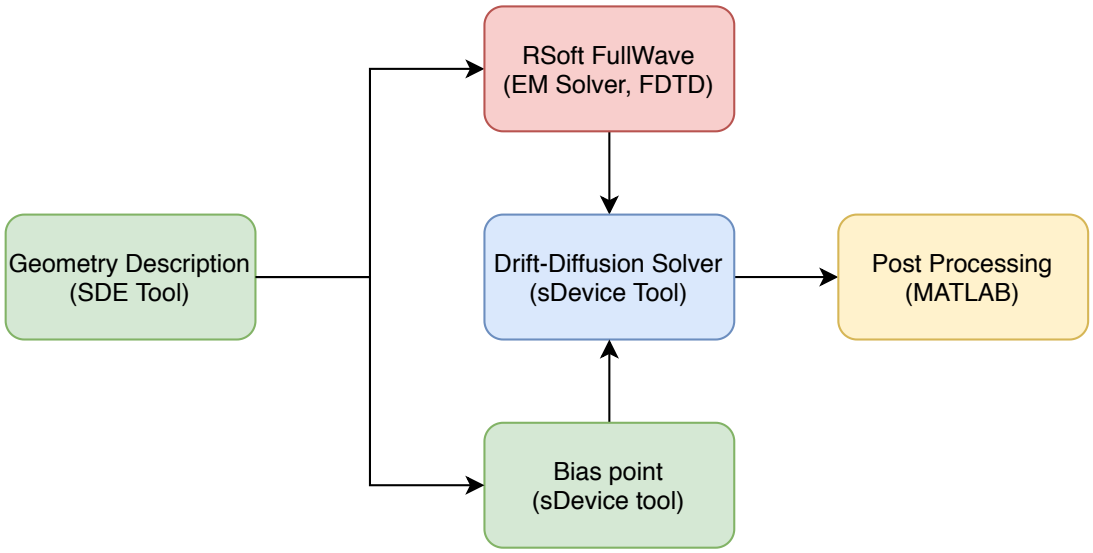


Figure 3.3: Block scheme of the solver used with indicated the name of the commercial tools. The electromagnetic problem is solved only once and used as input of the Drift Diffusion solver, that is solved self consistently

The chosen wavelengths for the simulations are $\lambda = 1.31\,\mu m$ and $\lambda = 1.55\,\mu m$, which are commonly referred to as *O band* and *C band* in optical communication systems. The analysis of these two wavelengths is crucial if compared with the transmission window of optical fibers. In fact, in the O band signal distortion inside the fiber is minimal, while the C band is nowadays broadly used in fiber optics communication system since is the amplification range of erbium doped fiber amplifiers (EDFAs)[18].

3.1 Approach used and methodology

We study both structures with the same approach using Synopsys TCAD Sentaurus. The main variables have been defined as parameters, such as the type of waveguide used (for evanescent coupling or butt coupling waveguide), the detector dimensions, ... In fact, the device parameters are the same for the two structures, with the exception of the waveguide, and only after the preliminary analysis some improvements have been explored.

A block scheme of the solver used is reported in figure 3.3. We simulated the electrical problem (i.e. finding the self consistent solution of Poisson's equation described in chapter 2 under a reverse bias) and the propagation of light inside the structures (solving Maxwell's equation). In this work, the multiphysics coupling is unidirectional, *i.e.,* the simulation is performed using the solution of the optical problem as a generation term for the drift-diffusion equations, without any self-consistent loop. This assumption can be justified considering that the input optical power is too low (tens of microwatt, from [2]) to affect the Ge absorption profile. Performing a complete self-consistent approach would require several days of computation since the solution of the Maxwell equation require by itself several hours. In this view, this simplification enables to achieve realistic results with an extreme reduction of the overall computational cost.

The geometry is generated with SDE tool, considering a worst case approach, i.e. the Ge on the Si substrate is considered bulk material, i.e. no strain is applied to the Ge, so there is no improvements in the absorption profile of the Ge at higher λ.

The geometry generated with SDE is used as input of the optical solver, RSoft FullWave, which simulates light propagation in the devices. A quite large computational box is needed, and a very dense mesh is required in order to get an accurate result. A uniform mesh in all directions is chosen, with very low Δr (i.e. distance between one point and the following one). Below 40 nm the solution of the optical problem obtained is invariant with respect to the chosen mesh size. Therefore, a mesh size of 25 nm is chosen. From the simulation perspective, this value is convenient since it is a multiple of all the dimensions of each component of the two configuration, eliminating the uncertainty at boundaries. As boundary condition for the optical problem, perfect matched layer (PML) boundary conditions are chosen. This kind of boundaries can be configured to absorb (ideally) all the light impinging on them, so that no reflected light from the boundary should be present.

Since metals contribute only with a power loss to the optical problem, they

have been described as perfect electrical conductors (PEC). Silicon is practically transparent at the two studied wavelength, in fact only the Ge detector is highly influenced by the optical field. Indeed, the Ge detector absorbs light and this triggers the generation process of electron-hole pairs.

From chapter 2 we recall that the optical power absorption per unit volume can be calculated from the divergence of the Poynting vector (S)

$$P_{abs} = -0.5\,\mathrm{real}(\nabla \cdot S). \tag{3.1}$$

It can be shown that the above formula is equivalent to

$$P_{abs} = -0.5\omega|E|^2 \text{ imag } (\varepsilon), \tag{3.2}$$

where ε is the dielectric constant, $\omega = 2\pi f$ (f is the optical frequency) and $|E|$ is the optical field intensity.

Since the problem variables are 3-dimensional, it is difficult to represent them clearly. In this view, we show integral averages over the Ge detector in the transverse plane to the light propagation direction, enabling a better understanding of the behaviour of light in the detecting region. This can be written as

$$\begin{aligned} P_{abs,z} &= \frac{1}{(x_2 - x_1)(y_2 - y:1)} \int_{x_1}^{x_2} \int_{y_1}^{y_2} P_{abs} \mathrm{d}x\mathrm{d}y = \\ &= \frac{1}{A} \int_{x_1}^{x_2} \int_{y_1}^{y_2} -0.5\omega|E|^2 \mathrm{imag}(\varepsilon)\, \mathrm{d}x\mathrm{d}y, \end{aligned} \tag{3.3}$$

where z is the light propagation direction, x_1, x_2, y_1, y_2 are the boundary of the Ge detector, whose transverse area is called A.

The mode light source used in the simulations is computed by the FDTD tool as a preliminary step, and then this is propagated in the waveguide/device. This procedure helps reducing noise and interference, since, if the mode is confined in the waveguide, practically no light may escape the waveguide and it is directly guided to the detector.

Once the optical power in the Ge detector is evaluated, it is used as a generation-recombination rate in the drift diffusion solver. RSoft provides an utility that converts automatically the results taking care of the different material parameters. The bias point is computed, as well as the dark current, with *sDevice* tool. A reverse bias voltage of 3 V is chosen and all other computation are based on this bias point. The responsivity is computed as the derivative of the power-current (PI) curve, but since it is linear, it has been simplified with the difference of two point corresponding to minimal and maximal input power respectively. Finally, the frequency response is found with a resistive load equal

to $50\,\Omega$ that is connected at one of the two pins of the photodiode, so that capacitive effects are noticeable.

3.2 Mode evolution photodetector

In a mode evolution photodetector the waveguide is parallel to the Ge detector. Since a distance of 0.1 µm is present between the substrate and the waveguide, there is no direct coupling, and the light is coupled through the evanescent field to the detector. This allows to achieve good performance and a good distribution of the input optical field in the Ge detector. The reference structure is reported in figure 3.2. The waveguide is divided in two sections: a first section is made up of a rectangular waveguide, then the second is tapered with a length equal to the Ge detector region. A taper is a region where the waveguide change its transverse dimensions orienting light in the wanted direction, in this case the detector. Light is coupled from the waveguide in the detector in an extent directly proportional to the length of device, so that the behaviour of the absorption is substantially different in the case of 1.31 µm and in the case of 1.55 µm. In fact, at a wavelength of 1.31 µm the absoption profile of the Ge is much higher with respect to 1.55 µm, leading to an higher absorption in the initial microns of the device. In figure 3.4 the average (mean) optical generation rate is shown along the direction of propagation of light z.

The optical generation rate is rapidly increasing at the beginning of the detector. This effect is related to the coupling of the optical field: the taper present at the end of the waveguide is orienting the light inside the detector along the z axis, so at the beginning very little portion of light is present inside the detector. This is crucial in order to reduce the screening effect of the high number of optical photocarriers generated by the detector region. In figure 3.4a we report the optical generation rate for a input wavelength of 1.31 µm with an input optical power of 200 µW. After the initial increase in the optical generation rate at the end of the detector the generation rate becomes very low, allowing optimisation of the device length. On the contrary, in figure 3.4b we show the optical generation rate for a input wavelength of 1.55 µm with an input optical power of 200 µW. In this case the absorption of light is distributed along the whole device. Two considerations can be done:

1. a lower responsivity is expected in the 1.55 µm case. Since the generation rate is not decreasing along the z axis, probably only a fraction of the light is absorbed, while the remaining part is reflected or transmitted trough the device;

2. a reduction of the device length is expected to lead to a decrease in the responsivity, because the total number of the photogeneration carriers should

be lower with respect to the case with a longer device.

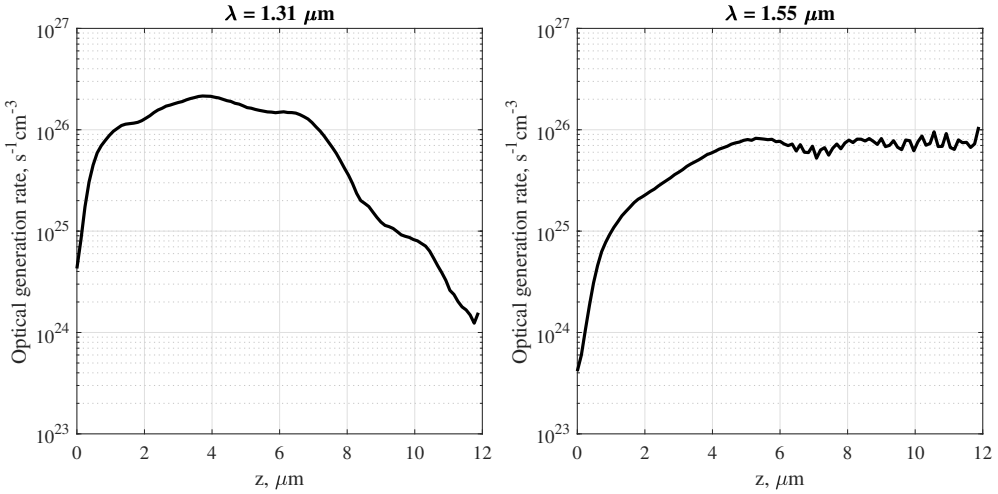


Figure 3.4: Optical generation rate of the mode evolution configuration at wavelength equal to 1.31µm (a) and 1.55 µm (b). Input power 200µW

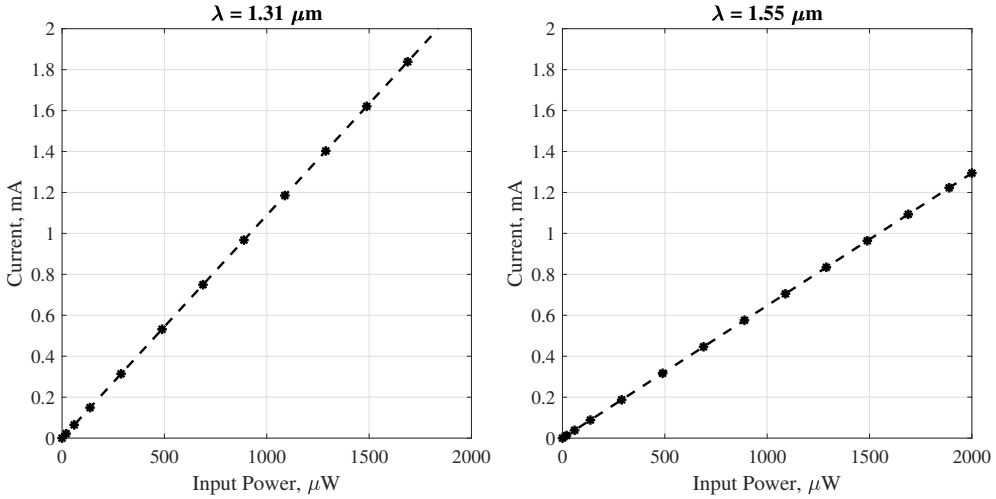
Let us now focus on the curve that represents the output current of the detector versus the input optical power. In figure 3.5a it is reported the curve for an input optical field with wavelength of 1.31 µm with different value of the input optical power. At low input optical power, no saturation effects is present, but also for high optical power the current is not saturating. This is consistent with the results obtained in the literature [2] that predict a saturation at tens of milliwatt. As expected from the optical generation rate, in the case with input optical wavelength of 1.55 µm, the curve has a lower slope with respect to the other wavelength. From these curves it possible to extract the responsivity values, which have been reported in 3.2.

	Responsivity
1.31 µm (simulation)	1.037 A/W
1.31 µm (ideal, from 1.20)	1.056 A/W
1.55 µm (simulation)	0.67 A/W
1.55 µm (ideal, from 1.20)	1.25 A/W

Table 3.2: Responsivity estimated for the mode evolution photodetector

The responsivity of the detector at 1.31 µm is close to the ideal value, whereas at 1.55 µm is very far from the ideal case, being approximately half of it. In real

devices this effect should be drastically different. Indeed here a worst case scenario has been used, since the strain of the Ge is not taken into account. In fact, a tensile strain applied in the Ge detector, should introduce an enhancement factor in the absorption profile of the Ge leading to similar results at both wavelength and this behaviour is visible in the literature [2].



(a) (b)

Figure 3.5: Power-Current (PI) plot of the mode evolution configuration at wavelength equal to 1.31 µm (a) and 1.55 µm (b)

The frequency response of the device is affected by screening effects related to the optical generation. In fact, at 1.31 µm, with higher input optical power, the cutoff frequency is reduced. The same does not occur at 1.55 µm, since the absorption of the Ge at that wavelength is considerably lower with respect to the other case, leading to an almost constant frequency response on the considered input optical power range. In figure 3.6a and 3.6b we show the complete low pass behaviour of the device with different input optical power applied to the detector. As expected, with lower wavelength the frequency is strongly affected by the input optical power, while for higher input wavelength the response is not changing for different input power.

In order to better understand the variation of the cutoff frequency of this configuration, they have been extracted from the figure 3.6 interpolating the −3 dB value. The results are reported in figure 3.7. It is clear from figure 3.7b that at higher wavelength the device is not affected by the input optical power range chosen for the simulation, while it decreases for lower wavelength as the input optical power increases.

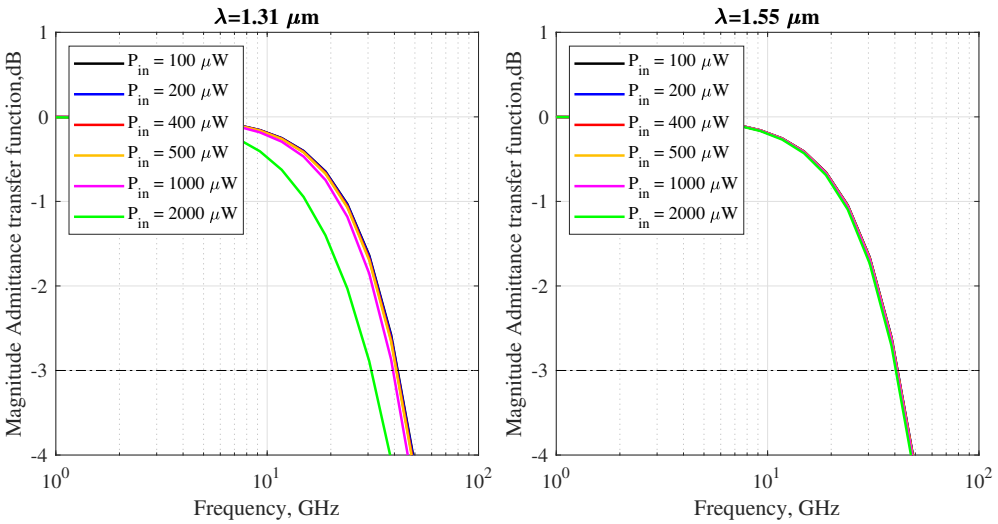


Figure 3.6: Frequency response plot of the mode evolution configuration at wavelength equal to 1.31 µm (a) and 1.55 µm (b)

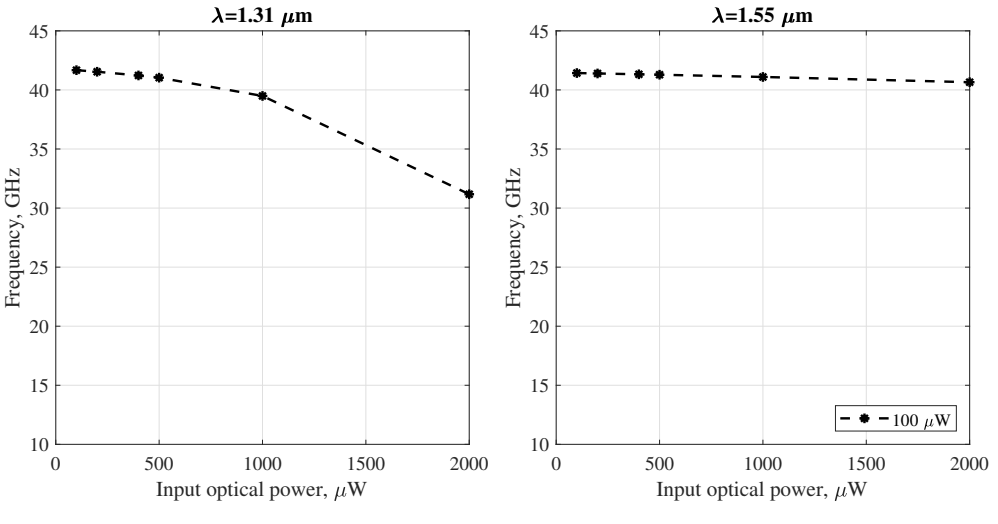


Figure 3.7: Cuttoff frequency plot of the mode evolution configuration at wavelength equal to 1.31 µm (a) and 1.55 µm (b)

3.3 Butt coupled photodetector

In mode evolution devices, the detector is fed by the waveguide through evanescent coupling. In butt-coupled detectors, the rectangular waveguide enters in the substrate. In this configuration, light is guided below the detector region. Even though this approach is expected to enhance light coupling, this leads also to higher screening effects of the number of photogenerated carriers.

It is possible to observe that the mean optical generation rate is very different at 1.31 µm in the two configurations. In the evanescent coupling configuration the very first micron presents an increase of the optical generation rate, with a slow decay in the following microns. Instead, in the butt coupling configuration one can observe a very fast increase of the generation rate and then a steep decay, until becoming negligible after few microns. In this view, in the mode evolution case, the overall optical generation rate is better distributed, being higher for a longer section of the device. Also the maximum value reached by the optical generation rate is quite different, in the butt coupling is slightly higher, leading to stronger photogenerated carriers screening effects.

The optical generation rate at 1.55 µm remains practically constant along the device length, so also in this configuration, as well as in the mode evolution configuration, a lower responsivity is expected since only a fraction of the light is absorbed, while the remaining part is reflected or radiated. This behaviour could be ascribed to the very fast oscillation pattern in the optical generation rate. In figure 3.8a and 3.8b the mean optical generation rates are reported.

The differences with respect to the previous case are significant from the responsivity and the frequency response standpoint. Starting from the responsivity, it is slightly higher at both considered wavelength, with values reported in table 3.3

	Responsivity
1.31µm (simulation)	1.037A/W
1.31µm (ideal, from 1.20)	1.056A/W
1.55µm (simulation)	0.71A/W
1.55µm (ideal, from 1.20)	1.25A/W

Table 3.3: Responsivity estimated for the mode evolution photodetector

The frequency response instead is different, since the screening effect at 1.31µm is higher with respect to the previous case. The cutoff frequency of this configuration are reported in figure 3.11. At higher wavelengths the influence of

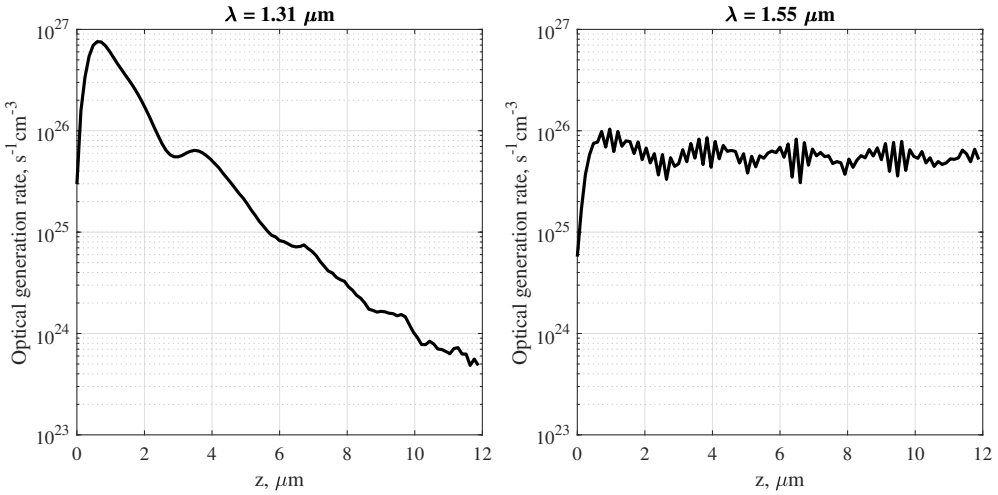
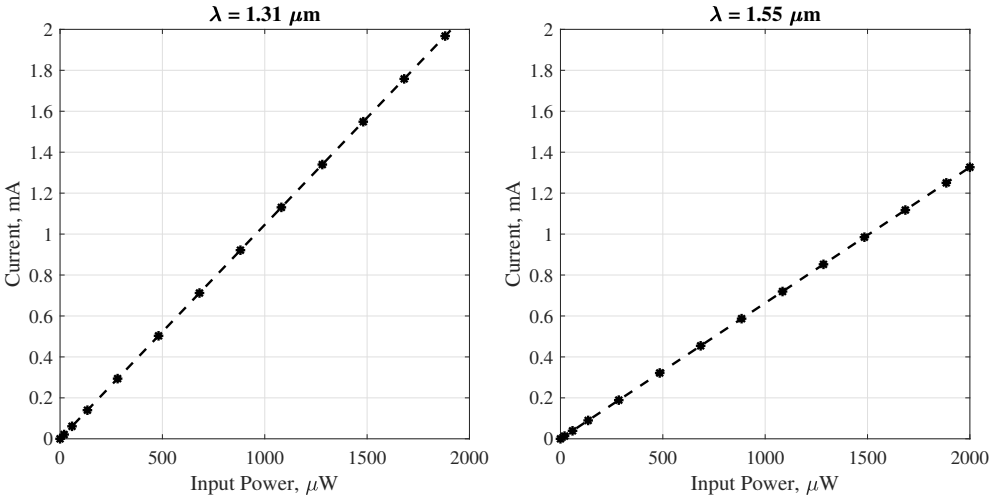


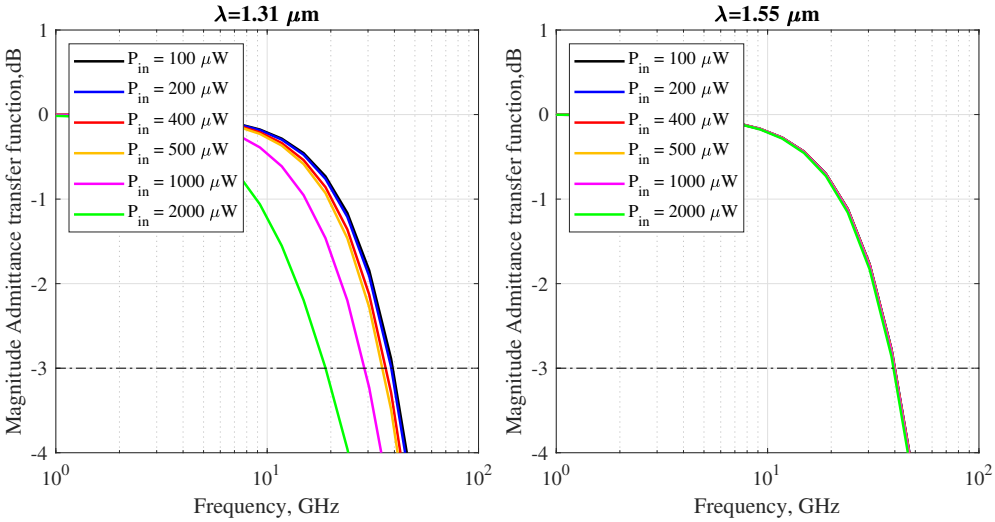
Figure 3.8: Optical generation rate of the mode evolution configuration at wavelength equal to 1.31µm (a) and 1.55µm (b)

the input optical power is practically negligible, while at lower wavelengths the dependence is practically linear with respect to the increase of the input optical power.



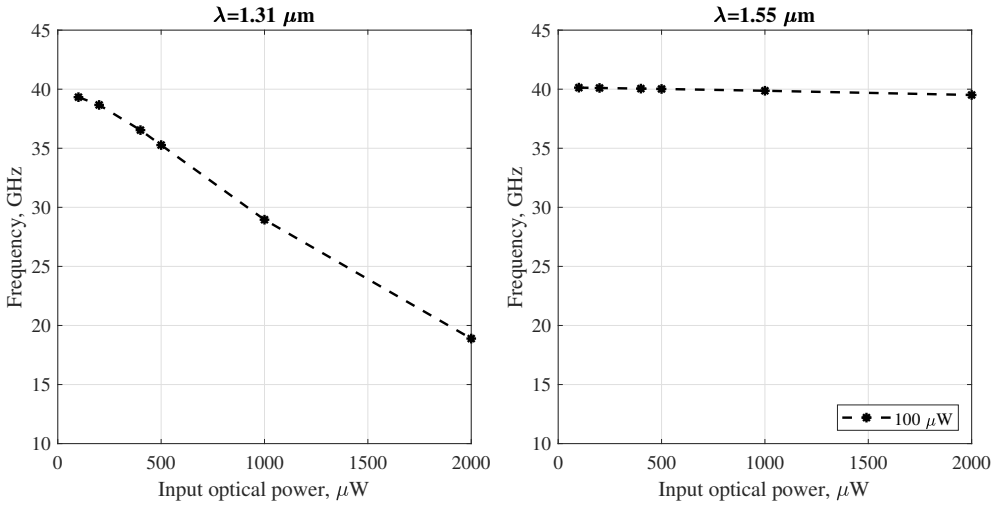
(a) (b)

Figure 3.9: Power-Current (PI) plot of the mode evolution configuration at wavelength equal to 1.31µm (a) and 1.55µm (b)



(a) (b)

Figure 3.10: Frequency response plot of the mode evolution configuration at wavelength equal to 1.31µm (a) and 1.55µm (b)



(a) (b)

Figure 3.11: Cutoff frequency plot of the mode evolution configuration at wavelength equal to 1.31µm (a) and 1.55µm (b)

3.4 Comparison and optimisation

After discussing the behaviour of the reference structure obtained from [2], some variations are applied to both configuration, in order to optimise and reduce the dimension of the device.

The experiments that are performed are carried out by trying to assess the effects of the different dimensions. As a first parameter, we investigate the length of the device, reducing it down to 5µm. Secondly, the width has been changed with values ranging from 3µm to 0.5µm. Finally, the height has been varied from 1.2µm to 0.4µm. All other parameters have been maintained unchanged.

Case	HGe (µm)	LGe (µm)	WGe (µm)	Ge doping width (µm)
1	1.2	12	1.5	1.0
2	1.0	12	1.5	1.0
3	0.8	12	1.5	1.0
4	0.6	12	1.5	1.0
5	0.4	12	1.5	1.0
6	0.8	10	1.5	1.0
7	0.8	8	1.5	1.0
8	0.8	6	1.5	1.0
9	0.8	5	1.5	1.0
10	0.8	12	0.5	0.40
11	0.8	12	1.0	0.75
12	0.8	12	2.0	1.5
13	0.8	12	3.0	2.0

Table 3.4: All the variation with respect to the reference case proposed

In the case of mode evolution detectors, the length of the waveguide is related to the length of the detector. The variations are applied in the same way to both configurations, so a total of 26 cases have been studied.

In figure 3.12 is reported the configuration of the case 9 of the table 3.4, corresponding to the minimum value of the of the length tested for this device. Figure 3.14 represents cases 1 and 4, corresponding to the extreme value of the height tested with respect to the reference case, as figure 3.13 corresponds to the extreme value (case 10 and case 13) of the width of the detector tested with respect to the reference case.

The new configuration with the variation are reported just for the butt coupling case, but the same variation have been tested also for the mode evolution

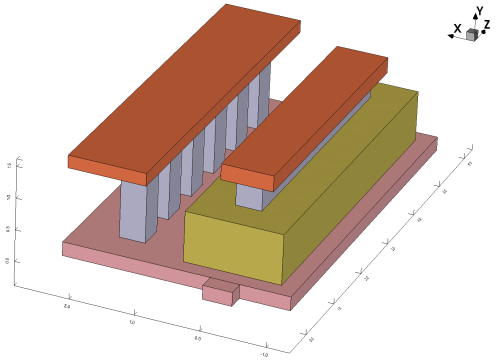
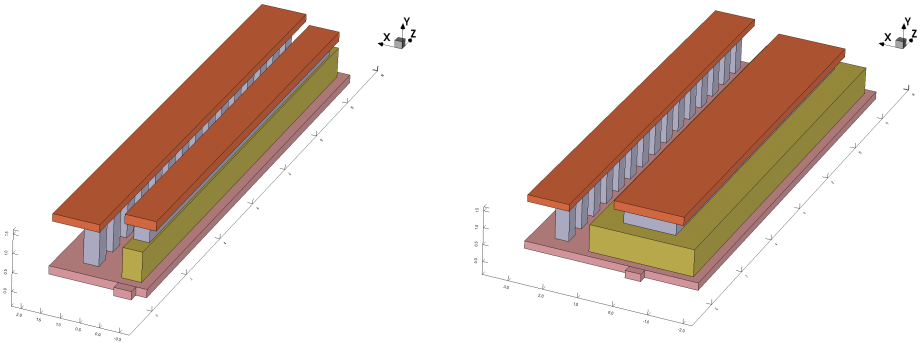


Figure 3.12: Butt coupling configuration with the minimum length used in the optimisation process

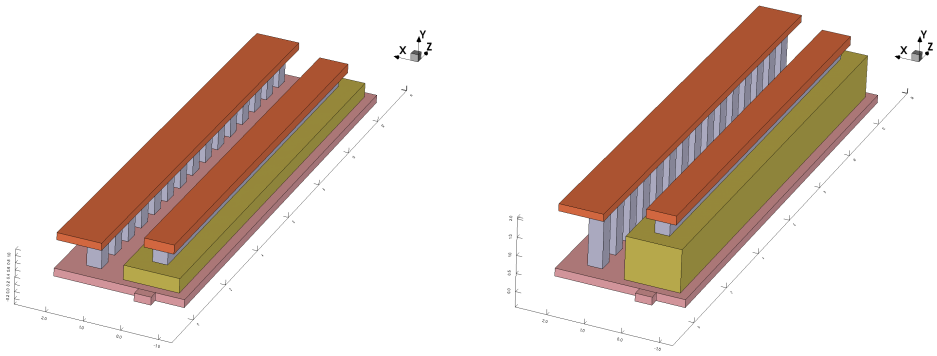


(a) Minimum value of the Ge width (b) Maximum value of the Ge width

Figure 3.13: Variation example of the Ge width in the butt coupling configuration

configuration.

For the optimisation process the quantities that we are considering are only the responsivity and the cutoff frequency of the different configurations. Let us start from the responsivity analysis with the variation of the length of Ge detector. Figs. 3.15a and 3.15b show the responsivity values at different detector length for both butt coupling and mode evolution configuration. At 1.31 µm the variation is very small in both configuration. This can be motivated recalling that the optical generation rate is steeply decreasing after few microns in the z axis, as shown in Fig. 3.8a and 3.4a. In fact, after few microns the two generation rate profiles have average values so low that cannot contribute to the current. At 1.55 µm the behaviour of the detectors are drastically different. In both butt coupling and mode evolution cases the optical generation rates (Fig. 3.8b and



(a) Minimum value of the Ge height (b) Maximum value of the Ge height

Figure 3.14: Variation example of the Ge height in the butt coupling configuration

3.4b) are almost constant along all the detector, so a decrease in the length of the device leads to a decrease in the responsivity. Similar considerations can be applied to the variation of the Ge width and height at 1.31 µm, reported in Fig. 3.15c, 3.15e, 3.15d and 3.15f. A more dramatic situation is for the variation of the width of the detector region at 1.55 µm in Fig. 3.15f. Probably the light is absorbed along the whole width of the device, and it is not localized at the beginning of the detector as in the butt coupling case, so the responsivity is drastically reduced.

From the cutoff frequency standpoint, the main dependence still remains on the input optical power. In fact, increasing the optical power, the cutoff frequency reduces, especially at 1.31 µm.

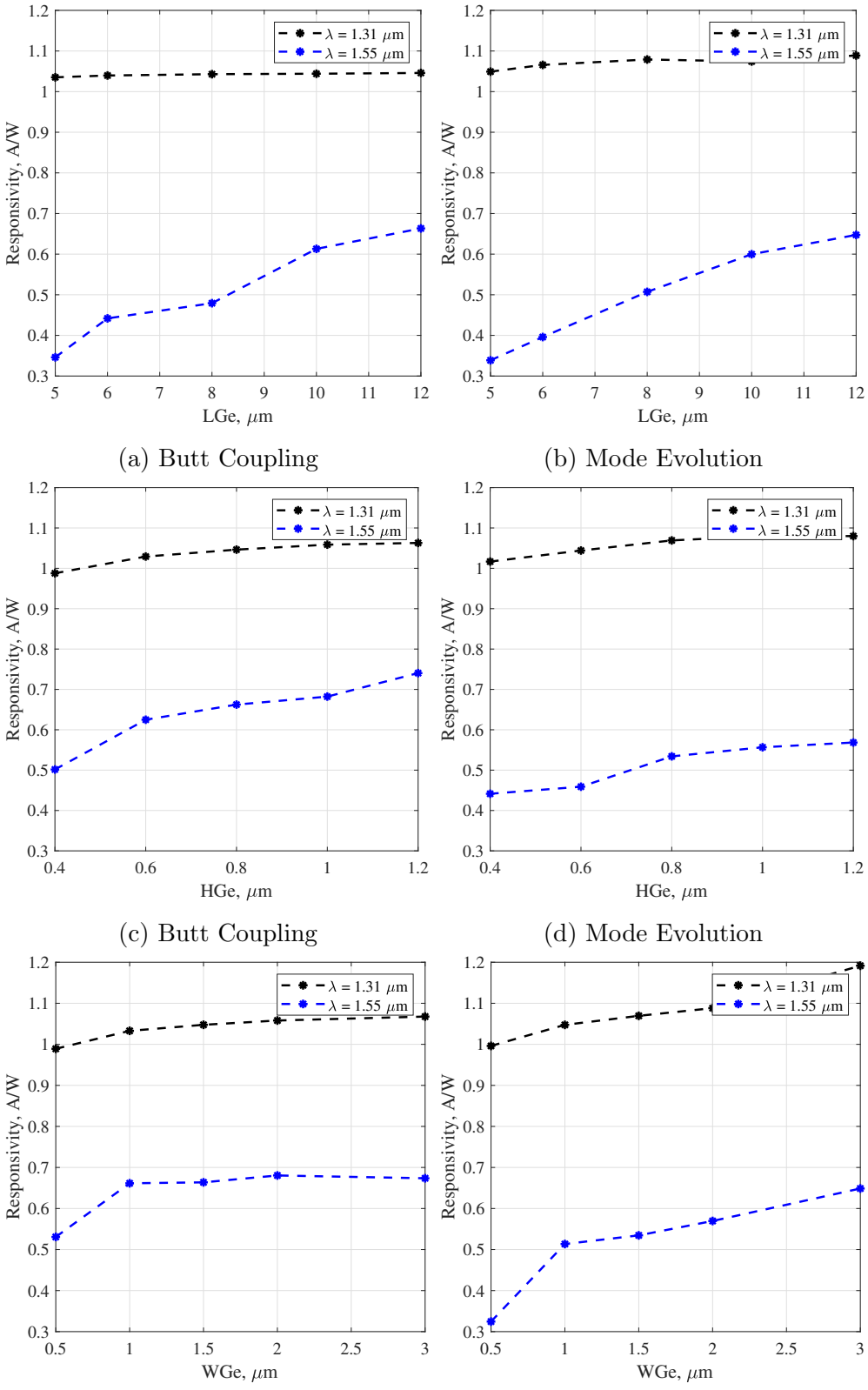
Let us start by analysing the cutoff frequency varying the length of the detector in Fig. 3.16. It is clear that the detector performance is not sensitive to the length and probably even shorter devices could be designed, if the specifications concern only the frequency response. In fact, only at high input optical power, and for the mode evolution configuration at 1.31 µm, there is a reduction of the cutoff frequency, while in all the other cases it remains practically invariant with respect to the detector length.

Then, we will focus only on the variation of the detector width. As for the variation of the Ge length, only at high input optical power there is significant variation, while in all other cases the sensitivity is very low. This behaviour is expected and it has can be derived from the discussion in 1, since the detector can be seen as a parallel plate capacitor, so only the distance of the plats is important, and since the detector is biased on top and from the bottom trough

the substrate, the important quantity for the frequency response is the height of the detector.

Finally, the variation of the height of the detector is reported in Fig. 3.18. It is clear that the height has a strong impact on the cutoff frequency, and its optimization might be crucial. Reducing the height of the device, the cutoff frequency increase has a behaviour like $1/x$, and probably even better performance might be obtained reducing even further the height. Also the power dependence becomes less evident when the detector is thin, and this can be described by means of the travel time of photogenerated carriers through the device towards the contacts. Since the device is thinner, carriers take less time to exit from the device, increasing the instantaneous response of the device.

These considerations emphasize the need of a trade-off for our final design goal: the objective is achieving an acceptable responsivity at both wavelengths, and a good frequency response, i.e., as high as possible and as independent of the input optical power as possible. In the following (and last section), a possible implementation has been presented.



(a) Butt Coupling

(b) Mode Evolution

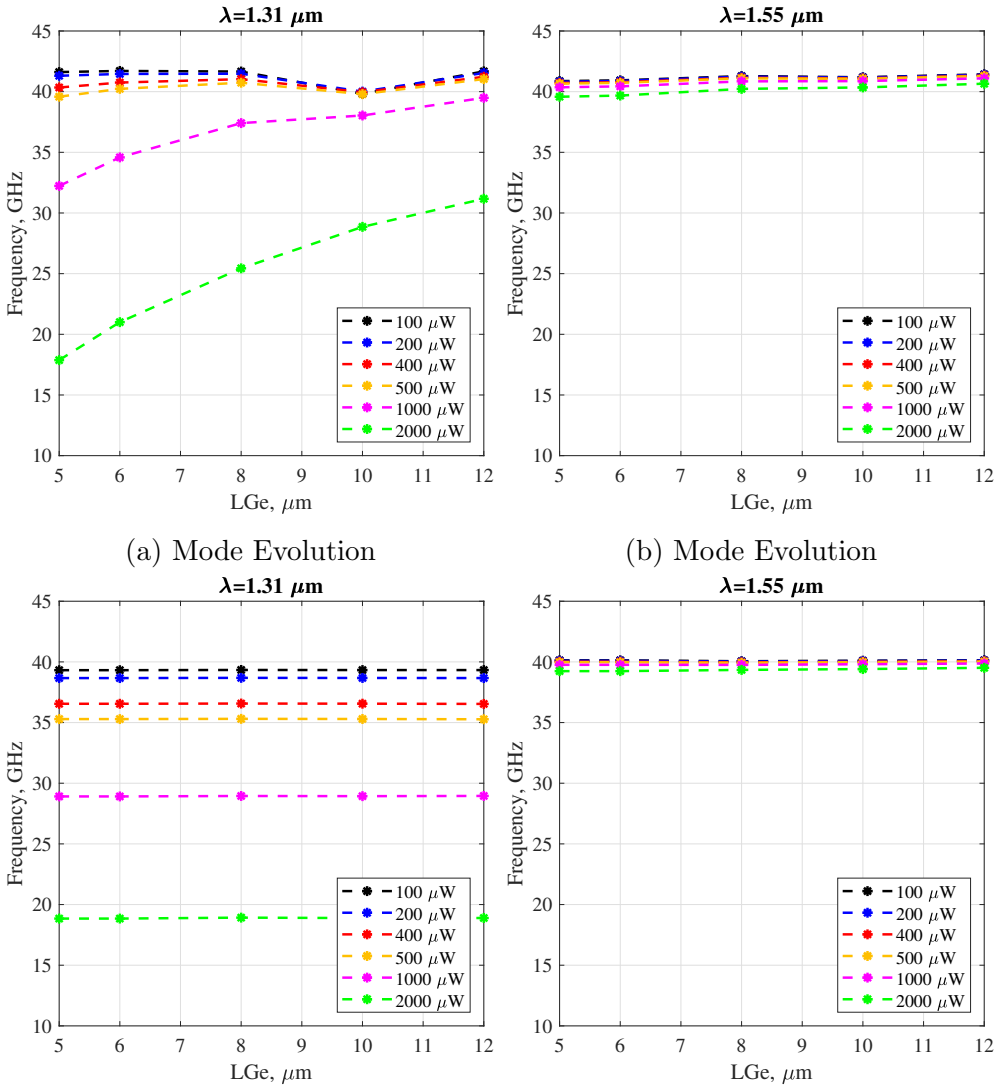
(c) Butt Coupling

(d) Mode Evolution

(e) Butt Coupling

(f) Mode Evolution

Figure 3.15: Responsivity evaluated with the variations of the length, width and height of the devices with respect to the reference case. Both configuration and wavelength are reported



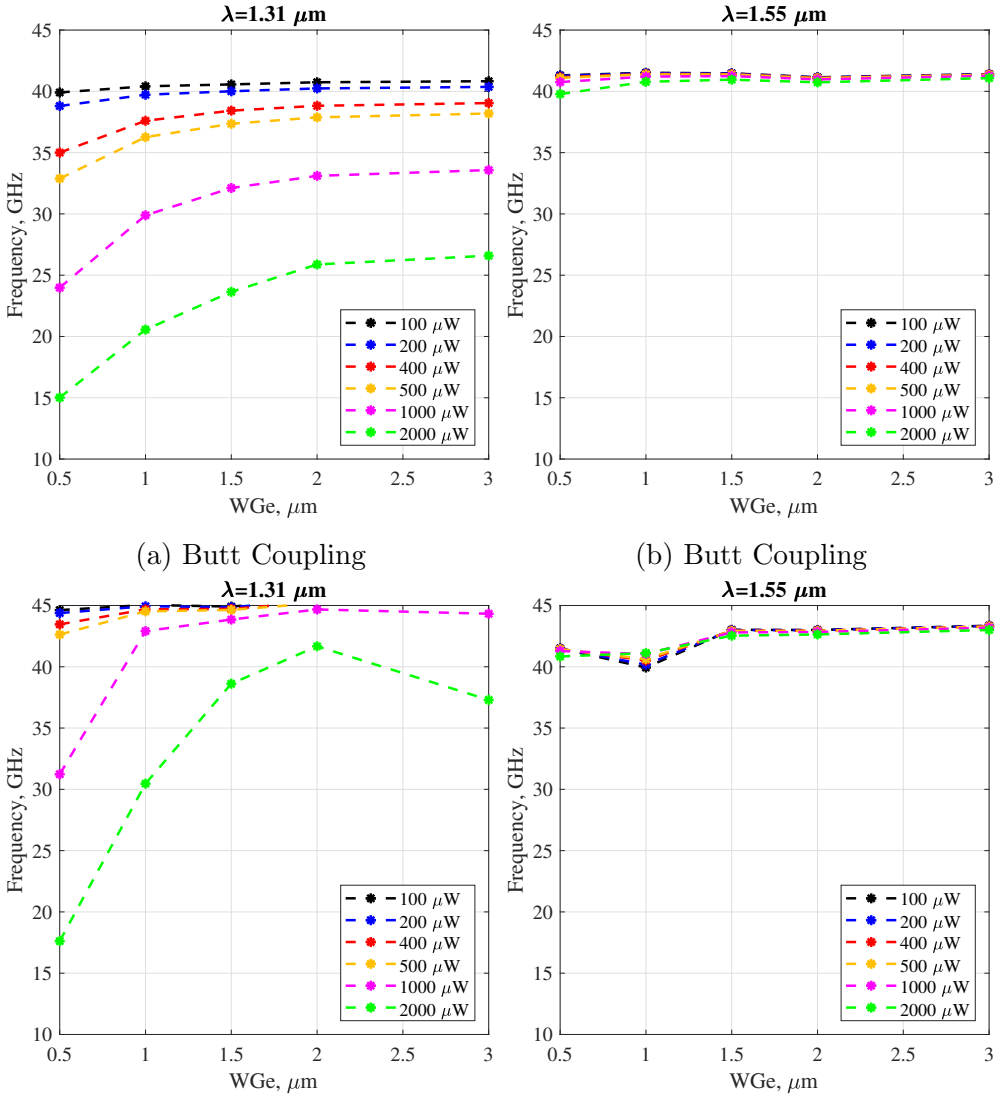
(a) Mode Evolution

(b) Mode Evolution

(c) Butt Coupling

(d) Butt Coupling

Figure 3.16: Cutoff frequency for both wavelength and with the variation of the length of the device



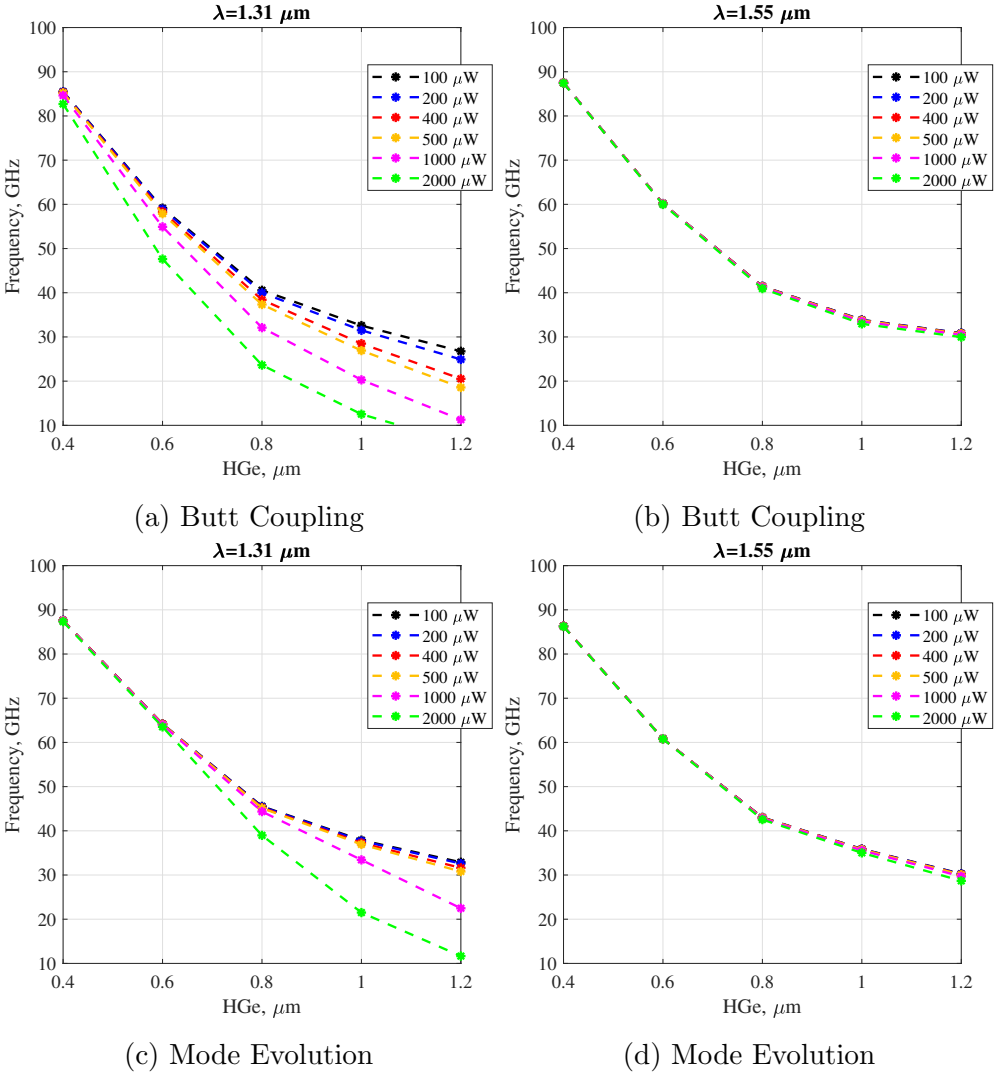
(a) Butt Coupling

(b) Butt Coupling

(c) Mode Evolution

(d) Mode Evolution

Figure 3.17: Cutoff frequency for both wavelength and with the variation of the width of the device



(a) Butt Coupling

(b) Butt Coupling

(c) Mode Evolution

(d) Mode Evolution

Figure 3.18: Cutoff frequency for both wavelength and with the variation of the height of the device

3.5 Conclusions

After the different values that have been tested in the previous section, a possible solution has been implemented. The detector is 6 µm long, 1 µm wide and 0.6 µm height. In Fig. 3.19 the butt coupling configuration is reported.

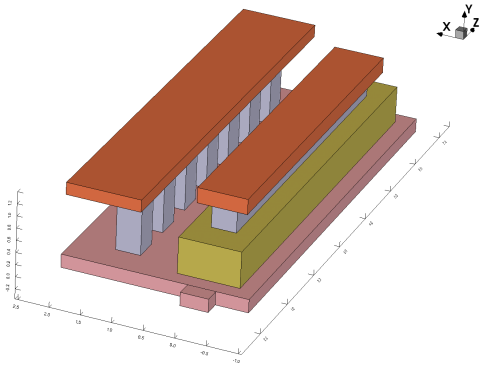
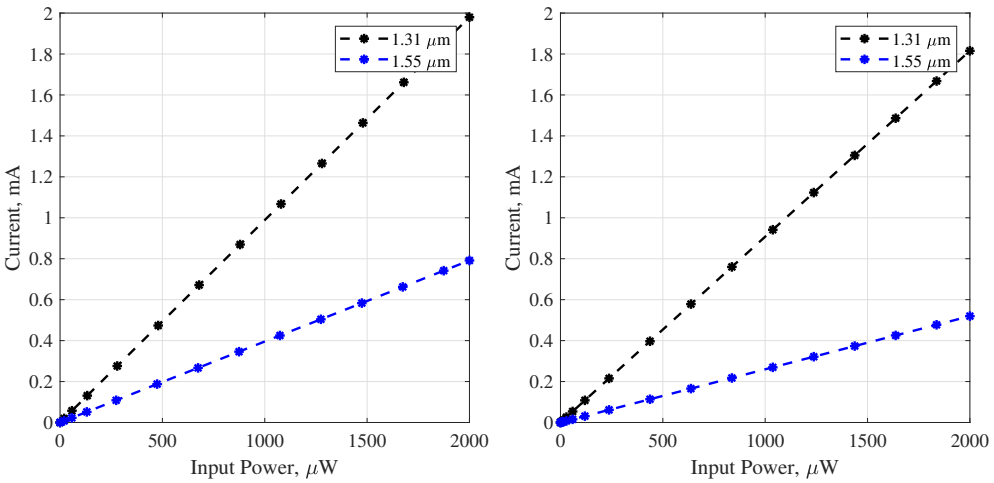


Figure 3.19: Butt coupling configuration with the new dimensions used

In Fig. 3.20 the PI curves of the optimized devices show that the responsivity at 1.31 µm is almost unchanged with respect to the reference case, with a values of around 1 A/W. Instead, a decrease in the 1.55 µm can be observed, leading to a responsivity of around 0.4 A/W in the butt coupling case and a responsivity of around 0.3 A/W in the mode evolution case. This seems a very low responsivity value, but, as we introduced at the beginning of the chapter, the analysis has been performed in a worst case scenario, leading to an underestimation of the responsivity a 1.55 µm, since in a real device strain of the Si on the Ge layer should strongly increase the absorption of the detector.

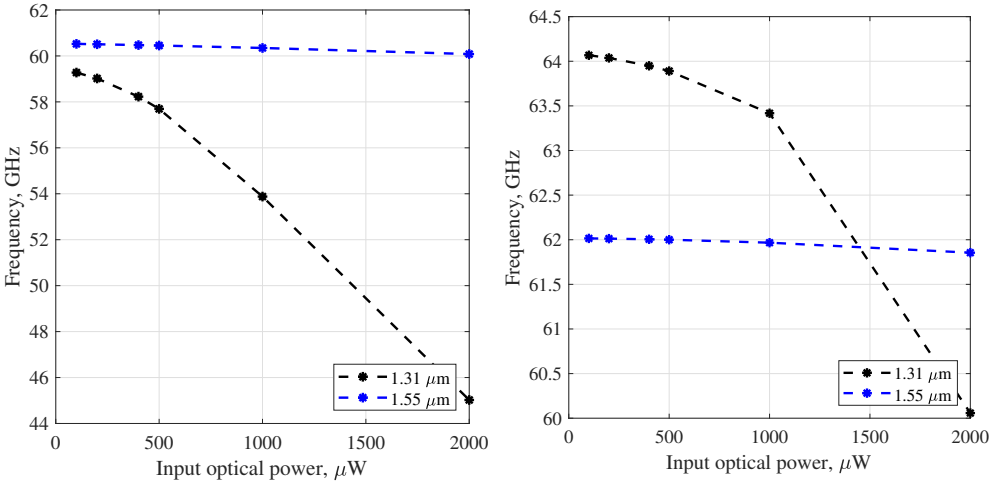
On the other hand, the frequency response improves for both wavelengths, reaching, for low input optical power, to a cutoff frequency higher that 60 GHz. The new device is also less sensitive to the input optical power, and in particular, in the mode evolution case, the cutoff frequency decreases only of few GHz increasing the optical power.

Probably, more sophisticated solutions are needed to further improve the device performance. Not only, the configurations proposed in this work are inspired from the literature, where experimental data are also present, and so these new devices could be fabricated with very similar processes.



(a) Butt Coupling (b) Mode Evolution

Figure 3.20: PI curve for both wavelength and with the new dimensions chosen for the detector



(a) Butt Coupling (b) Mode Evolution

Figure 3.21: Cutoff frequency for both wavelength and with the new dimensions chosen for the detector

Appendix A

Synopsys TCAD Sentaurus: design flow and tool description

The simulation software used is a CAD (computer aided design) software provided by Synopsys.

It is made by different tools,that have different tasks inside the simulation. To perform parametric simulation as it has been done in this thesis, the graphical user interface (GUI) is provided by *Sentaurus Workbench* (SWB), a software that is able to drive parametrically all the tools. The user has to choose the tools and the simulations are performed one after the other, substituting the parameters that the user has chosen. In fig. A.1 a screenshot of the software is reported.

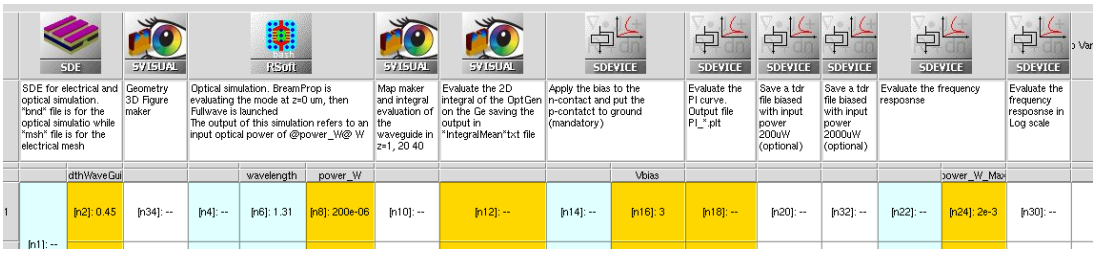


Figure A.1: Screenshot of the tool used in Synopsys TCAD Sentaurus

As it is possible to see, SWB is a big table where parameters can be inserted (in the figure reported the parameters are input optical power and waveguide width, in the thesis project there is also the Ge parameters, such as height, width and length). The simulations are performed from the left to the right and each tool is used to complete a specific task.

Let us now describe each of the tools one after the other. *SDE* (Sentaurus Device Editor) is the tool that takes care of the geometry, the contacts and the

mesh that will be used in the whole simulation. As input, the tool takes a command file written in *Scheme* scripting language (Scheme is its default scripting language), which is used to describe dimensions and position of each of the components of the device that it is considered (as for example the length of the Ge, as well as the dimension of the computational box, ...). Not only, inside the command file there are also present the position of the contacts, i.e. where in the following tools the device is connected to the external circuit (for example a source used to fixing the bias point), the material composition of each component and the doping profile of the device. It is possible also to specify the density of the mesh in certain region of the device. In the photodetector studied, very few points are used for the description of the oxide around the device, while a very high density mesh is used for the detector region, since it is crucial for understanding of the device, to have a precise description of all the involved quantities. The output of this tool are two files with *tdr* extension (an extension that implements HDF5). The first file is essentially a list of all the region of the device, with all the dimension, but without the contacts and the doping profile. This file is essential since it is the input file of the optical solver. The other file is the complete description of the geometry, with the mesh that will be used in the electrical simulation.

The following tool *sVisual* (Sentaurus Visual) is a visualisation tool of the tdr files that are the output files of all the other tools. It is possible to directly export plots and elaborate the quantities inside this tool, or, as it has been done in this thesis, it allows the export of 1D quantities, and a following MATLAB post processing can be used.

The optical solver is called RSoft FullWave and it provides a numerical solution of the Maxwell equations with the FDTD method (see chapter 2). It takes as input the first tdr file generated by the first SDE tool and it converts it with an internal utility to a RSoft project, where it is possible to set the monitored values and region, the boundary condition, the initial field distribution, the material description, ... The source that has been chosen in called *mode* source, where the first step that the solver does is to evaluate the mode of the optical waveguide, and then start to propagate it trough the device. In this way, almost all the light of the source is reaching the detector. The chosen boundary condition are called Perfect Matched Layer (PML), that are numerical boundary conditions that absorb all the light that reaches them, so no reflected light from the boundary is absorbed in the device. As output, it provides multiple files, according to the monitored values chosen during the setup process. In our case,

using the software utility *tdrutils*, the power absorbed in the whole detector is directly converted in a tdr file that can be read and used as generation rate in the electrical simulation.

Finally, *sDevice* (Sentaurus Device) is the tool related to electrical simulation. It is able to perform different kind of simulations, but in this thesis the *quasi-stationary* simulation has been used. After evaluating the equilibrium solution, modifying the boundary condition, it evaluates the new solution until reaching the goal that has been set (more details provided by the manual of the tool [15]). The sDevice tool is also able to introduce a netlist in order to add components connected to the device under study. So a $50\,\Omega$ ideal resistor is connected to the device under study and the frequency response is evaluated with the method that the tool calls *ac-coupled* [15].

Bibliography

[1] Max Born, Emil Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 7 edition, 1999.

[2] M. J. Byrd, E. Timurdogan, Z. Su, C. V. Poulton, N. M. Fahrenkopf, G. Leake, D. D. Coolbaugh, and M. R. Watts. Mode-evolution-based coupler for high saturation power Ge-on-Si photodetectors. *Opt. Lett.*, 42(4):851–854, feb 2017.

[3] M.L. Cohen and J.R. Chelikowsky. *Electronic Structure and Optical Properties of Semiconductors*. Springer Series in Solid-State Sciences. Springer Berlin Heidelberg, 2012.

[4] Simon M. Sze Gary S. May. *Fundamentals of Semiconductor Fabrication*. Wiley, 2003.

[5] Giovanni Ghione. *Dispositivi per la microelettronica*. Cambridge University Press, 2009.

[6] Giovanni Ghione. *High-speed semiconductor devices and circuits*. Cambridge University Press, 2009.

[7] ioffe.ru. Electronic properties of germanium (ge). http://www.ioffe.ru/SVA/NSM/Semicond/Ge/.

[8] ioffe.ru. Electronic properties of silicon (si). http://www.ioffe.ru/SVA/NSM/Semicond/Si/.

[9] Charles Kittel. *Introduction to Solid State Physics*. Wiley, 2004.

[10] P.A. Markowich, C.A. Ringhofer, and C. Schmeiser. *Semiconductor Equations*. Springer Vienna, 2012.

[11] Peter A. Markowich. *The Stationary Semiconductor Device Equations*. Springer Vienna, 1986.

[12] André Moliton. *Interactions of Electromagnetic Waves and Solid Semiconductors*. Springer New York, New York, NY, 2007.

[13] J.J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, 2017.

[14] D.M. Sullivan. *Electromagnetic Simulation Using the FDTD Method*. Wiley, 2013.

[15] Synopsys. *Sentaurus Manual v2017.09*.

[16] S.M. Sze and Kwok K. Ng. *Physics of Semiconductor Devices*. John Wiley & Sons, Inc., apr 2006.

[17] Wikipedia. Electronic band structure — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Electronic%20band%20structure&oldid=960055808, 2020. [Online; accessed 07-July-2020].

[18] Peter J. Winzer, David T. Neilson, and Andrew R. Chraplyvy. Fiber-optic transmission and networking: the previous 20 and the next 20 years. *Opt. Express*, 26(18):24190–24239, Sep 2018.

[19] Kane Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, 1966.

[20] W. Yu. *Electromagnetic Simulation Techniques Based on the FDTD Method*. Wiley Series in Microwave and Optical Engineering. Wiley, 2009.