Politecnico di Torino

Corso di Laurea Magistrale in INGEGNERIA ELETTRONICA
A.A. 2023/2024

Tesi di Laurea Magistrale

Next-generation Ge-on-Si waveguide photodetectors: modelling, design and optimization strategies

Relatori
Dr. Marco E. Vallone
Prof. Michele Goano
Dr. Alberto Tibaldi
Dr. Matteo G. C. Alasio

Candidato
Giuseppe Divincenzo

Abstract

Ge-on-Si waveguide photodetectors achieving remarkable bandwidths exceeding 250 GHz and compatible with mature CMOS technology have been demonstrated recently (S. Lischke et al., *Nat. Photon.* 2021). The present thesis leverages commercial 3D multiphysics simulation tools, specifically Synopsys TCAD Sentaurus and RSoft Photonic Device Tools, to develop a model for the computer-aided design of this novel concept of device. In the perspective of providing guidance to manufacturers, this model has been tailored to predict key performance metrics of the detector, namely modulation bandwidth and responsivity. After validating and calibrating the model versus experimental results in the literature, the performance of the detector has been correlated with its technological parameters, such as geometry and doping, and with various operating conditions, including power and wavelength of the light source and bias. Such physics-based model has also been used to synthesize a small-signal high-frequency analytical model to identify the mechanisms limiting the broadband operation of the device. The rest of the thesis explores strategies for optimizing the photodetector: an inductive peaking-based approach to extend the bandwidth further has been suggested and the design of a tapered waveguide to improve light coupling with the device has been proposed. The thesis is structured as follows: Chapter 1 provides a general introduction to photodetectors, Chapter 2 details the adopted simulation methodologies, Chapter 3 presents the outcome of the conducted analyses.

Summary

Over the last decade, cloud- and internet-based applications, such as video streaming, social networks, and internet search engines, have led to an exponential growth in data center traffic. More recently, the advent of artificial intelligence (AI) and machine learning (ML) has further accelerated this growth. Such a rapid increase in data traffic has driven the demand for high-speed optical transceivers, which are currently transitioning from 100 Gbps to 400 Gbps and beyond.

In this framework, silicon photonics (SiPh) has emerged as a transformative technology, able to overcome the limitations of traditional transceiver technologies in both long-haul and, especially, intra-data center connections. Silicon photonics allows for the seamless integration of optical components with traditional silicon-based electronics while surpassing the performance of traditional III-V semiconductor and photonic integrated circuit (PIC) platforms. The possibility to leverage well-established CMOS-compatible fabrication processes, which enables high-volume production at low cost, the significant increase in integration density, and the low power consumption are the driving forces behind the growing adoption of silicon photonics.

As key components in high-speed optoelectronics, photodetectors achieving bandwidths greater than 100 GHz have been a topic of intense research for several decades. This thesis explores a promising novel Ge-on-Si *pin* waveguide photodetector concept, recently presented by S. Lischke et al.[1]. This innovative design features a narrow, undoped germanium fin, about 100 nm wide, laterally sandwiched between complementary doped silicon regions. It has been demonstrated that this detector is able to achieve unprecedented electro-optic modulation bandwidths of 265 GHz, along with reasonable responsivity, outperforming state-of-the-art III-V detectors.

The aforementioned lateral Ge-on-Si photodetector has been modelled and studied by means of commercially available 3D multiphysics simulation tools from the Synopsys suite. A physics-based model for the computer-aided design

[1] S. Lischke, A. Peczek, J. S. Morgan, K. Sun, D. Steckler, Y. Yamamoto, F. Korndörfer, C. Mai, S. Marschmeyer, M. Fraschke, A. Krüger, A. Beling, and L. Zimmermann, "Ultra-fast germanium photodiode with 3-dB band- width of 265 GHz," *Nature Photon.*, vol. 15, pp. 925–931, Dec. 2021.

of these innovative devices, mainly aimed at predicting key performance metrics such as modulation bandwidth and responsivity, has been developed in the perspective of providing guidance to manufacturers, offering deeper insights into the potential of these detectors. Specifically, the following simulation approach has been adopted:

- RSoft Photonic Device Tools have been employed to handle the electromagnetic problem, i.e., the evanescent coupling and the propagation of light from the source to the detector; in particular, RSoft FullWAVE, a 3D finite-difference time-domain (FDTD) solver of Maxwell's equations, has been used to compute the optical generation rate within the device;
- TCAD Sentaurus has been used to manage the electrical problem; precisely, the charge transport problem has been solved through the Sentaurus Device tool, computing the coupled solution of the Poisson's equation with the semiconductor drift-diffusion transport equations, where the optical generation rate evaluated by RSoft has been incorporated as a source term.

The outcome of this work is here summarized. First, a small-signal high-frequency analytical model has been proposed for the photodetector, based on data extracted from AC simulations run for several geometries of the device in dark. The model has highlighted the impact of design choices on the RC bandwidth limitation, which is particularly relevant for an extremely narrow, fin-like device of this kind. Moving to simulations under illumination, it has been possible to evaluate the main figures of merit of the detector, namely the modulation bandwidth and the responsivity. The impact of variations of significant technological parameters, such as geometry and doping, on these performance metrics has been investigated, along with the behaviour of the device under various operating conditions. In particular, the effect of the applied reverse bias voltage and the high power performance have been considered. An inductive peaking-based approach has also been suggested to further extend the device bandwidth. The aforementioned experiments have been performed in the telecom O-Band ($\lambda = 1.31\,\mu m$), where the responsivity is quite high due to the favourable germanium absorption capability. Simulations in the C-Band ($\lambda = 1.55\,\mu m$) have been conducted as well later on. At last, it has been demonstrated that a properly designed tapered input waveguide improves the coupling of incident light with the device, enhancing the responsivity. Such an optimization turns out to be useful especially in the C-Band, affected by a lower germanium absorption coefficient compared to the O-Band.

The thesis is structured as follows:

1. Chapter 1 starts with a general introduction on photodetectors and their main figures of merit; then, it focuses on the case study of this thesis, i.e., lateral Ge-on-Si waveguide photodetectors;
2. Chapter 2 presents the adopted simulation tools and methods;

3. Chapter 3 reports details on the conducted analyses and the obtained results;

4. in Chapter 4 conclusions are drawn;

5. Appendix A includes, for the sake of completeness, the small-signal circuit parameters extracted from each performed AC simulation;

6. in Appendix B relevant material parameters used in simulations are reported.

Contents

Chapter 1

Introduction

This first Chapter introduces the research context of this work, provides a general overview on photodetectors and, in the end, focuses on the case study of this thesis: innovative ultra-wideband Ge-on-Si photodetectors.

1.1 Research context

In recent years, silicon photonics (SiPh) has attracted great attention because of the increasing demand for faster and more efficient transmission and reception of data, induced by the development of communication technology. Silicon photonics leverages decades of research that have led to robust, high-yield, low-cost silicon processing, integrating on a silicon substrate photonic components of different nature, such as optical waveguides, modulators and photodetectors (see figure 1.1) [1].

Silicon is a transparent material in the near infrared (NIR) optical communication bands, corresponding to the 1.3 - 1.6 µm wavelength range. Due to its 1.12 eV bandgap, silicon is not able to absorb wavelengths greater than 1.1 µm. While this is a desirable property for routing optical signals in a photonic integrated circuit (PIC), the need for long-wavelength detection requires the integration of another suitable material: germanium.

Germanium is a narrowgap material capable of absorbing light up to a 1.55 µm wavelength, corresponding to the C-Band of optical communications. It is an interesting material from an optical standpoint. In fact, it features a 0.66 eV indirect bandgap and a 0.8 eV direct bandgap, which implies that, due to direct optical transitions, absorption is particularly effective in the C-Band and, even more, in the O-Band, centred around 1.31 µm [2, Sec. 4.3] (see figure 1.2).

Progress in technology allows nowadays the selective epitaxial growth of

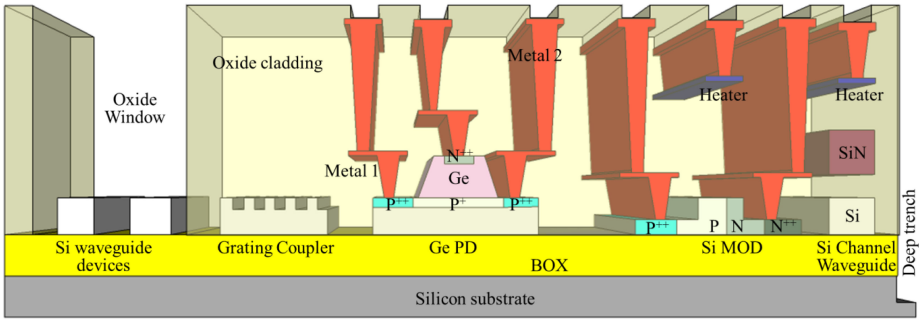


Figure 1.1: Cross-sectional view of Advanced Micro Foundry (AMF) standard Multi-Project Wafer (MPW) flow, an example of silicon photonics platform. Figure taken from [1].

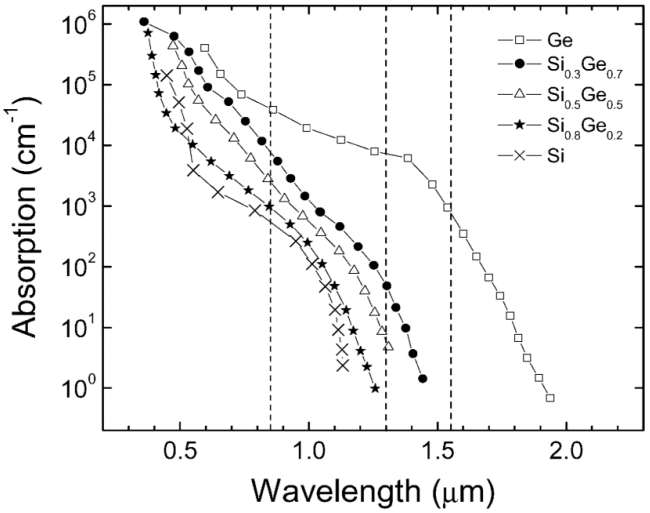


Figure 1.2: Absorption coefficient versus wavelength for SiGe with Ge concentration ranging from 0% (Si) to 100% (Ge). Figure taken from [3].

high-quality germanium layers on silicon, despite the 4.2% lattice mismatch between the two elements [4]. This has paved the way for the fabrication of Ge-on-Si photodetectors, key components in SiPh integrated circuits [5]. An example of germanium photodetector can be observed in figure 1.1, indicated as Ge PD.

Significant research effort has been dedicated to enhancing the performance of Ge-on-Si detectors in terms of bandwidth, i.e., the capability to respond to a quickly modulated optical signal. Enhancing the speed of photodetectors is highly desirable, given the anticipated symbol rates approaching 200 GBaud in modern optical transceivers. As a matter of fact, typical bandwidths for such detectors lie in the 50 - 70 GHz range [6–8]. On the contrary, detectors for InP-based photonics, representing the main competitor against SiPh, exhibit bandwidths larger than 150 GHz, primarily due to the flexibility offered by the use of III-V compound alloys such as InGaAs [9, 10].

However, recent literature [11, 12] has highlighted notable examples of Ge-on-Si photodetectors achieving bandwidths well beyond 100 GHz, with some even surpassing the performance of InP-based counterparts. Leveraging the benefits of SiPh, such as large-scale integration, cost-effective production, and CMOS compatibility, while maintaining high bandwidth, represents a significant advancement in the field. Such a development is poised to effectively meet the growing demand for high-speed optical communications, along with a number of potential applications, from high-speed signal processing to neuromorphic computing.

A novel concept of Ge-on-Si waveguide photodetector represents the case study on which this thesis is focused. Before proceeding with a detailed presentation of such device (section 1.3), a general overview on photodetectors follows.

1.2 Photodetectors: an overview

Photodetectors are semiconductor devices capable of converting optical signals into electrical signals. Their working principle is based on the optical generation of electron-hole pairs through the absorption of photons. Photogenerated carriers are separated by an electric field and collected by the device contacts, resulting in an output photocurrent. The electric field separating the photocarriers is usually induced by an external bias voltage in a reverse-biased junction, e.g. a *pn* or a *pin* junction.

1.2.1 Photodiodes

Photodetectors based on *pn* or *pin* junctions are called photodiodes. Photodiodes operate in reverse bias conditions to intensify the electric field in the depleted region. The overall output current of a photodiode I_{PD} is the super-

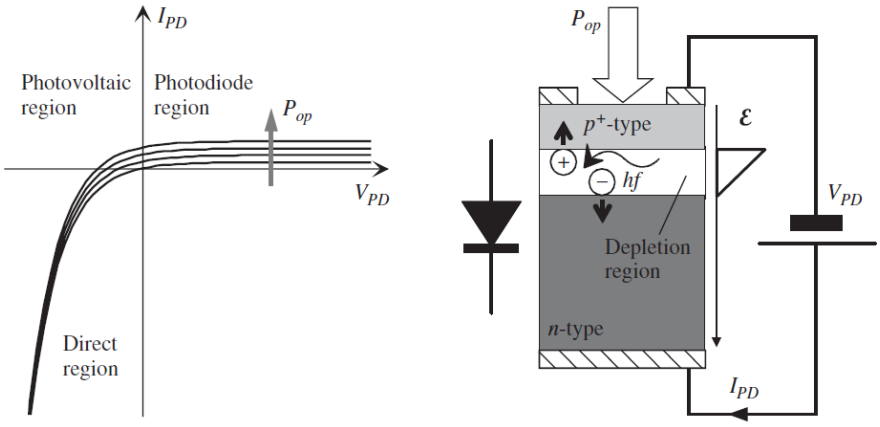


Figure 1.3: I-V characteristics of a photodiode (left) biased by an ideal voltage source and under illumination (right). Figure taken from [2, Sec. 4.6].

position of the dark current I_{dark}, i.e., the reverse saturation current of the diode, and the photocurrent I_{L}, which is the contribution due to photogenerated electron-hole pairs. The dark current is typically negligible, therefore the photocurrent practically represents the total output current of the device. The photocurrent is linearly dependent on the input optical power P_{op}, for low power levels. The input-output relation of a photodiode can be therefore approximated as

$$I_{\text{PD}} = I_{\text{dark}} + I_{\text{L}} \approx \mathcal{R} P_{\text{op}} \tag{1.1}$$

where $\mathcal{R}$ is the *responsivity* of the device, expressed in A/W units, which represents a fundamental figure of merit for a photodetector.

In *pn* photodiodes photocarriers are generated in both the depleted region and the *n*-side and *p*-side diffusion regions. Even though carriers generated in diffusion regions apparently increase the responsivity, the frequency response of the *pn* photodiode to a time-varying input optical power is quite inconvenient due to the low cutoff frequency introduced by the carriers lifetime [2, Sec. 4.7]. In other words, the detector is not fast enough to provide a photocurrent which effectively follows a rapidly modulated optical signal.

The speed of a photodetector is commonly expressed in terms of *modulation bandwidth* (also *electro-optic cutoff frequency*), corresponding to the modulation frequency at which the photocurrent is halved compared to its DC value. A wider depleted region, compared to diffusion lengths, would improve both the responsivity and the frequency response of a *pn* photodiode. However, this would require a too large reverse bias voltage. These are the main reasons leading to the adoption of *pin* photodiodes.

In a *pin* photodiode the generation of electron-hole pairs occurs in a large intrinsic or lightly doped region sandwiched between highly doped complemen-

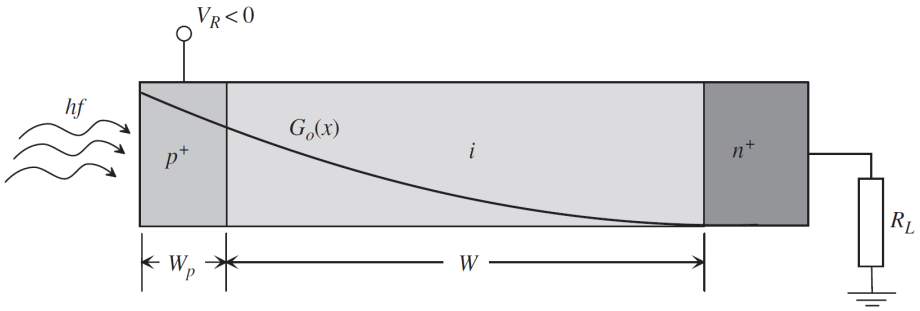


Figure 1.4: Schematic structure of a surface-illuminated *pin* photodiode. Figure taken from [2, Sec. 4.8].

tary layers. The electric field is almost uniform within the depleted region, which can be made large enough to enhance the absorption of photons. In addition, exploiting heterojunctions, i.e., using widegap materials for the n-doped and p-doped layers, absorption outside the depleted region can be suppressed, enhancing the bandwidth.

1.2.2 *pin* photodetectors

Many high-performance *pin* photodetectors exploit surface-illuminated structures. These typically exhibit a planar layout, like the one schematized in figure 1.4: the p- or n-side of the detector is directly illuminated and the incident light is progressively absorbed by the intrinsic region. Such devices are also referred to as *normal-incidence photodetectors*.

Responsivity and quantum efficiency

A simple model for evaluating the responsivty of *pin* photodetectors can be outlined. In a simple one-dimensional approximation along the x direction, assuming a device area A, a depleted region width W, and neglecting the width of doped layers, the detector photocurrent can be expressed as

$$I_{\mathrm{L}} = qA \int_0^W G_{\mathrm{opt}}(x)\mathrm{d}x \tag{1.2}$$

where q is the elementary charge and G_{opt} is the optical generation rate, i.e., the number of photogenerated electron-hole pairs per unit time and unit volume. Assuming an incident optical power P_{op} and a material absorption coefficient α, the power within the device exponentially decreases as

$$P(x) = (1 - R)P_{\mathrm{op}}\mathrm{e}^{-\alpha x} \tag{1.3}$$

where the term $(1 - R)$ accounts for the surface reflectivity. Introducing the internal quantum efficiency η_Q, i.e., the ratio between the number of photo-generated electron-hole pairs and the number of photons reaching the absorber (typically unitary), the absorbed power density, measured in $\mathrm{W/m^3}$ units, can be expressed as

$$u_{\mathrm{abs}}(x) = \eta_Q \frac{\alpha P(x)}{A} = \eta_Q (1 - R) \frac{\alpha P_{\mathrm{op}}}{A} \mathrm{e}^{-\alpha x} \tag{1.4}$$

Therefore, introducing the photon energy $E_{\mathrm{ph}} = hc/\lambda$, where h is the Planck constant, c is the speed of light and λ is the optical wavelength, the optical generation rate can be expressed as:

$$G_{\mathrm{opt}}(x) = \frac{u_{\mathrm{abs}}(x)}{E_{\mathrm{ph}}} = \frac{\lambda}{hc} u_{\mathrm{abs}}(x) \tag{1.5}$$

Plugging (1.4) into (1.5), and (1.5) into (1.2), the responsivity can be evaluated as

$$\mathcal{R} = \frac{I_{\mathrm{L}}}{P_{\mathrm{op}}} = q\eta_Q (1 - R) \frac{\lambda}{hc} \left(1 - \mathrm{e}^{-\alpha W}\right) \tag{1.6}$$

while the external quantum efficiency η_x, i.e., the ratio between the number of electron-hole pairs collected by contacts and the number of incident photons, can be evaluated as

$$\eta_x = \frac{I_{\mathrm{L}}/q}{P_{\mathrm{op}}/E_{\mathrm{ph}}} = \frac{hc}{q\lambda} \mathcal{R} = \eta_Q (1 - R) \left(1 - \mathrm{e}^{-\alpha W}\right) \tag{1.7}$$

Therefore, the width W of the active region of a *pin* detector should be at least greater than the absorption length $1/\alpha$ to allow for a significant efficiency.

Electro-optic frequency response

The thickness of the absorber does not only play a relevant role in the resulting efficiency of a *pin* photodetector. As a matter of fact, the frequency response of the device is strongly influenced by sizes. In particular, two main mechanisms limit the response of these detectors to a modulated optical signal:

1. the effect of the total diode capacitance, an *extrinsic* load-related cutoff mechanism, which determines the RC limit of the modulation bandwidth;
2. the transit time of carriers drifting across the depleted layer, an *intrinsic* cutoff mechanism, which determines the transit time limit of the modulation bandwidth.

In general, a photodetector can be modelled by the small-signal equivalent circuit reported in figure 1.5: I_{L} represents the short-circuit photocurrent, modelled as a current source, Y^i_{PD} is the detector intrinsic admittance, Y^x_{PD} is the

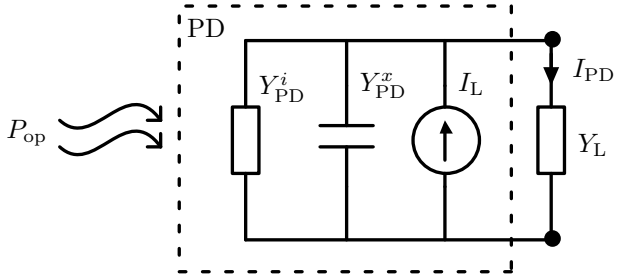


Figure 1.5: Small-signal equivalent circuit of a loaded photodetector.

parasitic (usually capacitive) admittance and Y_{L} is the load admittance (e.g., the input of a front-end amplifier).

The photocurrent I_{L} is proportional to the optical power P_{op} according to the *complex responsivity* $\mathfrak{R}(\omega)$:

$$I_{\mathrm{L}}(\omega) = \mathfrak{R}(\omega) P_{\mathrm{op}}(\omega) \tag{1.8}$$

where $\omega = 2\pi f$ represents the modulation frequency of the optical signal. The complex responsivity describes the *intrinsic* small-signal frequency response of the detector, i.e., it accounts for intrinsic mechanisms limiting the speed of the detector, such as the carriers transit time. It typically exhibits a low-pass behaviour versus the modulation frequency.

An analytical model for the complex responsivity frequency behaviour, accounting for the carrier transit time, has been proposed by [2, Sec. 4.9]:

$$\begin{aligned}\mathfrak{R}(\omega) = \mathfrak{R}(0) \Bigg[& \frac{\alpha W}{\alpha W - \mathrm{j}\omega\tau_{\mathrm{dr},h}} \left(\frac{1}{\alpha W} + \frac{1 - \mathrm{e}^{\mathrm{j}\omega\tau_{\mathrm{dr},h}}}{\mathrm{j}\omega\tau_{\mathrm{dr},h}} \frac{1}{\mathrm{e}^{\alpha W} - 1} \right) - \\ & - \frac{\alpha W}{\alpha W + \mathrm{j}\omega\tau_{\mathrm{dr},h}} \left(\frac{1}{\alpha W} + \frac{1 - \mathrm{e}^{\mathrm{j}\omega\tau_{\mathrm{dr},n}}}{\mathrm{j}\omega\tau_{\mathrm{dr},n}} \frac{\mathrm{e}^{\alpha W}}{\mathrm{e}^{\alpha W} - 1} \right) \Bigg]\end{aligned} \tag{1.9}$$

where α and W are the intrinsic layer absorption coefficient and width, respectively, while $\tau_{\mathrm{dr},n}$ and $\tau_{\mathrm{dr},h}$ represent the electron and hole transit times, defined as

$$\tau_{\mathrm{dr}} = \frac{W}{v_{\mathrm{sat}}} \tag{1.10}$$

where v_{sat} is the carrier saturation velocity, reached if a high enough electric field is assumed. It can be demonstrated that the $-3\,\mathrm{dB}$ cutoff frequency of the transfer function (1.9) is inversely proportional to the transit time and, therefore, to the absorber width. In particular, assuming v_{sat} is the same for electrons and holes, the following approximations hold for a thick ($\alpha W \gg 1$)

and a thin diode ($\alpha W \ll 1$), respectively:

$$f_{-3\mathrm{dB,tr}} \approx \frac{1}{2.2\tau_{\mathrm{dr}}}, \text{ if } \alpha W \gg 1 \tag{1.11a}$$

$$f_{-3\mathrm{dB,tr}} \approx \frac{3.5}{2\pi\tau_{\mathrm{dr}}}, \text{ if } \alpha W \ll 1 \tag{1.11b}$$

Clearly, a wider intrinsic layer slows the photodiode response down, as photocarriers require a longer time to cross the depleted region and be collected by contacts.

In general, taking into account both the frequency-dependent complex responsivity $\mathfrak{R}(\omega)$ and the small-signal equivalent circuit presented earlier, the transfer function linking the output current I_{PD} to the input optical power P_{op} can be written as

$$H(\omega) = \frac{I_{\mathrm{PD}}(\omega)}{P_{\mathrm{op}}(\omega)} = \frac{\mathfrak{R}(\omega) Y_{\mathrm{L}}(\omega)}{Y^{i}_{\mathrm{PD}}(\omega) + Y^{x}_{\mathrm{PD}}(\omega) + Y_{\mathrm{L}}(\omega)} \tag{1.12}$$

The expression above may also be factorized as

$$H(\omega) = \frac{I_{\mathrm{PD}}(\omega)}{P_{\mathrm{op}}(\omega)} = \frac{I_{\mathrm{L}}(\omega)}{P_{\mathrm{op}}(\omega)} \cdot \frac{I_{\mathrm{PD}}(\omega)}{I_{\mathrm{L}}(\omega)} = \mathfrak{R}(\omega) \cdot H_{RC}(\omega) \tag{1.13}$$

from which it is evident how both *intrinsic* (related to the transit time) and *extrinsic* (related to the circuit) mechanisms contribute to the overall electro-optic frequency response.

A simplified small-signal equivalent circuit for a *pin* photodiode is reported in figure 1.6: C_{j} represents the intrinsic diode capacitance, R_{D} is the parallel diode resistance, C_{p} is the external diode parasitic capacitance, R_{s} is the series parasitic diode resistance and R_{L} is the resistive load. Being $R_{\mathrm{D}} \gg R_{\mathrm{s}}, R_{\mathrm{L}}$, the transfer function H_{RC} defined in (1.12) and (1.13) can be written for the simplified circuit as

$$H_{RC}(\omega) = \frac{I_{\mathrm{PD}}(\omega)}{I_{\mathrm{L}}(\omega)} = \frac{1}{1 + \mathrm{j}\omega(R_{\mathrm{s}} + R_{\mathrm{L}})(C_{\mathrm{j}} + C_{\mathrm{p}})} = \frac{1}{1 + \mathrm{j}\omega RC} \tag{1.14}$$

resulting in a low-pass frequency behaviour with a cutoff frequency

$$f_{-3\mathrm{dB},RC} = \frac{1}{2\pi RC} \tag{1.15}$$

In practice, the broadband operation of a photodiode is also limited by its own capacitance, including both intrinsic and parasitic contributions, and by the load resistance, typically dominating on series parasitics. The intrinsic diode capacitance can be expressed, in a first approximation, as a parallel-plate capacitance:

$$C_{\mathrm{j}} \approx \frac{\varepsilon A}{W} \tag{1.16}$$

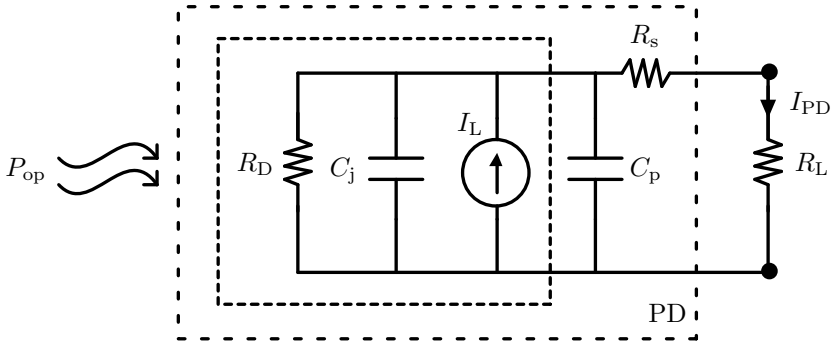


Figure 1.6: Simplified small-signal equivalent circuit of a loaded *pin* photodetector.

with ε dielectric permittivity of the absorber. Therefore, the RC cutoff frequency is approximately proportional to the absorber width W and inversely proportional to the device area A.

The resulting photodetector modulation bandwidth, i.e., the cutoff frequency of the transfer function $H(\omega)$ (1.12), can be evaluated as the combination of the cutoff frequencies of both $\mathfrak{R}(\omega)$ (1.9) and $H_{RC}(\omega)$ (1.14). Let $\mathfrak{R}(\omega)$ be approximated as a low-pass transfer function with a cutoff frequency $\omega_{-3\mathrm{dB,tr}}$:

$$\mathfrak{R}(\omega) \approx \frac{1}{1+\mathrm{j}\dfrac{\omega}{\omega_{-3\mathrm{dB,tr}}}} \tag{1.17}$$

Exploiting the definition of $-3\,\mathrm{dB}$ cutoff frequency

$$|H(\omega_{-3\mathrm{dB}})|^2 = \frac{1}{2} \tag{1.18}$$

and substituting the expressions from equations (1.13), (1.17) and (1.14)

$$|\mathfrak{R}(\omega_{-3\mathrm{dB}})|^2 \cdot |H_{RC}(\omega_{-3\mathrm{dB}})|^2 = \frac{1}{1+\left(\dfrac{\omega_{-3\mathrm{dB}}}{\omega_{-3\mathrm{dB,tr}}}\right)^2} \cdot \frac{1}{1+\left(\dfrac{\omega_{-3\mathrm{dB}}}{\omega_{-3\mathrm{dB},RC}}\right)^2} = \frac{1}{2} \tag{1.19}$$

the equation above can be solved for $\omega_{-3\mathrm{dB}}$, neglecting the higher order term, as

$$\omega_{-3\mathrm{dB}} \approx \frac{1}{\sqrt{\omega_{-3\mathrm{dB,tr}}^{-2} + \omega_{-3\mathrm{dB},RC}^{-2}}} \tag{1.20}$$

implying that

$$\omega_{-3\mathrm{dB}} < \min(\omega_{-3\mathrm{dB,tr}}, \omega_{-3\mathrm{dB},RC}) \tag{1.21}$$

Bandwidth-efficiency trade-off

Given the models for the efficiency and the bandwidth, it has been observed that

$$\eta_x \propto (1 - \mathrm{e}^{-\alpha W}) \tag{1.22a}$$

$$f_{-3\mathrm{dB,tr}} \propto \frac{1}{W} \tag{1.22b}$$

$$f_{-3\mathrm{dB},RC} \propto \frac{W}{A} \tag{1.22c}$$

$$f_{-3\mathrm{dB}} < \min(f_{-3\mathrm{dB,tr}}, f_{-3\mathrm{dB},RC}) \tag{1.22d}$$

Curves reported in figure 1.7 efficiently describe this whole framework quantitatively. It is possible to conclude that:

- the wider is the absorber, the higher is the achievable detector efficiency;
- the bandwidth of detectors having narrow absorbers is mainly RC-limited, and the maximum achievable bandwidth increases as the device area is reduced;
- wider intrinsic regions alleviate the RC limit but lead to a severe transit time limitation of the bandwidth.

Excellent high-frequency operation can be achieved only by small area photodiodes featuring thin intrinsic layers, at the expense of reponsivity. On the other hand, to maximize the responsivity a wide absorber is needed, at the expense of the transit time-limited bandwidth. In other words, a clear trade-off exists between bandwidth and efficiency.

The bandwidth-efficiency trade-off is ultimately due to the parallel directions along which both photons and photogenerated carriers travel within the detector. If these two directions are decoupled, the trade-off can be overcome: this is basically what happens in *waveguide photodetectors*.

1.2.3 *pin* waveguide photodetectors

In waveguide photodetectors light propagates along an intrinsic waveguide made of a narrowgap material, where optical generation occurs. The intrinsic waveguide is sandwiched between complementary doped regions which collect photogenerated carriers. An example of waveguide photodetector structure is reported in figure 1.8: thickness, width and length of the waveguide are indicated as d, a and W, respectively.

In waveguide photodetectors photons propagate along a direction orthogonal to the one along which photocarriers drift: this represents the key to remove the bandwidth-efficiency trade-off. As a matter of fact, the waveguide can be

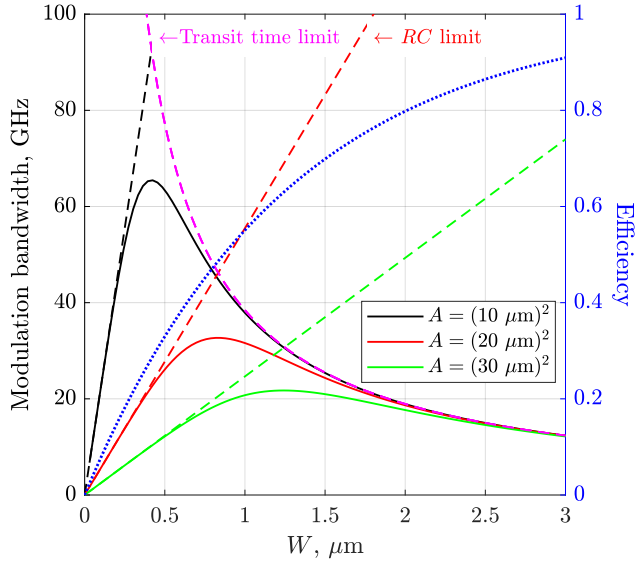


Figure 1.7: Bandwidth-efficiency trade-off for normal-incidence photodetectors. Solid curves describe the bandwidth behaviour against the absorber thickness W, for three different values of the device area A. The magenta dashed line marks the transit time limit of bandwidth, while the remaining dashed lines (in the same colors as the solid curves) indicate the area-dependent RC limit. The efficiency behaviour versus W is described by the blue dotted line.

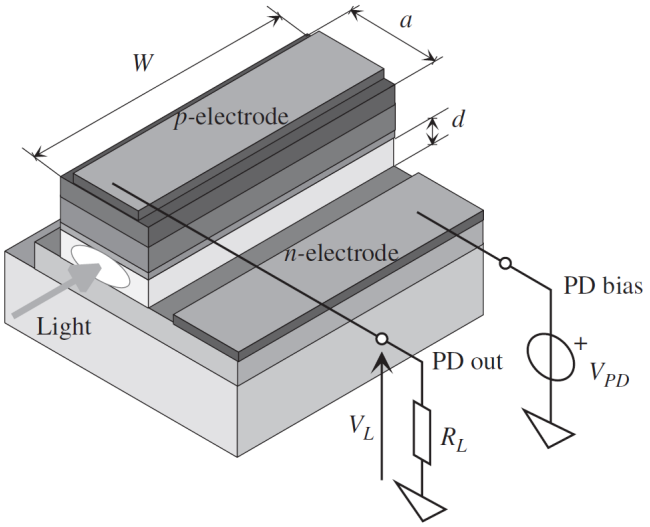


Figure 1.8: Schematic structure of a waveguide photodiode. Figure taken from [2, Sec. 4.10].

made long enough to maximize the efficiency without affecting the bandwidth. Specifically:

- the external quantum efficiency can be expressed as

$$\eta_x = \eta_Q(1-R)\left(1-\mathrm{e}^{-\Gamma_{\mathrm{ov}}\alpha W}\right) \tag{1.23}$$

  where the overlap integral $\Gamma_{\mathrm{ov}} < 1$ is introduced to account for the electromagnetic field within the absorbing part of the optical waveguide only; a device length $W \gg 1/\alpha$ allows for a virtually total absorption of the incident light;

- the transit time-related cutoff frequency is inversely proportional to d, i.e., the distance that photocarriers have to cross before being collected by contacts:

$$f_{-3\mathrm{dB,tr}} \propto \frac{1}{d} \tag{1.24}$$

- the RC-limited bandwidth, inversely proportional to the photodiode capacitance, increases as d increases, while it is reduced by a large detector area aW:

$$f_{-3\mathrm{dB},RC} \propto \frac{d}{aW} \tag{1.25}$$

Therefore, it is evident from the above relations that:

- a long absorber enhances the efficiency, while the resulting increased capacitance affects the RC limit of the bandwidth; on the other hand, the length W does not affect the transit time limit, which is the essential advantage of waveguide photodetectors, compared to normal-incidence ones;
- the bandwidth depends on the absorber thickness, as observed for normal-incidence detectors; the RC limit prevails for low d values, while high d values let the transit time limit dominate;
- at a fixed detector length W, large enough to improve the efficiency, a small waveguide width a helps to alleviate the RC bandwidth limit.

1.2.4 Ge-on-Si *pin* waveguide photodetectors

The focus will now be on Ge-on-Si *pin* waveguide photodetectors. Germanium has become the material of choice for detectors in the SiPh platform mainly because of its large absorption coefficient in the NIR wavelength range and its compatibility with CMOS processes.

The epitaxial growth of high-quality and thick crystalline germanium layers on silicon has been successfully demonstrated, despite the 4.2% lattice mismatch between the two materials. Specifically, a so-called LT/HT (Low Temperature/High Temperature) process is commonly adopted for growing germanium

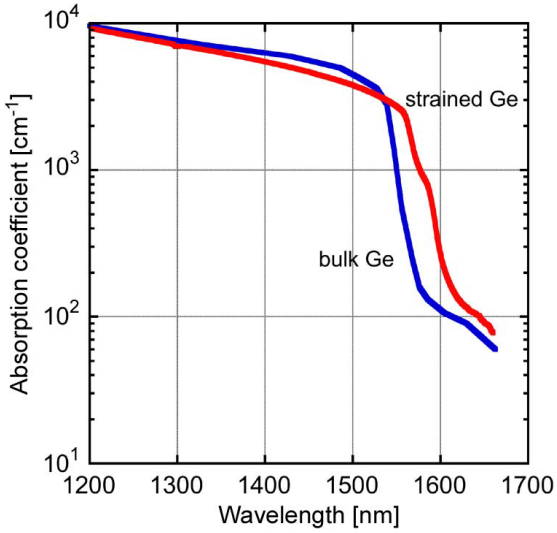


Figure 1.9: Optical absorption spectra for bulk germanium and tensile strained Ge-on-Si. Figure taken from [13].

on silicon substrates. A thin (30-60 nm) germanium buffer layer is grown first at low temperature (350°C), promoting the insertion of dislocations as a mechanism for strain relaxation. A high-temperature deposition (> 600°C) is then performed, followed by a cyclic thermal annealing (> 750°C), aimed at reducing the threading dislocation density and allowing for lower dark currents. After cooling, due to its higher thermal expansion coefficient, the germanium layer undergoes a tensile strain, despite having a lattice constant larger than silicon: the strain shifts the germanium absorption cutoff towards longer wavelengths (see figure 1.9) [13, 14].

Two main strategies have been developed to couple light into a Ge-on-Si photodetector: *evanescent coupling* and *butt coupling*, shown in figures 1.10b and 1.10c, respectively. Both schemes have demonstrated similar performances in terms of speed and responsivity. However, some differences may be underlined.

Butt coupling requires mode-matching conditions for waveguide and detector modes. If these are met, the optical power in the waveguide can be directly transferred into the detector, which can also remain relatively short. This configuration offers more robust coupling efficiency in terms of wavelength sensitivity and germanium thickness variations. Furthermore, this type of approach supports potentially flat surfaces, promoting hybrid integration.

Evanescent coupling relies on a gradual transfer of optical power from silicon to germanium. In fact, light can be easily coupled in an evanescent way from a lower-index material (silicon) to a higher-index material (germanium), as long as the index difference is small. This approach is often preferred because of the

easier integration into a CMOS process. Evanescent coupling typically requires longer germanium absorbers compared to butt coupling; however, this is typically not an issue, as limitations on the detector size are not so severe. This configuration offers better detection linearity with input optical power compared to butt coupling, as light is continuously distributed along the germanium layer, preventing serious absorption screening effects [5, 15].

Several kinds of electrical configuration for Ge-on-Si photodetectors have emerged. Three of them are schematically represented in figures 1.10d, 1.10e, 1.10f:

- figure 1.10d represents a *lateral homojunction* detector, with both n- and p-doped contacts directly realized within the germanium layer;
- figure 1.10e shows a *lateral double heterojunction* detector, with both doped regions formed in silicon;
- in figure 1.10f a *vertical single heterojunction* detector is depicted, where silicon includes a p-doped region while germanium is partially n-doped; in figure 1.11 an example of fabricated device of this kind is shown as well.

Both the lateral homojunction-based and the vertical *pin* detectors require heavily-doped germanium regions, as well as the direct contact of germanium with metal. However, being germanium processing far less mature compared to silicon, a photodetector that does not require doping or metallization of germanium is highly desirable [18]. Furthermore, doped germanium regions imply that photogeneration occurs outside the intrinsic region as well, with a negative impact on the detector frequency response. On the other hand, a lateral double heterojunction-based device, despite involving a challenging selective germanium epitaxial growth in a narrow silicon trench, exhibits a series of advantages: suppressed absorption outside the intrinsic region, due to the wider silicon bandgap; a reduction of technological steps, yielding lower fabrication cost and complexity; a decrease of the contact access resistance thanks to the use of silicides; an improvement of integration, being silicide-based contacts used for other optoelectronic devices as well, e.g., modulators [15].

1.3 Case study: ultra-wideband lateral Ge-on-Si waveguide photodetectors

A novel design of lateral heterojunction-based Ge-on-Si waveguide photodetector is exactly the case study on which this thesis is focused. Such innovative concept of CMOS-compatible germanium photodetector has been initially presented by S. Lischke et al. in 2020 [11], while further advancements have been published in 2021 [12]. A narrow fin-like germanium intrinsic region has been sandwiched between in-situ doped silicon regions, realising a lateral *pin* diode

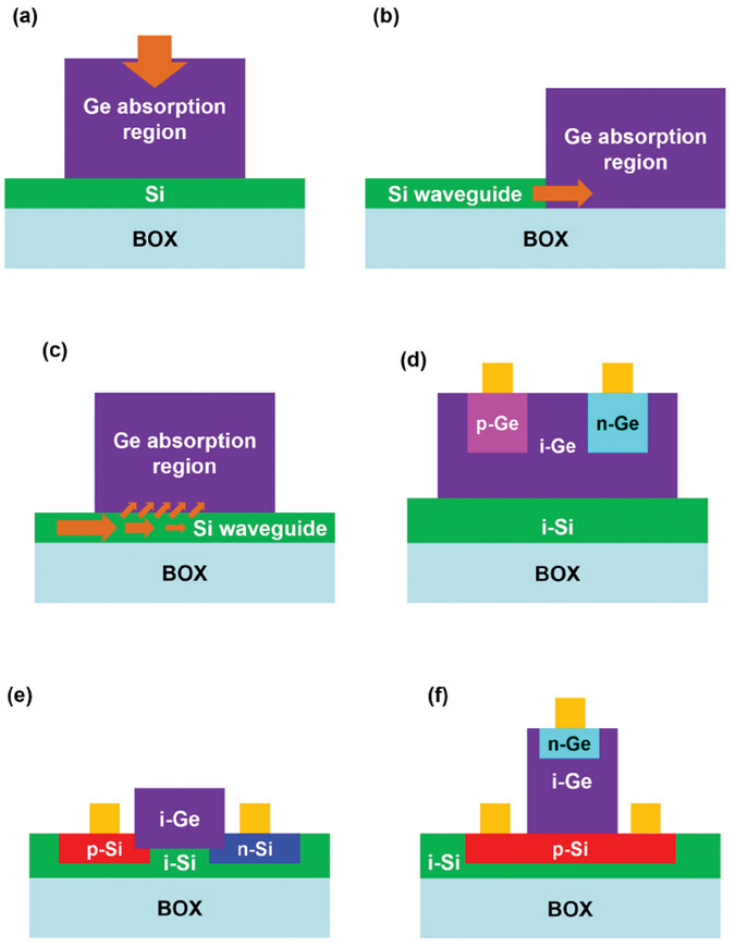


Figure 1.10: Normal incidence (a), butt coupling (b), and evanescent coupling (c) configurations of Ge-on-Si photodetectors. Ge-on-Si *pin* photodetectors with lateral homojunction (d), lateral heterojunction (e), vertical heterojunction (f). Figure taken from [16].

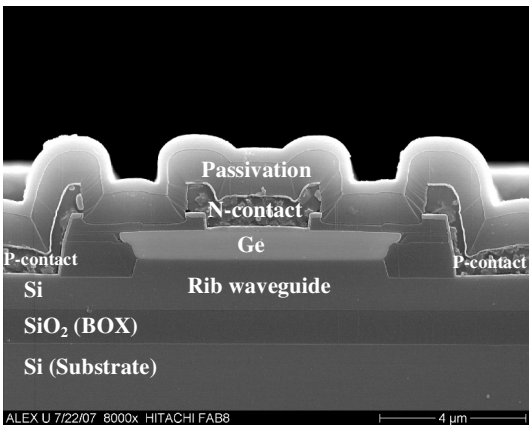


Figure 1.11: Cross-section SEM image of a vertical heterojunction-based Ge-on-Si waveguide photodetector. Figure taken from [17].

which has demonstrated extraordinary bandwidths, exceeding 265 GHz, outperforming state-of-the-art III-V detectors.

The procedure for the fabrication of these devices is schematized in figure 1.12. A few micrometer wide germanium layer, about 400 nm thick, is initially grown selectively on an SOI waveguide. The germanium layer thickness should be large enough to ensure a good crystalline quality, which is crucial for limiting the dark current due to defects. Germanium is protected from abrasive chemicals by a thin silicon layer. After deposition of a silicon dioxide hard-mask, one side of the germanium body is patterned by anisotropic dry etching and a trench is formed (figure 1.12a). An in-situ p-doped silicon layer is deposited by non-selective epitaxy, followed by an overfill with oxide (figure 1.12b). Planarization by chemical mechanical polishing (CMP) removes both oxide and p-doped silicon outside the trench (1.12c). Anisotropic dry etch of hard-mask oxide and germanium on the opposite side, followed by epitaxy of n-doped silicon and planarization, complete the fabrication process (1.12d). It has been demonstrated that this kind of photodetector can be fabricated with a high yield using this clever process.

A key advantage of the proposed fabrication procedure is the use of in-situ doping in place of ion implantation: the realized germanium fin truly remains undoped, which is beneficial for the device speed, as minority carriers diffusion currents are suppressed. High accuracy in the relative position of lithography masks and germanium dry etchings have allowed to shrink the nominal width of the germanium fin down to less than 100 nm, leading to a very intense electric field which minimized the photocarriers transit time, enhancing the device bandwidth.

Cross-sections of a fabricated device are shown by figures 1.13. It can be

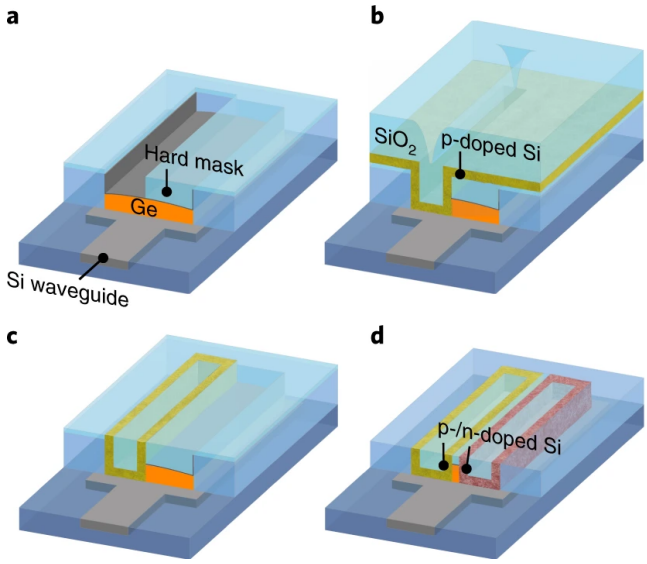


Figure 1.12: Process flow for the lateral fin-like Ge-on-Si *pin* photodetector design. Figure taken from [12].

observed that the germanium fin actually exhibits a biconcave shape. For the reported device, the narrowest point is only 60 nm wide. The epitaxial growth of such a thin germanium layer could hardly ensure a quality comparable to that of the fin here considered, which is obtained from the growth of a thick layer subsequently patterned.

The fabricated devices presented in [11,12] have been characterized in terms of modulation bandwidth and responsivity. Measured values of these two figures of merit are reported in table 1.1. The characterized devices present different widths of the germanium absorber. As it can be noticed, outstanding bandwidths of 240 GHz and 265 GHz have been measured for the narrowest devices, featuring 150 nm and 100 nm width, respectively. The experimental frequency response for these two devices is reported in figures 1.14. Such high cutoff frequencies are mainly enabled by the thin germanium fin, which yields a very short photocarriers transit time. It is evident how the bandwidth increases as the germanium width is decreased. On the other hand, the opposite trend is observed for the responsivity, decreasing with the fin width. Only a 0.3 A/W responsivity is achieved by the 100 nm wide device: though an improvement would be highly desirable, this is still an acceptable performance.

In conclusion, these innovative ultra-wideband Ge-on-Si photodetectors truly are promising devices for the next generation of SiPh-based receivers, which are required to comply with increasingly higher symbol rates of 400 GBaud or even beyond: ultra-fast photodetectors like these naturally enable a development in this direction.

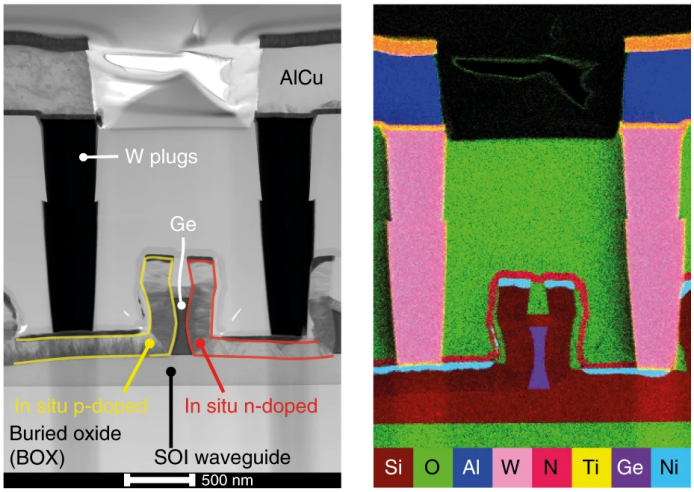


Figure 1.13: Cross-sectional images (left obtained by STEM, right by EDX) of a lateral fin-like Ge-on-Si *pin* photodetector with a 100 nm nominal width. Figure taken from [12].

Reference	[11]	[11]	[12]	[12]
Detector length, μm	20	20	10	10
Germanium nominal width, nm	350	300	150	100
Responsivity, A/W	0.74	0.64	0.45	0.30
Modulation bandwidth, GHz	> 100	> 110	240	265

Table 1.1: Responsivity and modulation bandwidth of some lateral fin-like Ge-on-Si *pin* photodetectors, as reported in [11, 12], measured at a -2 V bias, an optical wavelength $\lambda = 1.55\,\mu$m and a 1 mA photocurrent.

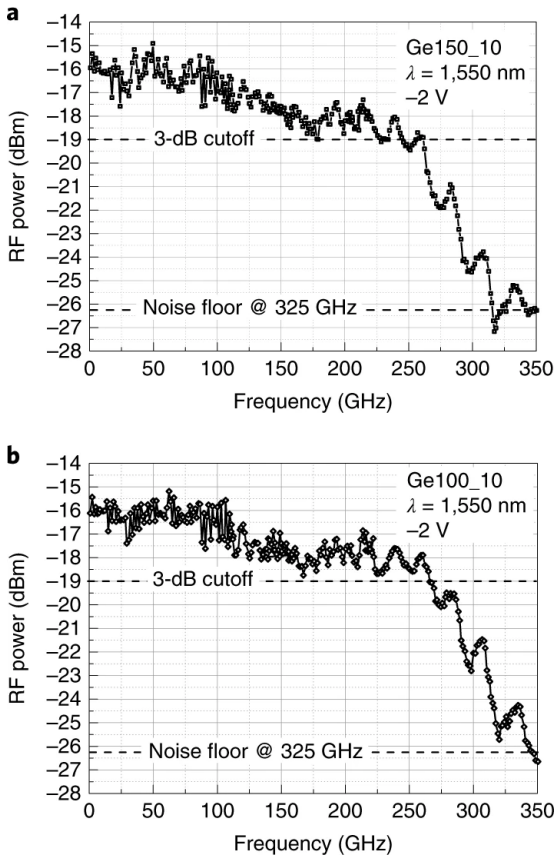


Figure 1.14: Small-signal electro-optic frequency response of lateral fin-like Ge-on-Si *pin* photodetectors with 150 nm (a) and 100 nm (b) nominal fin width. Measurement performed at a −2 V bias, $\lambda = 1.55\,\mu$m and a 1 mA photocurrent. Figure taken from [12].

Chapter 2

Methodologies

In this second Chapter the multiphysics modelling methodology and the adopted simulation tools for the study of the lateral fin-like Ge-on-Si *pin* waveguide photodetector are described in detail.

2.1 Multiphysics simulation

Multiphysics simulation is widely adopted for the analysis of optoelectronic devices, such as photodetectors. In fact, a multiphysics model allows for the self-consistent solution of problems of different nature: the electromagnetic (optical) problem and the electrical one, which is crucial for an accurate study of such devices.

In this work, the Synopsys TCAD Sentaurus suite has been used to develop a multiphysics model of the lateral Ge-on-Si photodetector under analysis.

First, the structure of the detector presented by literature has been reproduced by means of the Sentaurus Structure Editor tool, while the Sentaurus Mesh tool has been used for placing doping profiles and generating a properly refined mesh on the whole device structure. A thin buffer layer has been introduced between silicon and germanium, with a graded composition profile mimicking the low-temperature epitaxial growth process, mitigating the sharp discontinuity of the abrupt Si/Ge heterojunction in the band diagram (see figure 2.2) and providing a more realistic model, as discussed in [19]. An example of simulated device geometry is reported in figure 2.1.

Once the device structure has been defined, the following simulation approach has been adopted:

- RSoft Photonic Device Tools have been employed to handle the electromagnetic problem, i.e., the evanescent coupling and the propagation of

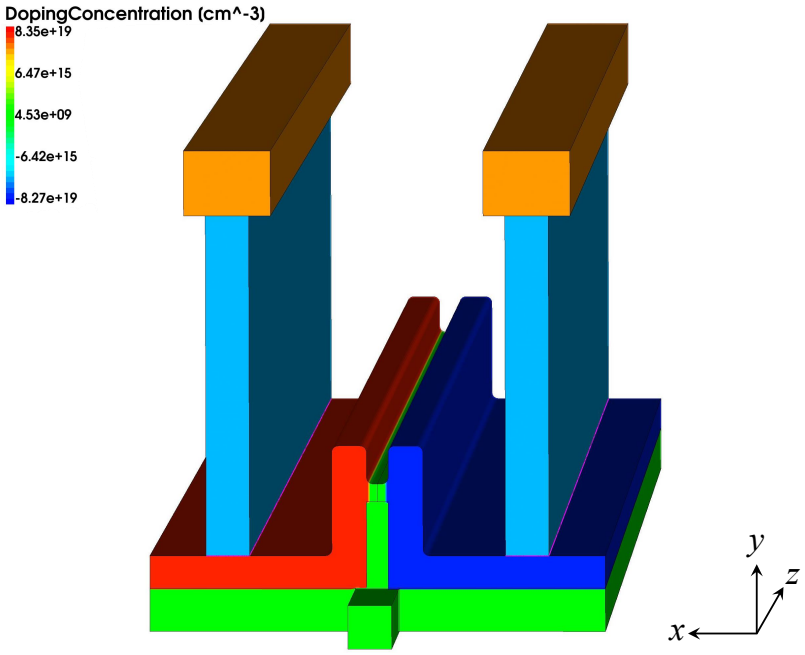
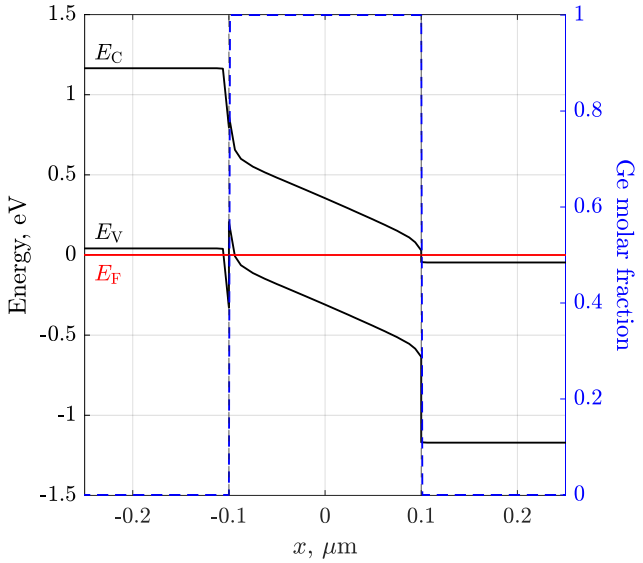
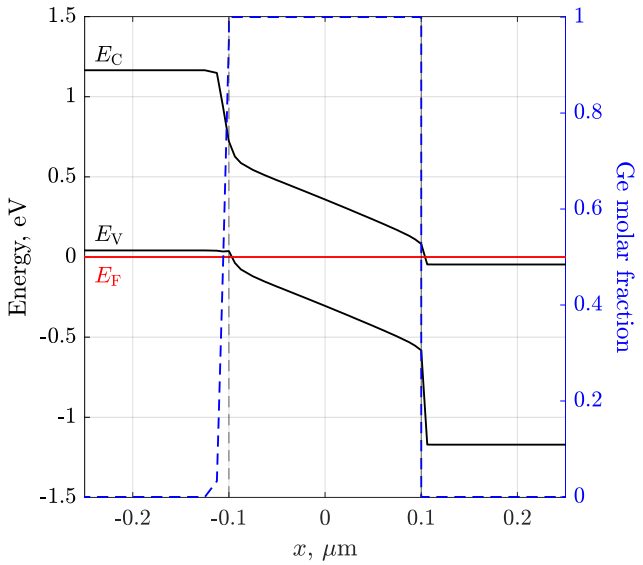


Figure 2.1: Lateral fin-like Ge-on-Si *pin* photodetector structure reproduced on TCAD Sentaurus. On the left the n-doped silicon contact (red), on the right the p-doped contact (blue); the intrinsic germanium fin (green) is sandwiched between complementary silicon sidewalls; silicon waveguide (green), tungsten (light blue) and copper (orange) contacts are shown as well; the cladding silicon oxide surrounding the whole device is not shown.



(a) Abrupt p-Si/Ge junction



(b) Graded p-Si/Ge junction

Figure 2.2: Comparison between band diagrams of lateral Ge-on-Si *pin* photodetectors without and with buffer layer at the p-Si/Ge heterojunction. An erf-like profile of the germanium molar fraction removes the valence band barrier arising from the abrupt junction.

light from the source to the detector; in particular, RSoft FullWAVE [20], a 3D finite-difference time-domain (FDTD) solver of Maxwell's equations, has been used to compute the optical generation rate G_{opt} within the device; further details about the solution of the optical problem are reported in section 2.3;

- TCAD Sentaurus has been used to manage the electrical problem; precisely, the charge transport problem has been solved through the Sentaurus Device tool [21], computing the coupled solution of the Poisson's equation with the semiconductor drift-diffusion transport equations, where the optical generation rate G_{opt} evaluated by RSoft FullWAVE has been incorporated as a source term; details about the solution of the electronic transport problem are reported in section 2.2.

The adopted multiphysics simulation flow is summarised by the block diagram reported in figure 2.3.

2.2 Electronic transport simulation

Details about the physical models implemented by Sentaurus Device for the solution of the charge transport problem in semiconductor devices are reported in this section, along with a description of the AC small-signal simulations which is possible to perform.

2.2.1 Physical model

Fundamental ingredients for the study of electronic transport in semiconductor devices are mobile charges (electrons and holes) and fixed charges (dopants). Charges determine and, at the same time, are influenced by the electrostatic potential, which is the essential quantity to be computed by any electrical device simulation.

The electrostatic potential ϕ is the solution of the Poisson equation [21, Ch. 7]:

$$\nabla \cdot (\varepsilon \nabla \phi) = -q \left(p - n + N_{\text{D}}^{+} - N_{\text{A}}^{-} \right) \tag{2.1}$$

where:

- ε is the semiconductor electrical permittivity;
- q is the elementary charge;
- n and p are the electron and hole densities;
- N_{D}^{+} and N_{A}^{-} are the ionized donors and acceptors concentrations.

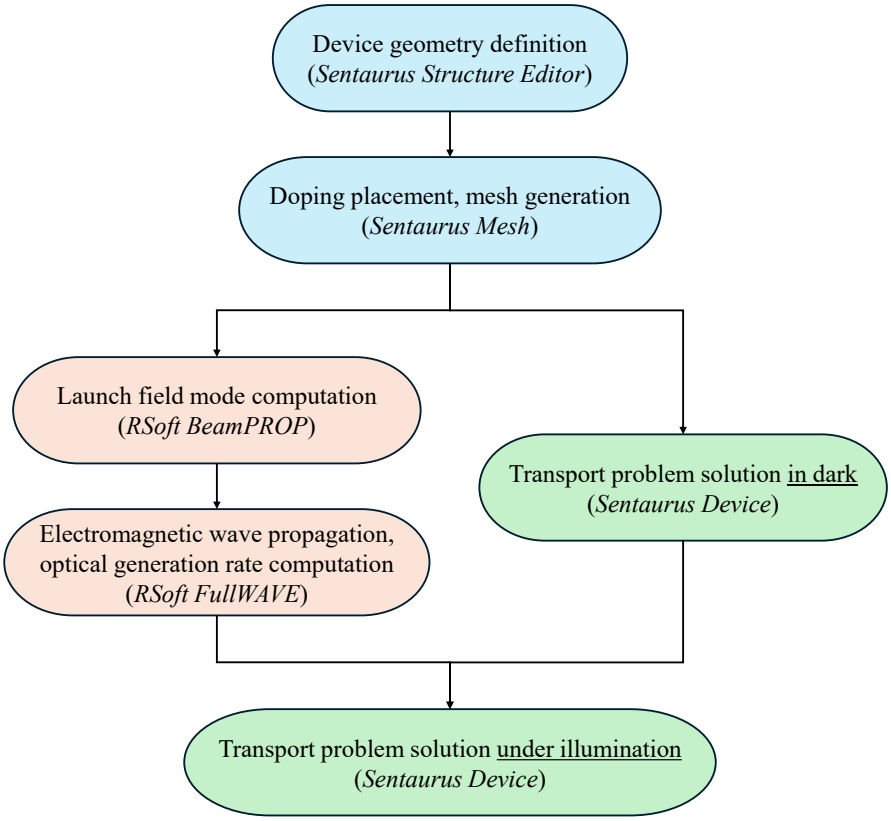


Figure 2.3: Block diagram summarising the adopted multiphysics simulation flow.

Defining the electric field

$$\mathcal{E} = -\nabla\phi \tag{2.2}$$

and the space charge density

$$\rho = q\left(p - n + N_{\mathrm{D}}^{+} - N_{\mathrm{A}}^{-}\right) \tag{2.3}$$

the Poisson equation (2.1) may be also written as the Gauss law:

$$\nabla \cdot (\varepsilon\mathcal{E}) = \rho \tag{2.4}$$

Assuming carriers are distributed in energy according to the Fermi-Dirac statistics [21, Ch. 7], which is more accurate especially for high carrier densities, n and p are expressed as

$$n = N_{\mathrm{C}} F_{1/2}\left(\frac{E_{\mathrm{F},n} - E_{\mathrm{C}}}{kT}\right) \tag{2.5a}$$

$$p = N_{\mathrm{V}} F_{1/2}\left(\frac{E_{\mathrm{V}} - E_{\mathrm{F},p}}{kT}\right) \tag{2.5b}$$

where:

- N_{C} and N_{V} are the conduction and valence band effective density of states;
- $F_{1/2}$ is the Fermi integral of order 1/2;
- $E_{\mathrm{F},n}$ and $E_{\mathrm{F},p}$ are the quasi-Fermi energies for electrons and holes;
- E_{C} and E_{V} are the conduction and valence band edges;
- k is the Boltzmann constant and T is the temperature.

The concentration of ionized dopants is expressed according to the Fermi-Dirac statistics [21, Ch. 13] as

$$N_{\mathrm{D}}^{+} = \frac{N_{\mathrm{D},0}}{1 + g_{\mathrm{D}} \exp\left(\dfrac{E_{\mathrm{F},n} - E_{\mathrm{D}}}{kT}\right)} \tag{2.6a}$$

$$N_{\mathrm{A}}^{-} = \frac{N_{\mathrm{A},0}}{1 + g_{\mathrm{A}} \exp\left(\dfrac{E_{\mathrm{A}} - E_{\mathrm{F},p}}{kT}\right)} \tag{2.6b}$$

where:

- $N_{\mathrm{D},0}$ and $N_{\mathrm{A},0}$ are the substitutional donor and acceptor concentrations;
- g_{D} and g_{A} are the degeneracy factors for the impurity levels;
- E_{D} and E_{A} are the donor and acceptor activation energies.

The framework is completed by the carrier transport equations [21, Ch. 8]. Carrier transport models can be all written in the form of continuity equations, which describe charge conservation:

$$\nabla \cdot \overrightarrow{J_n} = q\left(R_n - G_n\right) + q\frac{\partial n}{\partial t} \tag{2.7a}$$

$$-\nabla \cdot \overrightarrow{J_p} = q\left(R_p - G_p\right) + q\frac{\partial p}{\partial t} \tag{2.7b}$$

where:

- $\overrightarrow{J_n}$ and $\overrightarrow{J_p}$ are the electron and hole current densities;
- R_n and R_p are the electron and hole recombination rates;
- G_n and G_p are the electron and hole generation rates.

Details about generation-recombination mechanisms are provided later in this Chapter.

Transport models differ in the expressions of currents. In this work, the *drift-diffusion model* has been adopted. The drift-diffusion model, valid in quasi-equilibrium conditions, describes transport in terms of drift and diffusion currents:

$$\overrightarrow{J_n} = \overrightarrow{J_{n,\text{drift}}} + \overrightarrow{J_{n,\text{diff}}} = -qn\mu_n\nabla\phi + qD_n\nabla n \tag{2.8a}$$

$$\overrightarrow{J_p} = \overrightarrow{J_{p,\text{drift}}} + \overrightarrow{J_{p,\text{diff}}} = -qp\mu_p\nabla\phi - qD_p\nabla p \tag{2.8b}$$

where:

- μ_n and μ_p are the electron and hole mobilities;
- D_n and D_p are the electron and hole diffusion coefficients, related to mobilities through the Einstein relation, $D_n = kT\mu_n/q$ and $D_p = kT\mu_p/q$.

Details about mobility models are provided later in this Chapter.

The system of Poisson equation (2.1), continuity equations (2.7) and current density definitions (2.8) can be written in steady-state as

$$\begin{cases} \nabla \cdot (\varepsilon\nabla\phi) = -q\left(p - n + N_{\text{D}}^{+} - N_{\text{A}}^{-}\right) \\ \nabla \cdot (-n\mu_n\nabla\phi + D_n\nabla n) = (R_n - G_n) + \dfrac{\partial n}{\partial t} \\ \nabla \cdot (p\mu_p\nabla\phi + D_p\nabla p) = (R_p - G_p) + \dfrac{\partial p}{\partial t} \end{cases} \tag{2.9}$$

which is a system of partial differential equations (PDEs) in the unknowns ϕ, n and p. The computation of these three ingredients allows for a simplified but effective prediction of the device behaviour. The PDEs system, supplied with

boundary conditions (typically ohmic contacts), is discretized on the device mesh and solved by Sentaurus Device by means of a Newton solver, an iterative numerical algorithm solving a linearized system at each step until convergence (further details can be found in [21, Ch. 6]).

The models adopted for generation-recombination processes and carrier mobility are described in the following paragraphs. In general, these models are temperature-dependent. However, in this work all simulations have been performed at room temperature, therefore models and parameters here reported are only valid at $T = 300\,\mathrm{K}$.

Recombination and generation models

The generation-recombination processes modelled by Sentaurus Device are here detailed [21, Ch. 16].

- *Shockley-Read-Hall* (SRH) *recombination* is a trap-assisted process, i.e., electrons and holes recombine through defects energy levels. The SRH recombination rate is implemented as

$$R_{\mathrm{SRH}} = \frac{np - n_i^2}{\tau_p^{\mathrm{SRH}}(n + n_i) + \tau_n^{\mathrm{SRH}}(p + n_i)} \tag{2.10}$$

  where n_i is the effective intrinsic density, τ_p^{SRH} and τ_n^{SRH} are the carriers SRH lifetimes, reported in table B.2.

- *Radiative recombination* is a photon-assisted process and the corresponding recombination rate is implemented as

$$R_{\mathrm{rad}} = C^{\mathrm{rad}}(np - n_i^2) \tag{2.11}$$

  where C^{rad} is a material-dependent constant, reported in table B.2.

- *Auger recombination* is an electron- or hole-assisted process with a recombination rate implemented as

$$R_{\mathrm{Auger}} = (C_n^{\mathrm{Auger}} n + C_p^{\mathrm{Auger}} p)(np - n_i^2) \tag{2.12}$$

  with C_n^{Auger} and C_p^{Auger} material-dependent constants, reported in table B.2.

- *Avalanche generation* (or *impact ionization*) of electron-hole pairs takes place in presence of high-intensity electric fields within wide depleted regions. The avalanche generation rate is implemented as

$$G_{\mathrm{aval}} = \frac{1}{q}(\alpha_n \left|\overrightarrow{J_n}\right| + \alpha_p \left|\overrightarrow{J_p}\right|) \tag{2.13}$$

where α_n and α_p are ionization coefficients (reciprocal of carriers mean free path), expressed according to the Okuto-Crowell empirical model [22]:

$$\alpha(F) = a_{\text{aval}} F \exp\left[-\left(\frac{b_{\text{aval}}}{F}\right)^2\right] \tag{2.14}$$

where a_{aval} and b_{aval} are material-dependent coefficients, reported in table B.2, while F is the driving electric field.

- *Optical generation* occurs when photons are absorbed by the material, generating electron-hole pairs. In this work, the optical generation rate G_{opt} has been computed by means of RSoft FullWAVE, as described in section 2.3, and loaded on Sentaurus Device as an external source term.

Mobility models

Carrier mobility μ has been evaluated through the combination of two models:

- a *doping-dependent* mobility model, specifically the Masetti model [23], accounting for the mobility degradation due to impurity scattering:

$$\mu_{\text{low}} = \mu_{\text{min1}} \exp\left(-\frac{P_{\text{c}}}{N_{\text{tot}}}\right) + \frac{\mu_{\text{const}} - \mu_{\text{min2}}}{1 + (N_{\text{tot}}/C_{\text{r}})^{\alpha}} - \frac{\mu_1}{1 + (C_{\text{s}}/N_{\text{tot}})^{\beta}} \tag{2.15}$$

  where N_{tot} is the total doping concentration, while μ_{const}, μ_{min1}, μ_{min2}, μ_1, P_{c}, C_{r}, C_{s}, α and β are material-dependent parameters, reported in table B.3;

- a *high-field saturation* model, namely the Canali model [24], describing carrier velocity saturation under the effect of high electric fields:

$$\mu(F) = \frac{\mu_{\text{low}}}{\left[1 + \left(\frac{\mu_{\text{low}} F}{v_{\text{sat}}}\right)^{\beta}\right]^{1/\beta}} \tag{2.16}$$

  where μ_{low} is the low-field mobility evaluated by (2.15), F is the driving electric field, v_{sat} is the saturation velocity; v_{sat} and β values are reported in table B.4.

2.2.2 AC simulation

Small-signal AC simulations allow for the computation of the response of a device to small AC signals, superimposed to a DC bias, as a function of frequency.

Numerical method

In Sentaurus Device [21, Ch. 38], the AC response is obtained from the three basic semiconductor equations (2.1, 2.7). The simplified system of equations can be symbolically represented at node i of the computation mesh as:

$$\begin{cases} F_{\phi i}(\phi, n, p) = 0 \\ F_{ni}(\phi, n, p) = \dot{G}_{ni}(n) \\ F_{pi}(\phi, n, p) = \dot{G}_{pi}(p) \end{cases} \tag{2.17}$$

with F and G nonlinear functions of the arguments ϕ, n and p.

Vector functions are assumed to have the form $\xi_{\text{total}} = \xi_{\text{DC}} + \tilde{\xi} e^{\mathrm{i}\omega t}$, where ξ_{DC} is the value of ξ at the DC operating point, $\tilde{\xi}$ is the AC response magnitude and $\omega = 2\pi f$ is the AC excitation frequency. Expanding F and G in Taylor series around the DC operating point and keeping only the first-order terms, the system (2.17) can be rewritten as

$$\sum_j \left[\begin{array}{ccc} \dfrac{\partial F_{\phi i}}{\partial \phi_j} & \dfrac{\partial F_{\phi i}}{\partial n_j} & \dfrac{\partial F_{\phi i}}{\partial p_j} \\ \dfrac{\partial F_{ni}}{\partial \phi_j} & \dfrac{\partial F_{ni}}{\partial n_j} - \mathrm{i}\omega \dfrac{\partial G_{ni}}{\partial n_j} & \dfrac{\partial F_{ni}}{\partial p_j} \\ \dfrac{\partial F_{pi}}{\partial \phi_j} & \dfrac{\partial F_{pi}}{\partial n_j} & \dfrac{\partial F_{pi}}{\partial p_j} - \mathrm{i}\omega \dfrac{\partial G_{pi}}{\partial p_j} \end{array} \right]_{\text{DC}} \left[\begin{array}{c} \tilde{\phi}_j \\ \tilde{n}_j \\ \tilde{p}_j \end{array} \right] = 0 \tag{2.18}$$

The global AC matrix system is obtained by imposing AC boundary conditions and performing the summation. Therefore, the AC system becomes

$$[J + \mathrm{i}D]\tilde{X} = B \tag{2.19}$$

where J is the Jacobian matrix, D contains the contributions of the G functions to the matrix, B is a real vector dependent on the AC voltage drive and $\tilde{X}$ is the AC solution vector.

Once the AC system is solved, the AC displacement, electron and hole current density responses $\overrightarrow{\tilde{J}_D}$, $\overrightarrow{\tilde{J}_n}$ and $\overrightarrow{\tilde{J}_p}$ are computed as

$$\overrightarrow{\tilde{J}_D} = -\mathrm{i}\omega\varepsilon\nabla\tilde{\phi} \tag{2.20a}$$

$$\overrightarrow{\tilde{J}_n} = \left.\frac{\partial \overrightarrow{\tilde{J}_n}}{\partial \phi}\right|_{\text{DC}} \tilde{\phi} + \left.\frac{\partial \overrightarrow{\tilde{J}_n}}{\partial n}\right|_{\text{DC}} \tilde{n} + \left.\frac{\partial \overrightarrow{\tilde{J}_n}}{\partial p}\right|_{\text{DC}} \tilde{p} \tag{2.20b}$$

$$\overrightarrow{\tilde{J}_p} = \left.\frac{\partial \overrightarrow{\tilde{J}_p}}{\partial \phi}\right|_{\text{DC}} \tilde{\phi} + \left.\frac{\partial \overrightarrow{\tilde{J}_p}}{\partial p}\right|_{\text{DC}} \tilde{p} + \left.\frac{\partial \overrightarrow{\tilde{J}_p}}{\partial n}\right|_{\text{DC}} \tilde{n} \tag{2.20c}$$

Electrical AC analysis

An AC analysis evaluates the frequency-dependent admittance matrix Y, which fully describes the equivalent small-signal model of a device [21, Ch. 4].

For a given excitation frequency f, the following relation holds between the complex-valued vectors of voltage and current excitations δV and δI:

$$\delta I = Y \delta V \tag{2.21}$$

where the admittance matrix is expressed as

$$Y = A + \mathrm{i}2\pi f C = A + \mathrm{i}\omega C \tag{2.22}$$

with real-valued conductance matrix A measured in siemens Ω^{-1} and capacitance matrix C measured in farad $\mathrm{F} = (\Omega \cdot \mathrm{Hz})^{-1}$. A and C are exactly the output variables provided by an AC analysis on Sentaurus Device.

Optical AC analysis

An optical AC analysis allows for the computation of the quantum efficiency of a photodetector as a function of the modulation frequency of the input optical signal [21, Ch. 21].

A small perturbation of the incident optical power δP_0 results in a perturbation of the photogeneration rate as $G_{\mathrm{opt}} + \delta G_{\mathrm{opt}} e^{\mathrm{i}\omega t}$, where $\omega = 2\pi f$ is the modulation frequency and δG_{opt} is the amplitude of the perturbation. The perturbation of the photogeneration rate results in a complex perturbation of the small-signal device current δI. Given the total optical power perturbation δP_{tot} within the simulation domain, defined as

$$\delta P_{\mathrm{tot}} = \int_S \delta P_0 ds \tag{2.23}$$

Sentaurus Device provides as output variables of an optical AC analysis η and C_{opt}, defined as

$$\eta = \frac{\mathrm{Re}\left[\delta I\right]/q}{\delta P_{\mathrm{tot}}/\hbar\omega} \tag{2.24}$$

$$C_{\mathrm{opt}} = \frac{1}{\omega} \cdot \frac{\mathrm{Im}\left[\delta I\right]/q}{\delta P_{\mathrm{tot}}/\hbar\omega} \tag{2.25}$$

where $\delta P_{\mathrm{tot}}/\hbar\omega$ represents the perturbation of the total number of photons. The complex quantum efficiency $\mathfrak{Y}(\omega)$ can be thus reconstructed as

$$\mathfrak{Y}(\omega) = \frac{\delta I/q}{\delta P_{\mathrm{tot}}/\hbar\omega} = \eta + \mathrm{i}\omega C_{\mathrm{opt}} \tag{2.26}$$

and, in magnitude:

$$|\mathfrak{Y}(\omega)| = \sqrt{\eta^2 + (\omega C_{\mathrm{opt}})^2} = \sqrt{\eta^2 + (2\pi f C_{\mathrm{opt}})^2} \tag{2.27}$$

which can be normalized and expressed in dB as

$$|\mathfrak{y}(\omega)|_{\mathrm{dB}} = 20\log_{10}\left|\frac{\mathfrak{Y}(\omega)}{\mathfrak{Y}(0)}\right| \tag{2.28}$$

2.3 Electromagnetic simulation

The electromagnetic problem, aimed at computing the optical generation rate G_{opt} within the device, has been solved in three steps:

1. the input mode has been evaluated on the silicon input waveguide using a mode solver based on the RSoft BeamPROP tool, implementing the Beam Propagation Method (BPM); specifically, given the waveguide size and the refractive indices of materials, the field shape and propagation constant of the fundamental mode have been calculated;
2. the computed input mode has been used as launch field by the RSoft FullWAVE FDTD simulation, i.e., the mode has been propagated along the device and the spatial distribution of the electromagnetic field has been obtained;
3. the spatial distribution of the optical generation rate G_{opt} has been evaluated from the absorbed power density distribution.

Material refractive indices used by RSoft simulations are reported in tables B.5, B.6, B.7.

2.3.1 BPM for mode solving

The BPM is a simplified technique for approximating the exact wave equation for monochromatic waves propagating in arbitrary waveguiding geometries [25, Ch. 2]. This technique uses finite difference methods to solve the parabolic approximation of the Helmholtz wave equation.

Several mode-solving techniques have been developed that are based on BPM. Assuming a given incident field is launched into a geometry that is z-invariant, the BPM propagation can be equivalently described in terms of the modes and propagation constants of the guiding structure. Assuming a scalar incident field $\phi_{\mathrm{in}}(x)$ expanded in the modes of the structure as

$$\phi_{\mathrm{in}}(x) = \sum_m c_m \phi_m(x) \tag{2.29}$$

propagation along z can be expressed as

$$\phi(x,z) = \sum_m c_m \phi_m(x) \exp(\mathrm{i}\beta_m z) \tag{2.30}$$

By equating the propagating field obtained via BPM with (2.30) it is possible to extract mode information from BPM results directly.

2.3.2 FDTD method

The FDTD method provides a rigorous solution of Maxwell's equations, without any approximations or theoretical restrictions, and includes many more effects, such as reflection and interference, compared to other approximated methods [20, Ch. 2].

In general, Maxwell's equations read

$$\nabla \times \mathcal{H} = \frac{\partial \mathcal{D}}{\partial t} + \mathcal{J} \qquad \text{(Ampère-Maxwell law)} \tag{2.31a}$$

$$\nabla \times \mathcal{E} = -\frac{\partial \mathcal{B}}{\partial t} \qquad \text{(Faraday law for induction)} \tag{2.31b}$$

$$\nabla \cdot \mathcal{B} = 0 \qquad \text{(Gauss law for magnetism)} \tag{2.31c}$$

$$\nabla \cdot \mathcal{D} = \rho \qquad \text{(Gauss law)} \tag{2.31d}$$

with constitutive relations

$$\mathcal{D} = \varepsilon \mathcal{E} \tag{2.32a}$$

$$\mathcal{H} = \frac{1}{\mu} \mathcal{B} \tag{2.32b}$$

where $\mathcal{E}$ is the electric field, $\mathcal{B}$ is the magnetic field, $\mathcal{J}$ is the current density, ρ is the electric charge density, ε is the dielectric permittivity and μ is the magnetic permeability.

In absence of current sources ($\mathcal{J} = 0$) and charges ($\rho = 0$) and assuming a non-magnetic medium ($\mu = \mu_0$), Maxwell's curl equations reduce to

$$\nabla \times \mathcal{H} = \varepsilon \frac{\partial \mathcal{E}}{\partial t} \tag{2.33a}$$

$$\nabla \times \mathcal{E} = -\mu_0 \frac{\partial \mathcal{H}}{\partial t} \tag{2.33b}$$

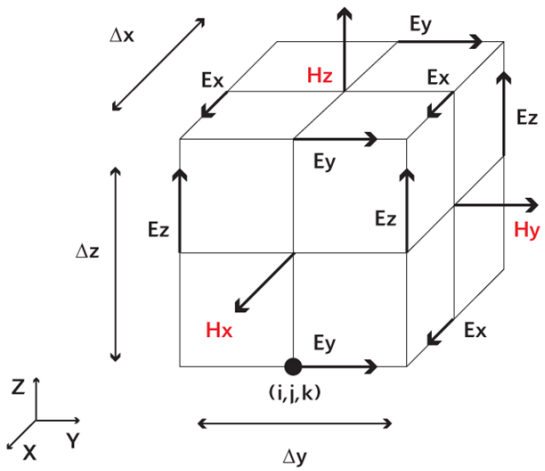


Figure 2.4: Yee cell of dimensions Δx, Δy, Δz. Figure taken from [20].

and can be written in cartesian coordinates as six simple scalar equations:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu_0}\left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y}\right) \tag{2.34a}$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu_0}\left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\right) \tag{2.34b}$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu_0}\left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}\right) \tag{2.34c}$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right) \tag{2.34d}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right) \tag{2.34e}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon}\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right) \tag{2.34f}$$

The FDTD method solves Maxwell's equations by first discretizing the equations via central differences in time and space, and then numerically solving these equations. The most common method to solve these equations is based on Yee's algorithm [26]. The spatial simulation domain is a grid with points spaced Δx, Δy, Δz apart (see figure 2.4), while time is divided into discrete steps Δt.

In order to produce an accurate simulation, the spatial grid spacing must be able to resolve the wavelength λ in time, i.e., a condition such as

$$\max(\Delta x, \Delta y, \Delta z) < \lambda/10 \tag{2.35}$$

should be met.

At the same time, in order to get a stable simulation, the temporal step size Δt should meet the Courant condition

$$c\Delta t < \frac{1}{\sqrt{1/\Delta x^2 + 1/\Delta y^2 + 1/\Delta z^2}} \tag{2.36}$$

where c is the speed of light.

The field at a given mesh point, denoted by integers (i, j, k), is computed according to discretized Maxwell's equations, solved iteratively, alternating between computations of electric field components at times $t = n\Delta t$ and magnetic field components at $t = (n + 1/2)\Delta t$ (n is an integer number representing the computation step). For example, equations (2.34a) and (2.34d) are discretized as

$$H^{n+1/2}_{x(i,j,k)} = H^{n-1/2}_{x(i,j,k)} + \frac{\Delta t}{\mu_0 \Delta z}\left(E^n_{y(i,j,k)} - E^n_{y(i,j,k-1)}\right) - \\ - \frac{\Delta t}{\mu_0 \Delta y}\left(E^n_{z(i,j,k)} - E^n_{z(i,j-1,k)}\right) \tag{2.37a}$$

$$E^{n+1}_{x(i,j,k)} = E^n_{x(i,j,k)} + \frac{\Delta t}{\varepsilon \Delta y}\left(H^{n+1/2}_{z(i,j+1,k)} - H^{n+1/2}_{z(i,j,k)}\right) - \\ - \frac{\Delta t}{\varepsilon \Delta z}\left(H^{n+1/2}_{y(i,j,k+1)} - H^{n+1/2}_{y(i,j,k)}\right) \tag{2.37b}$$

The solution of the electromagnetic problem requires boundary conditions. If perfectly matched layer (PML) [27] boundary conditions are applied, energy within the simulation domain is absorbed without inducing reflections, which is the most effective choice for device simulations of the kind here performed.

The electromagnetic simulation stops when convergence of the results is detected, i.e., the steady-state field distribution has been reached.

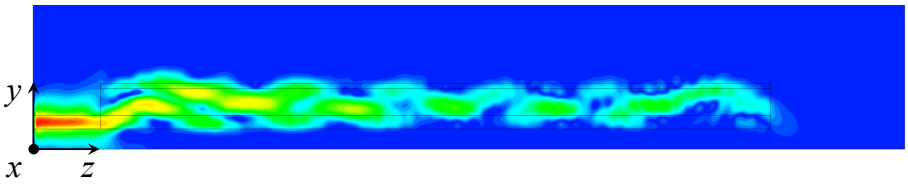
The steady-state spatial distribution of power within the simulated Ge-on-Si photodetector can be observed in figures 2.5: the evanescent coupling between the silicon waveguide and the germanium absorber is fully captured by the FDTD simulation, along with reflection and interference effects which generate evident periodic patterns.

2.3.3 Optical generation rate evaluation

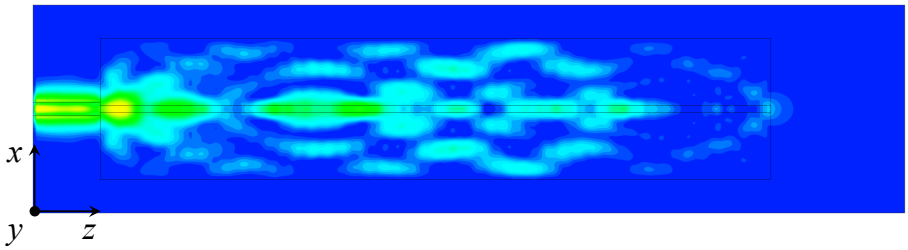
Once the steady-state electromagnetic field spatial distribution is available from the FDTD simulation, this can be used to compute the optical generation rate G_{opt} as a function of the position $\overrightarrow{r} = (x, y, z)$. In fact, it can be demonstrated that the absorbed power per unit volume u_{abs} can be expressed as

$$u_{\text{abs}}(\overrightarrow{r}) = \omega \cdot \text{Im}[\varepsilon(\overrightarrow{r})] \cdot |\mathcal{E}(\overrightarrow{r})|^2 \tag{2.38}$$

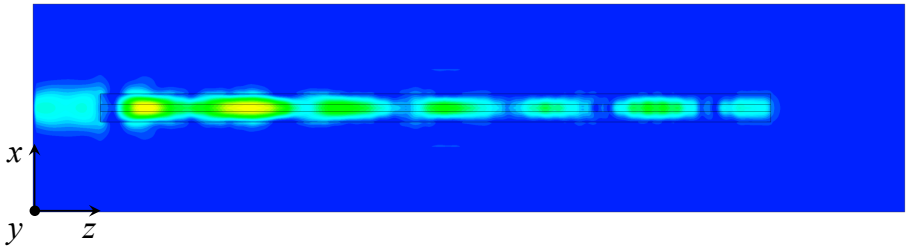
It can be observed that u_{abs} depends on the frequency of photons $\omega = 2\pi c/\lambda$, on the imaginary part of the dielectric permittivity ε (related to losses) and on the field distribution, through the squared magnitude of the electric field $\mathcal{E}$.



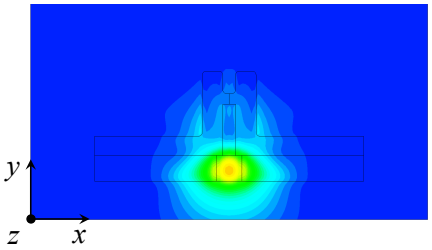
(a) $x = 0$ (center of the absorber)



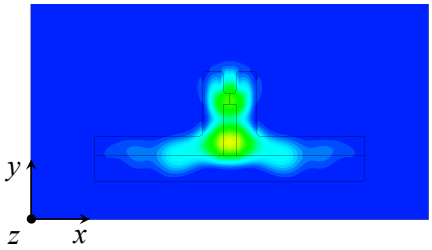
(b) $y = -0.15\,\mu\text{m}$ (interface between silicon waveguide and absorber)



(c) $y = 0.05\,\mu\text{m}$ (center of the absorber)



(d) $z = 0$ (front face of the absorber)



(e) $z = 1\,\mu\text{m}$

Figure 2.5: Steady-state spatial distribution of power within the simulated lateral Ge-on-Si photodetector.

Under the assumption of a unitary internal quantum efficiency, the optical generation rate G_{opt} coincides with the absorbed photon density and it can be directly evaluated from the absorbed power density as

$$G_{\mathrm{opt}}(\overrightarrow{r}) = \frac{u_{\mathrm{abs}}(\overrightarrow{r})}{E_{\mathrm{ph}}} = \frac{u_{\mathrm{abs}}(\overrightarrow{r})}{hc/\lambda} \tag{2.39}$$

where $E_{\mathrm{ph}} = hc/\lambda$ is the energy of photons.

Chapter 3

Analysis and results

This third Chapter presents the analyses conducted on the lateral fin-like Ge-on-Si *pin* waveguide photodetector through the developed physics-based model.

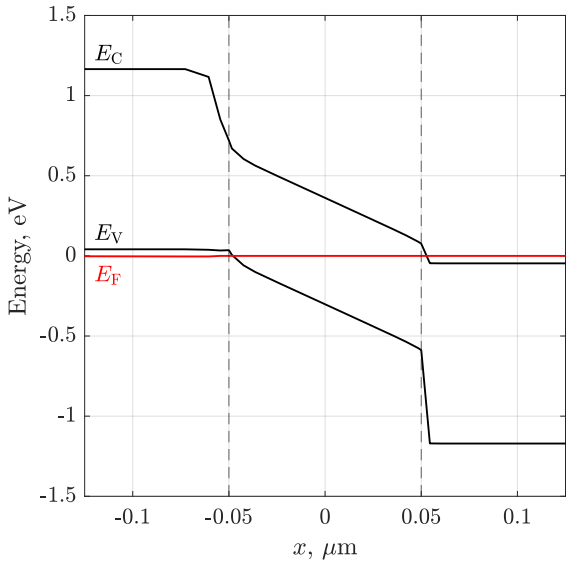
3.1 Fundamental device analysis

A preliminary electrostatic analysis of a reference Ge-on-Si lateral *pin* photodetector, featuring a germanium fin width $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, has been performed in dark conditions, both at equilibrium and in reverse bias. A 10^{20} cm^{-3} doping concentration has been adopted for the silicon contacts.

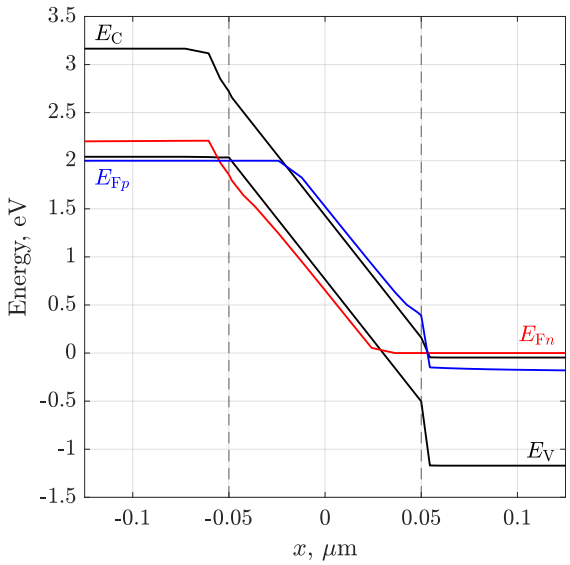
The energy band diagrams obtained for the reference device at equilibrium and for a $-2\,\mathrm{V}$ reverse bias are reported in figures 3.1. The high doping level makes the silicon contacts degenerate, as it can be deduced from the position of the Fermi level.

The electric field in the intrinsic germanium region, which is approximately constant, is further intensified by the application of a reverse bias voltage, as shown in figure 3.2a. At the same time, the intrinsic region is further depleted, as it can be noticed from the carrier density distribution in figure 3.2b.

Figure 3.3 shows some I-V curves, obtained for several values of the germanium fin width W_{Ge} and in reverse bias conditions. The dark current, i.e. the diode reverse saturation current, is proportional to W_{Ge}. In the considered voltage range, I-V curves for larger devices exhibit the typical saturation behaviour, while for very narrow fins an inflection point can be observed. This is due to the avalanche generation mechanism, arising when the electric field is particularly intense, according to equation (2.14): this occurs for lower voltage values when the germanium layer is thinner.

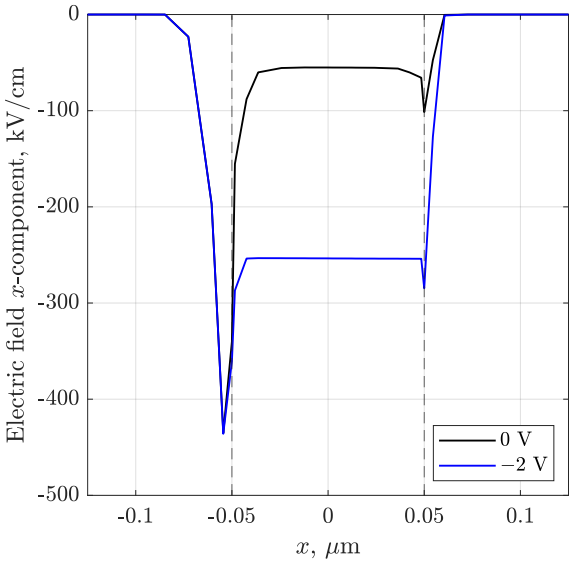


(a) Equilibrium

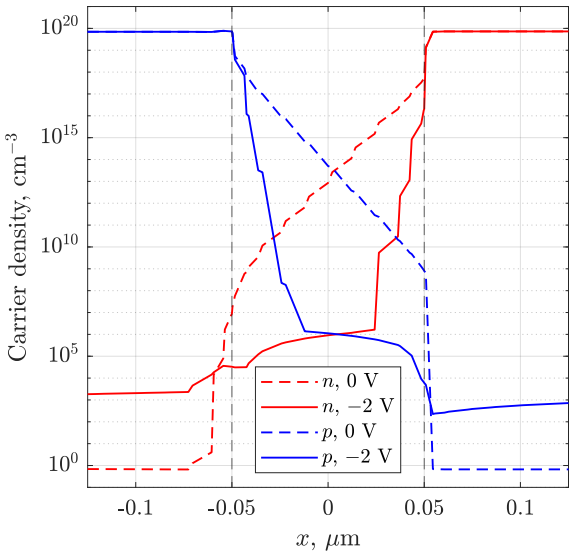


(b) Reverse bias ($-2\,\mathrm{V}$)

Figure 3.1: Reference Ge-on-Si lateral *pin* photodetector band diagram, at equilibrium and in reverse bias conditions.



(a) Electric field x-component



(b) Carrier density distribution

Figure 3.2: Electric field x-component and carrier density distribution in the reference Ge-on-Si lateral *pin* photodetector, at equilibrium and in reverse bias conditions.

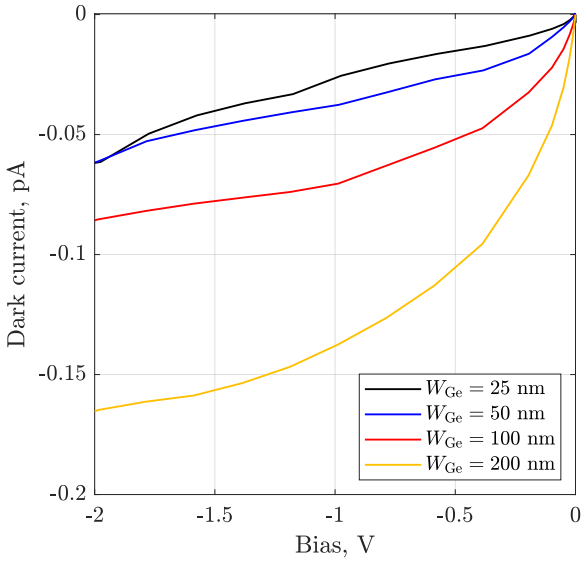


Figure 3.3: I-V curves obtained for Ge-on-Si lateral *pin* photodetectors featuring different W_{Ge} values in reverse bias conditions.

3.2 Small-signal analytical model

In this second section the method adopted for the extraction of a small-signal analytical model of the Ge-on-Si lateral *pin* photodetector is described and the proposed model is presented.

3.2.1 Preliminary small-signal AC analysis

As a starting point, a small-signal AC analysis (see subsection 2.2.2) has been performed using Sentaurus Device on a reference device, whose cross-section is shown in figure 3.4. Relevant dimensions of the simulated reference device are:

1. the germanium absorber width $W_{Ge} = 100\,\text{nm}$;
2. the germanium absorber height $H_{Ge} = 400\,\text{nm}$;
3. the germanium absorber length $L_{Ge} = 10\,\mu\text{m}$;
4. the distance between tungsten contacts and silicon sidewalls $d = 0.38\,\mu\text{m}$.

Henceforward, for brevity, the device having these features will always be referred to as *reference device*. The doping concentration of the silicon contacts has been kept constant to 10^{20} cm^{-3} in the analyses presented in this section.

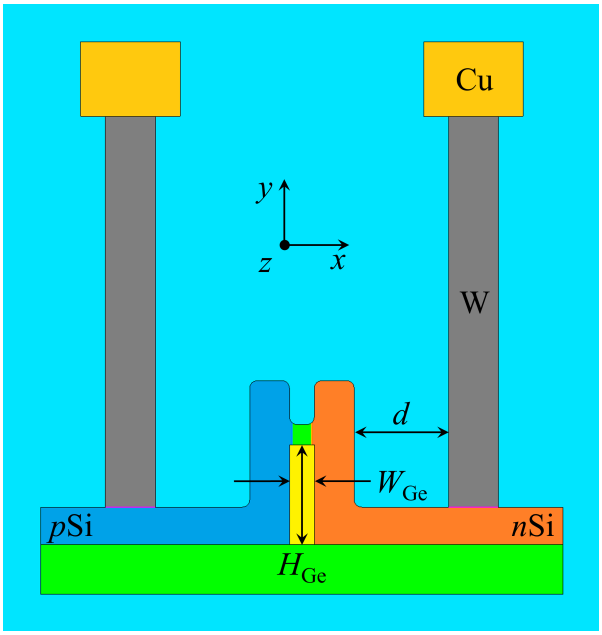


Figure 3.4: Simulated Ge-on-Si lateral *pin* photodetector cross-section in the *xy* plane. Key dimensions such as the absorber width W_{Ge}, the absorber height H_{Ge} and the distance d between tungsten contacts and silicon sidewalls are indicated.

The frequency-dependent small-signal admittance between the n and p contacts nodes has been computed in the frequency range between 0.1 GHz and 1000 THz and in reverse bias, with an applied voltage of -2 V and in dark conditions.

Results of this preliminary simulation are shown in figures 3.5. The frequency behaviour of the equivalent admittance can be described as follows:

1. at frequencies below 10^2 GHz the device is characterized by a very low conductance (below 1 mS), which increases with frequency; the low-frequency capacitance is equal to 9.3 fF;
2. above 10^2 GHz the equivalent capacitance starts decreasing, until it stabilizes to about 1.25 fF for frequencies above 10^4 GHz;
3. above 10^3 GHz the conductance stabilizes around a value of $0.1\,\Omega^{-1}$ (corresponding to an equivalent resistance of the order of $10\,\Omega$), while above 10^5 GHz it further increases.

3.2.2 Admittance fitting procedure and equivalent circuit model

It is evident from the obtained results that the equivalent admittance transfer function is characterized by poles and zeros. A first fitting of experimental data, henceforward referred to as *ideal fitting*, has been performed using the MATLAB function `rationalfit`, included in the RF Toolbox [28]. The function uses vector fitting with complex frequencies to perform rational fitting on a complex frequency-dependent data, satisfying the form:

$$F(s) = \sum_{k=1}^{n} \frac{C_k}{s - A_k} + D$$

with $s = \mathrm{i}2\pi f$. The ideal fitting results in the following impedance transfer function in the Laplace domain:

$$\begin{aligned} &Z_{\text{ideal}}(s) = \\ &= \frac{0.07526(s + 9.727 \times 10^{15})(s + 9.638 \times 10^{12})(s + 1.741 \times 10^{12})}{(s + 6.348 \times 10^{13})(s + 1.793 \times 10^{12})} \cdot \\ &\cdot \frac{(s + 1.047 \times 10^{11})(s + 1.294 \times 10^{10})}{(s + 1.05 \times 10^{11})(s + 1.294 \times 10^{10})(s + 647.5)} \end{aligned} \tag{3.1}$$

A comparison between Z_{ideal} and simulation data is shown in figures 3.6. An excellent agreement between the fitting and the data can be observed up to 10^5 GHz.

Some simplifications can be performed on the obtained transfer function. In fact, three couples of zeros and poles are close enough to compensate the effect of each other, i.e.:

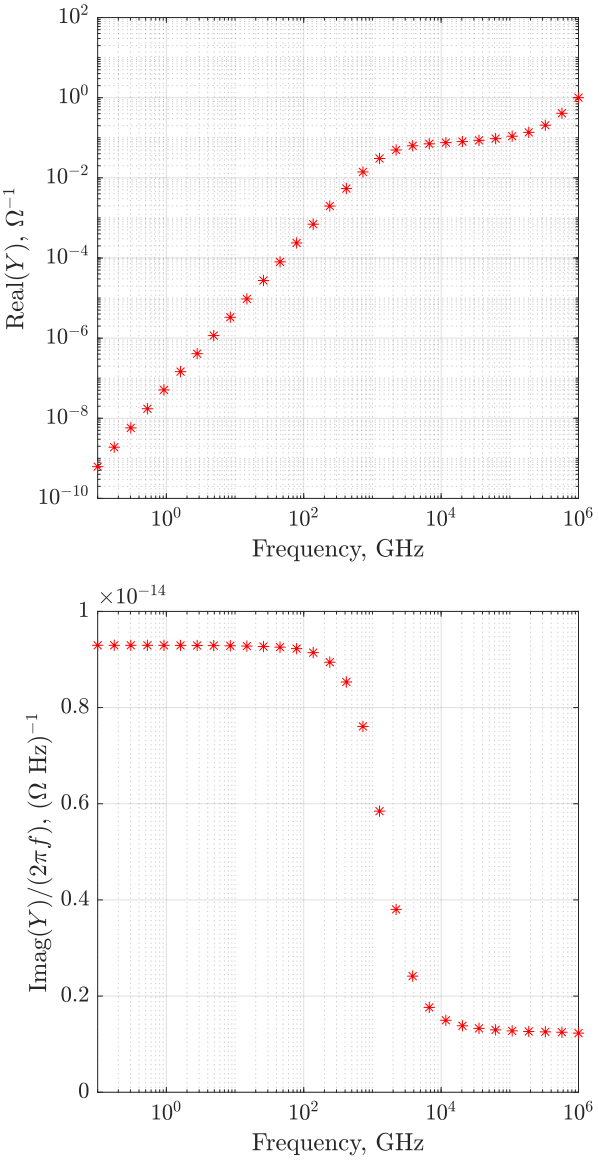


Figure 3.5: Reference device admittance at a $-2\,\mathrm{V}$ bias.

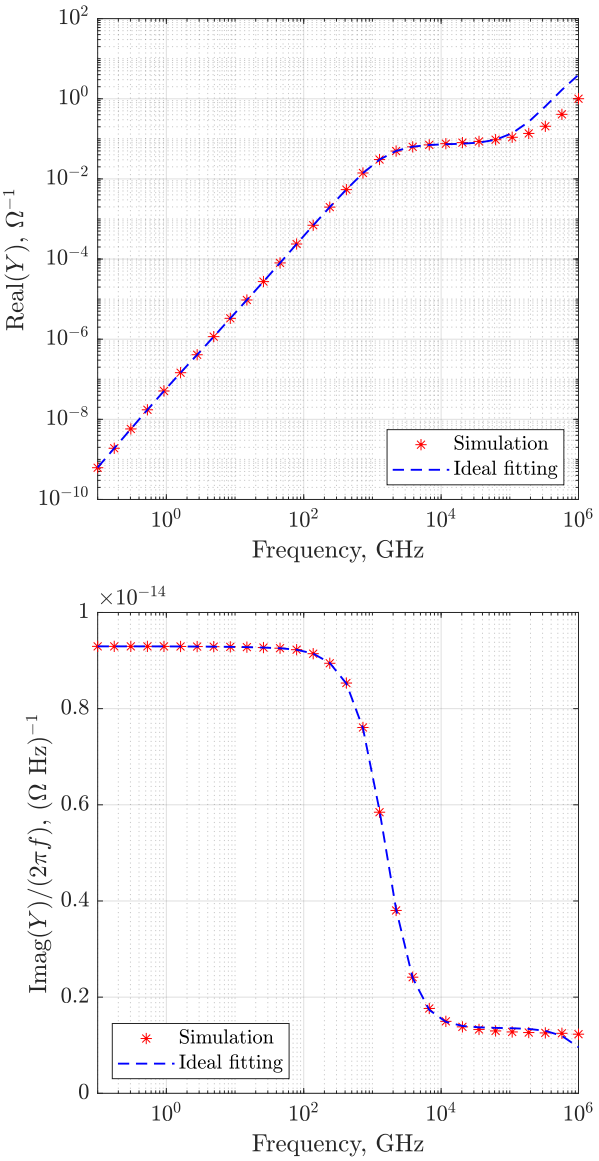


Figure 3.6: Ideal fitting (3.1) of the reference device admittance data at a $-2\,\mathrm{V}$ bias.

1. the zero located at 1.741×10^{12} rad/s and the pole located at 1.793×10^{12} rad/s;
2. the zero located at 1.047×10^{11} rad/s and the pole located at 1.05×10^{11} rad/s;
3. the zero located at 1.294×10^{10} rad/s and the pole located at 1.294×10^{10} rad/s.

In addition, the pole located at 647.5 rad/s is far enough from other poles and zeros to be approximated with a pole at 0 rad/s. Therefore, the impedance expression simplifies (*simplified fitting*) as:

$$Z_{\text{simplified}}(s) = \frac{0.07526(s + 9.727 \times 10^{15})(s + 9.638 \times 10^{12})}{s(s + 6.348 \times 10^{13})} \tag{3.2}$$

A comparison between $Z_{\text{simplified}}$ and simulation data can be observed in figures 3.7. The simplification introduces some error in the low-frequency fitting. Nevertheless, it still reproduces the frequency behaviour of the admittance quite accurately, meaning that it is reasonable to choose a fitting circuit model based on this transfer function. A successive optimization of circuit parameters will eventually lead to a modification of zeros, poles and gain, resulting in a better fitting.

The fitting circuit model has been chosen starting from the reference circuit which is typically found in the literature, described in subsection 1.2.2. In order to reproduce the transfer function (3.2), the selected circuit topology is the one represented in figure 3.8. The equivalent impedance of the device can be expressed in the Laplace domain as:

$$Z(s) = R_1 + \cfrac{1}{sC_1 + \cfrac{1}{R_2 + \cfrac{1}{sC_2}}} \tag{3.3}$$

which can be explicited as:

$$Z(s) = \frac{s^2 R_1 R_2 C_1 C_2 + sR_1(C_1 + C_2) + sR_2 C_2 + 1}{s(sR_2 C_1 C_2 + C_1 + C_2)} \tag{3.4}$$

It is evident the similarity between expressions (3.4) and (3.2).

In order to fit simulation data, circuit elements R_1, R_2, C_1 and C_2 have to be properly sized. An initial guess of their values can be made based on the experimental admittance data, reasoning on what they physically represent:

1. the capacitance C_2 represents the intrinsic diode capacitance, which is the dominant contribution to the low-frequency capacitance; a reasonable initial guess is $C_2 = 10\,\text{fF}$;

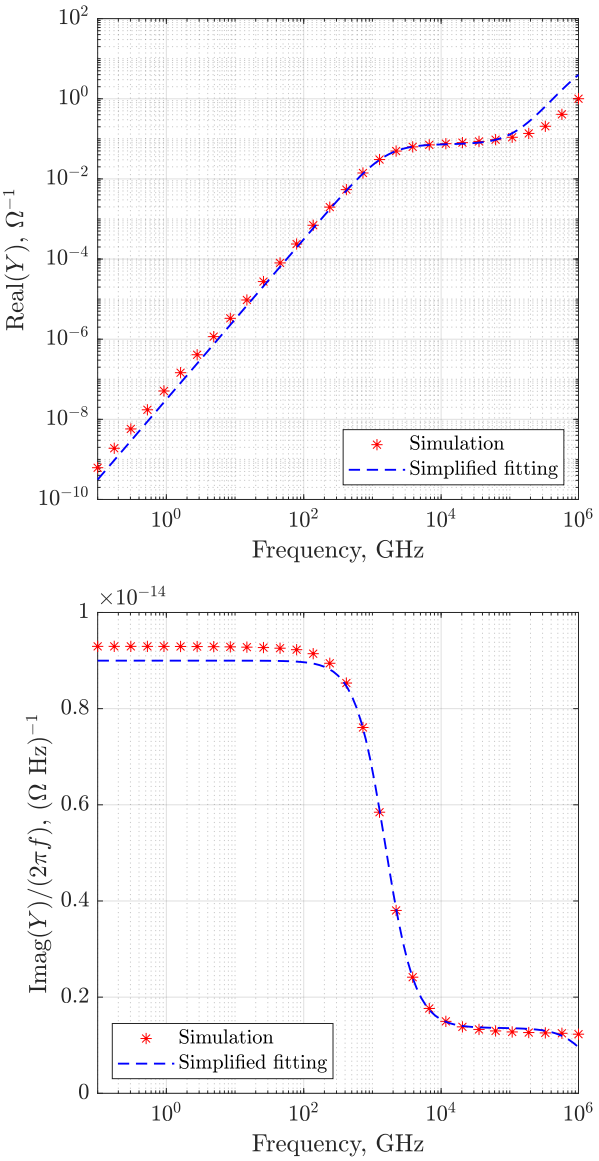


Figure 3.7: Simplified fitting (3.2) of the reference device admittance data at a $-2\,\mathrm{V}$ bias.

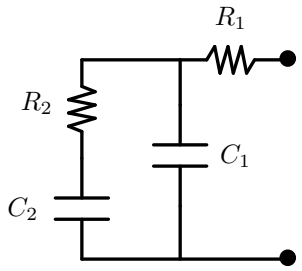


Figure 3.8: Fitting small-signal circuit model of the Ge-on-Si lateral *pin* photodetector.

2. the capacitance C_1 represents the parasitic capacitance of the device, which is essentially the high-frequency capacitance; in fact, the impedance of C_2, larger than C_1, becomes negligible above the pole located around 10^4 GHz, introduced by C_2 itself; a reasonable initial guess is $C_1 = 1\,\mathrm{fF}$;

3. the resistance R_2 represents the series parasitic diode resistance, which is practically the equivalent conductance once the impedance of C_1 becomes negligible, i.e., around 10^4 GHz; a reasonable initial guess is $R_2 = 10\,\Omega$;

4. the resistance R_1 represents another parasitic resistance, responsible of the high-frequency zero located around 10^6 GHz (it can be observed from the expression (3.4) that if $R_1 = 0$ only a single zero is present); a reasonable initial guess is $R_1 = 10\,\mathrm{m}\Omega$.

Using the guess values reported above, the frequency behaviour of the equivalent admittance of the device is the one reported in figures 3.9, where it is also compared with the simulation data.

An optimization of the values of circuit elements has to be performed in order to ensure a better fitting. This has been done by means of a `Zfit` function from the MATLAB Central File Exchange [29]. Given a circuit topology and initial guess values (possibly bounded) for circuit parameters, the function allows for the fitting of the circuit on the available impedance data. The fitting procedure has been performed repeatedly until convergence of the values of circuit elements. The final *optimized fitting* results in an impedance:

$$Z_{\mathrm{optimized}}(s) = \frac{0.01724(s + 4.782 \times 10^{16})(s + 7.982 \times 10^{12})}{s(s + 6.168 \times 10^{13})} \tag{3.5}$$

Comparing the expressions of $Z_{\mathrm{simplified}}$ and $Z_{\mathrm{optimized}}$, it is possible to observe that the frequencies of zeros and poles have been corrected, preserving their order of magnitude, allowing for the fitting shown in figures 3.10.

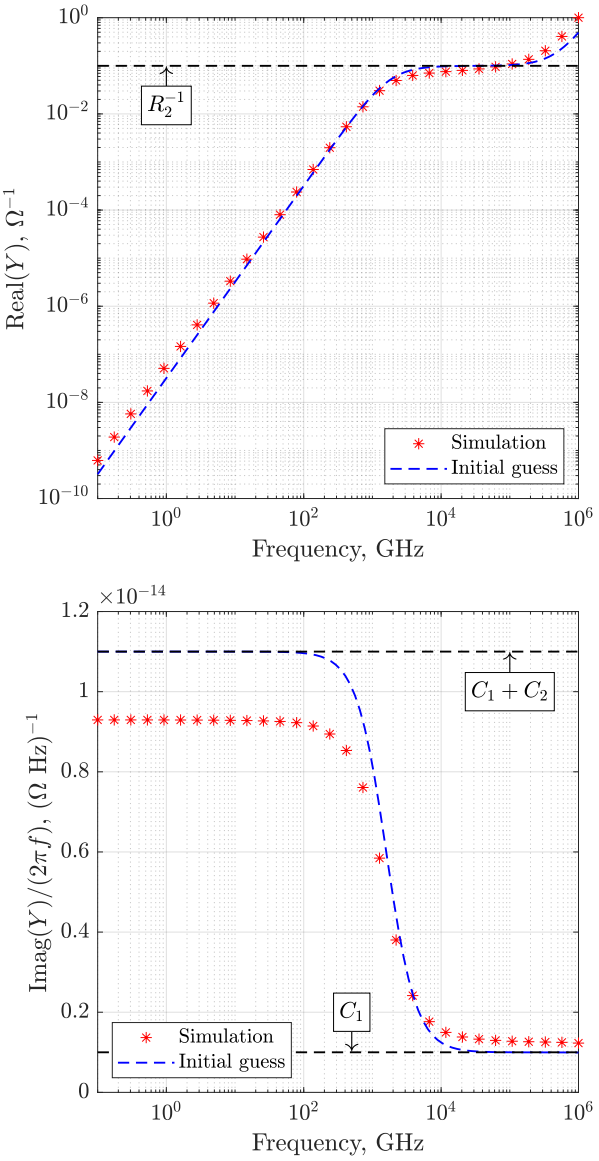


Figure 3.9: Fitting of the reference device admittance data at a $-2\,\mathrm{V}$ bias, based on the expression (3.4), with guess values $R_1 = 10\,\mathrm{m\Omega}$, $R_2 = 10\,\Omega$, $C_1 = 1\,\mathrm{fF}$, $C_2 = 10\,\mathrm{fF}$.

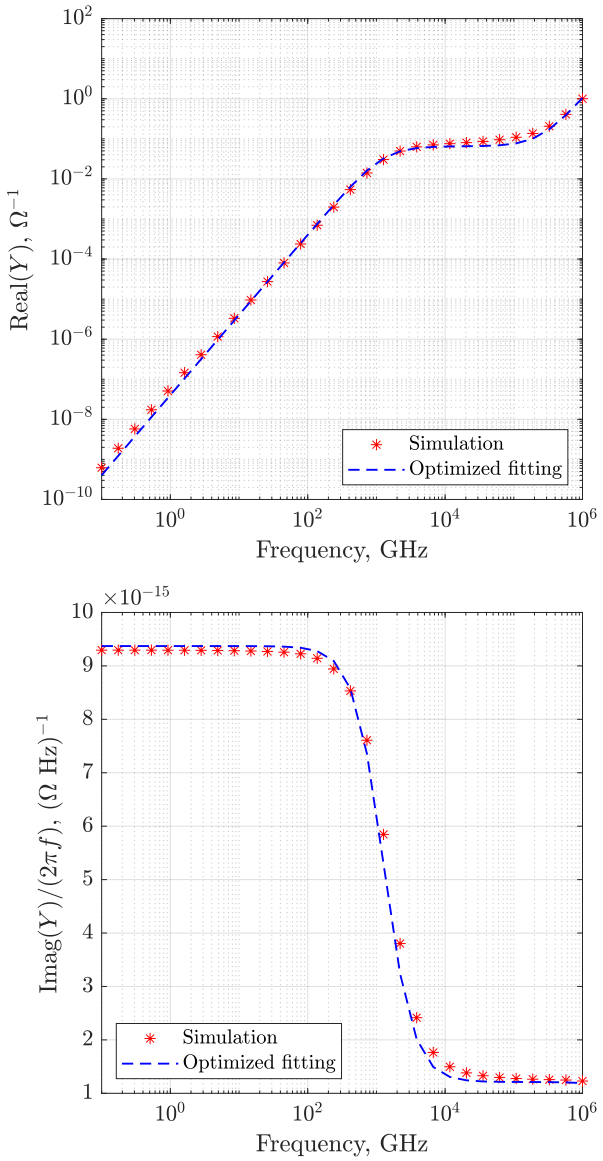


Figure 3.10: Optimized fitting (3.5) of the reference device admittance data at a $-2\,\mathrm{V}$ bias.

A comparison between the three described fittings is reported in figures 3.11.

The corresponding values found for the circuit parameters are reported in table 3.1 and are consistent with the equivalent resistance and capacitance measured by the literature [12], i.e., $R = 59\,\Omega$ (including the $50\,\Omega$ load) and $C = 6.5\,\text{fF}$ for the same device ($W_{\text{Ge}} = 100\,\text{nm}$, $L_{\text{Ge}} = 10\,\mu\text{m}$) with the same $-2\,\text{V}$ applied bias.

3.2.3 Small-signal AC response of microscopic quantities

The observed admittance frequency behaviour also finds confirmation in the AC response of microscopic quantities, such as the electrostatic potential or the current densities, computed as described in subsection 2.2.2. Figures 3.12 and 3.13 are cross-sections of the device in the xy plane showing such significant microscopic quantities at the extremes of the simulated frequency range, i.e., $f = 0.1\,\text{GHz}$ (*low frequency*) and $f = 1000\,\text{THz}$ (*high frequency*), respectively.

The first interesting effect which can be observed, for example, is the different AC response of the electrostatic potential at low and high frequency:

1. at low frequency the potential is constant in the whole silicon contacts and equal to the value set by the respective metal contacts (see figure 3.12a);
2. when the frequency gets too high, instead, only metal contacts are properly polarized, while a voltage drop is observed along the silicon contacts (see figure 3.13a).

This condition suggests the reason why a different equivalent capacitance is exhibited by the device at low and high frequency: while at low-frequency the intrinsic diode capacitance dominates, due to the well-polarized silicon contacts, at high frequency only the parasitic capacitance survives. This is also confirmed by the displacement current density response, which is higher where capacitive effects are more significant:

1. at low frequency the largest displacement current density can be observed within the intrinsic germanium layer, while significant contributions are also visible in the silicon intrinsic waveguide and in the oxide cladding, especially in the region between the metal contacts (see figure 3.12b);

R_1	$17.24\,\text{m}\Omega$
R_2	$15.34\,\Omega$
C_1	$1.214\,\text{fF}$
C_2	$8.157\,\text{fF}$

Table 3.1: Extracted small-signal parameters for the reference device at a $-2\,\text{V}$ bias.

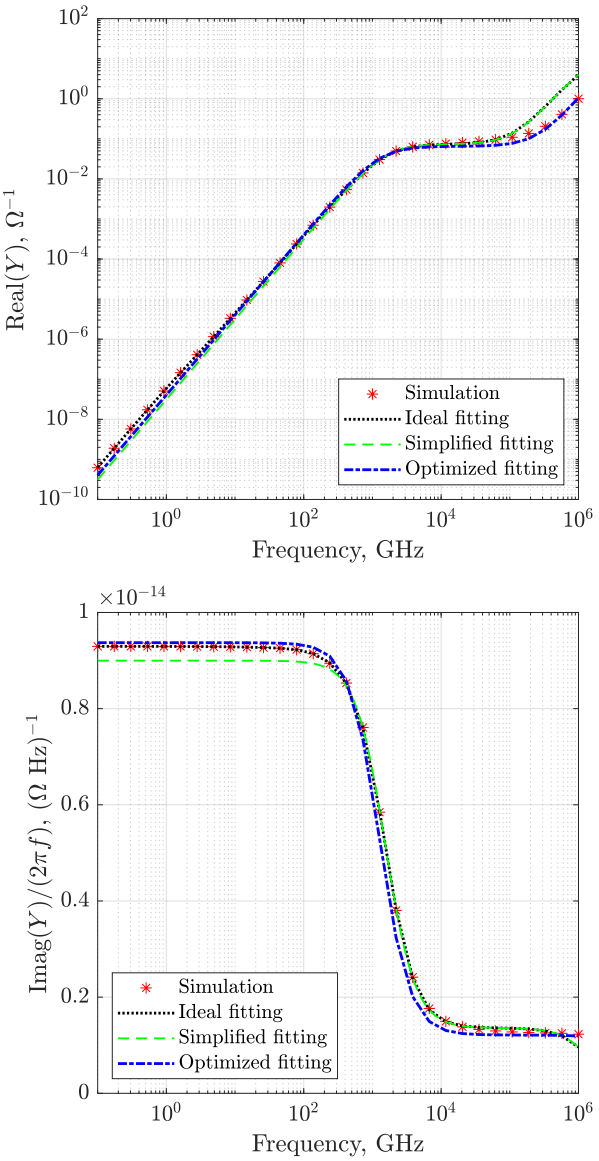
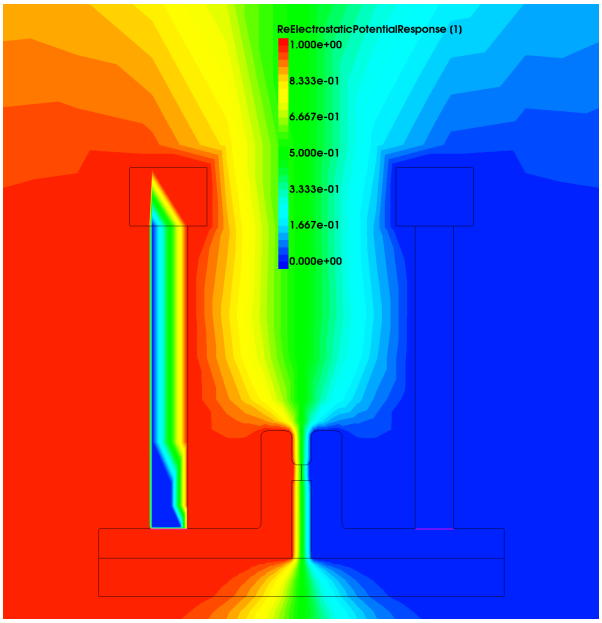
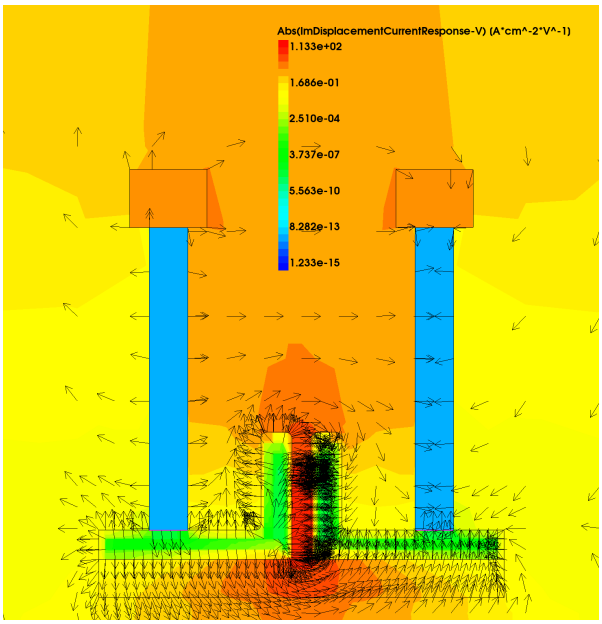


Figure 3.11: Comparison between ideal (3.1), simplified (3.2) and optimized (3.5) fittings of the reference device admittance data at a $-2\,\mathrm{V}$ bias.

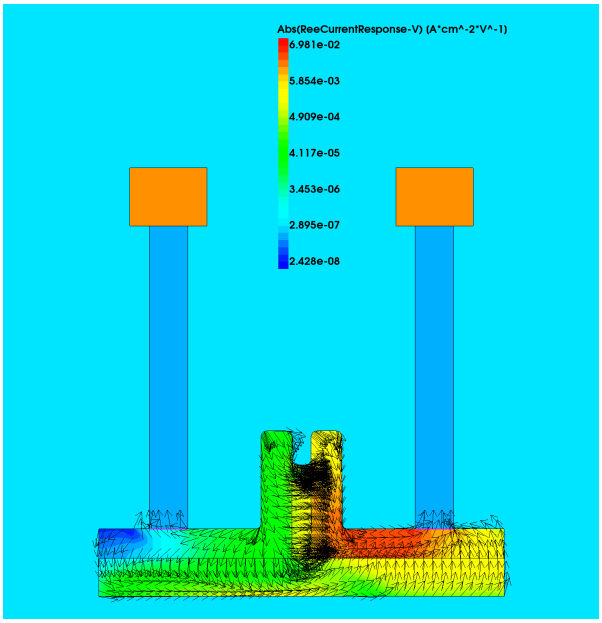


(a) Real part of the electrostatic potential AC response

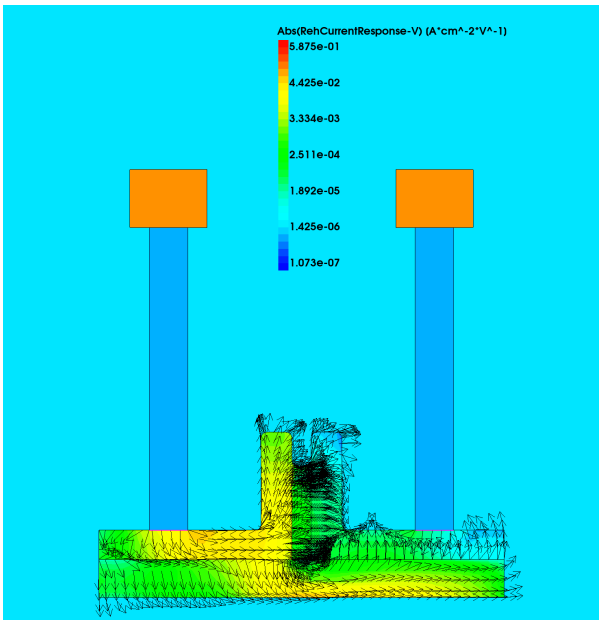


(b) Imaginary part of the displacement current density AC response

Figure 3.12: AC response of microscopic quantities at $f = 0.1\,\mathrm{GHz}$ and a $-2\,\mathrm{V}$ bias. Cross-sections of the reference device in the xy plane.

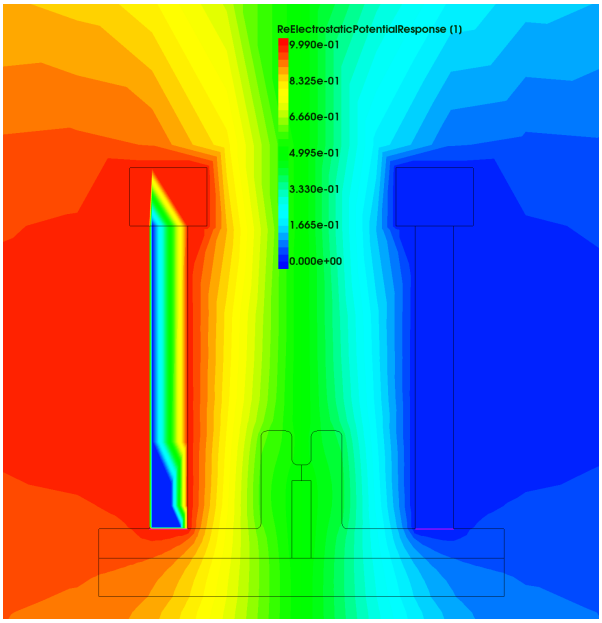


(c) Real part of the electron current density AC response

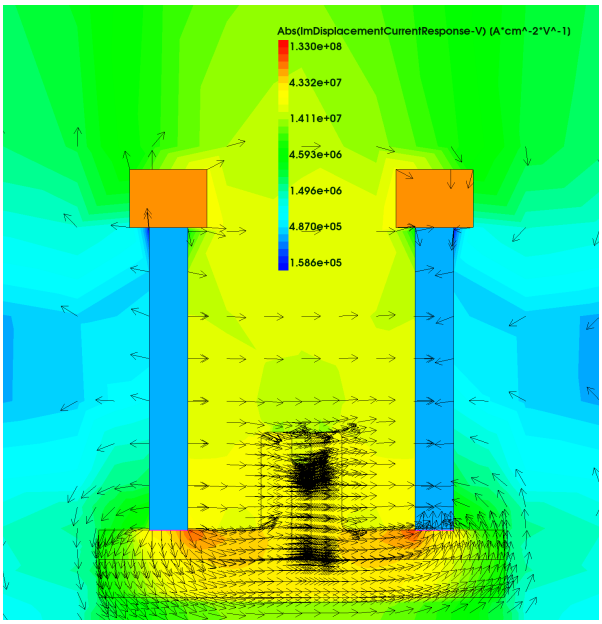


(d) Real part of the hole current density AC response

Figure 3.12: AC response of microscopic quantities at $f = 0.1\,\mathrm{GHz}$ and a $-2\,\mathrm{V}$ bias. Cross-sections of the reference device in the xy plane.

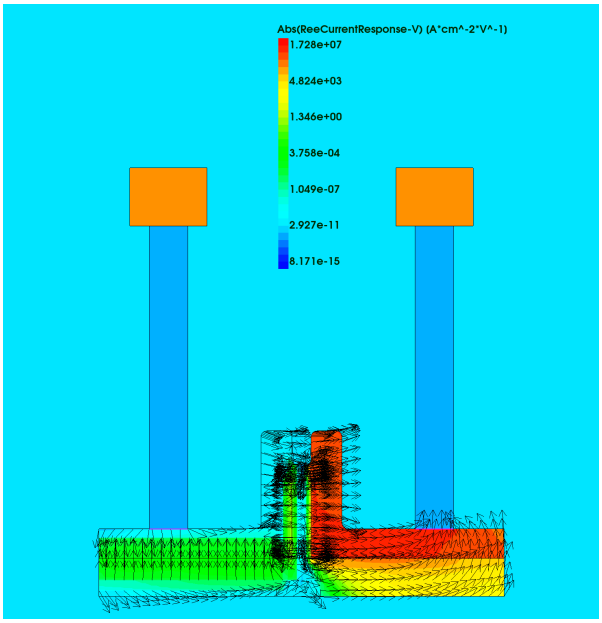


(a) Real part of the electrostatic potential AC response

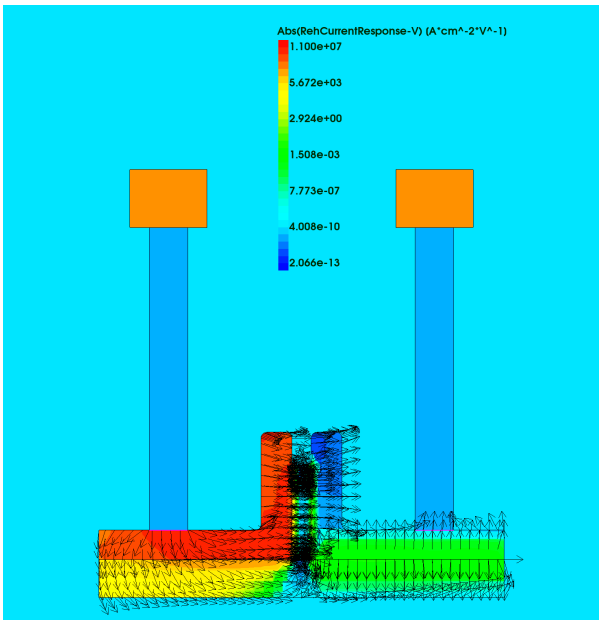


(b) Imaginary part of the displacement current density AC response

Figure 3.13: AC response of microscopic quantities at $f = 1000\,\mathrm{THz}$ and a $-2\,\mathrm{V}$ bias. Cross-sections of the reference device in the xy plane.



(c) Real part of the electron current density AC response



(d) Real part of the hole current density AC response

Figure 3.13: AC response of microscopic quantities at $f = 1000\,\mathrm{THz}$ and a $-2\,\mathrm{V}$ bias. Cross-sections of the reference device in the xy plane.

2. at high frequency the situation is quite different; in fact, since the silicon contacts are not properly polarized (see figure 3.13a), the displacement current is almost uniformly distributed not only in the oxide and in the silicon waveguide but also within the core of the device, i.e., the whole *pin* diode (see figure 3.13b); in practice, the intrinsic diode capacitance is not able to respond to the high frequency AC perturbation and a unique lower parasitic capacitance is observed between the two metal contacts.

As for the conductance, the AC response of electron and hole current density justifies the observed frequency behaviour, i.e., a difference of about 9 orders of magnitude can be observed between the peak current density values at $f = 0.1\,\mathrm{GHz}$ (see figures 3.12c and 3.12d) and $f = 1000\,\mathrm{THz}$ (see figures 3.13c and 3.13d), which is coherent with the conductance difference obtained from simulations (from $\sim 1 \times 10^{-9}\ \Omega^{-1}$ at $f = 0.1\,\mathrm{GHz}$ to $\sim 1\ \Omega^{-1}$ at $f = 1000\,\mathrm{THz}$).

3.2.4 Parametric small-signal AC analyses

Once a fitting procedure of simulated admittance data on a chosen circuit topology has been set up, AC analyses have been performed at five equally spaced bias points between $-2\,\mathrm{V}$ and $0\,\mathrm{V}$, for several versions of the device structure. In particular, the role played by the three dimensions of the germanium absorber (width W_{Ge}, height H_{Ge} and length L_{Ge}) has been investigated, as well as the influence of the presence and the position of metal contacts (the distance d between the tungsten contacts and the silicon sidewalls has been varied). Variations of such parameters (also indicated in figure 3.4) have allowed for the study of dependencies of the circuit parameters extracted from experimental data on the key dimensions of the device. It is important to remark that, due to the virtually infinite number of possible combinations, only one parameter at a time has been varied within a reasonable range, keeping all the remaining dimensions constant.

Effect of the bias voltage

Before proceeding with variations on the device structure, the effect of the applied bias voltage on the equivalent admittance of the device has been observed first. AC simulations have been run on the reference device at five equally spaced bias points between $-2\,\mathrm{V}$ and $0\,\mathrm{V}$, obtaining the results shown in figures 3.14. While the conductance and the parasitic (high-frequency) capacitance is practically independent of the bias, the intrinsic (low-frequency) capacitance exhibits a significant bias dependency. This is consistent with the theory: in fact, the intrinsic diode capacitance is mainly due to the intrinsic layer sandwiched between the p and n contacts, but it still includes a residual bias-dependent junction differential capacitance due to the presence of charged depleted regions [30, Sec. 7.1].

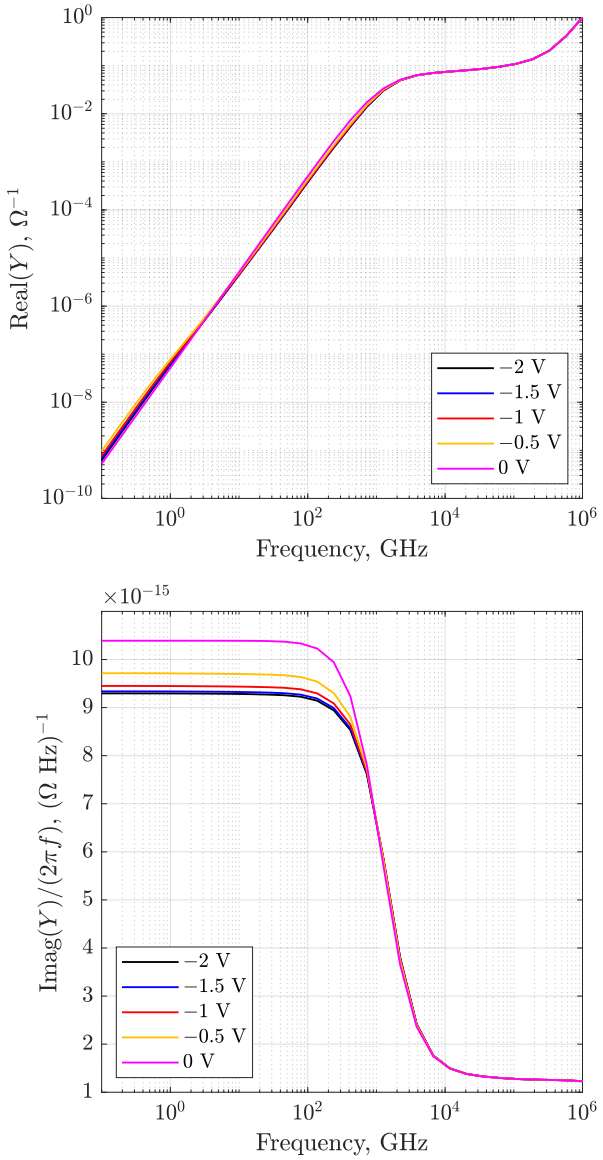


Figure 3.14: Effect of the bias voltage on the reference device admittance.

Effect of metal contacts

In order to understand the effect of metal contacts on the device equivalent admittance, AC simulations have been run on the usual reference device in three different versions:

1. in presence of both the copper and tungsten contacts (as shown in figure 3.4);
2. in presence of the tungsten contacts only (the copper ones have been substituted by the oxide cladding);
3. in absence of metal contacts (both tungsten and copper substituted by the oxide cladding).

Results of simulations at a $-2\,\mathrm{V}$ bias are shown in figures 3.15. Metal contacts only have a high-frequency effect on the conductance: the zero introduced by R_1 seems to be shifted towards higher frequencies. On the other hand, metal contacts increase the equivalent capacitance of the device in the whole frequency range, contributing to the parasitic term which sums to the device intrinsic capacitance.

The fitting procedure described in subsection 3.2.2 has been applied on the data obtained from simulations for all the bias values. The extracted circuit parameters are reported in figures A.1. As previously stated, R_1, R_2 and C_1, apart from negligible numerical errors introduced by the fitting procedure, are practically constant versus the bias, while C_2, which includes the junction differential capacitance, exhibits a dependency on the bias. It can be observed that:

1. metal contacts contribute significantly to R_1, while only marginally to R_2;
2. C_1 is more than doubled by the presence of metal contacts; in fact, C_1 represents the parasitic capacitance and it is increased by metal contacts, which form a sort of parallel-plate capacitor having as dielectric the cladding oxide;
3. C_2 is practically unaffected by metal contacts, being it the intrinsic diode capacitance.

Effect of the position of metal contacts

The position of metal contacts also affects the small-signal model. In order to understand its role, AC simulations have been run, fixing $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$ and $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, for two W_{Ge} values, i.e., $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$ and $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, and for four values of the distance d between the tungsten contacts and the silicon

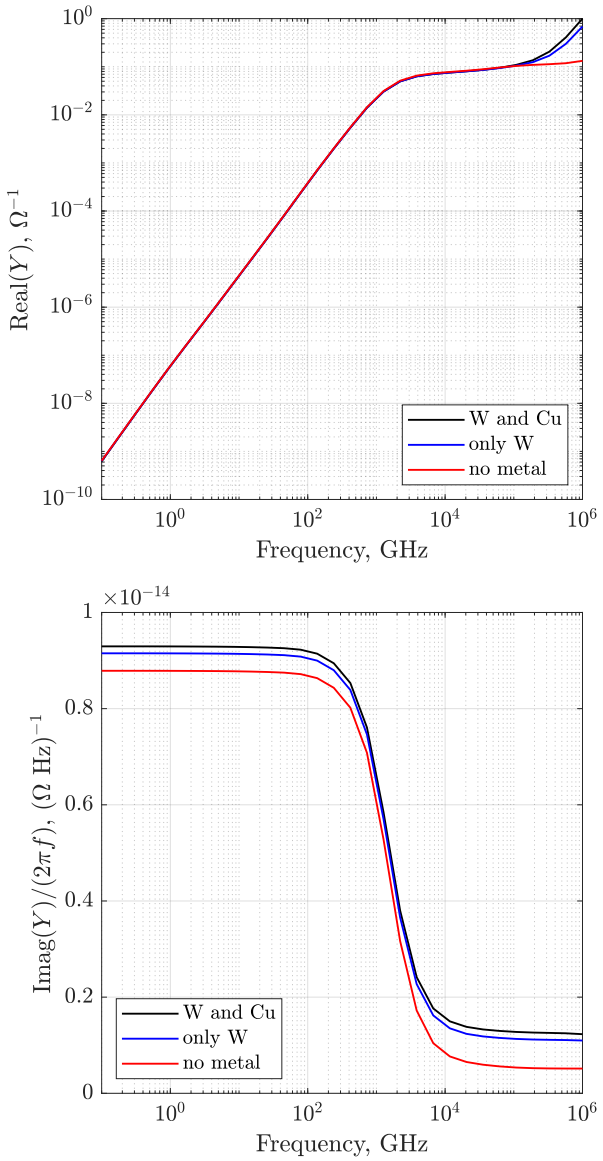


Figure 3.15: Effect of metal contacts on the reference device admittance at a -2 V bias.

sidewalls, equally spaced between 0.19 µm and 0.76 µm. The width of the silicon waveguide W_{SiWG} has been varied accordingly as:

$$W_{\mathrm{SiWG}} = 2\,\mu\mathrm{m} + W_{\mathrm{Ge}} + 2(d - 0.38\,\mu\mathrm{m})$$

Results of simulations at a $-2\,\mathrm{V}$ bias are shown in figures 3.16 and 3.17, for $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$ and $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, respectively. The relevant effect on the conductance can be observed in the range between 10^4 and 10^5 GHz, corresponding to the reciprocal of R_2, where it is inversely proportional to d: in fact, d impacts on the equivalent length of R_2, which is due to the silicon contacts. On the other hand, the effect on the equivalent device capacitance is limited to the parasitic one, which increases as d decreases: in fact, the thickness of the cladding oxide between the metal contacts, which is mainly contributing to the parasitic capacitance, is approximately $2d$, neglecting the width of the device core.

The extracted circuit parameters are reported in figures A.2 and A.3, for $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$ and $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, respectively. While no significant trend can be identified for R_1 and C_2, R_2 and C_1 follow the behaviour predicted from the obtained admittance results: the former is directly proportional to d, the latter is inversely proportional to d.

Effect of the absorber width

The effect of the germanium absorber width on the small-signal model has been investigated as well. AC simulations have been run on the same reference device ($H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$) for four values of W_{Ge}, equally spaced between 50 nm and 200 nm. The obtained admittance can be observed in figures 3.18. It is evident the large impact that W_{Ge} has on the low-frequency capacitance, i.e. on the intrinsic diode capacitance, which increases as W_{Ge} is reduced. The effect is reflected on the frequency behaviour of the conductance, i.e., the zero-pole couple is shifted towards lower frequencies due to the increased capacitance.

The extracted circuit parameters are reported in figures A.4. Some trends can be observed:

1. R_2 is slightly increased by a larger W_{Ge};
2. C_1 and C_2 are both inversely proportional to W_{Ge}; the effect on C_1 is only marginal, since the parasitic capacitance formed by metal contacts and the cladding is mainly influenced by the position of the contacts (the distance d, as discussed previously), while the intrinsic capacitance is strongly dependent on W_{Ge}, which is thickness of the intrinsic layer itself.

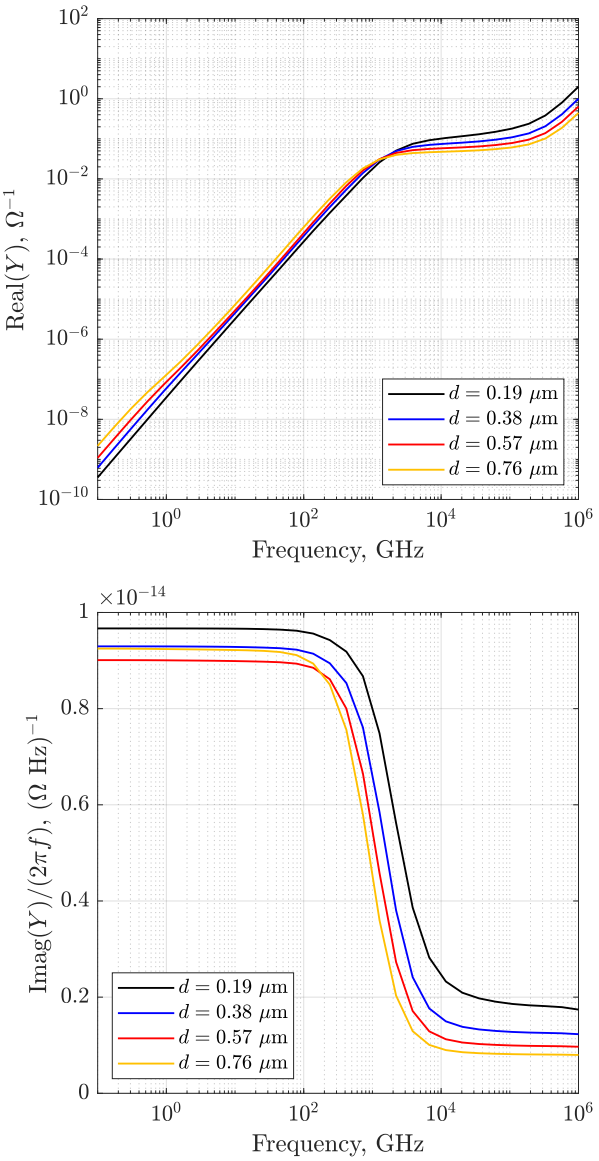


Figure 3.16: Effect of the position of metal contacts on the device admittance, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

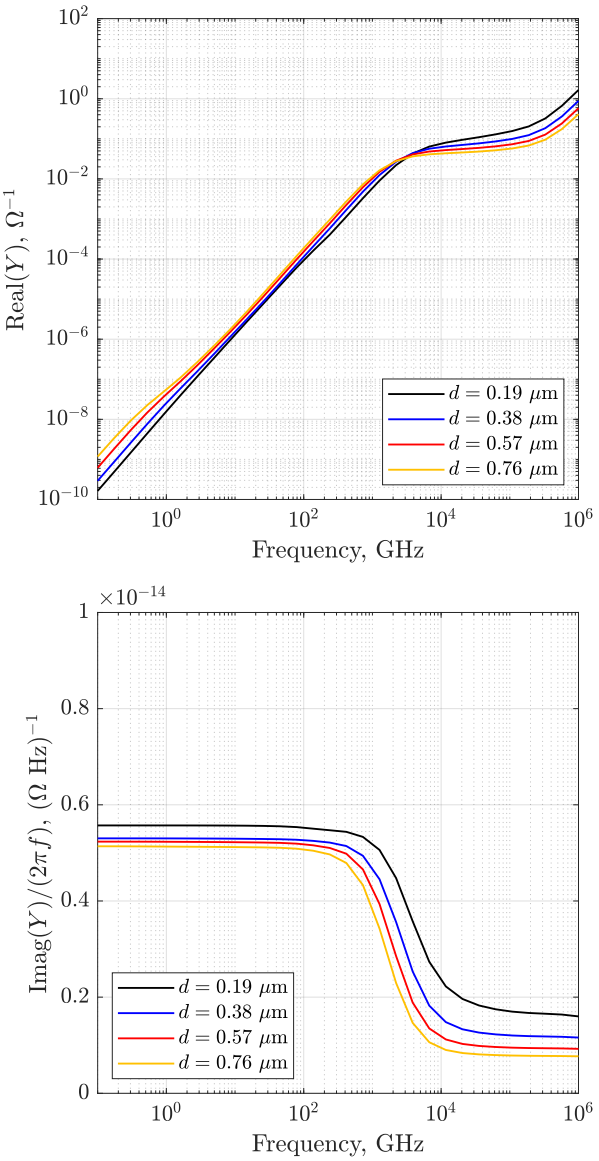


Figure 3.17: Effect of the position of metal contacts on the device admittance, with $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

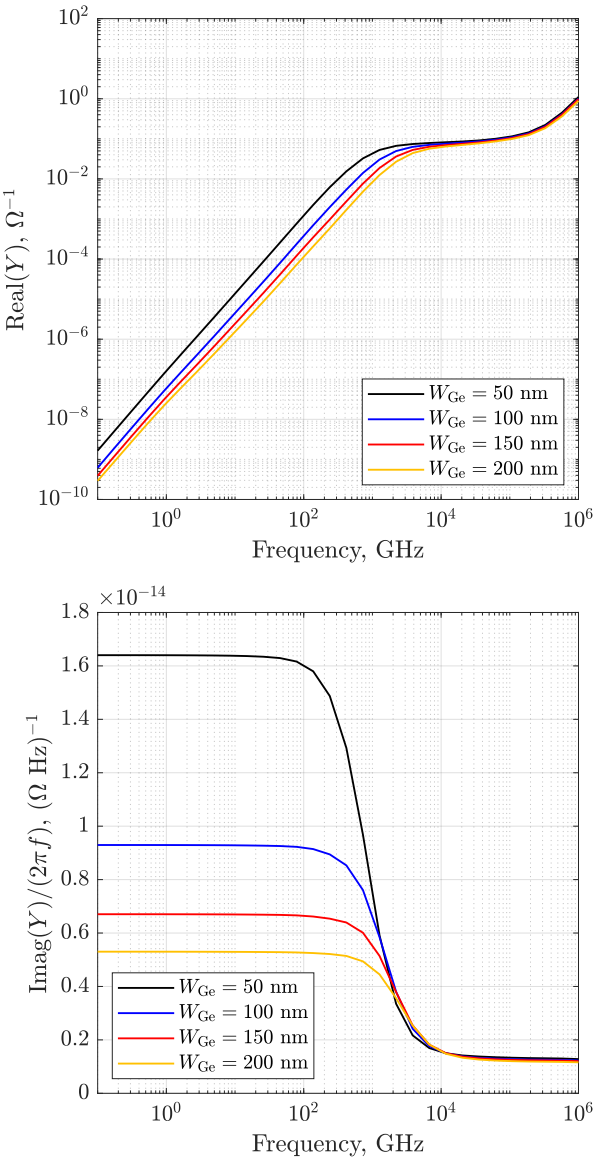


Figure 3.18: Effect of the absorber width W_{Ge} on the device admittance, with $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

Effect of the absorber height

The effect of the height of the germanium absorber has been analysed as well. AC simulations have been run, fixing $d = 0.38\,\mu\text{m}$ and $L_{\text{Ge}} = 10\,\mu\text{m}$, for two W_{Ge} values, i.e., $W_{\text{Ge}} = 100\,\text{nm}$ and $W_{\text{Ge}} = 200\,\text{nm}$, and for four values of H_{Ge}, equally spaced between 300 nm and 600 nm. Results of simulations at a $-2\,\text{V}$ bias are shown in figures 3.19 and 3.20, for $W_{\text{Ge}} = 100\,\text{nm}$ and $W_{\text{Ge}} = 200\,\text{nm}$, respectively. The main effect can be observed on the capacitance, directly proportional to H_{Ge}: in fact, the area of the intrinsic diode capacitance is given by the product $H_{\text{Ge}} \cdot L_{\text{Ge}}$.

The extracted small-signal circuit parameters are reported in figures A.5 and A.6, for $W_{\text{Ge}} = 100\,\text{nm}$ and $W_{\text{Ge}} = 200\,\text{nm}$, respectively. A strong dependency of C_2 on H_{Ge} can be observed. The effect on R_1 and C_1 is negligible and clear trends cannot be identified. As for R_2, a slight proportionality on H_{Ge} can be noticed as well.

Effect of the absorber length

At last, for the sake of completeness, the effect of the absorber length L_{Ge} on the admittance has been observed as well. AC simulations have been run on the same reference device ($W_{\text{Ge}} = 100\,\text{nm}$, $H_{\text{Ge}} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$) for three different device lengths: $5\,\mu\text{m}$, $10\,\mu\text{m}$, $15\,\mu\text{m}$. Results are reported in figures 3.21. As expected, both the conductance and the capacitance are directly proportional to L_{Ge}: in fact, a longer device results in larger areas, leading on one hand to a larger capacitance, on the other hand to a smaller resistance (or larger conductance).

This behaviour can be also observed in the extracted small-signal circuit parameters, reported in figures A.7.

3.2.5 Proposed small-signal analytical model

A small-signal analytical model based on the circuit introduced in subsection 3.2.2, the microscopic analysis hinted in subsection 3.2.3 and the parametric results shown in subsection 3.2.4 has been developed, in order to allow for a direct computation of the circuit parameters, given the device key dimensions d, W_{Ge}, H_{Ge}, L_{Ge}, for a reverse bias of $-2\,\text{V}$.

Model for R_1

The series parasitic resistance R_1 has been shown to be approximately constant when d, W_{Ge} and H_{Ge} are varied, while a significant dependency on L_{Ge} has

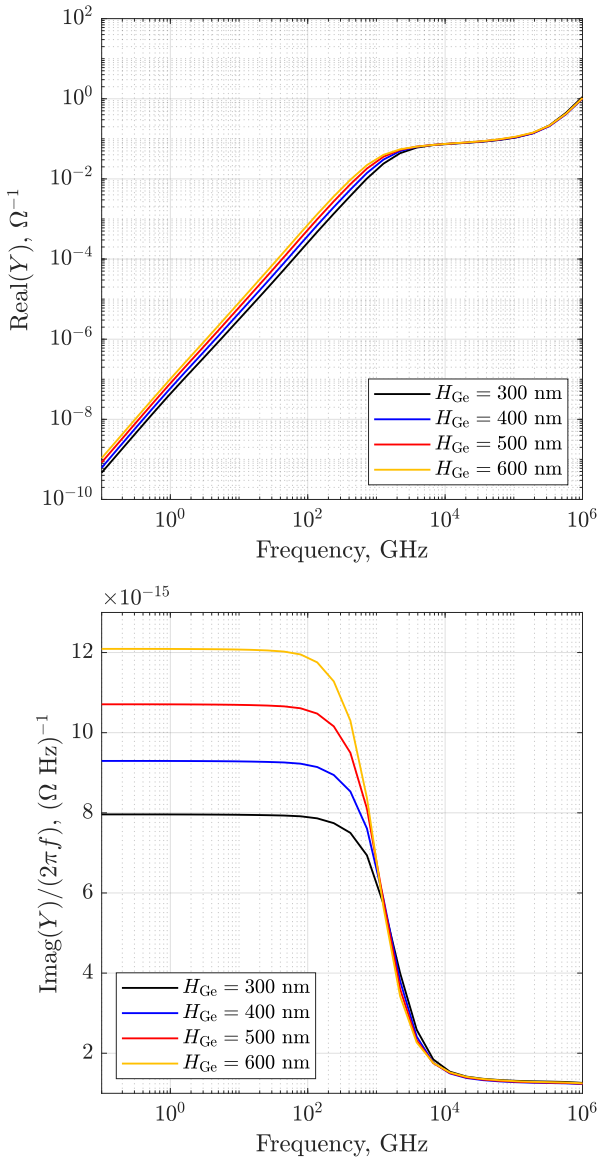


Figure 3.19: Effect of the absorber height H_{Ge} on the device admittance, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

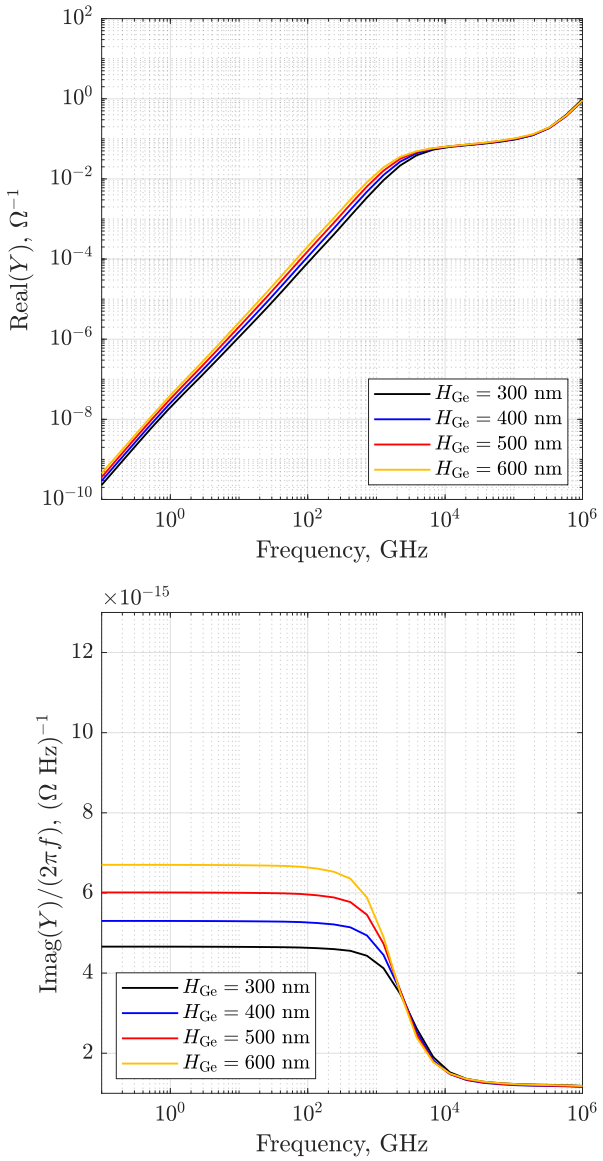


Figure 3.20: Effect of the absorber height H_{Ge} on the device admittance, with $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

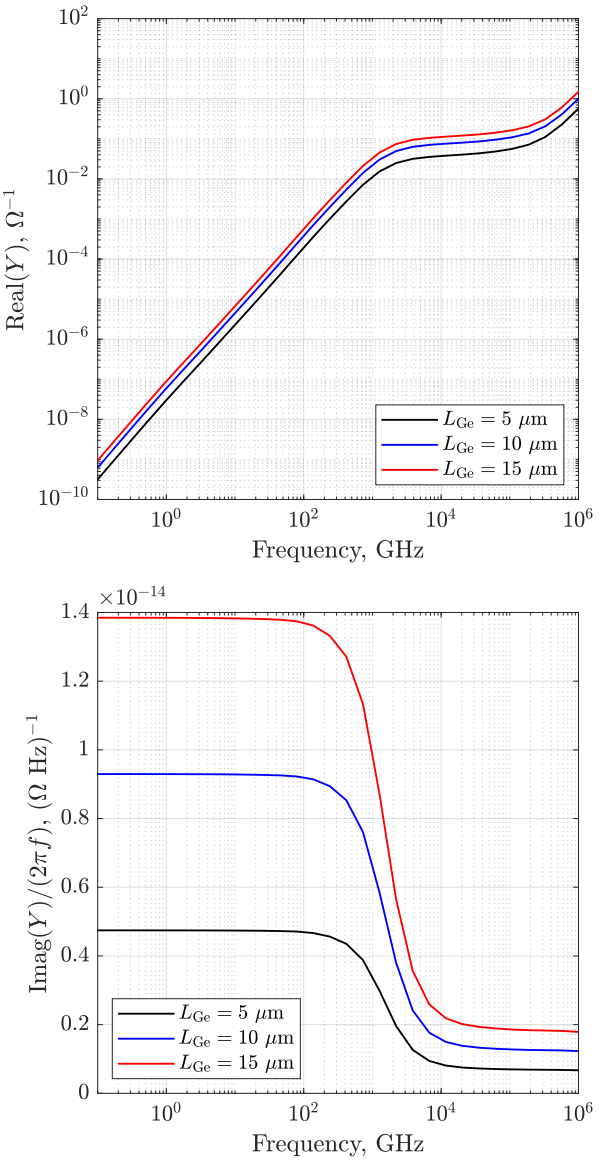


Figure 3.21: Effect of the absorber length L_{Ge} on the device admittance, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

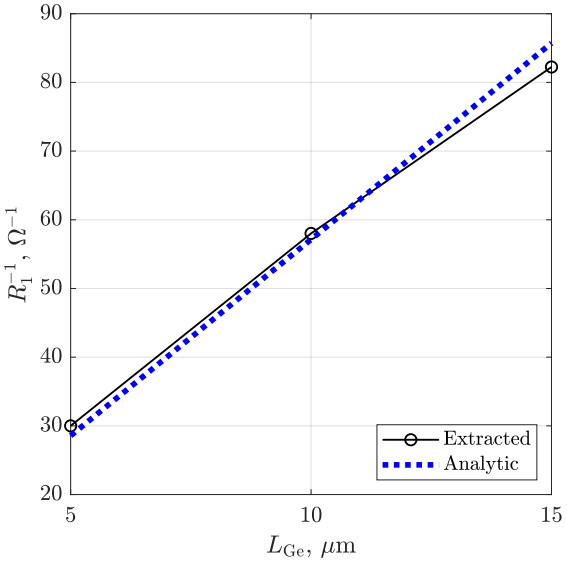


Figure 3.22: Comparison between the analytic model for R_1 (3.6) and the extracted values for $W_{Ge} = 100\,\text{nm}$, $H_{Ge} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$, $-2\,\text{V}$ bias (A.7a).

been observed. Therefore, the proposed model for R_1 is:

$$R_1(L_{Ge}) = \frac{K}{L_{Ge}} \tag{3.6}$$

with $K = 1.75 \times 10^{-7}\ \Omega \cdot \text{m}$ and L_{Ge} expressed in metres. The value of the constant K corresponds to the product of the average value found for R_1 in all the simulations performed for fixed $L_{Ge} = 10\,\mu\text{m}$, i.e., $17.5\,\text{m}\Omega$, and $L_{Ge} = 10\,\mu\text{m}$. A good agreement between this simple model and the values extracted from simulations (for $W_{Ge} = 100\,\text{nm}$, $H_{Ge} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$, $-2\,\text{V}$ bias, as reported in figure A.7) in the L_{Ge} range between $5\,\mu\text{m}$ and $15\,\mu\text{m}$ can be observed in figure 3.22.

Model for R_2

The series parasitic diode resistance R_2, which is due to the heavily doped p and n silicon contacts, exhibits several dependencies on the device dimensions. These can be justified by considering the device structure and the definition of resistance as:

$$R_2 = \rho \frac{L}{A}$$

where ρ, L and A are the resistivity, the equivalent length and the equivalent area of the resistance, respectively.

n	7.268×10^{19} cm^{-3}
p	7.004×10^{19} cm^{-3}
μ_n	60.864 cm^2/(V $\cdot$ s)
μ_p	49.862 cm^2/(V $\cdot$ s)

Table 3.2: Carrier density and mobility values extracted from simulations, used to evaluate the resistivity ρ of silicon contacts, needed to evaluate R_2.

1. The resistivity ρ has been evaluated as the sum of the resistivity of both the p and n silicon contacts, ρ_n and ρ_p, respectively:

$$\rho = \rho_n + \rho_p$$

The resistivity ρ_n (ρ_p) corresponds to the reciprocal of the conductivity σ_n (σ_p), evaluated as follows:

$$\rho_n = \frac{1}{\sigma_n} = \frac{1}{qn\mu_n} \tag{3.7a}$$

$$\rho_p = \frac{1}{\sigma_p} = \frac{1}{qp\mu_p} \tag{3.7b}$$

where q is the elementary charge, n (p) and μ_n (μ_p) are the average electron (hole) density and the average electron (hole) mobility within the silicon contacts, respectively. The values extracted from simulations, reported in table 3.2, have been used for the quantities above. The evaluated resistivity is $\rho = 3.2 \times 10^{-5}\ \Omega \cdot \text{m}$.

2. The equivalent length L of the resistance R_2 includes several contributions due to the device geometrical shape. As observed from the microscopic analysis (e.g., figure 3.12), the current density is mainly concentrated between the tungsten contact and the silicon sidewall: the resistive region is L-shaped, therefore the resistance length definition is not straightforward. In order to get a good fitting of the extracted R_2 values, some fitting parameters have been introduced as well. The final expression for L is the following one:

$$L = W_{\text{SiSW}} + 1.4d + 0.15H_{\text{Ge}} + 1.5W_{\text{Ge}} + (5 \times 10^{-8}\ \text{m}) \tag{3.8}$$

where W_{SiSW} is the width of the silicon sidewalls, as indicated in figure 3.23, and W_{SiSW}, d, H_{Ge} and W_{Ge} are all expressed in metres. Since in the whole analysis W_{SiSW} has always been kept constant to 0.16 µm, the expression of L can be rewritten as

$$L = 1.4d + 0.15H_{\text{Ge}} + 1.5W_{\text{Ge}} + (2.1 \times 10^{-7}\ \text{m}) \tag{3.9}$$

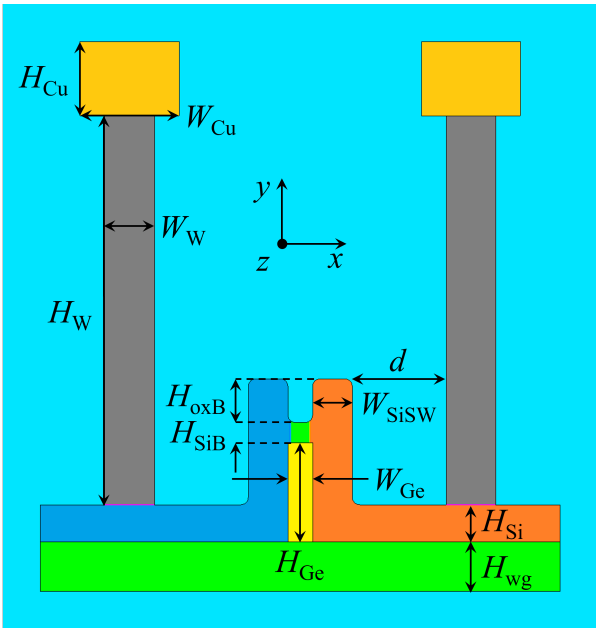


Figure 3.23: Simulated Ge-on-Si lateral *pin* photodetector cross-section in the *xy* plane. Key dimensions of the device involved in the proposed small-signal analytic model are indicated.

3. The equivalent area A of R_2 has been assumed to be simply the area of the tungsten contacts:
$$A = W_{\mathrm{W}} L_{\mathrm{Ge}}$$
where W_{W} is the width of the tungsten contacts, kept constant to 0.2 µm.

Therefore, the proposed model for R_2 is

$$\begin{aligned} &R_2(d, W_{\mathrm{Ge}}, H_{\mathrm{Ge}}, L_{\mathrm{Ge}}) = \\ &= \left(\frac{1}{qn\mu_n} + \frac{1}{qp\mu_p}\right) \frac{W_{\mathrm{SiSW}} + 1.4d + 0.15H_{\mathrm{Ge}} + 1.5W_{\mathrm{Ge}} + (5 \times 10^{-8}\ \mathrm{m})}{W_{\mathrm{W}} L_{\mathrm{Ge}}} \end{aligned} \tag{3.10}$$

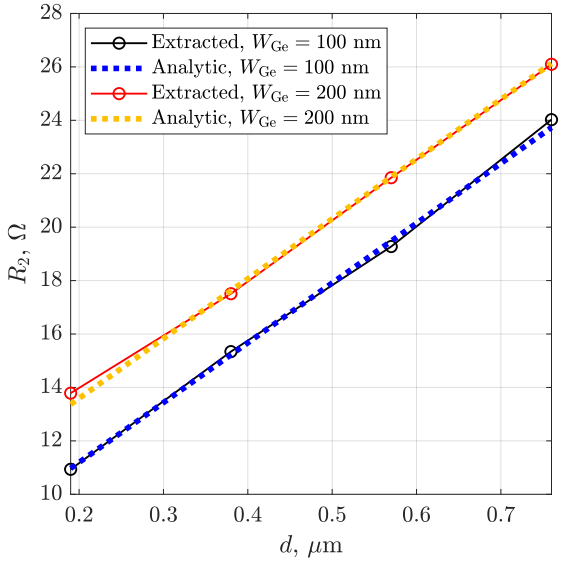
which can be rewritten as follows if the values of the parameters kept constant are substituted into the expression

$$R_2(d, W_{\mathrm{Ge}}, H_{\mathrm{Ge}}, L_{\mathrm{Ge}}) = \frac{1.4d + 0.15H_{\mathrm{Ge}} + 1.5W_{\mathrm{Ge}} + K_1}{K_2 \cdot L_{\mathrm{Ge}}} \tag{3.11}$$

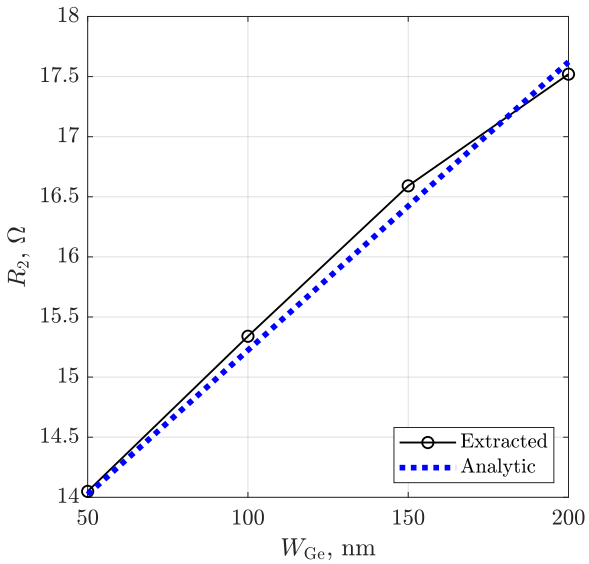
with all the dimensions expressed in metres, $K_1 = 2.1 \times 10^{-7}$ m and $K_2 = 6.25 \times 10^{-3}\ \Omega^{-1}$. The agreement between this model and the values extracted from simulations can be observed in figures 3.24.

Model for C_1

The parasitic capacitance C_1 is the high-frequency capacitance of the device. As it has been observed in the microscopic analysis and, in particular, in the displacement current density high-frequency response (figure 3.13b), at high frequency the device behaves like a sort of parallel-plate capacitor with a varying dielectric. The proposed model for C_1 is based exactly on this idea: the capacitance has been expressed as the parallel of several capacitances, some of which are in turn expressed as the series of more than one capacitance. The dimensions involved in the evaluation of C_1 are all indicated in figure 3.23. The

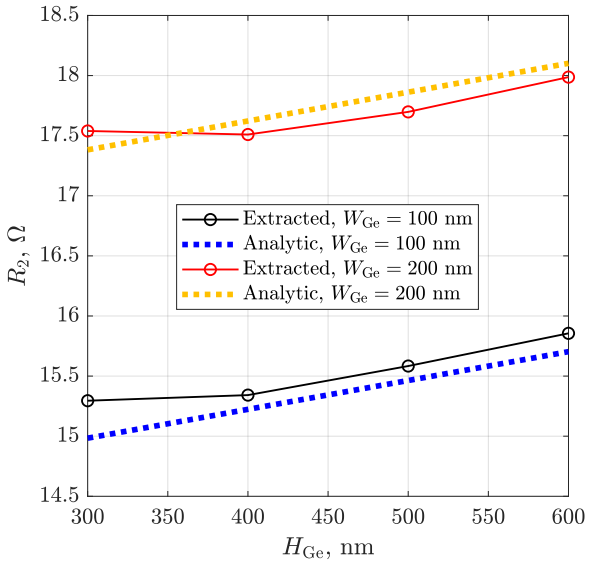


(a) Extracted values for $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$ (A.2b, A.3b).

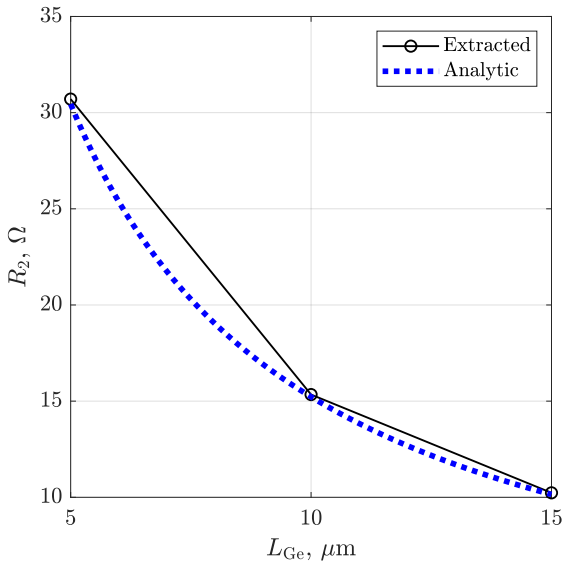


(b) Extracted values for $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$ (A.4b).

Figure 3.24: Comparison between the analytic model for R_2 (3.11) and the extracted values for a $-2\,\mathrm{V}$ bias.



(c) Extracted values for $d = 0.38\,\mu m$, $L_{Ge} = 10\,\mu m$ (A.5b, A.6b).



(d) Extracted values for $W_{Ge} = 100\,nm$, $H_{Ge} = 400\,nm$, $d = 0.38\,\mu m$ (A.7b).

Figure 3.24: Comparison between the analytic model for R_2 (3.11) and the extracted values for a $-2\,V$ bias.

proposed expression is the following one:

$$
\begin{aligned}
&C_1(d, W_{\mathrm{Ge}}, H_{\mathrm{Ge}}, L_{\mathrm{Ge}}) = \\
&= L_{\mathrm{Ge}} \cdot \left\{ \frac{\varepsilon_{\mathrm{ox}}}{\pi} \log\left(1 + \frac{2W_{\mathrm{Cu}}}{2(d + W_{\mathrm{SiSW}}) + W_{\mathrm{Ge}} - (W_{\mathrm{Cu}} - W_{\mathrm{W}})}\right) + \right. \\
&+ \frac{\varepsilon_{\mathrm{ox}} H_{\mathrm{Cu}}}{2(d + W_{\mathrm{SiSW}}) + W_{\mathrm{Ge}} - (W_{\mathrm{Cu}} - W_{\mathrm{W}})} + \\
&+ \frac{\varepsilon_{\mathrm{ox}} \left[H_{\mathrm{W}} - H_{\mathrm{oxB}} - H_{\mathrm{SiB}} - (H_{\mathrm{Ge}} - H_{\mathrm{Si}})\right]}{2(d + W_{\mathrm{SiSW}}) + W_{\mathrm{Ge}}} + \\
&+ H_{\mathrm{oxB}} \left(\frac{2d + W_{\mathrm{Ge}}}{\varepsilon_{\mathrm{ox}}} + \frac{2W_{\mathrm{SiSW}}}{\varepsilon_{\mathrm{Si}}}\right)^{-1} + \\
&+ H_{\mathrm{SiB}} \left(\frac{2d}{\varepsilon_{\mathrm{ox}}} + \frac{2W_{\mathrm{SiSW}} + W_{\mathrm{Ge}}}{\varepsilon_{\mathrm{Si}}}\right)^{-1} + \\
&+ (H_{\mathrm{Ge}} - H_{\mathrm{Si}}) \left(\frac{2d}{\varepsilon_{\mathrm{ox}}} + \frac{2W_{\mathrm{SiSW}}}{\varepsilon_{\mathrm{Si}}} + \frac{W_{\mathrm{Ge}}}{\varepsilon_{\mathrm{Ge}}}\right)^{-1} + \\
&+ H_{\mathrm{Si}} \left(\frac{2(d + W_{\mathrm{SiSW}})}{\varepsilon_{\mathrm{Si}}} + \frac{W_{\mathrm{Ge}}}{\varepsilon_{\mathrm{Ge}}}\right)^{-1} + \\
&+ \frac{\varepsilon_{\mathrm{Si}} H_{\mathrm{wg}}}{2(d + W_{\mathrm{SiSW}}) + W_{\mathrm{Ge}}} + \\
&+ \left. K_1 \log\left(1 + \frac{K_2}{d}\right) \right\}
\end{aligned}
\tag{3.12}
$$

with all the dimensions expressed in metres, $K_1 = 4.665 \times 10^{-12}\ \mathrm{F \cdot m^{-1}}$ and $K_2 = 3.824 \times 10^{-5}$ m, and where several parameters, i.e., the dielectric constants and some fixed device dimensions, have been kept constant to the values reported in table 3.3.

The terms in equation (3.12), i.e., the capacitances put in the parallel, are described in the details as follows. Starting from the top:

1. the coplanar capacitance formed by the top of the copper contacts and the cladding oxide has been expressed as inspired by [31], i.e., given two coplanar plates wide W and distant D, separated by a material with dielectric constant ε, the corresponding capacitance can be computed as

$$C = \frac{\varepsilon}{\pi} \log\left(1 + \frac{2W}{D}\right) \tag{3.13}$$

   in this case, $W = W_{\mathrm{Cu}}$ is the width of the copper contacts, while the distance between them can be written as

$$D = 2(d + W_{\mathrm{SiSW}}) + W_{\mathrm{Ge}} - (W_{\mathrm{Cu}} - W_{\mathrm{W}})$$

2. the parallel-plate capacitance formed by the sides of the copper contacts and the cladding oxide; H_{Cu} is the height of the copper contacts, which is

$\varepsilon_{ox}/\varepsilon_0$	3.9
$\varepsilon_{Si}/\varepsilon_0$	11.7
$\varepsilon_{Ge}/\varepsilon_0$	16.2
H_{Cu}	0.300 µm
W_{Cu}	0.400 µm
H_W	1.570 µm
W_W	0.200 µm
H_{oxB}	0.176 µm
H_{SiB}	0.082 µm
W_{SiSW}	0.160 µm
H_{Si}	0.150 µm
H_{wg}	0.200 µm

Table 3.3: Materials dielectric constants and fixed device dimensions used to evaluate C_1.

the width of the plates; the dielectric thickness corresponds to the distance between the copper contacts, expressed as above;

3. the parallel-plate capacitance formed by the sides of the tungsten contacts and the cladding oxide; the width of the plates is expressed as

$$H_W - H_{oxB} - H_{SiB} - (H_{Ge} - H_{Si})$$

where H_W is the height of the tungsten contacts, H_{oxB} is the height of the oxide between the top of the silicon sidewalls, H_{SiB} is the height of the silicon bridge above the germanium absorber and H_{Si} is the height of the silicon contacts; the dielectric thickness is

$$2(d + W_{SiSW}) + W_{Ge}$$

4. the parallel-plate capacitance formed by the sides of the tungsten contacts, the cladding oxide and the upper part of the silicon sidewalls, expressed as the series of two capacitances; the plates of such a capacitance are H_{oxB} wide; the oxide contributes with a thickness of $2d + W_{Ge}$, while silicon with $2W_{SiSW}$;

5. the parallel-plate capacitance formed by the sides of the tungsten contacts, the cladding oxide, part of the silicon sidewalls and the silicon bridge above the germanium absorber, expressed as the series of two capacitances; the plates of such a capacitance are H_{SiB} wide; the oxide contributes with a thickness of $2d$, while silicon with $2W_{SiSW} + W_{Ge}$;

6. the parallel-plate capacitance formed by the sides of the tungsten contacts, the cladding oxide, part of the silicon sidewalls and part of the germanium absorber, expressed as the series of three capacitances; the plates of such

Term	$d = 0.38\,\mu m$	$d = 0.57\,\mu m$
1	0.054	0.054
2	0.087	0.081
3	0.256	0.251
4	0.058	0.052
5	0.031	0.027
6	0.080	0.077
7	0.111	0.108
8	0.145	0.142
9	0.178	0.211

Table 3.4: Relative weights of the nine terms contributing to the analytical expression of C_1 (3.12), computed for $W_{Ge} = 100\,nm$, $H_{Ge} = 400\,nm$, $L_{Ge} = 10\,\mu m$, two different values of d, $-2\,V$ bias.

a capacitance are $H_{Ge} - H_{Si}$ wide; the oxide contributes with a thickness of $2d$, silicon with $2W_{SiSW}$, germanium with W_{Ge};

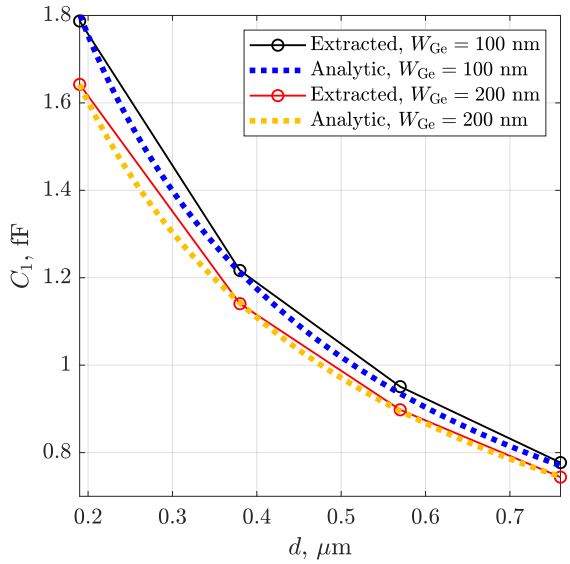
7. the capacitance formed by the tungsten contacts, the silicon contacts and part of the germanium absorber, which can be approximated as a parallel-plate capacitance, due to the horizontal field lines observed in figure 3.13b; such a capacitance is H_{Si} wide and can be expressed as the series of two capacitances; silicon contributes with a thickness of $2(d + W_{SiSW})$, while germanium with W_{Ge};

8. the capacitance formed by the tungsten contacts and the silicon waveguide, which can be approximated as a parallel-plate capacitance, due to the horizontal field lines observed in figure 3.13b; such a capacitance is H_{wg} wide, where H_{wg} is the thickness of the silicon waveguide, and $2(d + W_{SiSW}) + W_{Ge}$ thick;

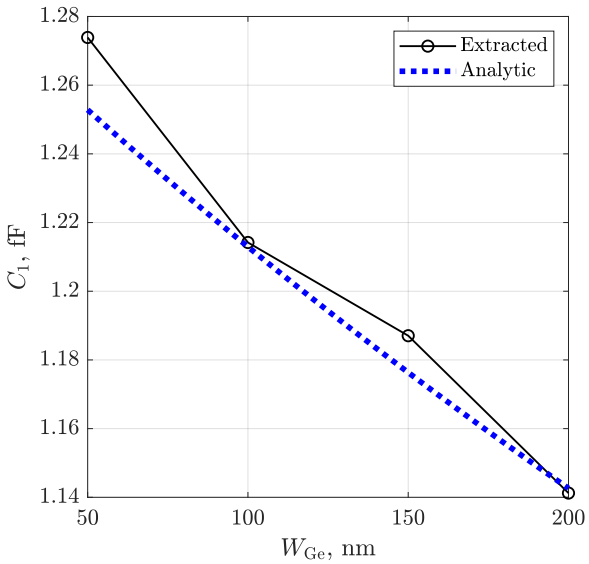
9. a fitting term, depending on d with a behaviour similar to equation (3.13), added in order to ensure a better fitting of the extracted results and accounting for the observed fringing field not included in previous terms.

The agreement between this model and the values extracted from simulations can be observed in figures 3.25 and 3.26.

Furthermore, the relative weights of each of the nine terms contributing to the analytical expression of C_1 are reported in table 3.4 for the reference device and a second device, identical to the reference one, except for $d = 0.57\,\mu m$. It can be observed that major contributions are due to terms 3 (parallel-plate capacitance formed by the sides of the tungsten contacts and the cladding oxide), 8 (capacitance formed by the tungsten contacts and the silicon waveguide) and 9 (fitting term).

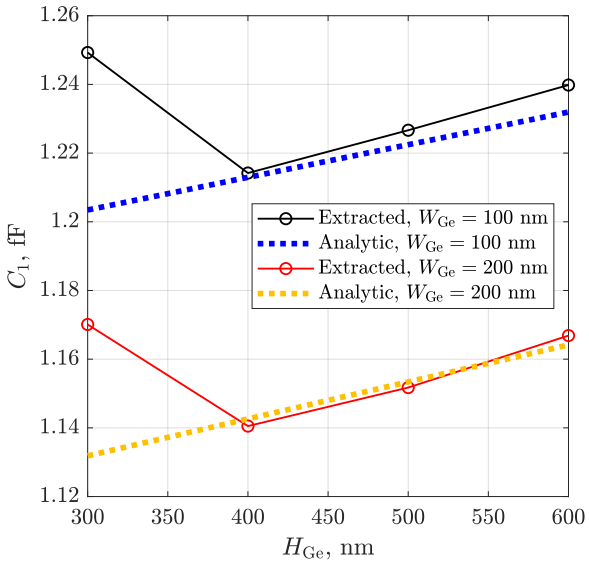


(a) Extracted values for $H_{Ge} = 400\,\text{nm}$, $L_{Ge} = 10\,\mu\text{m}$ (A.2c, A.3c).



(b) Extracted values for $H_{Ge} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$, $L_{Ge} = 10\,\mu\text{m}$ (A.4c).

Figure 3.25: Comparison between the analytic model for C_1 (3.12) and the extracted values for a $-2\,\text{V}$ bias.



(c) Extracted values for $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$ (A.5c, A.6c).

Figure 3.25: Comparison between the analytic model for C_1 (3.12) and the extracted values for a $-2\,\mathrm{V}$ bias.

Model for C_2

A model for the intrinsic diode capacitance C_2 has been developed with the same approach used for C_1, i.e., observing microscopic results (figure 3.12b) and trying to include every relevant contribution analytically. The proposed model for C_2 is

$$\begin{aligned} &C_2(W_{\mathrm{Ge}}, H_{\mathrm{Ge}}, L_{\mathrm{Ge}}) = \\ &= L_{\mathrm{Ge}} \cdot \left\{ \frac{\varepsilon_{\mathrm{ox}}}{\pi} \log\left(1 + \frac{2W_{\mathrm{SiSW}}}{W_{\mathrm{Ge}}}\right) + \right. \\ &+ \frac{\varepsilon_{\mathrm{ox}} H_{\mathrm{oxB}} + \varepsilon_{\mathrm{Si}} H_{\mathrm{SiB}} + \varepsilon_{\mathrm{Ge}} H_{\mathrm{Ge}}}{W_{\mathrm{Ge}}} + \\ &\left. + \frac{\varepsilon_{\mathrm{Si}}}{\pi} \log\left(1 + \frac{K}{W_{\mathrm{Ge}}}\right) \right. \end{aligned} \tag{3.14}$$

with all the dimensions expressed in metres, $K = 1 \times 10^{-6}$ m and fixed parameters reported in table 3.3. The model simply expresses the parallel of five capacitances. Starting from the top:

1. the coplanar capacitance formed by the top of the silicon sidewalls, wide W_{SiSW} and distant W_{Ge}, expressed using equation (3.13);

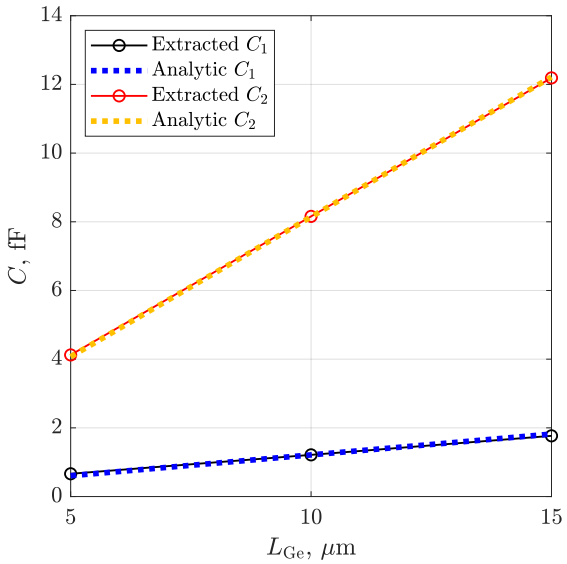


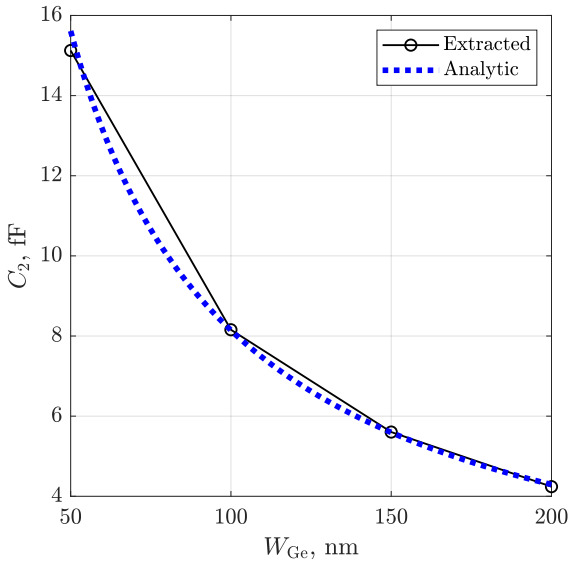
Figure 3.26: Comparison between the analytic models for C_1 (3.12) and C_2 (3.14) and the extracted values for $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

2. the parallel-plate capacitance formed by the upper part of the silicon sidewalls and the cladding oxide, wide H_{oxB} and thick W_{Ge};
3. the parallel-plate capacitance formed by part of the silicon sidewalls and the intrinsic silicon bridge above the germanium absorber, wide H_{SiB} and thick W_{Ge};
4. the parallel-plate capacitance formed by part of the silicon sidewalls and the intrinsic germanium absorber, wide H_{Ge} and thick W_{Ge};
5. the coplanar capacitance formed by the bottom of the silicon contacts, expressed using equation (3.13), assuming a width of the plates of $0.5\,\mu\mathrm{m}$ and a distance W_{Ge}.

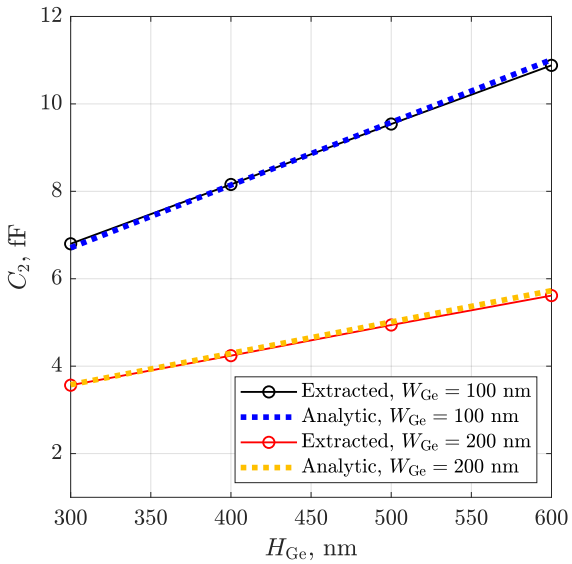
The coherence between the model and the extracted results can be observed in figures 3.27 and 3.26.

3.2.6 Small-signal AC analysis under illumination

The small-signal analytical model previously proposed has been developed from AC simulations in dark conditions. AC simulations have been performed under illumination as well on the same reference device ($H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$,



(a) Extracted values for $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$ (A.4d).



(b) Extracted values for $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$ (A.5d, A.6d).

Figure 3.27: Comparison between the analytic model for C_2 (3.14) and the extracted values for a $-2\,\mathrm{V}$ bias.

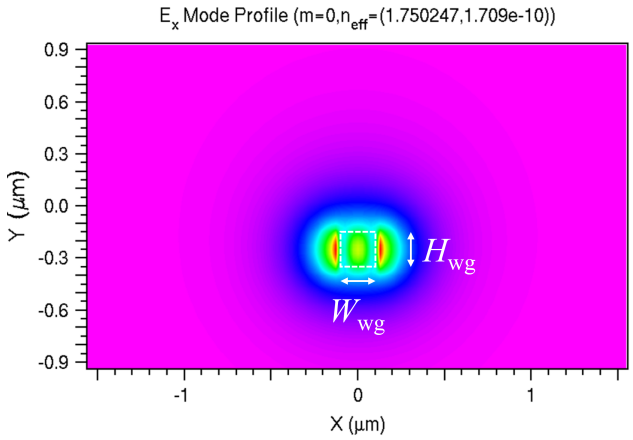


Figure 3.28: Fundamental mode profile (E_x contour map) at $\lambda = 1.31\,\mu\text{m}$ computed on the silicon input waveguide (squared cross-section, $W_{\text{wg}} = H_{\text{wg}} = 0.2\,\mu\text{m}$). The effective refractive index of the mode is $n_{\text{eff}} = 1.750$.

$L_{\text{Ge}} = 10\,\mu\text{m}$) for four values of W_{Ge}, at a $-2\,\text{V}$ bias. Specifically, a light source with a low input optical power $P_{\text{op}} = 200\,\mu\text{W}$ at a wavelength $\lambda = 1.31\,\mu\text{m}$ (O-Band) has been coupled to the device through a $1\,\mu\text{m}$ long input segment of the silicon waveguide, having a squared cross-section, with a 200 nm side. The mode profile computed on the input waveguide and then propagated by the electromagnetic simulation is shown in figure 3.28.

The obtained equivalent device admittance is shown in figures 3.29.

It can be observed that illumination marginally affects the equivalent device capacitance, i.e., a small decrease of the low-frequency capacitance occurs under illumination, especially for narrow devices. On the other hand, illumination introduces a low-frequency constant conductance, which is absent in dark conditions. Such conductance is inversely proportional to W_{Ge} and it can be simply included in the equivalent small-signal circuit model as a shunt resistance R_{ill} between the contact nodes, as shown in figure 3.30.

It is reasonable to assume that both the decreased capacitance and the increased conductance observed for thin devices are linked to avalanche generation phenomena, arising when the electric field in the germanium layer is large enough. Avalanche generation leads to an increase in carrier density, resulting in a screening of the electric field, compatible with the slightly decreased capacitance, and in an enhanced conductance.

Interpolating the low-frequency conductance obtained varying W_{Ge}, a simple

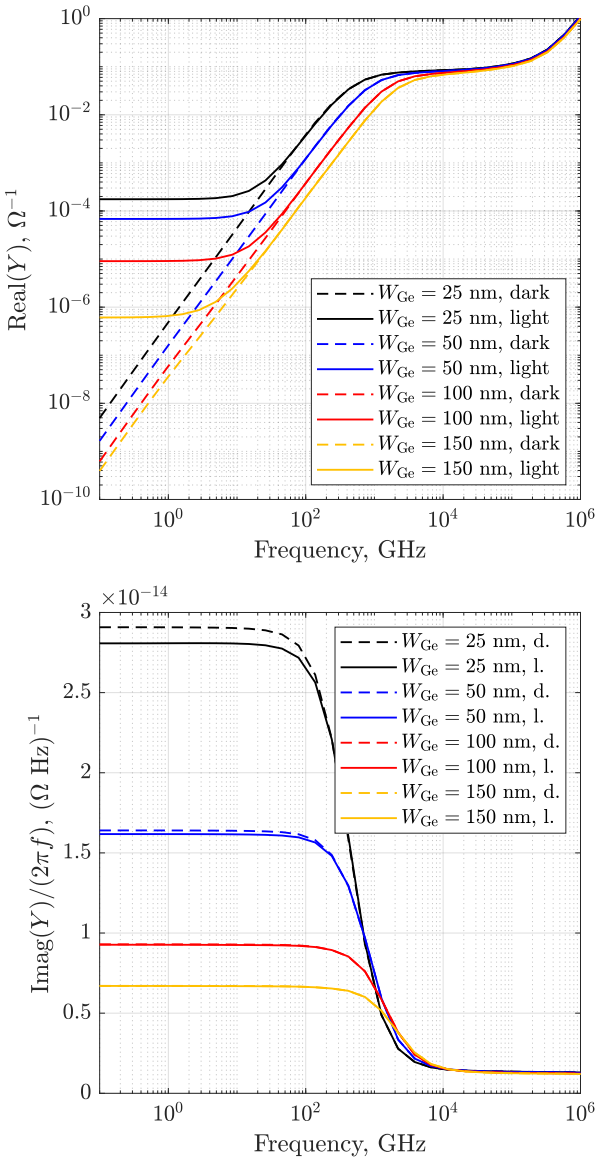


Figure 3.29: Effect of illumination on the device admittance, varying W_{Ge}, with $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $-2\,\mathrm{V}$ bias.

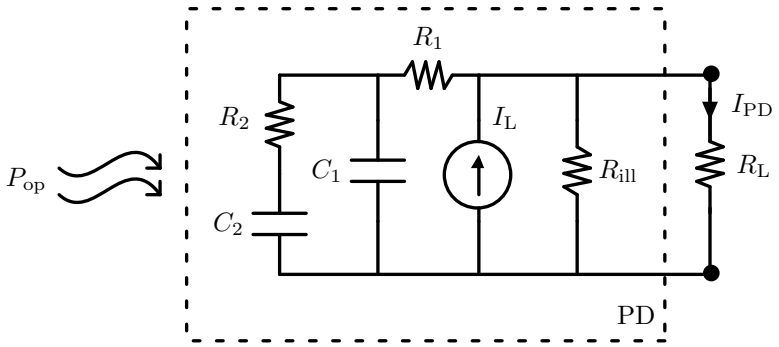


Figure 3.30: Fitting small-signal circuit model of the Ge-on-Si lateral *pin* photodetector under illumination.

analytical expression for R_{ill} can be obtained:

$$R_{ill}(W_{Ge}) = \left[10^{K_1+K_2(W_{Ge}-K_3)}\right]\Omega \tag{3.15}$$

with $K_1 = 3.756$, $K_2 = 1.967 \times 10^7$ m^{-1} and $K_3 = 25 \times 10^{-9}$ m.

A comparison between the model (3.15) and extracted data is shown in figure 3.31.

3.3 Performance evaluation and design guidelines

In this third section the evaluation of the performance of the Ge-on-Si lateral *pin* photodetector in terms of responsivity and modulation bandwidth is discussed. The effect of design choices on such figures of merit has been investigated, in order to provide some useful guidelines. In particular, variations of the three dimensions of the germanium absorber (W_{Ge}, H_{Ge}, L_{Ge}) and of the position of metal contacts (distance d) have been performed, as previously done in section 3.2. In addition to that, the role of doping of the silicon contacts has been examined.

All the following analysis has been performed for a low input optical power $P_{op} = 200\,\mu W$, at a wavelength $\lambda = 1.31\,\mu m$ (O-Band) and two different values of applied bias voltage, i.e., $-1\,V$ and $-2\,V$. The input segment of the silicon waveguide used to couple the light source with the device has a squared cross-section, with a 200 nm side, and is 1 µm long. The mode profile computed on the input waveguide and then propagated by the electromagnetic simulation is shown in figure 3.28.

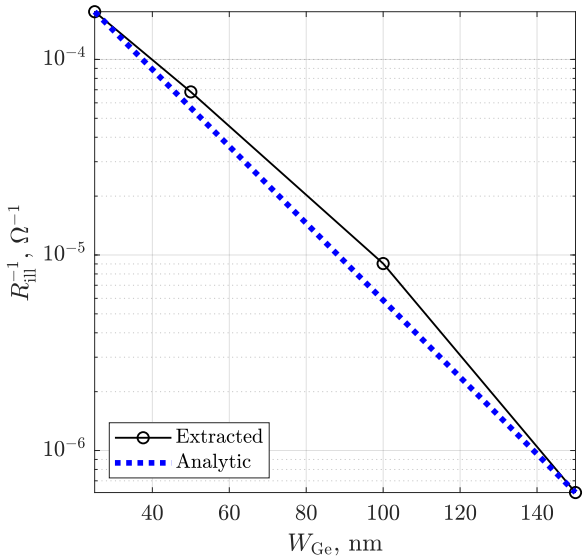


Figure 3.31: Comparison between the analytic model for R_{ill} (3.15) and the extracted values for a $-2\,\mathrm{V}$ bias.

3.3.1 Responsivity evaluation

From a simulation standpoint, the responsivity has been evaluated as follows:

1. the electromagnetic problem has been solved and the optical generation rate within the device has been computed (see section 2.3);
2. the applied voltage has been ramped from equilibrium to the desired bias value ($-1\,\mathrm{V}$ or $-2\,\mathrm{V}$) in dark ($P_{\mathrm{op}} = 0$), solving the charge transport problem (see section 2.2);
3. the input optical power has been ramped from zero to the desired value ($P_{\mathrm{op}} = 200\,\mu\mathrm{W}$), i.e., the previously computed optical generation rate has been included in the solution of the charge transport problem;
4. the responsivity has been evaluated as the ratio between maximum output current and maximum input optical power, which in practice coincides with the slope of the obtained P-I curve (input optical power vs. output current); P-I curves, in fact, are linear for low optical power levels and the dark current in the order of $0.1\,\mathrm{pA}$ can be safely neglected.

Examples of P-I curves, obtained for the reference device are reported in figure 3.32, with the corresponding responsivity values reported in table 3.5.

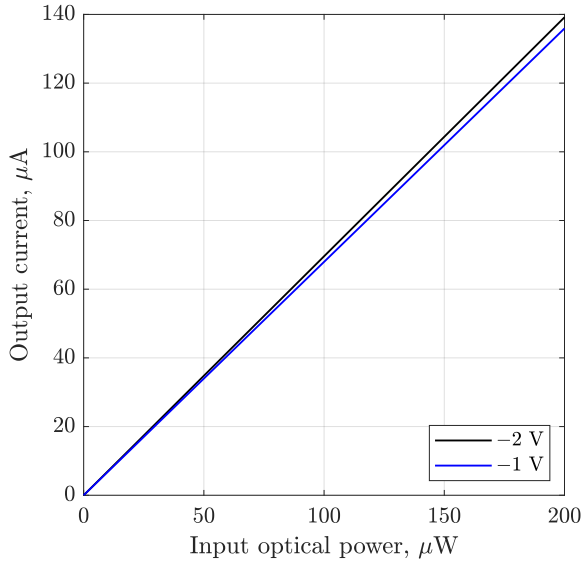


Figure 3.32: P-I curves obtained at $\lambda = 1.31\,\mu\text{m}$ and two different bias values for the reference device.

Bias, V	Responsivity, A/W	Modulation bandwidth, GHz
−2	0.696	269.22
−1	0.680	275.91

Table 3.5: Responsivity and modulation bandwidth obtained at $\lambda = 1.31\,\mu\text{m}$, $P_{\text{op}} = 200\,\mu\text{W}$ and two different bias values for the reference device.

Effect of the bias voltage

It can be observed from the results shown in figure 3.32 and table 3.5 that a higher applied bias voltage yields a slightly larger responsivity. This can be justified considering that, due to the narrow absorber (small W_{Ge}), even a relatively low bias voltage, such as the ones considered here, results in a very large electric field (see figure 3.33a), favouring impact ionization phenomena (see figure 3.33b). In other words, avalanche multiplication takes place within the device, boosting the responsivity. The impact ionization rate increases with the electric field, according to equation (2.14): in figure 3.34 an exponential increase of the avalanche generation rate can be observed for bias values larger (in magnitude) than $-1\,\mathrm{V}$, consistently with the responsivity gain.

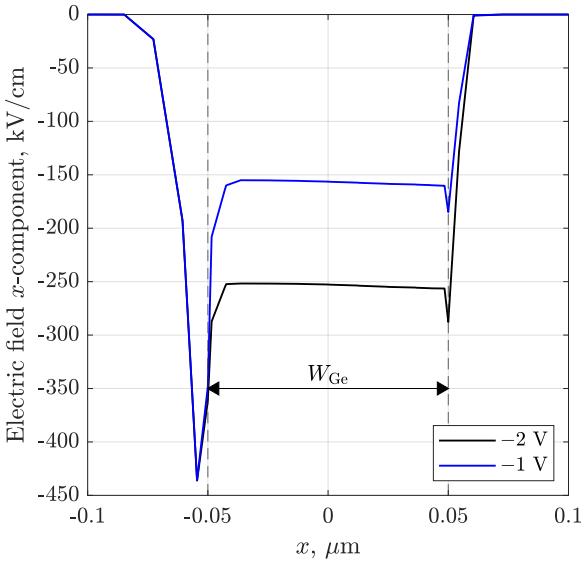
3.3.2 Modulation bandwidth evaluation

The evaluation of the modulation bandwidth requires a further step in addition to the ones needed for the computation of the responsivity, i.e., an optical AC analysis (see subsection 2.2.2). Once the device is biased to $-1\,\mathrm{V}$ or $-2\,\mathrm{V}$ and illuminated by an optical power $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, a modulation of the optical signal has to be simulated, in order to obtain the electro-optic frequency response of the device and evaluate its modulation bandwidth. All the following results have been obtained by running optical AC analyses in the modulation frequency range between $0.1\,\mathrm{GHz}$ and $400\,\mathrm{GHz}$, large enough to allow for the evaluation of the $-3\,\mathrm{dB}$ cutoff frequency, assuming the device to be loaded by a resistance $R_{\mathrm{L}} = 50\,\Omega$.

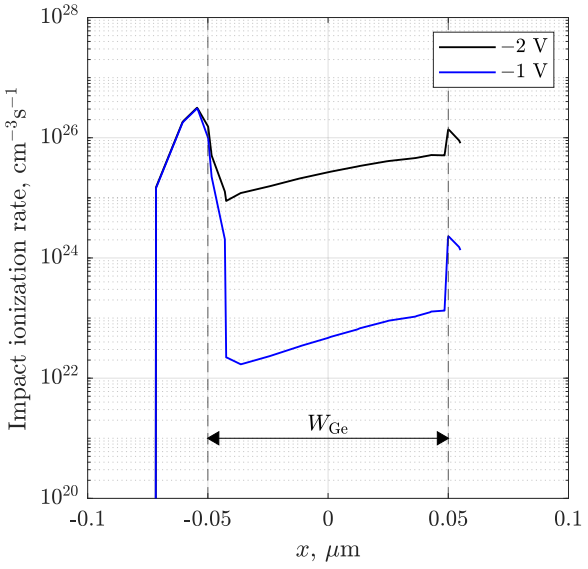
Examples of frequency response, obtained for the same reference device as before, are reported in figure 3.35, with the corresponding modulation bandwidth values reported in table 3.5. The device frequency response exhibits the typical low-pass behaviour, i.e., only optical signals modulated up to a certain cutoff frequency are efficiently detected by the device.

Effect of the bias voltage

It can be observed from the results shown in figure 3.35 and table 3.5 that a higher applied bias voltage yields a slightly lower modulation bandwidth, which is the opposite of the trend observed for the responsivity. The reason of this bandwidth penalty lies, again, in the avalanche generation mechanism: if, on one hand, avalanche multiplication enhances the responsivity, on the other hand an avalanche build-up delay adds to the transit time that carriers undergo before being collected by the contacts, which is detrimental for the speed of the device [2, Sec. 4.13].



(a)



(b)

Figure 3.33: Effect of the applied bias voltage on the x-component of the electric field and the impact ionization rate, obtained for the reference device, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$. 1D-cut at $y = -0.075\,\mu\mathrm{m}$, $z = 5.5\,\mu\mathrm{m}$.

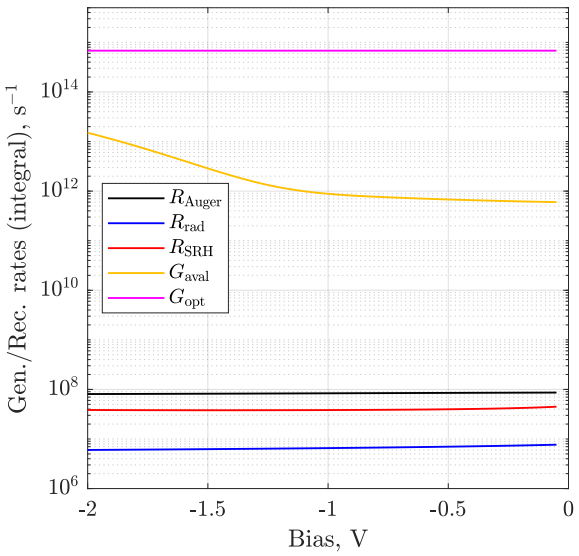


Figure 3.34: Effect of the applied bias voltage on the generation and recombination rates integrated on the absorber volume, obtained for the reference device, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$. In the order shown by the legend: Auger recombination, radiative recombination, SRH recombination, impact ionization (avalanche generation), optical generation rates.

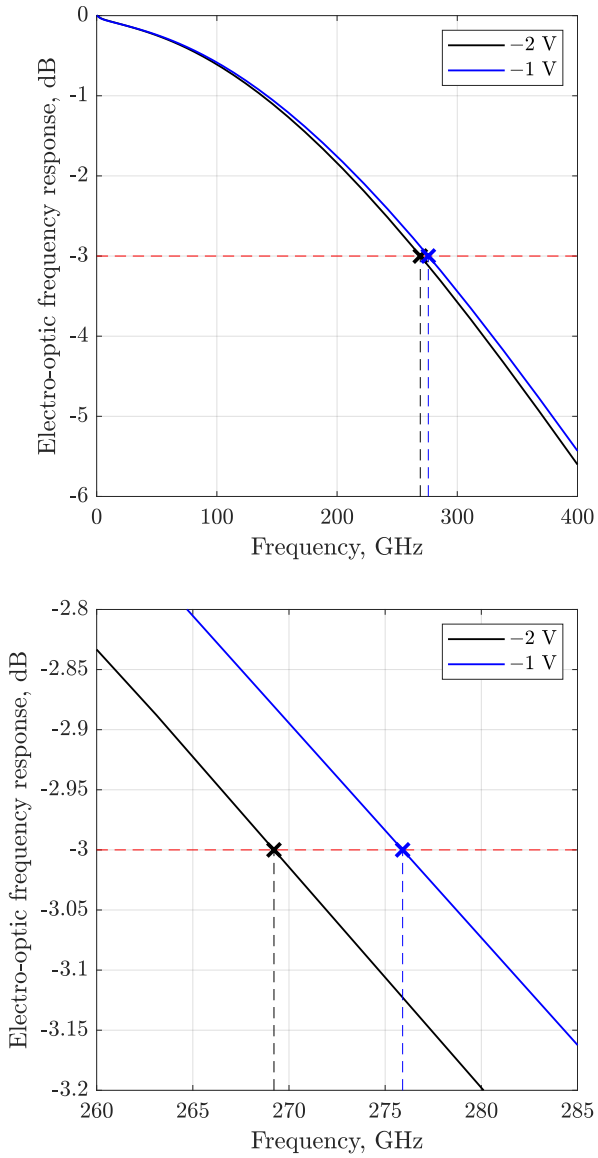


Figure 3.35: Electro-optic frequency response obtained at $\lambda = 1.31\,\mu m$, $P_{op} = 200\,\mu W$ and two different bias values for the reference device.

3.3.3 Parametric analyses of the performance

Responsivity and modulation bandwidth have been evaluated for several versions of the device, varying one key dimension at a time within a reasonable range, as previously done for the small-signal analyses presented in subsection 3.2.4.

Effect of the absorber width

The width of the germanium absorber W_{Ge} has a significant impact on both the responsivity and the modulation bandwidth. Results are reported in figures 3.36 and 3.37 for a reference device with dimensions $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $d = 0.38\,\mu\mathrm{m}$, for a W_{Ge} range between 25 nm and 250 nm.

For small W_{Ge} values, i.e., below 125 nm, the responsivity depends on the bias conditions: in fact, by applying a $-2\,\mathrm{V}$ bias a significant gain is obtained due to avalanche multiplication mechanisms. Such a gain disappears when W_{Ge} is increased and the electric field is too weak to effectively induce impact ionization. Apart from these considerations, it can be observed that, in general, a larger absorber is beneficial for the responsivity of the device, simply because the volume where photogeneration occurs is larger. However, it has to be remarked that responsivity also depends on the coupling of light between the silicon waveguide and the germanium absorber. These results have been obtained for a fixed height of the silicon waveguide $H_{\mathrm{wg}} = 0.2\,\mu\mathrm{m}$, while its width W_{SiWG} has been varied according to the following equation:

$$W_{\mathrm{SiWG}} = 2\,\mu\mathrm{m} + W_{\mathrm{Ge}} \tag{3.16}$$

The input segment width has been fixed to $W_{\mathrm{wg}} = 0.2\,\mu\mathrm{m}$, which may not be the optimum size for all the considered W_{Ge} range. Optimizations of the input waveguide have been explored in section 3.7.

As for the modulation bandwidth, for W_{Ge} values larger than 75 nm a monotonically decreasing trend can be observed versus W_{Ge}, which is mainly due to the transit time limit. For example, a bandwidth drop exceeding 100 GHz is observed comparing 100 nm and 200 nm widths. The RC limit is relevant for low W_{Ge} values only, when the small-signal capacitance is larger (e.g., see figure 3.18): excluding data for 50 nm, an increasing trend is observed between 25 nm and 75 nm, which is coherent with the equivalent capacitance decrease and the broader RC limit. In addition, the bandwidth penalty due to the avalanche build-up time can be observed in general when the bias is increased and the absorber is narrow.

Leveraging the small-signal analytical model proposed in section 3.2, the RC limit of the modulation bandwidth has been evaluated for the reference device in the same W_{Ge} range explored in simulations, according to equation (1.12). The transit time limit has been computed as well, using the model (1.9). The modulation bandwidth has been obtained according to equation (1.20), combining

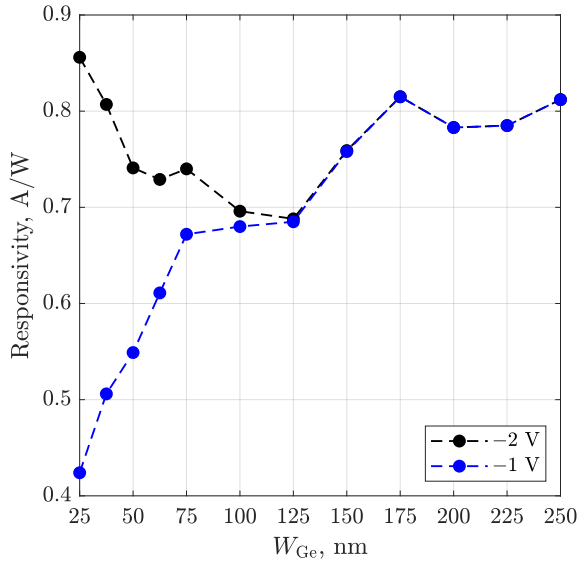


Figure 3.36: Effect of the absorber width W_{Ge} on the responsivity of the device, with $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $d = 0.38\,\mu\mathrm{m}$, for two different bias values, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

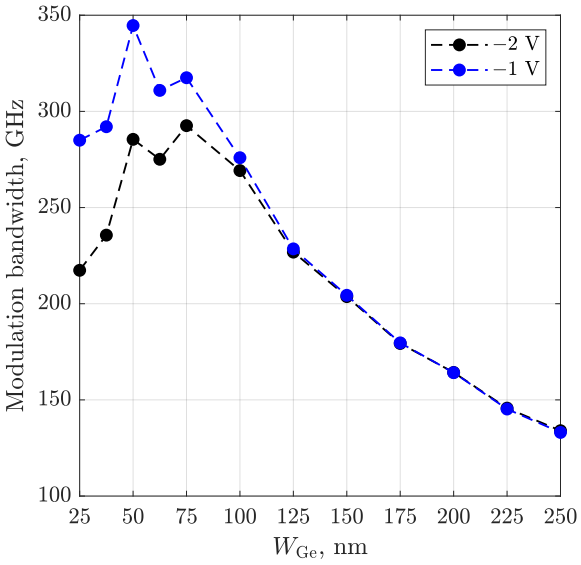


Figure 3.37: Effect of the absorber width W_{Ge} on the modulation bandwidth of the device, with $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $d = 0.38\,\mu\mathrm{m}$, for two different bias values, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

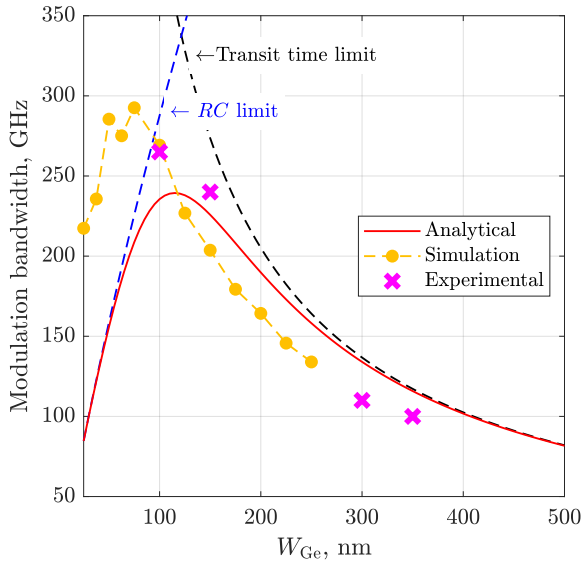


Figure 3.38: Effect of the absorber width W_{Ge} on the modulation bandwidth of the device. Comparison between analytical, simulated and experimental data.

both the *RC* and the transit time limits. The bandwidth evaluated analytically is compared in figure 3.38 with the one provided by simulations (corresponding to the $-2\,V$ curve in figure 3.37) and to the available measurements reported by literature, i.e., the ones in table 1.1. As it can be observed, the adopted analytical model for the evaluation of the bandwidth is not able to reproduce simulation and experimental data properly. In particular, the analytical *RC* limitation is heavier than the one resulting from simulations, while the analytical transit time limit turns out to be slightly optimistic. This issue would require further investigation. As a matter of fact, this analytical model may lose validity for deeply-scaled devices like these, with absorber widths in the order of 100 nm and bandwidths well above 100 GHz, where other physical effects which can be neglected in larger devices may become important.

In conclusion, a narrow germanium fin allows for the achievement of spectacular bandwidths, practically limited by the *RC* cutoff only. If a sufficiently high bias voltage is applied, the avalanche multiplication mechanism allows for an appreciable responsivity as well, despite introducing a bandwidth penalty. On the other hand, a wider absorber enhances by itself the responsivity, even though speed is heavily limited by the carriers transit time.

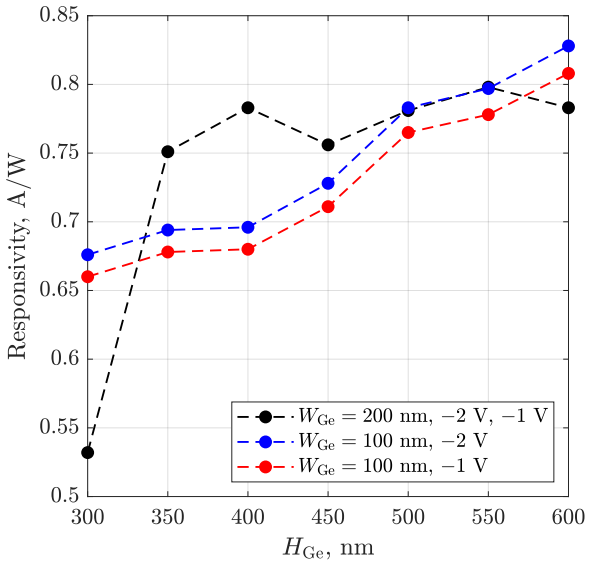


Figure 3.39: Effect of the absorber height H_{Ge} on the responsivity of the device, with two different W_{Ge} values, $L_{Ge} = 10\,\mu m$, $d = 0.38\,\mu m$, for two different bias values, $P_{op} = 200\,\mu W$, $\lambda = 1.31\,\mu m$.

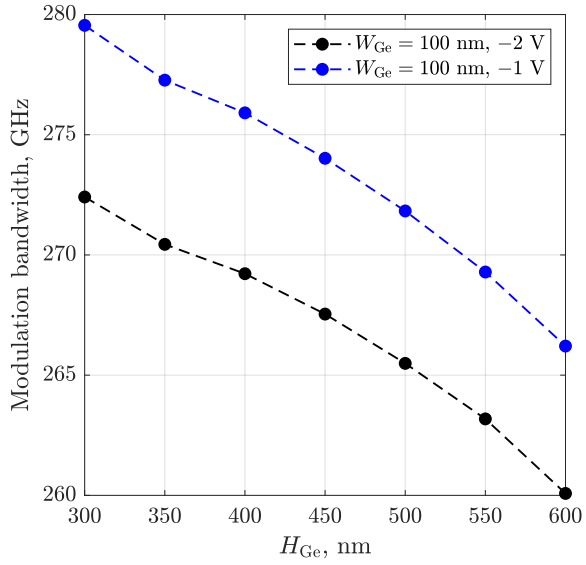
Effect of the absorber height

Variations of the absorber height H_{Ge} have been performed in a range between 300 nm and 600 nm on a reference device, for two W_{Ge} values, i.e., 100 nm and 200 nm, keeping constant $L_{Ge} = 10\,\mu m$ and $d = 0.38\,\mu m$.

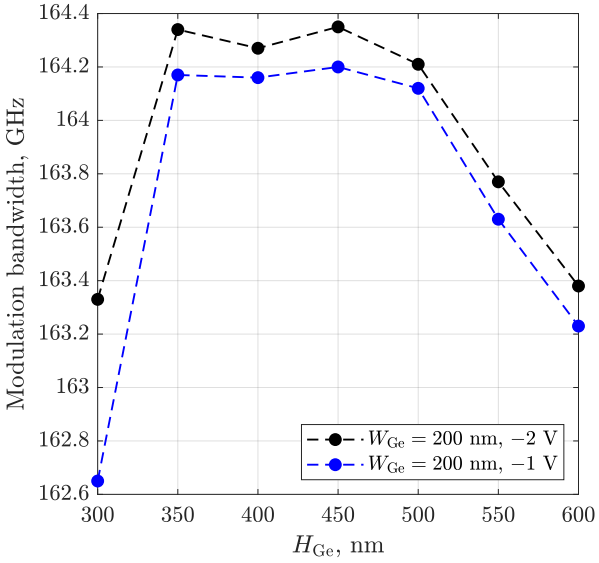
The effect on the responsivity is reported in figure 3.39: a larger H_{Ge} is generally favourable because it increases the absorber volume.

In figures 3.40 the effect on the modulation bandwidth can be observed. The decreasing trend can be justified: a larger H_{Ge} results in a larger equivalent small-signal capacitance (e.g., see figure 3.19) and, as a consequence, a lower RC limit of the bandwidth, which is relevant especially for $W_{Ge} = 100\,nm$. However, the decrease in bandwidth is limited, compared to the one observed when W_{Ge} is increased: reductions of about 13 GHz for $W_{Ge} = 100\,nm$ and only 1 GHz for $W_{Ge} = 200\,nm$ have been reported in the considered H_{Ge} range.

Therefore, a clear trade-off exists between responsivity and bandwidth in relation to the height of the germanium fin: a few GHz of bandwidth may be sacrificed to gain some responsivity or vice versa.



(a) $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$.



(b) $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$.

Figure 3.40: Effect of the absorber height H_{Ge} on the modulation bandwidth of the device, with two different W_{Ge} values, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $d = 0.38\,\mu\mathrm{m}$, for two different bias values, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

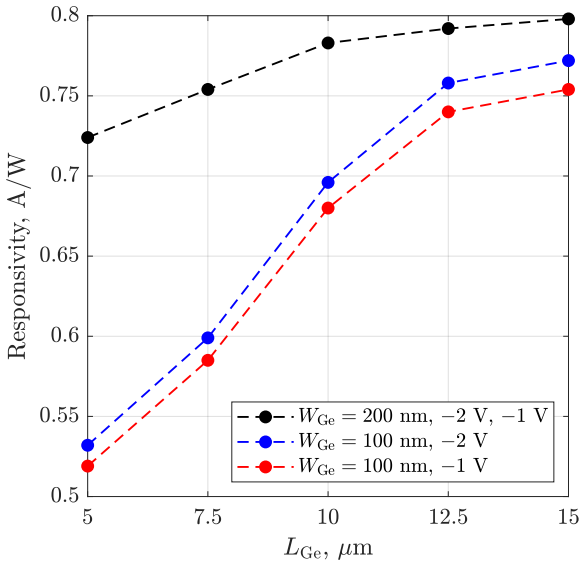


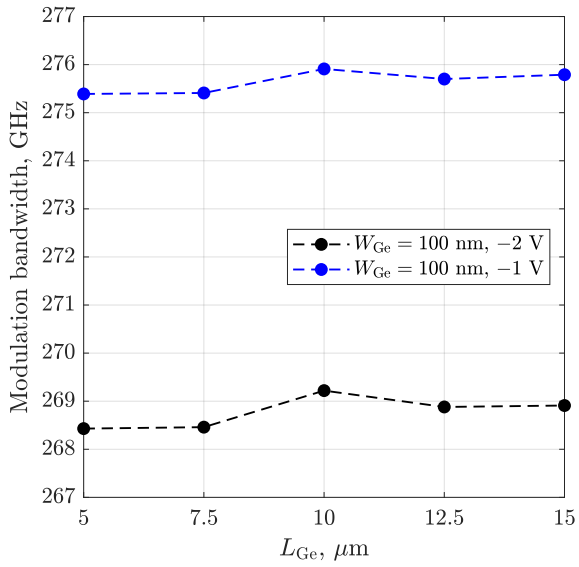
Figure 3.41: Effect of the absorber length L_{Ge} on the responsivity of the device, with two different W_{Ge} values, $H_{Ge} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$, for two different bias values, $P_{op} = 200\,\mu\text{W}$, $\lambda = 1.31\,\mu\text{m}$.

Effect of the absorber length

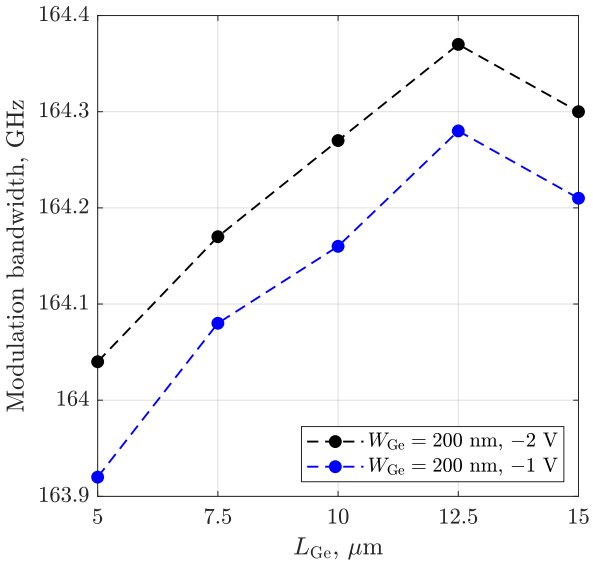
The length of the device L_{Ge} has been varied in the range between $5\,\mu\text{m}$ and $15\,\mu\text{m}$, for two W_{Ge} values, i.e., $100\,\text{nm}$ and $200\,\text{nm}$, keeping constant $H_{Ge} = 400\,\text{nm}$ and $d = 0.38\,\mu\text{m}$.

The effect of L_{Ge} on the responsivity can be observed in figure 3.41. Clearly, the longer is the device, the higher is the achievable responsivity, even though a saturation occurs once the device is long enough to allow for a virtually total absorption of the incident optical power (see equation (1.23)).

As for the modulation bandwidth, only negligible variations (lesser than $1\,\text{GHz}$) have been observed in the results, reported in figures 3.42. This is a peculiar advantage of waveguide photodetectors: compatibly with scaling requirements, the device can be made as long as needed to improve the efficiency without excessively affecting the RC-limited bandwidth, since the capacitance is kept small enough by the height of the absorber H_{Ge}, as discussed in subsection 1.2.3.



(a) $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$.



(b) $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$.

Figure 3.42: Effect of the absorber length L_{Ge} on the modulation bandwidth of the device, with two different W_{Ge} values, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, for two different bias values, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

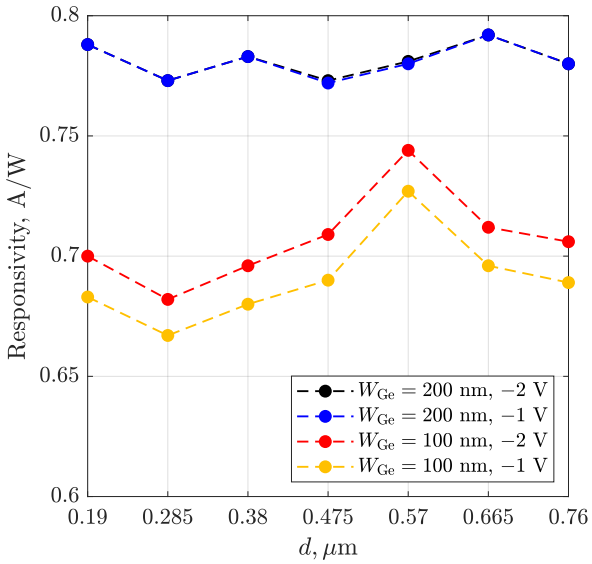


Figure 3.43: Effect of the silicon waveguide width (varied along with d according to (3.17)) on the responsivity of the device, with two different W_{Ge} values, $H_{Ge} = 400\,\text{nm}$, $L_{Ge} = 10\,\mu\text{m}$, for two different bias values, $P_{op} = 200\,\mu\text{W}$, $\lambda = 1.31\,\mu\text{m}$.

Effect of the position of metal contacts and of the silicon waveguide width

The role of the position of metal contacts has been investigated by varying the distance between the silicon sidewalls and the tungsten contacts d in the range between $0.19\,\mu\text{m}$ and $0.76\,\mu\text{m}$ (half and twice the reference $0.38\,\mu\text{m}$ value), for two W_{Ge} values, i.e., $100\,\text{nm}$ and $200\,\text{nm}$, keeping constant $H_{Ge} = 400\,\text{nm}$ and $L_{Ge} = 10\,\mu\text{m}$, and varying the width of the silicon waveguide W_{SiWG} according to the following expression:

$$W_{SiWG} = 2\,\mu\text{m} + W_{Ge} + 2(d - 0.38\,\mu\text{m}) \tag{3.17}$$

The responsivity, reported in figure 3.43, exhibits variations against d, especially for $W_{Ge} = 100\,\text{nm}$. Actually, the width of the silicon waveguide W_{SiWG}, varied along with d, has influenced the computed responsivity. In fact, among other factors, the coupling of light from the waveguide to the device depends on the size of the waveguide itself, which can be optimized to maximize the efficiency. Figure 3.44 shows the interference pattern created by the photon absorption density in two of the simulated devices: a 9% increase of the responsivity has been achieved simply by widening the waveguide.

The modulation bandwidth is influenced by the position of the contacts as

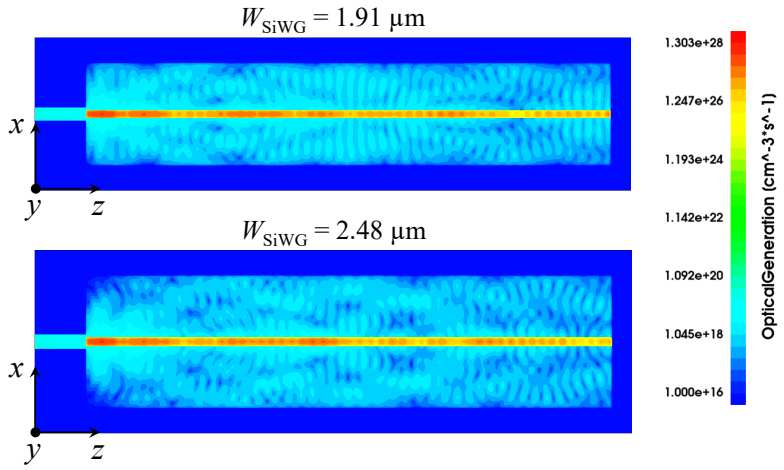


Figure 3.44: Effect of the silicon waveguide width on the optical generation rate within the device, for two different W_{SiWG} values (corresponding to $d = 0.285\,\mu\mathrm{m}$ and $d = 0.57\,\mu\mathrm{m}$), $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$. 2D-cut at $y = -0.15\,\mu\mathrm{m}$, at the interface between the waveguide and the device.

well, as shown by figures 3.45. In this case, the distance d itself has an impact on the cutoff frequency, being one of the parameters influencing the equivalent admittance of the device (e.g., see figure 3.16), which contributes to the RC limit of the bandwidth, crucial especially for small W_{Ge} values. As a matter of fact, fluctuations of the computed bandwidth are more significant for $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$ (about 40 GHz) than for $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$ (about 10 GHz).

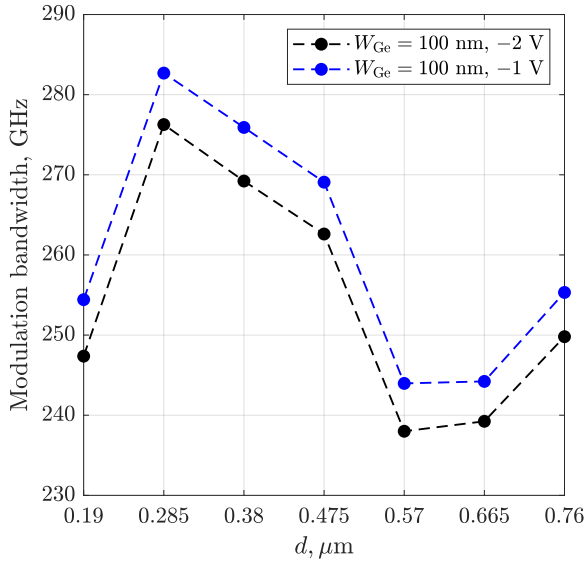
Effect of the doping of silicon contacts

Up to this point, the doping concentration in the silicon contacts has always been kept constant to 10^{20} cm^{-3}. The effect of a lower doping level has been investigated, down to 10^{18} cm^{-3}.

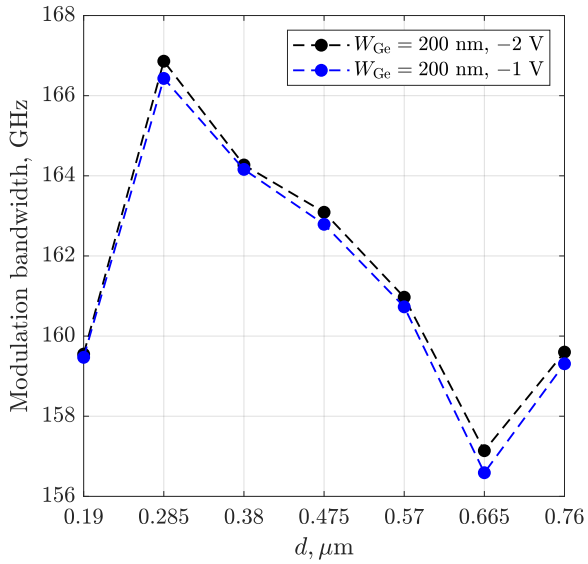
A lower doping results in a higher resistivity of the silicon contacts: as a consequence, the potential drop on the increased contacts resistance (see figure 3.46a) decreases the horizontal component of the electric field in the germanium absorber (see figure 3.46b) and, along with that, the impact ionization rate (see figure 3.47a). This is detrimental for the responsivity, reported in figure 3.48, which remains quite high anyway.

It can be observed in figure 3.49 that a low doping level has a negative impact on modulation bandwidth as well, much more severe compared to the one on responsivity. The device frequency response for a doping concentration of 10^{18} cm^{-3}, reported in figure 3.50, is visibly dominated by a cutoff in the order of only a few tens of GHz, determined by the diffusion of carriers within the silicon contacts. In fact, in presence of a lower doping level, the lower recombination rate (see figure 3.47b) results in longer lifetimes and diffusion lengths, affecting the broadband operation of the device [2, Sec. 4.7]. In addition, the reduced electric field within the absorber increases the transit time as well, limiting the achievable bandwidth. As for the RC limit, the decreased contact resistivity modifies the small-signal admittance of the device, as shown in figures 3.51. However, the reduced capacitance yielded by a lower doping concentration does not actually limit the bandwidth more than happens with high doping.

In conclusion, a very large doping concentration allows for both the best responsivity and modulation bandwidth. Decreasing the doping is detrimental especially for the bandwidth, dominated by the lifetime cutoff.

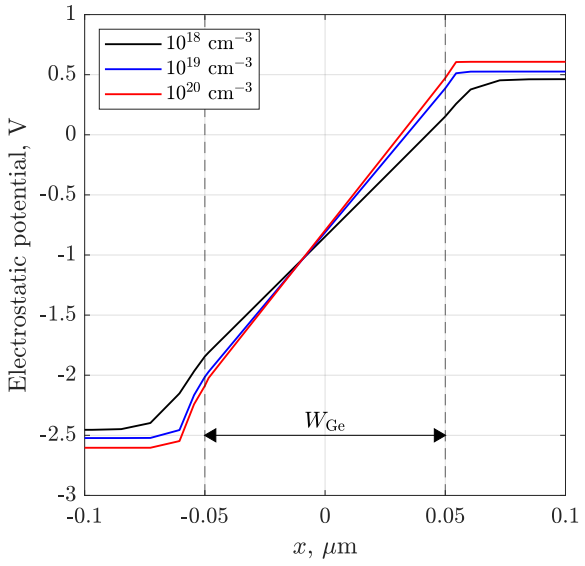


(a) $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$.

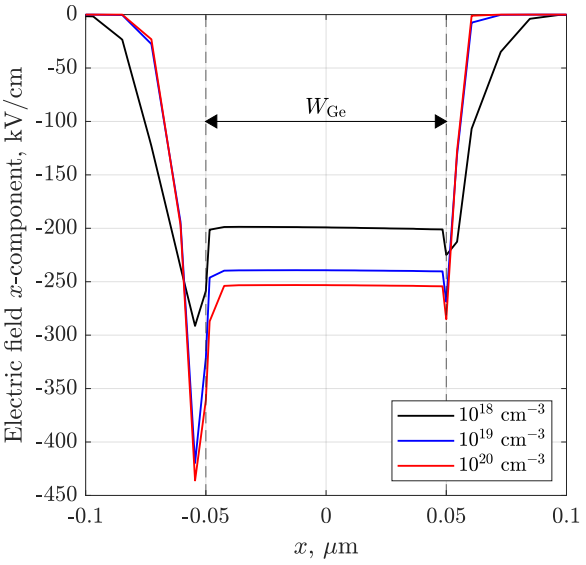


(b) $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$.

Figure 3.45: Effect of the position of metal contacts on the modulation bandwidth of the device, with two different W_{Ge} values, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$, for two different bias values, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

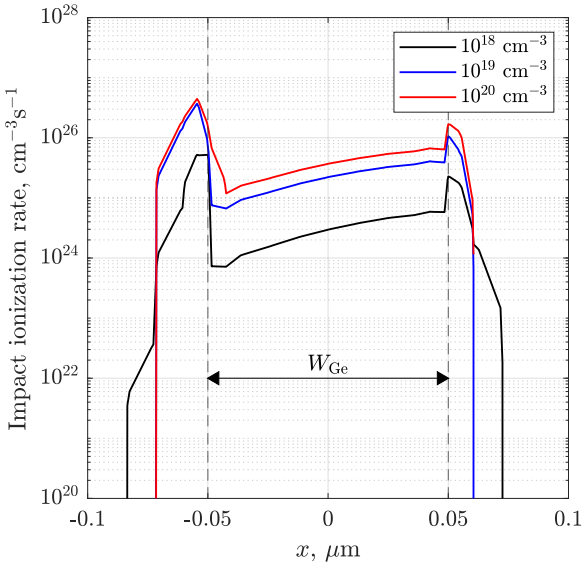


(a)

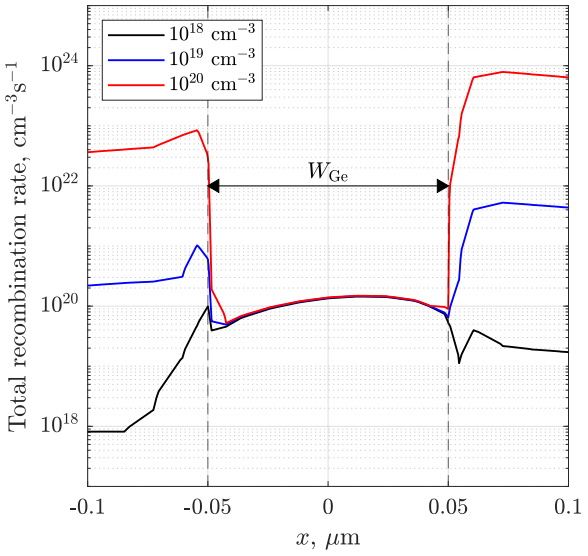


(b)

Figure 3.46: Effect of the doping concentration of silicon contacts on the electrostatic potential and the x-component of the electric field, obtained for the reference device at a -2 V bias, $P_{\mathrm{op}} = 200$ µW, $\lambda = 1.31$ µm. 1D-cut at $y = 0$ µm, $z = 5$ µm.



(a)



(b)

Figure 3.47: Effect of the doping concentration of silicon contacts on the impact ionization rate and the total recombination rate, obtained for the reference device at a $-2\,\mathrm{V}$ bias, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$. 1D-cut at $y = 0\,\mu\mathrm{m}$, $z = 5\,\mu\mathrm{m}$.

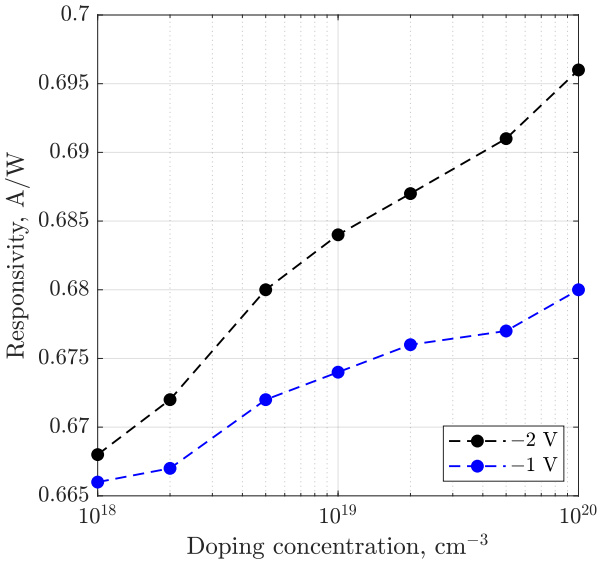


Figure 3.48: Effect of the doping concentration of silicon contacts on the responsivity of the reference device for two different bias values, $P_{\text{op}} = 200\,\mu\text{W}$, $\lambda = 1.31\,\mu\text{m}$.

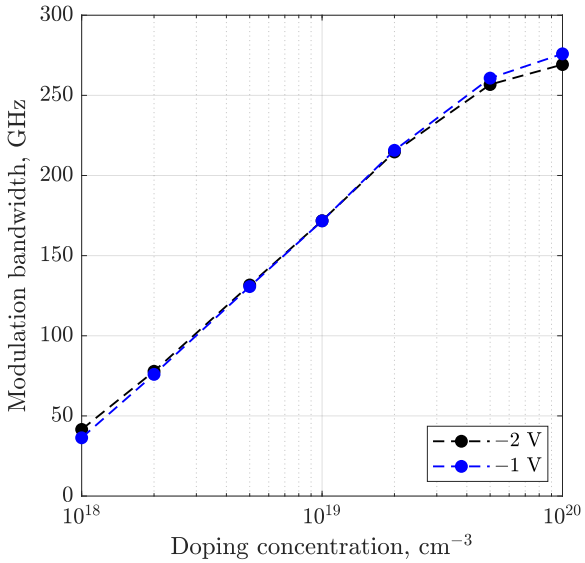


Figure 3.49: Effect of the doping concentration of silicon contacts on the modulation bandwidth of the reference device for two different bias values, $P_{\text{op}} = 200\,\mu\text{W}$, $\lambda = 1.31\,\mu\text{m}$.

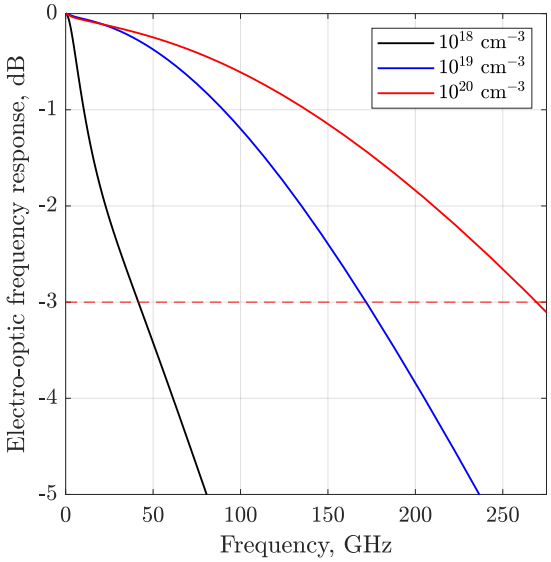


Figure 3.50: Effect of the doping concentration of silicon contacts on the electro-optic frequency response of the reference device for a $-2\,\mathrm{V}$ bias, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

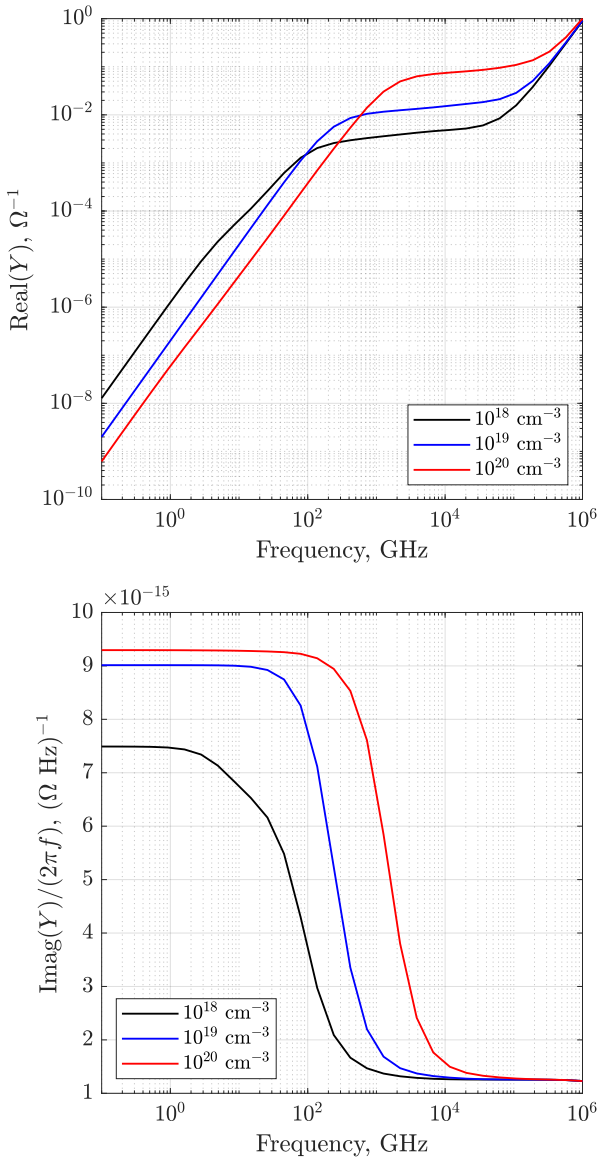


Figure 3.51: Effect of the doping concentration of silicon contacts on the reference device admittance at a $-2\,\mathrm{V}$ bias.

3.4 High optical power performance

Inspired by [32], the effect of the input optical power level on the behaviour of the device has been investigated. The reference device has been simulated in the optical power range between 200 µW and 200 mW, at a wavelength $\lambda = 1.31$ µm (O-Band) and for two different values of applied bias voltage, i.e., -1 V and -2 V.

The responsivity results are shown in figure 3.52. These values have been evaluated, as stated in subsection 3.3.1, as the slope of *P-I* curves. The behaviour of responsivity may be described as follows:

1. up to about 10 mW the responsivity remains constant and equal to the low-power value;
2. between 10 mW and 100 mW a small increase is observed, as if some gain has been achieved;
3. a final drop occurs for the highest simulated power level, i.e. 200 mW.

As a matter of fact, when high power levels have been considered, the obtained *P-I* curves were not exactly linear (see figure 3.53). For power levels above 10 mW the output current visibly deviates from the low-power linearity, with a peak deviation of 3 mA around 100 mW. This can be effectively assumed as a responsivity gain, attributed to the avalanche multiplication mechanism. Towards the end of the considered range, i.e., 200 mW, the curve approaches the linear approximation again, resulting in the computed responsivity drop. It has to be remarked that the solution of the charge transport problem faced some convergence problems, unlike other simulations, as it can be also noticed from the many dots composing the high power *P-I* curve, which correspond to each step of the quasi-stationary power ramp. This was expected, being the high power operation of the device substantially different from the low power behaviour, due to the significant space charge effects [2, Sec. 4.9.2].

When the incident optical power is increased, a high charge injection takes place within the intrinsic germanium layer. The horizontal electric field, which is supposed to drift photocarriers, is actually screened by the space charge and is not uniform along the photodetector, as it can be observed from figures 3.54 and 3.55. The behaviour of such microscopic quantities at different power levels can be described as follows:

1. the field profiles in dark ($P_{\mathrm{op}} = 0$) and for low power levels (0.2 mW) are nearly coincident, being the space charge still negligible;
2. for increasing optical power (2 mW, 20 mW), a field screening becomes clearly visible, especially in the initial section of the detector, where the optical generation is higher; since the optical generation rate progressively

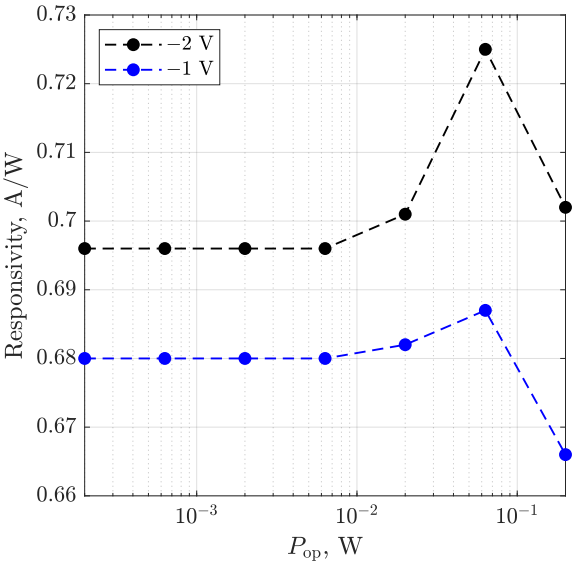


Figure 3.52: Effect of the input optical power on the responsivity of the reference device for two different bias values and $\lambda = 1.31\,\mu\mathrm{m}$.

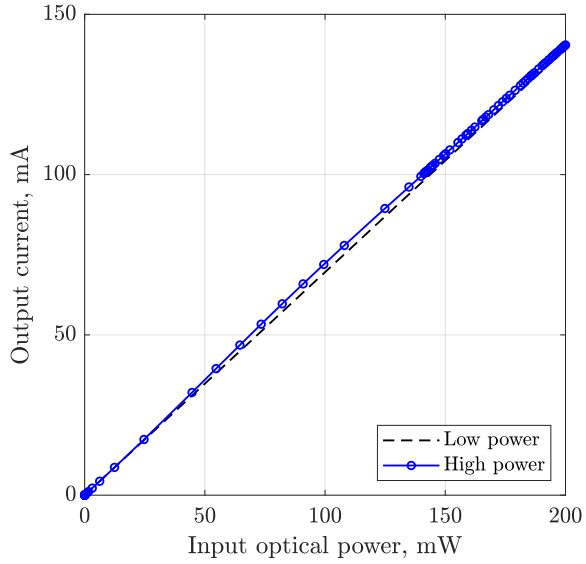


Figure 3.53: Comparison between the extension of the P-I curve obtained for the reference device for a maximum $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$ (low power) and the curve obtained up to $P_{\mathrm{op}} = 200\,\mathrm{mW}$ (high power), both for $\lambda = 1.31\,\mu\mathrm{m}$ and a $-2\,\mathrm{V}$ bias.

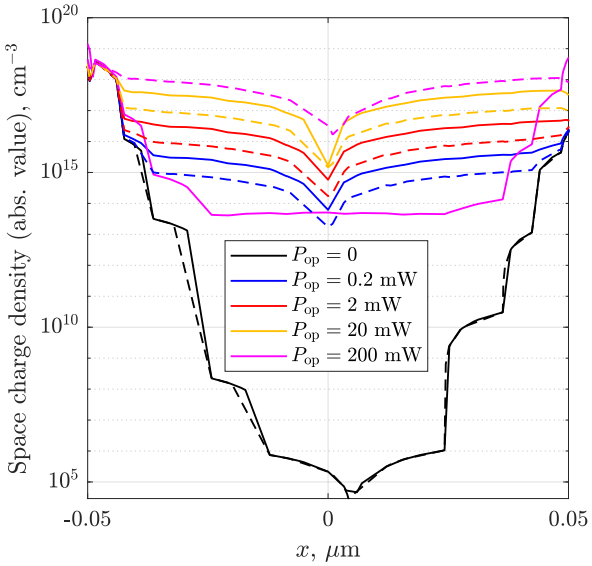


Figure 3.54: Effect of the input optical power on the space charge density (reported in absolute value) within the reference device for a $-2\,\mathrm{V}$ bias and $\lambda = 1.31\,\mu\mathrm{m}$. 1D-cuts at $y = 0.05\,\mu\mathrm{m}$ and two different z values, i.e., $z = 1\,\mu\mathrm{m}$ (solid lines) and $z = 6\,\mu\mathrm{m}$ (dashed lines).

decreases along the absorber, the space charge is reduced as well and the field screening is alleviated;

3. at high power levels (200 mW), the field screening effect in the initial segment of the detector is dramatic and other mechanisms get involved; as a matter of fact, as it is shown by figure 3.56, both generation and recombination rates grow almost in parallel until power levels are low, while for $P_{op} = 200\,\mathrm{mW}$ carrier injection is so high that recombination processes undergo a steep increase, with major contributions of Auger and radiative recombination; the situation is partially recovered further along the detector, where the electric field is still screened by the space charge but recombination is less effective.

The screening of the electric field affects the modulation bandwidth of the detector as well, as reported in figure 3.57. Below 10 mW, the bandwidth is virtually unaffected by the increase of P_{op}, while a dramatic decrease is observed for higher power levels. This is due to the weaker electric field, which is not able to saturate the velocity of photocarriers, with severe consequences on the transit time limit of the bandwidth. The electro-optic frequency response for increasing P_{op} is shown in figure 3.58. While up to 20 mW the typical single-pole low-pass behaviour can be observed, two separated poles are clearly visible

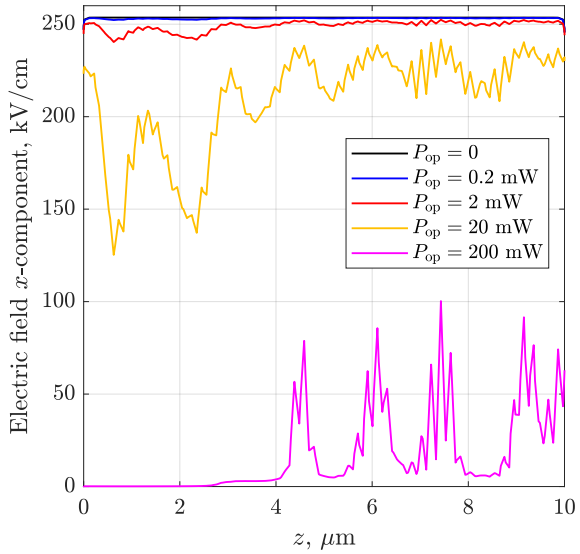


Figure 3.55: Effect of the input optical power on the x-component of the electric field within the reference device for a $-2\,\text{V}$ bias and $\lambda = 1.31\,\mu\text{m}$. 1D-cut at $x = 0\,\mu\text{m}$, $y = 0.05\,\mu\text{m}$, at the center of the germanium absorber.

in the frequency response obtained for a 200 mW incident power, with a cutoff frequency of barely 2 GHz. Recombination processes are responsible of this poor performance, as discussed previously. The bandwidth is ultimately limited by the lifetime of photocarriers [2, Sec. 4.7].

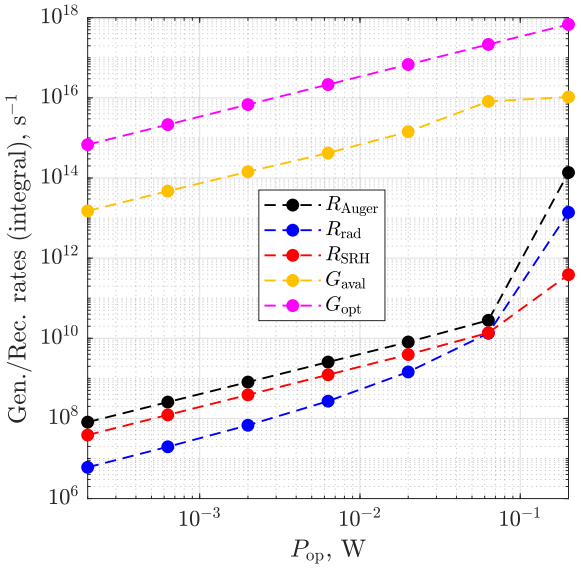


Figure 3.56: Effect of the input optical power on the generation and recombination rates integrated on the absorber volume, obtained for the reference device at a $-2\,\mathrm{V}$ bias and $\lambda = 1.31\,\mu\mathrm{m}$. In the order shown by the legend: Auger recombination, radiative recombination, SRH recombination, impact ionization (avalanche generation), optical generation rates.

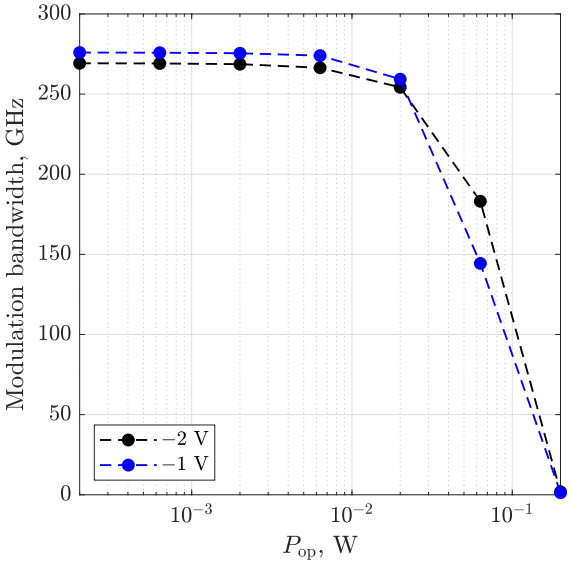


Figure 3.57: Effect of the input optical power on the modulation bandwidth of the reference device for two different bias values and $\lambda = 1.31\,\mu\mathrm{m}$.

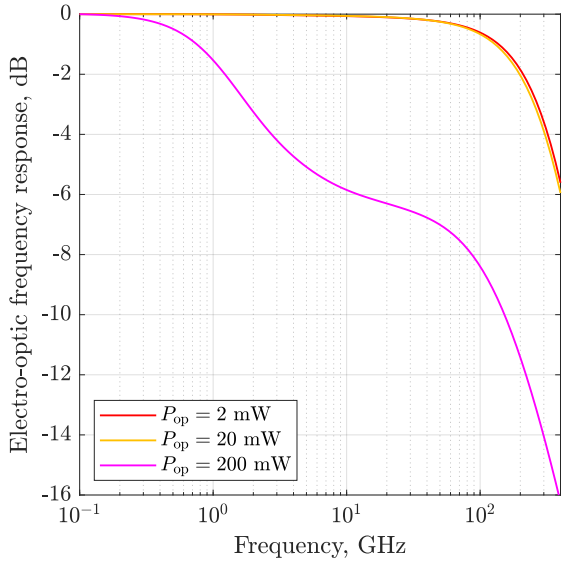


Figure 3.58: Effect of the input optical power on the electro-optic frequency response of the reference device for $\lambda = 1.31\,\mu\mathrm{m}$ and a $-2\,\mathrm{V}$ bias.

3.5 Inductive peaking-based bandwidth extension

Inductive peaking, also known as gain peaking, is a widely used technique in CMOS and photoreceiver amplification design to extend the operating bandwidth [33, 34]. Bandwidth enhancement of vertical germanium photodetectors based on this principle has been demonstrated in the literature [35, 36]. As a matter of fact, high-Q inductors can be efficiently integrated in close proximity with the detectors, leveraging mature silicon photonics processes.

Here a series inductive peaking configuration has been adopted to further enhance the broadband operation of the lateral Ge-on-Si photodetector. Simulations have been run on the same reference device considered in the previous sections.

3.5.1 Inductive peaking: theoretical introduction

The series inductive peaking working principle is rather easy. Consider the circuit in figure 3.59a: an inductance L is placed in series with both the load resistance R and the capacitance C responsible of limiting the broadband operation of the circuit [37, Sec. 10.6].

The peaking inductance should be chosen in such a way to intervene around the expected RC cutoff frequency, possibly introducing a peak in the frequency response (see figure 3.59b). The following input-output transfer function can be written for the circuit in figure 3.59a:

$$\frac{V_o}{I_{in}}(s) = \frac{R}{s^2LC + sRC + 1} = \frac{1}{mR^2C^2} \cdot \frac{R}{s^2 + s/mRC + 1/mR^2C^2} \tag{3.18}$$

with $L = mR^2C$, where m is a dimensionless parameter that determines the poles location and the kind of frequency response of the circuit [38]. The complex

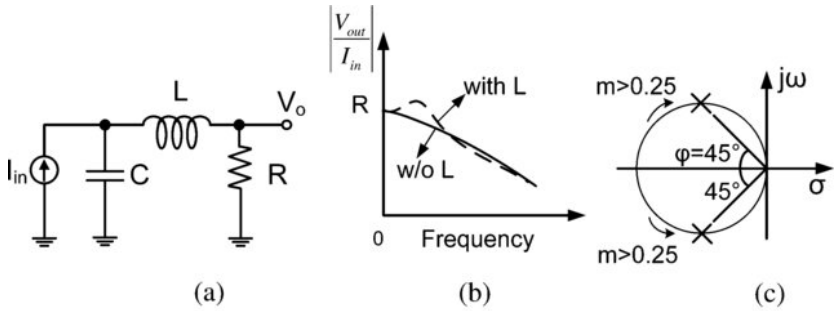


Figure 3.59: Series inductive peaking circuit. Figure taken from [38].

conjugates poles can be expressed as:

$$s_{1,2} = \frac{1}{2mRC}\left(-1 \pm j\sqrt{4m-1}\right) \tag{3.19}$$

Depending on the value of m, three different situations arise:

1. if $0 < m < 0.25$ poles are real and distinct (overdamped response);
2. if $m = 0.25$ poles are real and coincident (critically damped response);
3. if $m > 0.25$ poles are complex conjugates (underdamped response).

If R and C are assumed to be fixed, it is evident that the frequency response of the circuit only depends on the chosen inductance L, which directly determines the m parameter. It is also common to rewrite the transfer function (3.18) in the form

$$\frac{V_\mathrm{o}}{I_\mathrm{in}}(s) = R \cdot \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{3.20}$$

where

$$\zeta = \frac{R}{2}\sqrt{\frac{C}{L}} \tag{3.21}$$

is the damping factor and

$$\omega_n = 2\pi f_n = \frac{1}{\sqrt{LC}} \tag{3.22}$$

is the resonance (natural) frequency. The three different types of response may be also expressed in terms of damping factor ζ [39, Sec. 11.14.4]:

1. overdamped response for $\zeta > 1$;
2. critically damped response for $\zeta = 1$;
3. underdamped response for $0 < \zeta < 1$.

The $-3\,\mathrm{dB}$ bandwidth $\omega_{-3\mathrm{dB}}$ can be expressed as

$$\omega_{-3\mathrm{dB}}^2 = (2\pi f_{-3\mathrm{dB}})^2 = \left[-\left(2\zeta^2 - 1\right) + \sqrt{\left(2\zeta^2 - 1\right)^2 + 1}\right]\omega_n^2 \tag{3.23}$$

and it can be observed that a damping factor $\zeta = 1/\sqrt{2} \approx 0.707$, corresponding to an underdamped response, yields the maximum achievable bandwidth $\omega_{-3\mathrm{dB}} = \omega_n = \sqrt{2}/(RC) \approx 1.414/(RC)$, corresponding to an increase by about 40%. Considering the previously introduced m parameter, it is trivial to verify that

$$\zeta = \frac{1}{2\sqrt{m}}. \tag{3.24}$$

Therefore, $\zeta = 1/\sqrt{2}$ corresponds to $m = 0.5$ and the inductance required for maximizing the bandwidth of the circuit can be found as

$$L = \frac{R^2 C}{2}. \tag{3.25}$$

3.5.2 Simulation and results

As observed from the small-signal analysis of the reference device, in the frequency range between 0.1 GHz and 100 GHz the equivalent device admittance (see figure 3.5) features a negligible conductance (lesser than 1 mS) and a 9.3 pF capacitance, corresponding to the sum of C_1 and C_2, from the small-signal circuit described in subsection 3.2.2 and reported in figure 3.8. Therefore, in a first approximation, a current source in parallel with a capacitance $C = 9.3\,\text{pF}$ represent a reasonable small-signal model. Loading the device with the series of an ideal inductance L and a resistance $R = 50\,\Omega$, the simple circuit of figure 3.59a is obtained, with $C = 9.3\,\text{pF}$ and $R = 50\,\Omega$. As a matter of fact, this approximate circuit model provides a legitimate rule of thumb for designing the required inductance to effectively enhance the bandwidth of the detector. Substituting the values for R and C in (3.25), an optimum inductance $L = 11.6\,\text{pH}$ is obtained.

Considered the starting estimation of the required peaking inductor, simulations were run on the reference device with a $-2\,\text{V}$ bias applied and an incident optical power $P_{\text{op}} = 200\,\mu\text{W}$ at $\lambda = 1.31\,\mu\text{m}$, loaded by ideal inductors in the range between 1 pH and 50 pH.

The modulation bandwidth computed by simulations is reported in figure 3.60 as a function of the chosen inductance L. The observed behaviour can be described as follows:

1. starting from the 269.22 GHz value computed in absence of the inductance, the bandwidth undergoes a steep increase up to 702.51 GHz, obtained for $L = 5.5\,\text{pH}$;
2. a plateau around an average of 700 GHz is observed for inductance values between 5.5 pH and 8 pH;
3. for increasing L values the bandwidth decreases again; the 50 pH inductance chosen as upper limit of this analysis still yields a 324.47 GHz bandwidth, higher than the original one.

As expected from the preliminary computation, the optimal inductance lays between 5.5 pH and 8 pH, quite close to the estimated 11.6 pH value. In particular, the $L = 5.5\,\text{pH}$ seems to yield a bandwidth equal to the resonance frequency f_n. The theoretical cutoff frequency $f_{-3\text{dB}}$, computed according to equation (3.23) (with $R = 50\,\Omega$ and $C = 9.3\,\text{pF}$), effectively predicts the simulated values, save for the 700 GHz plateau.

3.5.3 Analysis of results

Deeper insights into the origin of the L-dependent bandwidth behaviour have been provided by the shape of the simulated frequency response, reported in

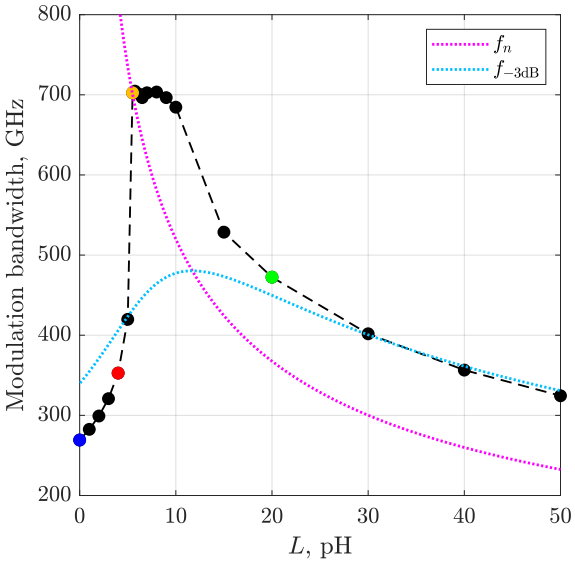


Figure 3.60: Effect of the peaking inductance L on the modulation bandwidth of the reference device, obtained for a $-2\,\mathrm{V}$ bias, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$. The resonance frequency f_n (3.22) and the theoretical bandwidth $f_{-3\mathrm{dB}}$ (3.23) are reported.

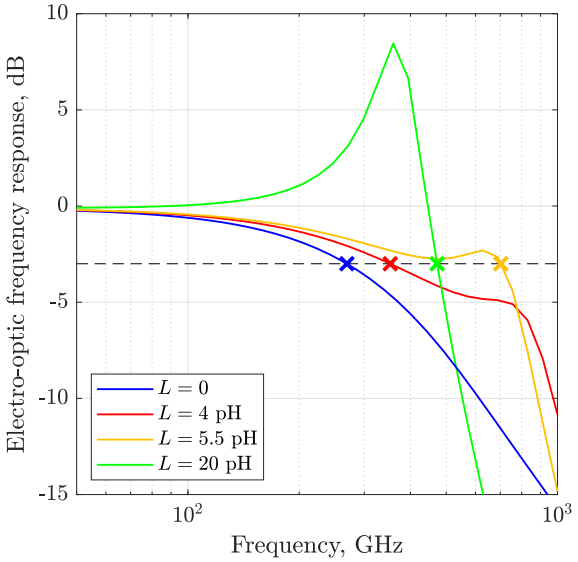


Figure 3.61: Effect of the peaking inductance L on the electro-optic frequency response of the reference device, obtained for a $-2\,\mathrm{V}$ bias, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$. The $-3\,\mathrm{dB}$ cutoff frequencies are highlighted (values reported in table 3.6).

figure 3.61. The three considered inductance values and the corresponding modulation bandwidth are reported in table 3.6. Three types of response can be observed:

1. $L = 4\,\mathrm{pH}$, which is a sub-optimum inductance value, yields a double-pole response, without any peaking;
2. for $L = 5.5\,\mathrm{pH}$ the maximum bandwidth is achieved, coincident with the resonance frequency $f_n = [2\pi\sqrt{L(C_1 + C_2)}]^{-1}$, and a moderate peaking arises (poles are complex conjugates but the imaginary part is small);
3. with $L = 20\,\mathrm{pH}$ peaking becomes clearly visible (poles are complex conjugates with a significant imaginary part).

In order to get a better understanding of the different situations which may arise depending on the exact value chosen for L, the frequency behaviour of the equivalent admittance of the device has been investigated including the effect of the inductance. As a matter of fact, as outlined in subsection 1.2.2 by equation (1.12), the detector frequency response depends on the device equivalent admittance and on the load impedance, which in this case is fixed to $50\,\Omega$. Considering the small-signal circuit shown in figure 3.62, the equivalent impedance

L, pH	Modulation bandwidth, GHz
0.0	269.22
4.0	352.75
5.5	702.51
20.0	472.42

Table 3.6: Effect of the peaking inductance L on the modulation bandwidth of the reference device, obtained for a $-2\,\mathrm{V}$ bias, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.31\,\mu\mathrm{m}$.

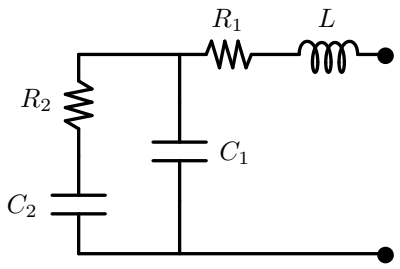


Figure 3.62: Small-signal circuit model of the Ge-on-Si lateral *pin* photodetector with a series peaking inductor.

seen by the terminals has been computed as

$$Z(s) = sL + R_1 + \cfrac{1}{sC_1 + \cfrac{1}{R_2 + \cfrac{1}{sC_2}}} \tag{3.26}$$

where $s = \mathrm{i}2\pi f$ and the reference device circuit parameters extracted in subsection 3.2.1 and reported in table 3.1 have been substituted.

The zero-pole compensation introduced by the peaking inductor can be highlighted by a numerical evaluation of the equivalent impedance of the dipole. The reference device small-signal equivalent impedance, without any inductor, reads, as previously reported by expression (3.5)

$$Z(s)|_{L=0} = \frac{0.01724(s + 4.782 \times 10^{16})(s + 7.982 \times 10^{12})}{s(s + 6.168 \times 10^{13})} \tag{3.27}$$

while, e.g., for $L = 20\,\mathrm{pH}$, the expression modifies to

$$Z(s)|_{L=20\,\mathrm{pH}} = \frac{2 \times 10^{-11}(s + 6.109 \times 10^{13})(s^2 + 5.867 \times 10^{11} s + 5.387 \times 10^{24})}{s(s + 6.168 \times 10^{13})} \tag{3.28}$$

where the zero located at 6.109×10^{13} rad/s is indeed very close to the pole located at 6.168×10^{13} rad/s.

The effect of the chosen inductance on the frequency behaviour of the equivalent admittance can be observed in figures 3.63. The conductance reaches a peak value of about 0.1 mS at the resonance frequency f_n (indicated by the dashed vertical lines), while the imaginary part of the admittance vanishes around this frequency. Below f_n, the circuit is approximately equivalent to a capacitance, while above f_n the inductive nature of the admittance is revealed by its negative imaginary part. These features suggest a possible interpretation of the effect of the chosen L on the overall frequency response:

1. for $L = 4\,\text{pH}$ the increase in the admittance is the main responsible for the bandwidth enhancement, while the capacitance remains practically unvaried;
2. for $L = 5.5\,\text{pH}$ the 0.1 mS conductance combined with the almost null admittance imaginary part allows for the remarkable bandwidth extension;
3. for $L = 20\,\text{pH}$ the bandwidth improvement can be attributed to the combination of the increased conductance and the inductive behaviour of the dipole.

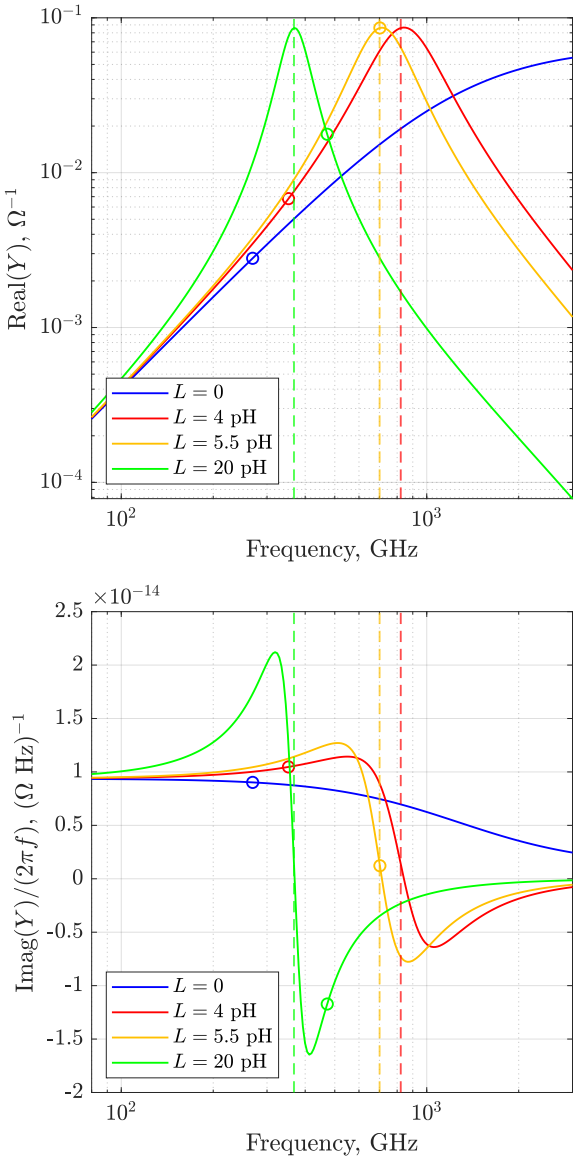


Figure 3.63: Effect of the peaking inductance L on the equivalent admittance seen by the terminals of the circuit in figure 3.62 (impedance expression (3.26), circuit parameters from table 3.1). Dashed vertical lines highlight the resonance frequency f_n (3.22), circles correspond to the simulated cutoff frequency value (as reported in table 3.6).

3.6 Spectral response

All the experiments and simulations on the lateral Ge-on-Si photoedetector described in the previous sections of this Chapter have been performed in the optical communications O-Band (*Original* band), centred around a $\lambda = 1.31\,\mu m$ wavelength ($1.26\,\mu m$ to $1.36\,\mu m$).

Optical fibers in the O-Band exhibit minimum dispersion, which is why this band was historically the first used for optical communications. Today optical fibers feature the lowest loss in the C-Band (*Conventional* band), centred around $\lambda = 1.55\,\mu m$ ($1.53\,\mu m$ to $1.565\,\mu m$), which is commonly used in long-haul transmission systems [40].

As far as the Ge-on-Si photodetector under analysis is concerned, the chosen operating band has a major impact on the responsivity achievable by the device. As a matter of fact, the germanium absorption coefficient, on which efficiency and responsivity depend (see equation (1.23)), undergoes a significant decrease in the considered spectral range: the absorption capability of germanium is much more favourable in the O-Band than in the C-Band. Specifically, the absorption coefficient α can be evaluated as

$$\alpha = \frac{4\pi k}{\lambda} \tag{3.29}$$

where k is the material extinction coefficient (imaginary part of the refractive index), which decreases as the wavelength λ increases. Values of the germanium refractive index and of the computed absorption coefficient are reported in table B.7.

In order to analyse the spectral response of the detector, simulations have been run on the reference device for a low input optical power $P_{op} = 200\,\mu W$ in the wavelength range between $1.2\,\mu m$ and $1.6\,\mu m$, for two different values of applied bias voltage, i.e., $-1\,V$ and $-2\,V$. The usual square-section silicon input waveguide has been used to couple the light source with the device (200 nm wide, $1\,\mu m$ long).

The computed responsivity is reported in figure 3.64, together with the germanium absorption coefficient. The responsivity behaviour follows the one of the absorption coefficient, i.e., it decreases monotonically in the considered spectral range. A comparison between the optical generation rates within the device at $\lambda = 1.31\,\mu m$, effectively absorbed by the germanium, and $\lambda = 1.55\,\mu m$, yielding a lower photogeneration, can be observed in figure 3.65. A responsivity peak value of $0.72\,A/W$ is reported for $\lambda = 1.28\,\mu m$ while a $0.105\,A/W$ responsivity is obtained at $\lambda = 1.55\,\mu m$, compatible with the $0.12\,A/W$ value reported by the literature [12].

For the sake of completeness, also the computed modulation bandwidth is reported as a function of the wavelength in figure 3.66. As it can be observed, it is quite stable against the wavelength, only slightly decreasing in the considered spectral range.

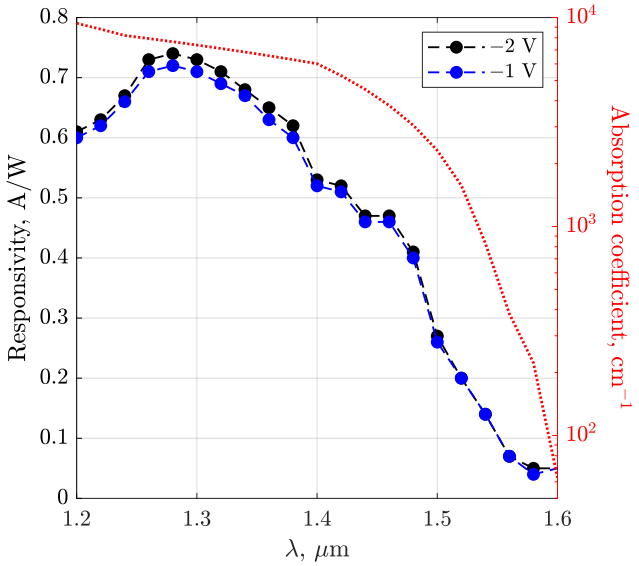


Figure 3.64: Effect of the incident light wavelength on the responsivity of the reference device, obtained for two different bias values and $P_{\text{op}} = 200\,\mu\text{W}$. In red, the germanium absorption coefficient.

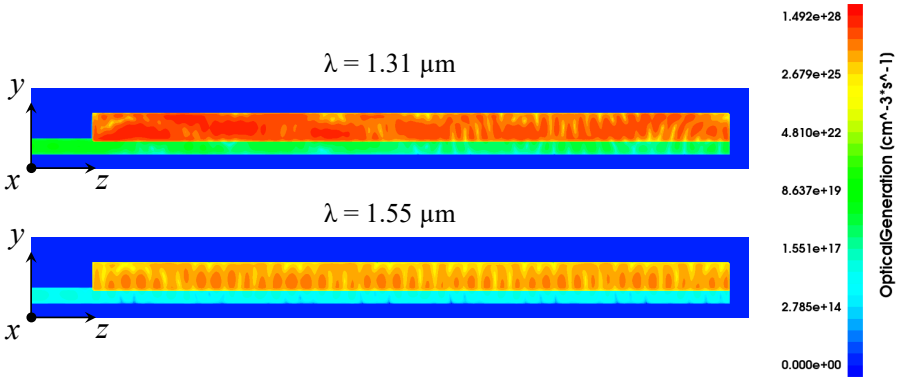


Figure 3.65: Effect of the incident light wavelength on the optical generation rate within the reference device, obtained for $P_{\text{op}} = 200\,\mu\text{W}$. 2D-cut at $x = 0$.

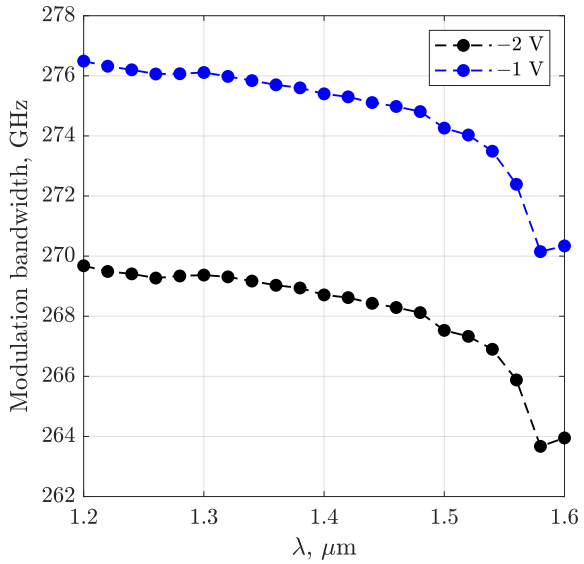


Figure 3.66: Effect of the incident light wavelength on the modulation bandwidth of the reference device, obtained for two different bias values and $P_{\text{op}} = 200\,\mu\text{W}$.

3.6.1 Performance in the C-Band

The focus will now be on the operation in the C-Band, specifically at a $\lambda = 1.55\,\mu\text{m}$ wavelength.

Input waveguide optimization

During the simulations performed in the spectral range between $1.2\,\mu\text{m}$ and $1.6\,\mu\text{m}$, it has been observed that a silicon input waveguide with a squared cross-section and a $200\,\text{nm}$ side does not efficiently guide fundamental modes at all the considered wavelengths. For example, with such a waveguide, at $\lambda = 1.31\,\mu\text{m}$ the fundamental mode exhibits an effective refractive index $n_{\text{eff}} = 1.750$, while at $\lambda = 1.55\,\mu\text{m}$ an effective index $n_{\text{eff}} = 1.540$ is achieved. An effective index close to the silicon oxide refractive index, i.e., $n = 1.44$, is symptom of poor guiding. Therefore, widening of the input waveguide has been considered to allow for a better mode confinement. The waveguide height has been fixed to $H_{\text{wg}} = 0.2\,\mu\text{m}$ and the fundamental mode effective index has been evaluated for width W_{wg} values varying between $0.2\,\mu\text{m}$ and $0.3\,\mu\text{m}$. Results are shown in figure 3.67. It is evident that the wider is the waveguide, the higher is the mode confinement. Based on the observed results, a waveguide width $W_{\text{wg}} = 0.25\,\mu\text{m}$ has been chosen, achieving $n_{\text{eff}} = 1.724$ (close to the reference value of $n_{\text{eff}} = 1.750$

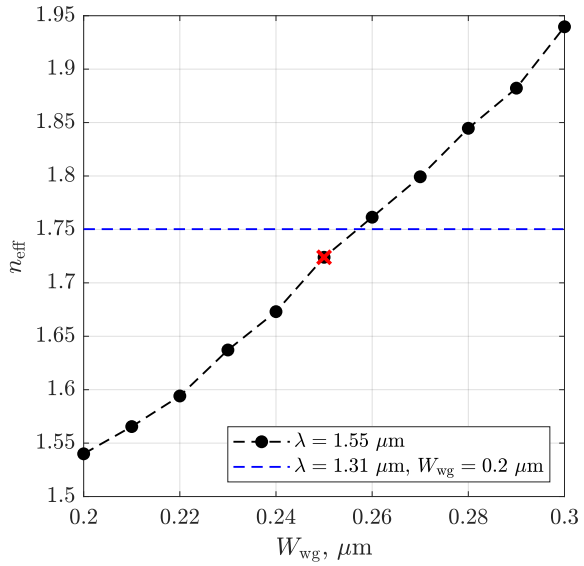


Figure 3.67: Effect of the silicon input waveguide width on the fundamental mode effective refractive index, obtained for a fixed $H_{\text{wg}} = 0.2\,\mu\text{m}$, at $\lambda = 1.55\,\mu\text{m}$. In blue, the reference value $n_{\text{eff}} = 1.750$ obtained at $\lambda = 1.31\,\mu\text{m}$ for $W_{\text{wg}} = H_{\text{wg}} = 0.2\,\mu\text{m}$.

obtained at $\lambda = 1.31\,\mu\text{m}$ for the square-section waveguide). The fundamental mode profile for the widened waveguide is shown in figure 3.68. This waveguide size has been used for all the simulations at $\lambda = 1.55\,\mu\text{m}$.

Electro-optic frequency response

The electro-optic frequency response of the detector has been evaluated at $\lambda = 1.55\,\mu\text{m}$ in order to compare the simulated response with the measured one, reported in the literature [12]. In addition to the reference device, the one having a germanium fin width $W_{\text{Ge}} = 150\,\text{nm}$ has also been simulated, being experimental data available for this device as well. Simulations have been run with a $-2\,\text{V}$ bias applied and a low optical power $P_{\text{op}} = 200\,\mu\text{W}$. In figure 3.69 the agreement between simulation results and available measurements can be noticed.

Effect of the absorber length

The effect of the device length on the device figures of merit has been examined at $\lambda = 1.55\,\mu\text{m}$ as well, as previously done for $\lambda = 1.31\,\mu\text{m}$ in subsection 3.3.3. The length of the device L_{Ge} has been varied in the range between $5\,\mu\text{m}$ and

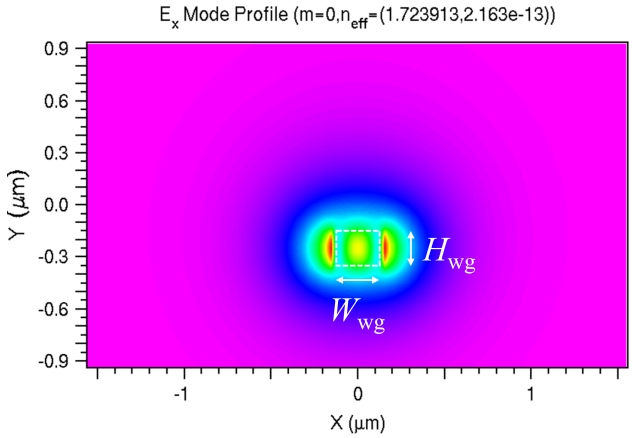


Figure 3.68: Fundamental mode profile (E_x contour map) at $\lambda = 1.55\,\mu\text{m}$ computed on the silicon input waveguide (rectangular cross-section, $W_{\text{wg}} = 0.25\,\mu\text{m}$, $H_{\text{wg}} = 0.2\,\mu\text{m}$). The effective refractive index of the mode is $n_{\text{eff}} = 1.724$.

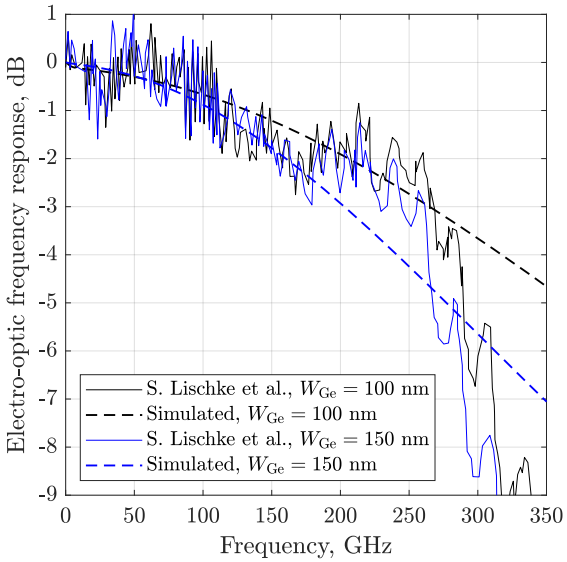


Figure 3.69: Comparison between simulated and measured [12] electro-optic frequency response, obtained for the reference device, having $W_{\text{Ge}} = 100\,\text{nm}$, and a second device with $W_{\text{Ge}} = 150\,\text{nm}$, at a $-2\,\text{V}$ bias and $\lambda = 1.55\,\mu\text{m}$ ($P_{\text{op}} = 200\,\mu\text{W}$ for simulated data).

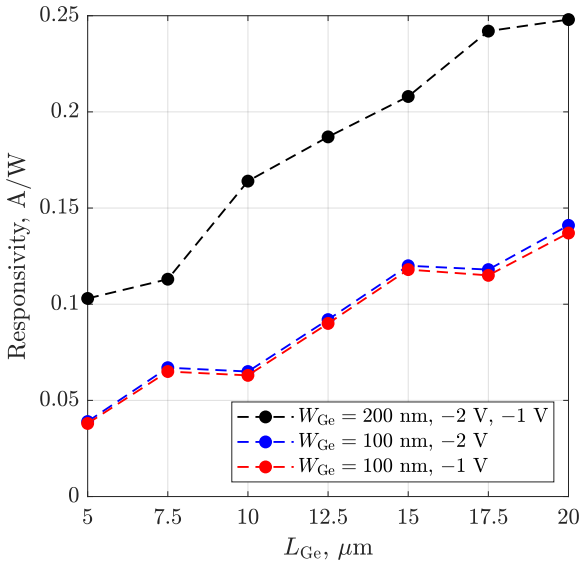
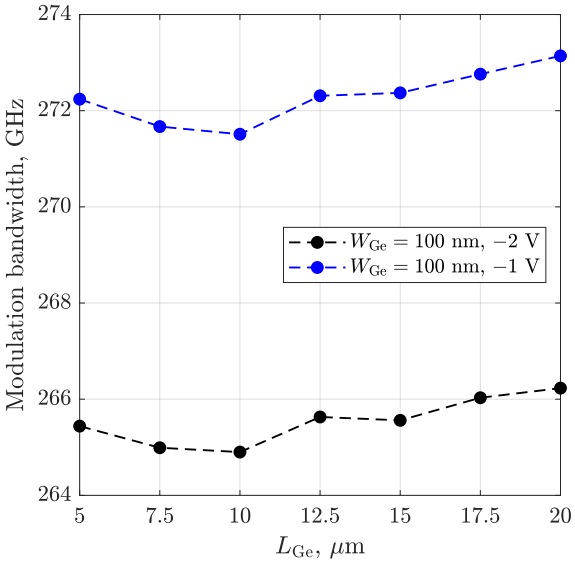


Figure 3.70: Effect of the absorber length L_{Ge} on the responsivity of the device, with two different W_{Ge} values, $H_{Ge} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$, for two different bias values, $P_{op} = 200\,\mu\text{W}$, $\lambda = 1.55\,\mu\text{m}$.

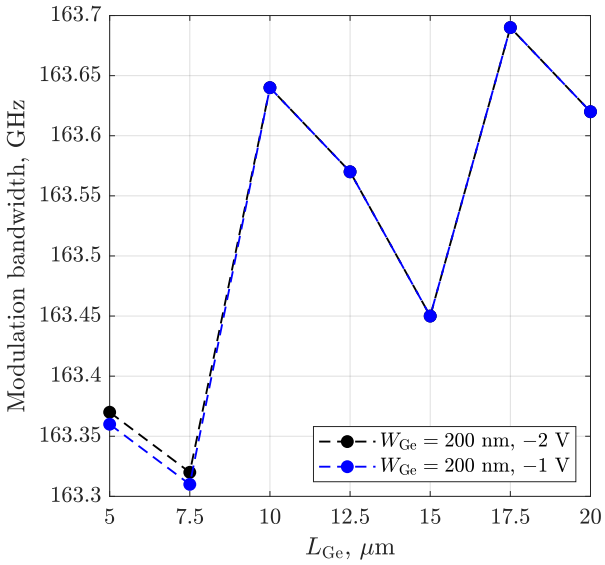
20 µm, for two W_{Ge} values, i.e., 100 nm and 200 nm, keeping constant $H_{Ge} = 400\,\text{nm}$ and $d = 0.38\,\mu\text{m}$. Results are reported for two bias values, i.e., $-1\,\text{V}$ and $-2\,\text{V}$, and for a low power $P_{op} = 200\,\mu\text{W}$.

In general, a longer detector allows for an improvement of the efficiency and, therefore, of the responsivity (see equation (1.23)). The advantage is particularly visible here at $\lambda = 1.55\,\mu\text{m}$ (see results in figure 3.70), since the germanium absorption capability is not so favourable. For example, considering $W_{Ge} = 100\,\text{nm}$, doubling L_{Ge} from 10 µm to 20 µm doubles the responsivity as well.

Modulation bandwidth data are reported as well, as a function of L_{Ge}, in figures 3.71. Only negligible fluctuations occur when the device length is varied, which is expected from a waveguide photodetector, as already discussed for $\lambda = 1.31\,\mu\text{m}$.



(a) $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$.



(b) $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$.

Figure 3.71: Effect of the absorber length L_{Ge} on the modulation bandwidth of the device, with two different W_{Ge} values, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, for two different bias values, $P_{\mathrm{op}} = 200\,\mu\mathrm{W}$, $\lambda = 1.55\,\mu\mathrm{m}$.

3.7 Efficiency optimization

The analysis of the lateral Ge-on-Si photodetector in the C-Band ($\lambda = 1.55\,\mu m$) has highlighted the relatively low responsivity of the device, mainly affected by the poor germanium absorption coefficient. As it has been observed in subsection 3.3.3, an improvement of responsivity is achievable acting on the key dimensions of the absorber. An increase in the active region volume is naturally beneficial for the photocurrent. However, the modulation bandwidth reduces dramatically when the germanium fin width is increased, due to the severe transit time limit. Increasing either the fin height or the device length may be considered to enhance responsivity without significantly sacrificing bandwidth.

A more direct approach to address the issue without modifying the device would be to adopt a tapered input waveguide. Tapered waveguides are commonly used in silicon photonics, as they allow for smooth transitions between components presenting different modal sizes, minimizing insertion losses and reflections while enhancing mode matching [41]. By designing the taper appropriately, the coupling between the incident light and the detector active region can be significantly improved.

The simulated tapered waveguide geometry can be observed in figure 3.72, where significant dimensions are indicated. A standard input waveguide having width W_{wg} and length L_{wg} is followed by a linear taper with length L_{taper} and final width $W_{wg} + \Delta W_{taper}$. The thickness of the whole silicon waveguide is H_{wg}.

Taper designs have been proposed in the O-Band ($\lambda = 1.31\,\mu m$) for the reference device only ($W_{Ge} = 100\,nm$), while in the C-Band ($\lambda = 1.55\,\mu m$) three germanium widths have been considered ($W_{Ge} = 100\,nm$, $125\,nm$, $150\,nm$). First, electromagnetic simulations have been performed, scanning several possible designs and evaluating the *coupling efficiency* as P_{Ge}/P_{in}, i.e., as the fraction of input optical power P_{in} ending up in the germanium active region P_{Ge}. Then the responsivity has been evaluated by means of electrical simulations for optimal designs only.

3.7.1 Taper design in the O-Band

In order to design an optimized taper for an operating wavelength of $\lambda = 1.31\,\mu m$, the coupling efficiency of the reference device ($W_{Ge} = 100\,nm$) has been evaluated for ΔW_{taper} ranging from $-0.1\,\mu m$ (*inverse taper*) to $0.2\,\mu m$. Taper and input waveguide lengths have been both fixed to $L_{wg} = L_{taper} = 1\,\mu m$, the input waveguide width to $W_{wg} = 0.2\,\mu m$ and the waveguide thickness to $H_{wg} = 0.2\,\mu m$.

The obtained results are reported in figure 3.73. While an inverse taper does not seem to improve efficiency, a taper width variation of $\Delta W_{taper} = 0.08\,\mu m$ yields a responsivity enhancement of about 12%. As it can be observed in the

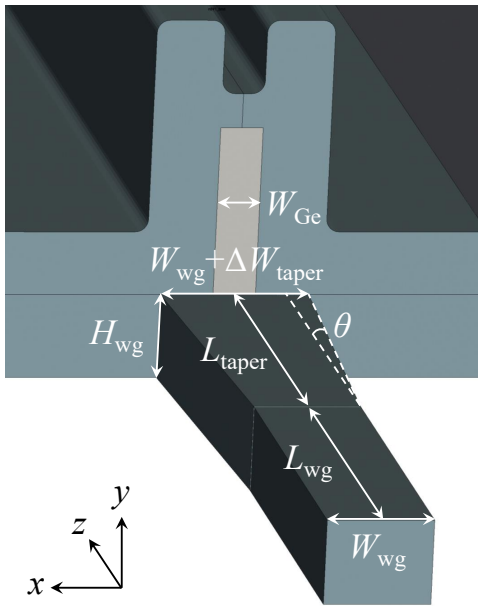


Figure 3.72: Silicon tapered input waveguide structure. Key dimensions varied during design experiments are highlighted.

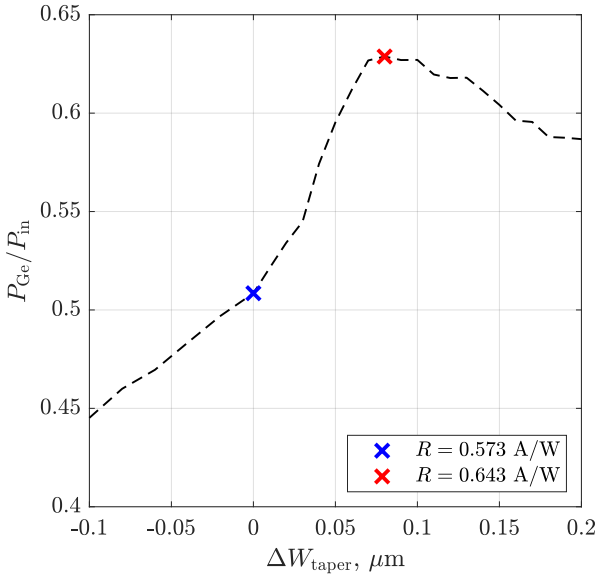


Figure 3.73: Effect of the taper width variation on the reference device ($W_{Ge} = 100\,\text{nm}$) coupling efficiency, obtained at $\lambda = 1.31\,\mu\text{m}$ for $W_{wg} = 0.2\,\mu\text{m}$, $L_{wg} = L_{taper} = 1\,\mu\text{m}$, $H_{wg} = 0.2\,\mu\text{m}$.

2D color maps reported in figure 3.74, the taper modifies the mode coupling between the input waveguide and the main waveguide, resulting in different interference patterns. An optimum design effectively helps to improve light coupling with the device active region.

3.7.2 Taper design in the C-Band

Moving to the C-Band ($\lambda = 1.55\,\mu\text{m}$), a taper has been designed for the reference device first ($W_{Ge} = 100\,\text{nm}$). Taper and input waveguide lengths have been both fixed to $L_{wg} = L_{taper} = 1\,\mu\text{m}$, the input waveguide width to $W_{wg} = 0.25\,\mu\text{m}$ and the waveguide thickness to $H_{wg} = 0.2\,\mu\text{m}$. ΔW_{taper} has been varied in the range between 0 and $0.2\,\mu\text{m}$.

Results shown in figure 3.75 have allowed for the choice of $\Delta W_{taper} = 0.1\,\mu\text{m}$ as optimum taper size (corresponding to a $\theta = 2.86°$ angle). In fact, electrical simulations have confirmed a responsivity improvement by about 20%.

A device identical to the previous one save for the germanium width changed to $W_{Ge} = 125\,\text{nm}$ has been then considered. Keeping the same input waveguide sizes used for the previous device, electromagnetic simulations have highlighted an optimum taper width variation $\Delta W_{taper} = 0.12\,\mu\text{m}$ (corresponding to a $\theta = 3.43°$ angle), yielding a responsivity enhancement by nearly 33%.

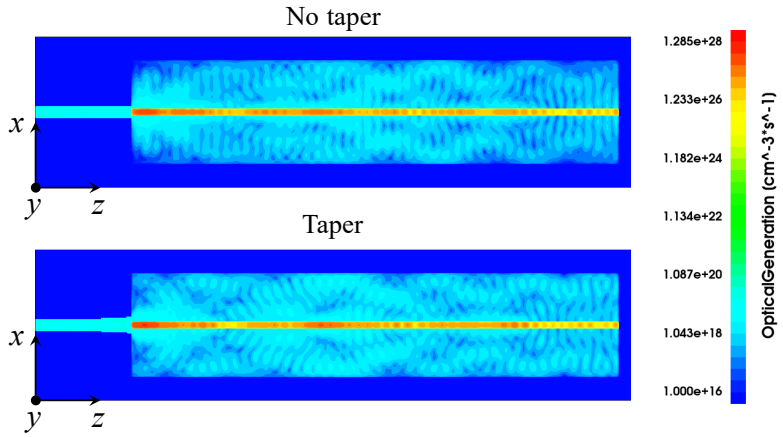


Figure 3.74: Effect of a tapered input waveguide (with $\Delta W_{\text{taper}} = 0.08\,\mu\text{m}$) on the optical generation rate within the reference device, obtained for $P_{\text{op}} = 200\,\mu\text{W}$ at $\lambda = 1.31\,\mu\text{m}$. 2D-cut at $y = -0.15\,\mu\text{m}$, at the interface between the waveguide and the device.

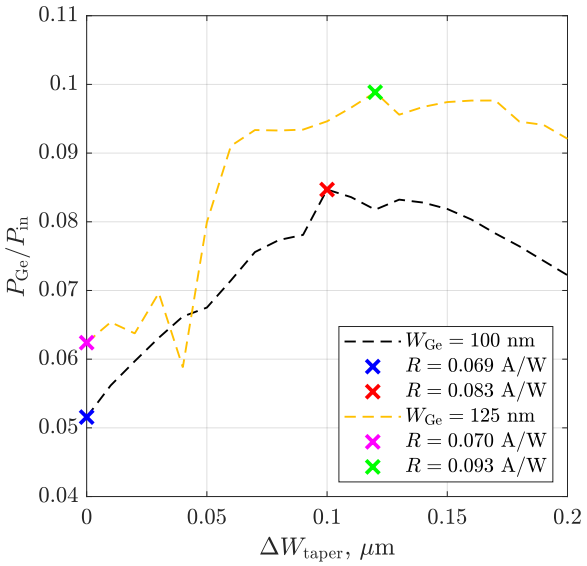


Figure 3.75: Effect of the taper width variation on the coupling efficiency of devices having $W_{Ge} = 100\,\text{nm}$ and $W_{Ge} = 125\,\text{nm}$ widths, obtained at $\lambda = 1.55\,\mu\text{m}$ for $W_{wg} = 0.25\,\mu\text{m}$, $L_{wg} = L_{taper} = 1\,\mu\text{m}$, $H_{wg} = 0.2\,\mu\text{m}$.

It can be observed from the two considered cases and the obtained results that a wider absorber requires a wider taper as well.

A taper design has also been proposed for a device with the germanium width increased to $W_{Ge} = 150\,\mu\text{m}$. First, a ΔW_{taper} scan has been performed keeping the same waveguide sizes adopted for the narrower devices. In particular, lengths have been kept constant to $L_{wg} = L_{taper} = 1\,\mu\text{m}$. The trend observed for $W_{Ge} = 100\,\text{nm}$ and $W_{Ge} = 125\,\text{nm}$, i.e., the need for a wider taper as the germanium fin widens, has been confirmed for $W_{Ge} = 150\,\text{nm}$ as well. However, as it can be noticed from figure 3.76, the coupling efficiency does not exhibit an evident maximum versus the taper width: a sort of saturation is observed instead. In addition, the ΔW_{taper} values required to effectively increase the efficiency are too large, compared to the taper length $L_{taper} = 1\,\mu\text{m}$. As a matter of fact, e.g., a $\Delta W_{taper} = 0.3\,\mu\text{m}$ corresponds to a $\theta = 8.53°$ angle, greater than the $\theta \approx 3°$ angles observed in previous optimizations. The variation of the taper width should be slow enough to allow for a proper evolution of the propagating mode: abrupt changes could cause scattering or reflections, leading to significant losses. An electrical simulation with such an abrupt taper, with $\Delta W_{taper} = 0.33\,\mu\text{m}$, has been run anyway, resulting, as expected, in a scarce responsivity improvement ($\sim 4\%$).

In light of the poor performance of the abrupt taper, the ΔW_{taper} scan has been repeated reducing the input waveguide length to $L_{wg} = 0.5\,\mu\text{m}$ and

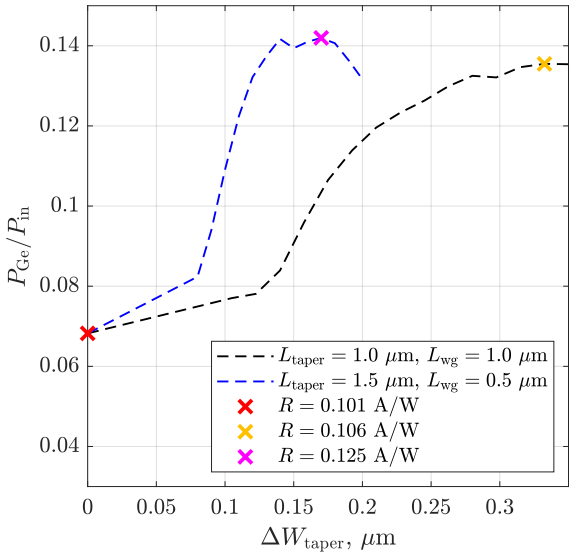


Figure 3.76: Effect of the taper width variation on the coupling efficiency of the device having a $W_{Ge} = 200\,\text{nm}$ width, obtained at $\lambda = 1.55\,\mu\text{m}$ for $W_{wg} = 0.25\,\mu\text{m}$, $H_{wg} = 0.2\,\mu\text{m}$ and two different combinations of L_{wg} and L_{taper}.

increasing the taper length to $L_{taper} = 1.5\,\mu\text{m}$, i.e., keeping $L_{wg} + L_{taper} = 2\,\mu\text{m}$, like in previous simulations. Using a longer taper, the coupling efficiency versus ΔW_{taper} recovers a behaviour similar to the one observed in 3.75, i.e., a maximum occurs for $\Delta W_{taper} = 0.17\,\mu\text{m}$, corresponding to a $\theta = 3.24°$ angle, which is adequate. According to the electrical simulation, the taper improves reponsivity by about 24%.

As a final experiment, the possibility of increasing the thickness of the whole waveguide has been considered to improve light coupling. Returning to the reference device ($W_{Ge} = 100\,\text{nm}$) and to equal lengths of taper and input waveguide ($L_{wg} = L_{taper} = 1\,\mu\text{m}$), the waveguide thickness has been increased to $H_{wg} = 0.25\,\mu\text{m}$, while the input waveguide width has been decreased to $W_{wg} = 0.2\,\mu\text{m}$, i.e., the dimensions of the input waveguide have been swapped. The fundamental mode profile reported in figure 3.77 has been computed at the input. It has to be remarked that this mode is very weakly guided, unlike the previous ones. A scan on ΔW_{taper} has been then performed, finding a maximum efficiency at $\Delta W_{taper} = 0.125\,\mu\text{m}$ ($\theta = 3.58°$), as reported in figure 3.78. Comparing the responsivity previously obtained for $H_{wg} = 0.2\,\mu\text{m}$, $W_{wg} = 0.25\,\mu\text{m}$, $\Delta W_{taper} = 0.1\,\mu\text{m}$, with the one now computed for $H_{wg} = 0.25\,\mu\text{m}$, $W_{wg} = 0.2\,\mu\text{m}$, $\Delta W_{taper} = 0.125\,\mu\text{m}$, an improvement by 5% is observed: even though the enhancement is very limited, it still suggests that increasing H_{wg} may represent a path to be pursued.

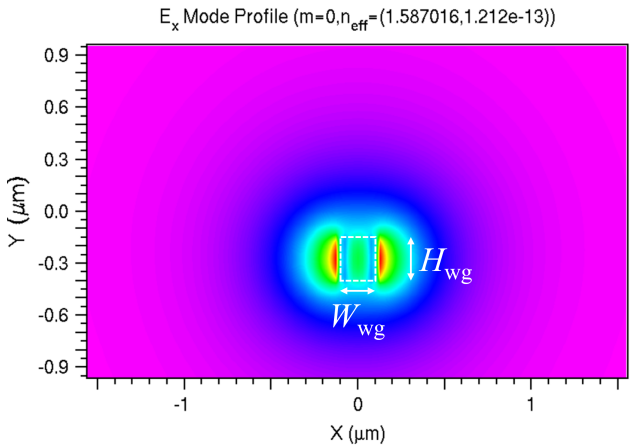


Figure 3.77: Fundamental mode profile (E_x contour map) at $\lambda = 1.55\,\mu\text{m}$ computed on the silicon input waveguide (rectangular cross-section, $W_{\text{wg}} = 0.2\,\mu\text{m}$, $H_{\text{wg}} = 0.25\,\mu\text{m}$). The effective refractive index of the mode is $n_{\text{eff}} = 1.587$.

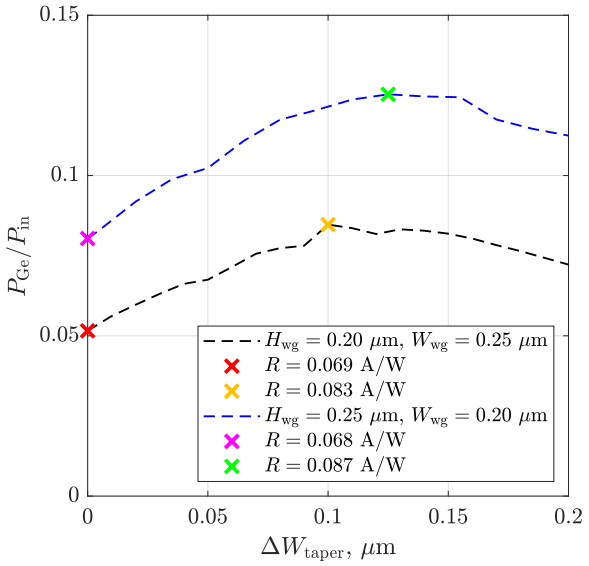


Figure 3.78: Effect of the taper width variation on the coupling efficiency of the device having a $W_{\text{Ge}} = 100\,\text{nm}$ width, obtained at $\lambda = 1.55\,\mu\text{m}$ for $L_{\text{wg}} = L_{\text{taper}} = 1\,\mu\text{m}$ and two different combinations of W_{wg} and H_{wg}.

Chapter 4

Conclusions

The developed physics-based model for the lateral Ge-on-Si *pin* waveguide photodetector has allowed for an extensive study of the device performance, providing deeper insights into its operation conditions. The model has been validated against measurements on fabricated samples, suggesting useful guidelines for the design of optimized devices. Multiphysics simulation has offered the possibility of investigating the effect of variations of several fundamental parameters, such as geometry, doping, bias, as well as the power and the wavelength of the optical source. Parametric analyses of the kind performed in this work would result in unfeasible time and costs for the fabrication and the characterization of real devices.

The small-signal analytical model has highlighted the role of geometry variations in determining the equivalent capacitance and resistance of the device, which ultimately limit the broadband operation of particularly thin detectors. The choice of the applied bias voltage has turned out to be fundamental for narrow devices, since their scarce responsivity can be boosted by avalanche generation phenomena. Depending on requirements in terms of bandwidth and efficiency, sizes of the germanium absorber should be chosen according to a proper compromise. The importance of a high doping of silicon contacts has been justified to achieve the best performance. The developed model has been brought to limits of numerical convergence, analysing the device operation in high optical power conditions, though a comparison with high-power experimental data would be certainly constructive. The bandwidth extension based on the inductive peaking technique would also require validation against real measurements, or at least further simulations considering real inductors, though it is reasonable to assume that this well-established approach would truly unlock the device potential in terms of speed. Operation in the O-Band and the C-Band have been compared, proposing optimizations for the latter, affected by a lower responsivity due to the limited absorption capability of germanium. In particular, attention has been drawn to the problem of light coupling between

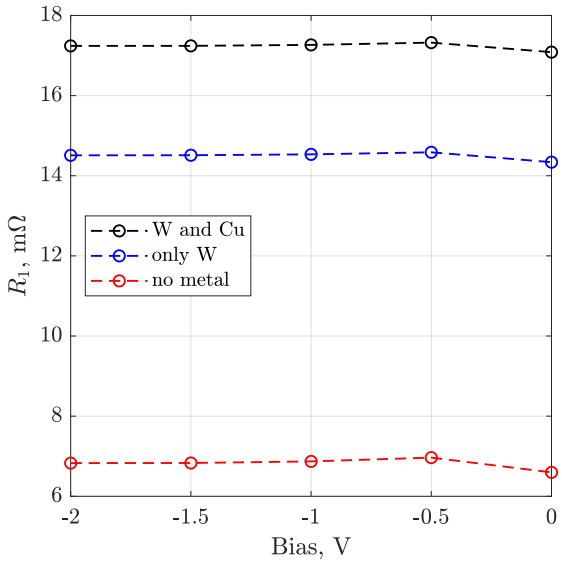
the source and the device: in this sense, the impact of the silicon waveguide sizes and of a possible tapered input waveguide have been demonstrated.

In conclusion, the comprehensive study of this innovative fin-like Ge-on-Si waveguide photodetector has further revealed its role as a perfect candidate for the next generation of optical receivers, enabling unprecedented speed in the detection of optical signals.

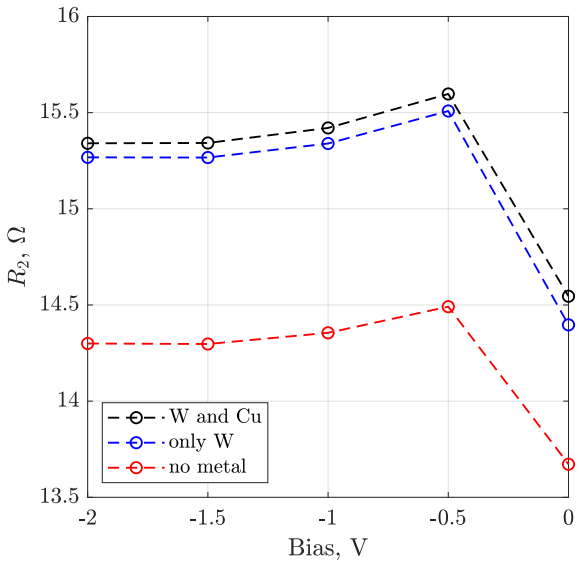
Appendix A

Small-signal parameters

This Appendix includes the small-signal circuit parameters extracted from each performed AC simulation described in section 3.2.

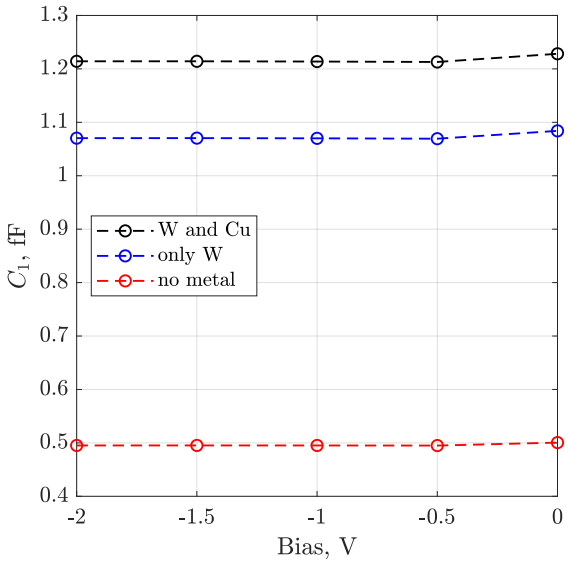


(a)

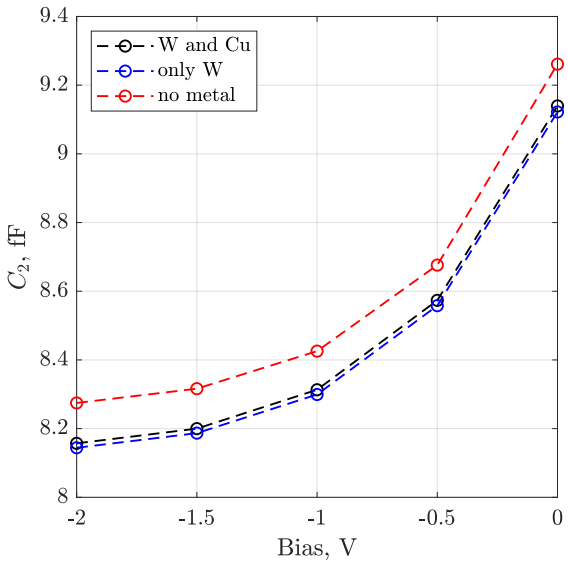


(b)

Figure A.1: Effect of metal contacts on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

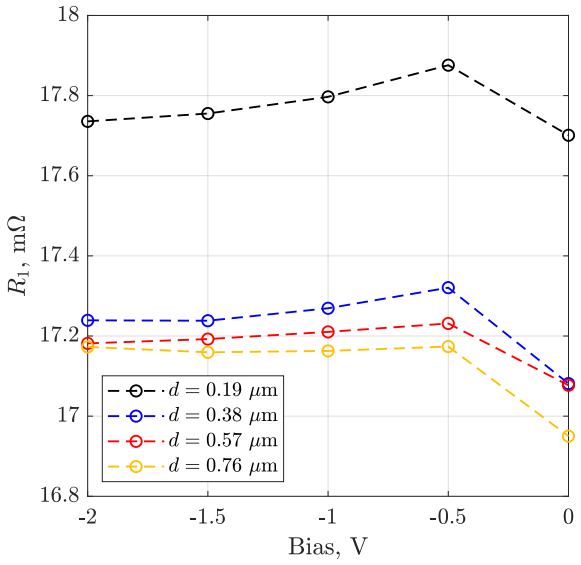


(c)

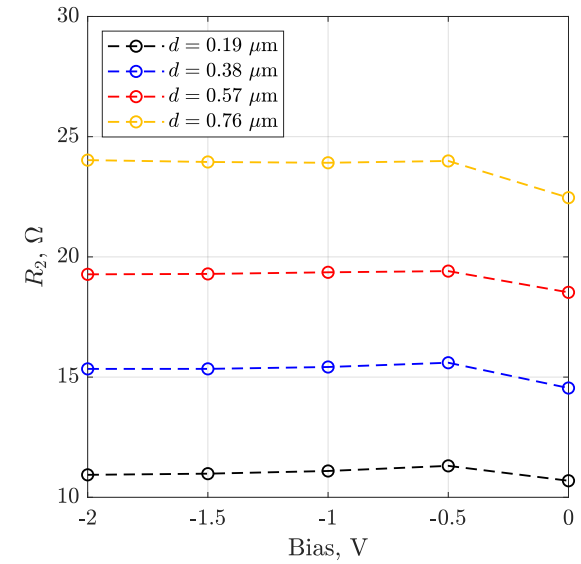


(d)

Figure A.1: Effect of metal contacts on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

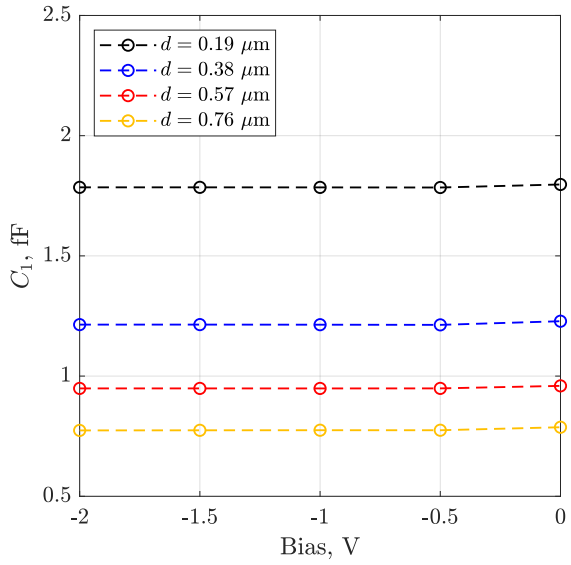


(a)

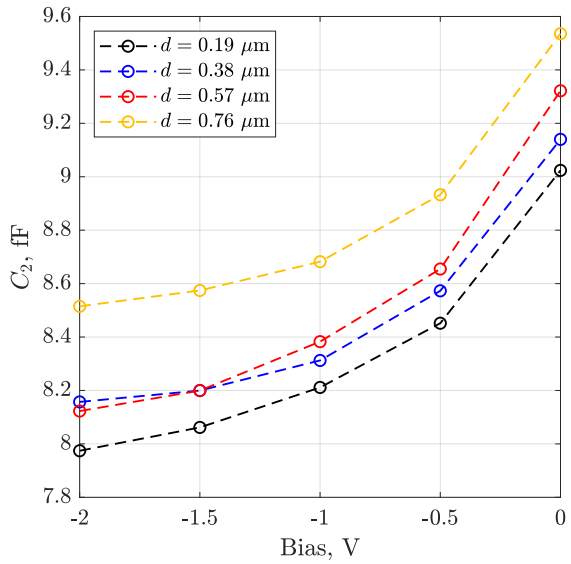


(b)

Figure A.2: Effect of the position of metal contacts on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

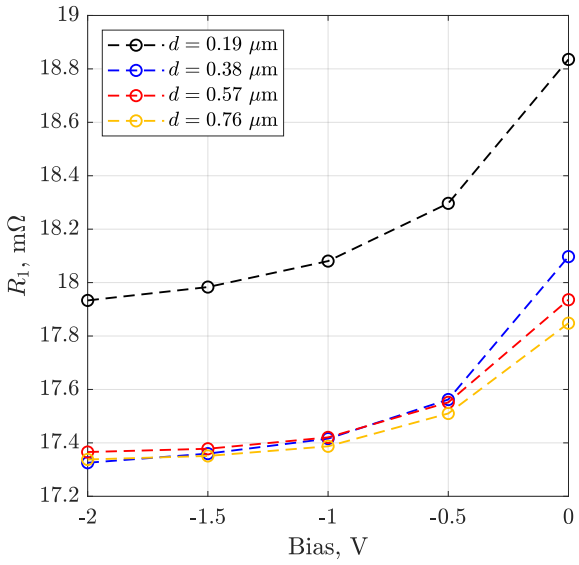


(c)

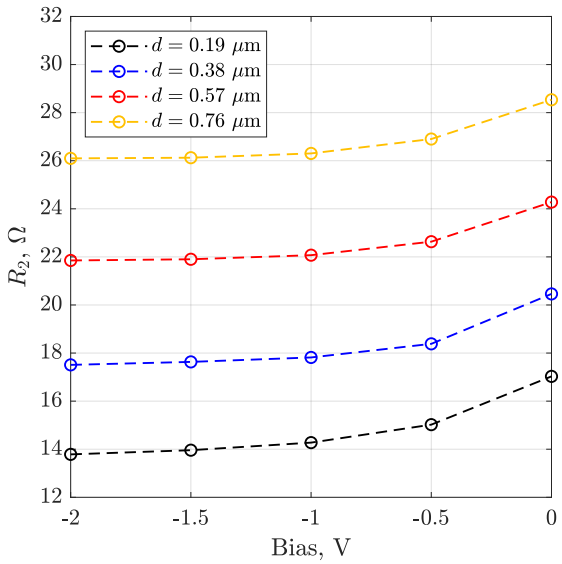


(d)

Figure A.2: Effect of the position of metal contacts on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

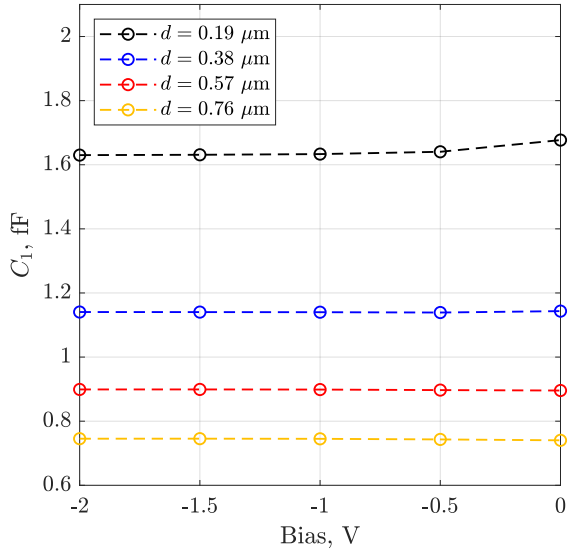


(a)

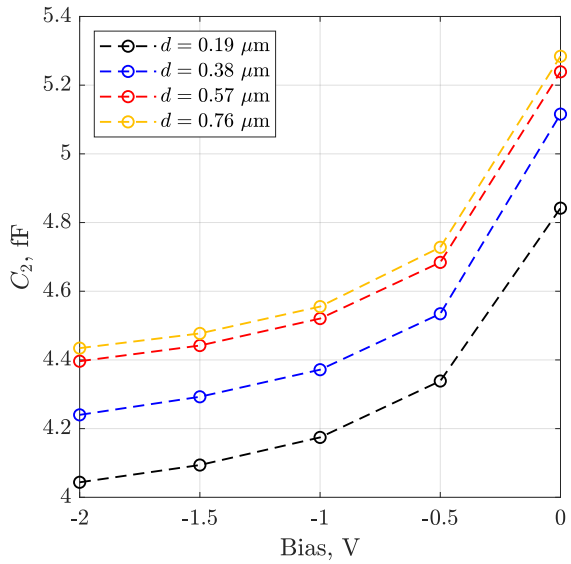


(b)

Figure A.3: Effect of the position of metal contacts on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

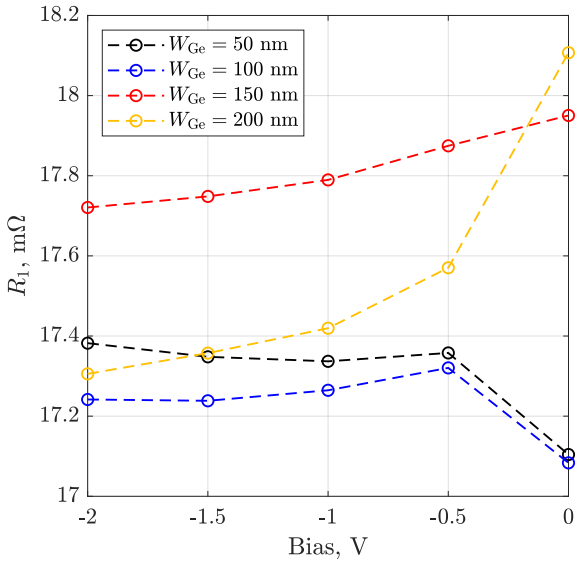


(c)

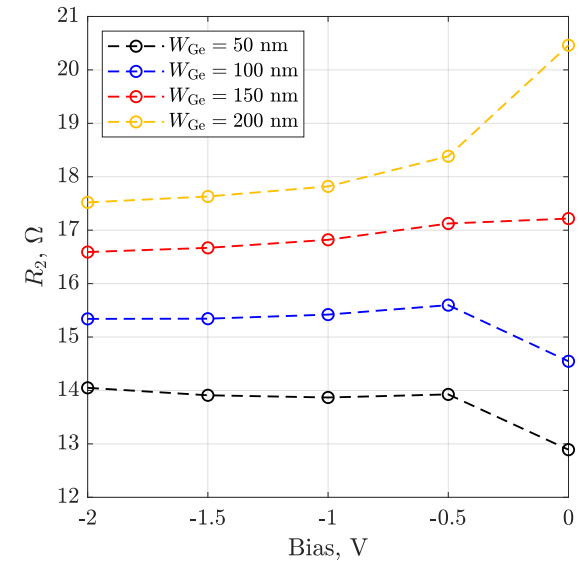


(d)

Figure A.3: Effect of the position of metal contacts on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

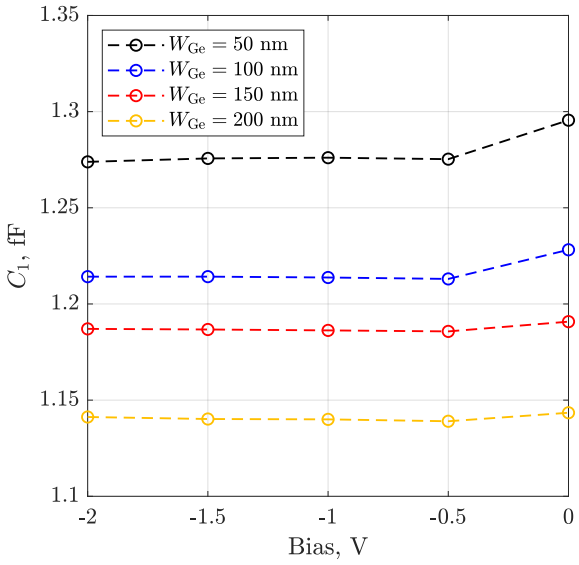


(a)

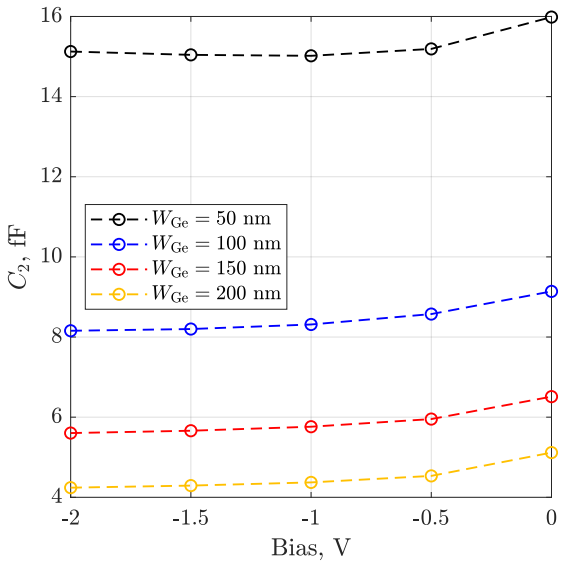


(b)

Figure A.4: Effect of the absorber width W_{Ge} on the extracted small-signal parameters, with $H_{Ge} = 400\,\text{nm}$, $d = 0.38\,\mu\text{m}$, $L_{Ge} = 10\,\mu\text{m}$.

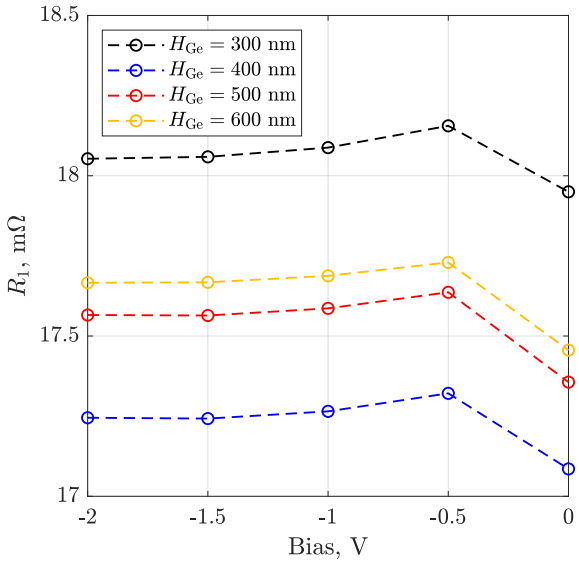


(c)

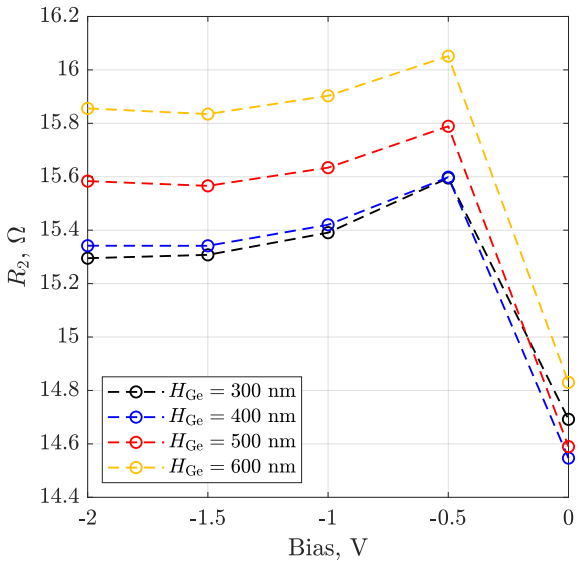


(d)

Figure A.4: Effect of the absorber width W_{Ge} on the extracted small-signal parameters, with $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

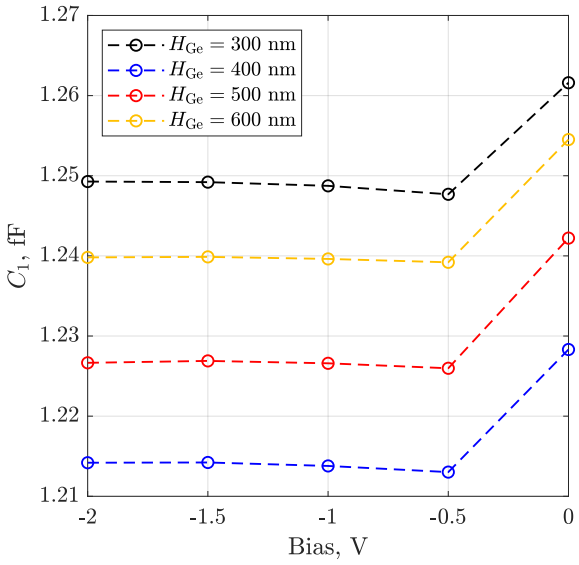


(a)

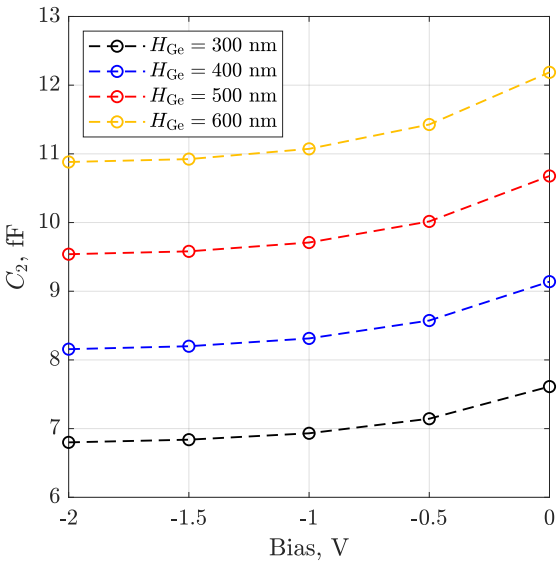


(b)

Figure A.5: Effect of the absorber height H_{Ge} on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

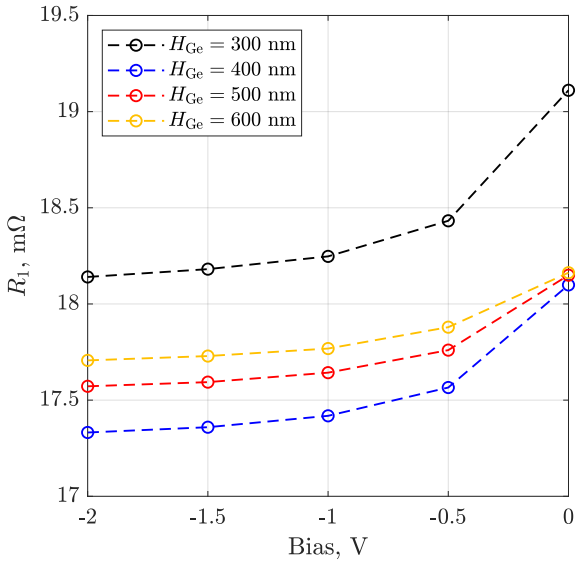


(c)

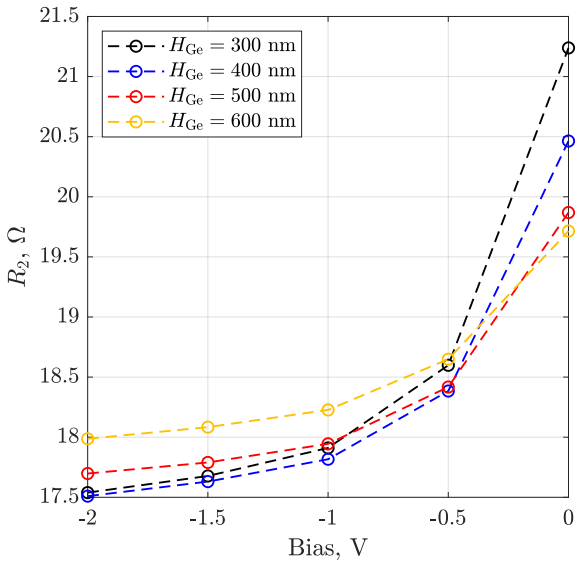


(d)

Figure A.5: Effect of the absorber height H_{Ge} on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

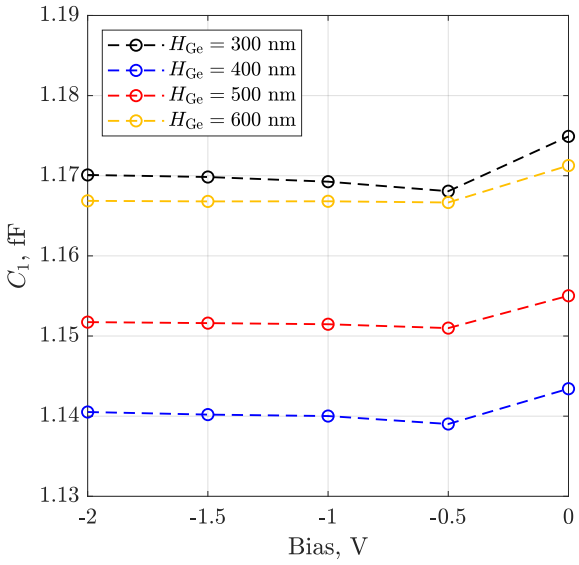


(a)

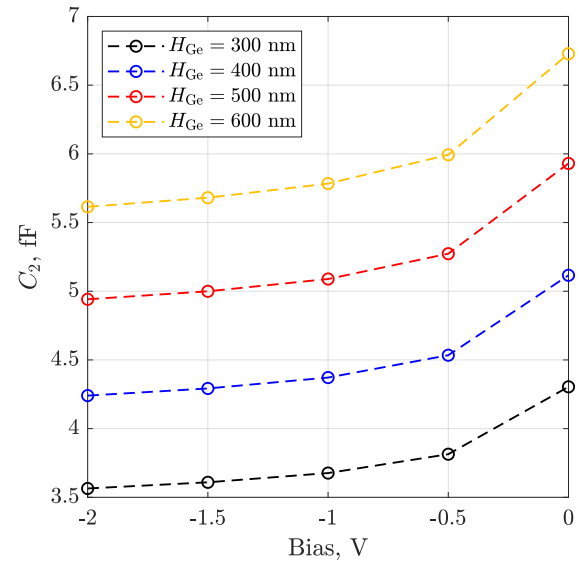


(b)

Figure A.6: Effect of the absorber height H_{Ge} on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

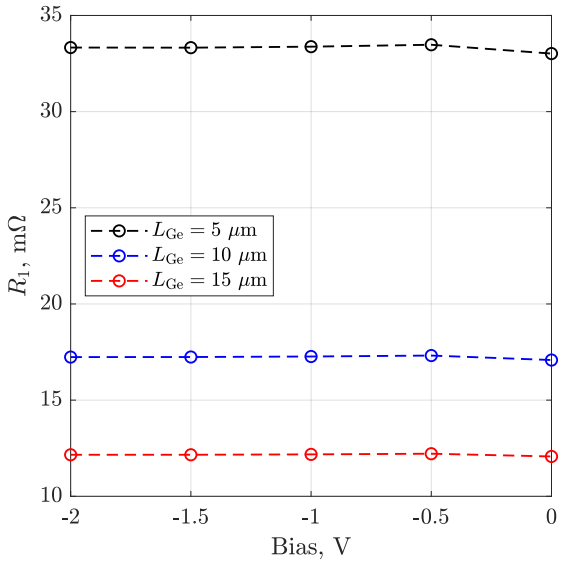


(c)

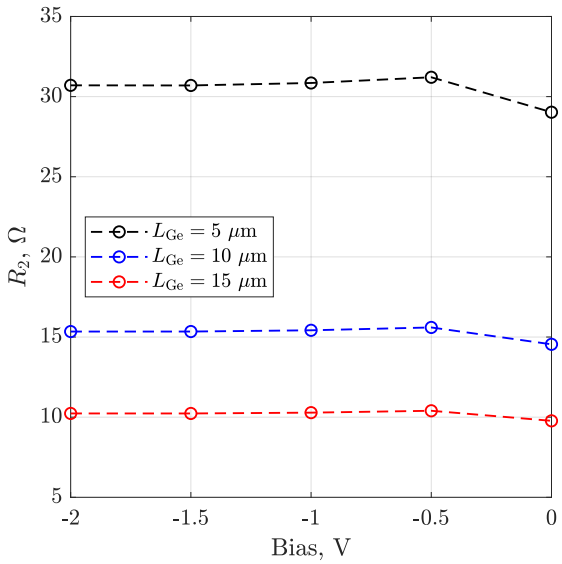


(d)

Figure A.6: Effect of the absorber height H_{Ge} on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 200\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$, $L_{\mathrm{Ge}} = 10\,\mu\mathrm{m}$.

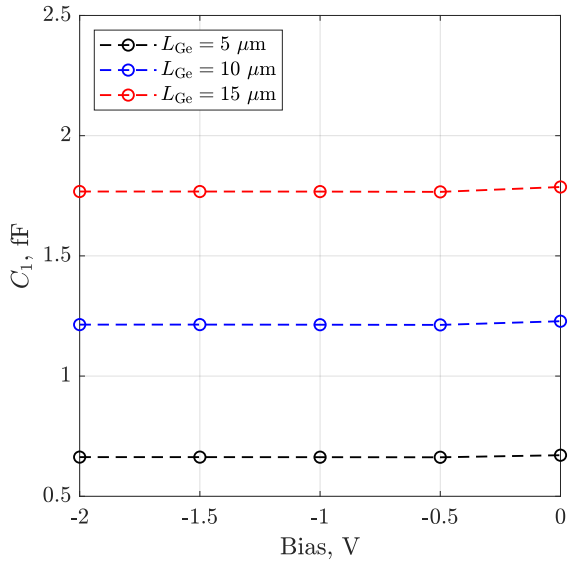


(a)

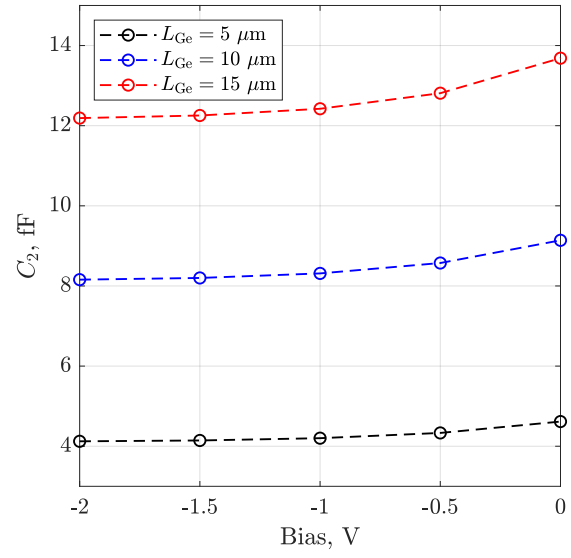


(b)

Figure A.7: Effect of the absorber length L_{Ge} on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$.



(c)



(d)

Figure A.7: Effect of the absorber length L_{Ge} on the extracted small-signal parameters, with $W_{\mathrm{Ge}} = 100\,\mathrm{nm}$, $H_{\mathrm{Ge}} = 400\,\mathrm{nm}$, $d = 0.38\,\mu\mathrm{m}$.

Appendix B

Material parameters

Material parameters adopted in simulations are reported in this Appendix.

	Silicon		Germanium		Unit
	Electron	Hole	Electron	Hole	
E_g	1.12416		0.6638		eV
χ	4.0727		4		eV
$\varepsilon/\varepsilon_0$	11.7		16.2		1

Table B.1: Material properties of silicon and germanium.

	Silicon		Germanium		Unit
	Electron	Hole	Electron	Hole	
τ^{SRH}	10^{-5}	3×10^{-6}	10^{-5}	3×10^{-6}	s
C^{rad}	0		6.4×10^{-14}		cm^3/s
C^{Auger}	2.8×10^{-31}	10.1×10^{-32}	1.5×10^{-34}	1.05×10^{-32}	cm^6/s
a_{aval}	0.426	0.243	0.569	0.559	1/V
b_{aval}	4.81×10^5	6.53×10^5	3.32×10^5	2.72×10^5	V/cm

Table B.2: Generation-recombination models (2.10, 2.11, 2.12, 2.13, 2.14) parameters for silicon and germanium.

	Silicon		Germanium		Unit
	Electron	Hole	Electron	Hole	
μ_{const}	1417	470.5	3900	1900	$\text{cm}^2/(\text{V}\cdot\text{s})$
μ_{min1}	52.2	44.9	60	60	$\text{cm}^2/(\text{V}\cdot\text{s})$
μ_{min2}	52.2	0	0	0	$\text{cm}^2/(\text{V}\cdot\text{s})$
μ_1	43.4	29	20	40	$\text{cm}^2/(\text{V}\cdot\text{s})$
P_{c}	0	9.23×10^{16}	1×10^{17}	9.23×10^{16}	cm^{-3}
C_{r}	9.68×10^{16}	2.23×10^{17}	8×10^{16}	2×10^{17}	cm^{-3}
C_{s}	3.43×10^{20}	6.1×10^{20}	3.43×10^{20}	1×10^{20}	cm^{-3}
α	0.68	0.719	0.55	0.55	1
β	2	2	2	2	1

Table B.3: Mobility Masetti model (2.15) parameters for silicon and germanium.

	Silicon		Germanium		Unit
	Electron	Hole	Electron	Hole	
β	1.109	1.213	1.109	1.213	1
v_{sat}	1.07×10^7	8.37×10^6	7.43×10^6	7.43×10^6	cm/s

Table B.4: Mobility Canali model (2.16) parameters for silicon and germanium.

Wavelength λ, µm	n	k
1.200	3.5193	1.4709×10^{-6}
1.372	3.5007	1.3041×10^{-6}
1.400	3.4876	1.2770×10^{-6}
1.532	3.4784	1.1490×10^{-6}
1.600	3.4710	1.0831×10^{-6}

Table B.5: Silicon refractive index (from RSoft materials library).

Wavelength λ, µm	n
1.200	1.448043947
1.300	1.446910193
1.400	1.445770877
1.500	1.444607762
1.600	1.443407657

Table B.6: Silicon dioxide refractive index (from RSoft materials library).

Wavelength λ, µm	n	k	α, cm^{-1}
1.1300	4.3850	0.1030	1.145×10^4
1.2400	4.3250	0.0810	8.208×10^3
1.4000	4.2380	0.0670	6.013×10^3
1.5500	4.2750	0.0057	462.1
1.6000	4.1730	0.0008	62.83

Table B.7: Germanium refractive index and absorption coefficient (from RSoft materials library).

Bibliography

[1] S. Y. Siew, B. Li, F. Gao, H. Y. Zheng, W. Zhang, P. Guo, S. W. Xie, A. Song, B. Dong, L. W. Luo, *et al.*, "Review of silicon photonics technology and platform development," *J. Lightwave Technol.*, vol. 39, pp. 4374–4389, July 2021.

[2] G. Ghione, *Semiconductor Devices for High-Speed Optoelectronics*. Cambridge, U.K.: Cambridge University Press, 2009.

[3] S. J. Koester, J. D. Schaub, G. Dehlinger, and J. O. Chu, "Germanium-on-SOI infrared detectors for integrated photonic applications," *IEEE J. Select. Topics Quantum Electron.*, vol. 12, pp. 1489–1502, Nov. 2006.

[4] H.-C. Luan, D. R. Lim, K. K. Lee, K. M. Chen, J. G. Sandland, K. Wada, and L. C. Kimerling, "High-quality Ge epilayers on Si with low threading-dislocation densities," *Appl. Phys. Lett.*, vol. 75, pp. 2909–2911, 11 1999.

[5] J. Michel, J. Liu, and L. C. Kimerling, "High-performance Ge-on-Si photodetectors," *Nature Photon.*, vol. 4, no. 8, pp. 527–534, 2010.

[6] H. Chen, P. Verheyen, P. D. Heyn, G. Lepage, J. D. Coster, S. Balakrishnan, P. Absil, W. Yao, L. Shen, G. Roelkens, and J. V. Campenhout, "-1 V bias 67 GHz bandwidth Si-contacted germanium waveguide p-i-n photodetector for optical links at 56 Gbps and beyond," *Opt. Express*, vol. 24, pp. 4622–4631, Mar. 2016.

[7] D. Benedikovic, L. Virot, G. Aubin, F. Amar, B. Szelag, B. Karakus, J.-M. Hartmann, C. Alonso-Ramos, X. L. Roux, P. Crozat, E. Cassan, D. Marris-Morini, C. Baudot, F. Boeuf, J.-M. Fédéli, C. Kopp, and L. Vivien, "25 Gbps low-voltage hetero-structured silicon-germanium waveguide pin photodetectors for monolithic on-chip nanophotonic architectures," *Photon. Res.*, vol. 7, pp. 437–444, Apr. 2019.

[8] Y. Zhu, Z. Liu, C. Niu, Y. Pang, D. Zhang, X. Liu, J. Zheng, Y. Zuo, H. Xue, and B. Cheng, "High-speed and high-power germanium photodetector based on a trapezoidal absorber," *Opt. Lett.*, vol. 47, pp. 3263–3266, July 2022.

[9] N. Shimizu, N. Watanabe, T. Furuta, and T. Ishibashi, "InP-InGaAs uni-traveling-carrier photodiode with improved 3-dB bandwidth of over 150 GHz," *IEEE Photon. Technol. Lett.*, vol. 10, no. 3, pp. 412–414, 1998.

[10] E. Rouvalis, M. Cthioui, F. van Dijk, M. J. Fice, G. Carpintero, C. C. Renaud, and A. J. Seeds, "170 GHz Photodiodes for InP-based photonic integrated circuits," in *IEEE Photonics Conference 2012*, pp. 88–89, 2012.

[11] S. Lischke, A. Peczek, F. Korndörfer, C. Mai, H. Haisch, M. Koenigsmann, M. Rudisile, D. Steckler, F. Goetz, M. Fraschke, S. Marschmeyer, A. Krüger, Y. Yamamoto, D. Schmidt, U. Saarow, P. Heinrich, A. Kroh, M. A. Schubert, J. Katzer, P. Kulse, A. Trusch, and L. Zimmermann, "Ge Photodiode with -3 dB OE Bandwidth of 110 GHz for PIC and ePIC Platforms," in *2020 IEEE International Electron Devices Meeting*, pp. 7.3.1–7.3.4, 2020.

[12] S. Lischke, A. Peczek, J. S. Morgan, K. Sun, D. Steckler, Y. Yamamoto, F. Korndörfer, C. Mai, S. Marschmeyer, M. Fraschke, A. Krüger, A. Beling, and L. Zimmermann, "Ultra-fast germanium photodiode with 3-dB bandwidth of 265 GHz," *Nature Photon.*, vol. 15, pp. 925–931, Dec. 2021.

[13] L. Colace and G. Assanto, "Germanium on silicon for near-infrared light sensing," *IEEE Photon. J.*, vol. 1, pp. 69–79, Aug. 2009.

[14] M. Tada, J.-H. Park, D. Kuzum, G. Thareja, J. R. Jain, Y. Nishi, and K. C. Saraswat, "Low temperature germanium growth on silicon oxide using boron seed layer and in situ dopant activation," *J. Electrochem. Soc.*, vol. 157, p. H371, Feb. 2010.

[15] L. Vivien, C. Baudot, F. Bœuf, B. Szelag, C. Alonso-Ramos, D. Benedikovic, D. Marris-Morini, E. Cassan, S. Guerber, M. Douix, L. Virot, P. Rodriguez, F. Nemouchi, C. Jany, B. Ben Bakir, X. Le Roux, D. Perez-Galacho, M. Berciano, G. Marcaud, I. Charlet, L. Deniel, C. Lafforgue, J. Zhang, S. Serna, P. Damas, P. T. Do, D. Doser, J. Durel, E. Ghegin, V. Vakarin, J.-M. Ramirez, S. Monfray, S. Cremer, E. Duran Valdeiglesias, L. Sanchez, F. Fournel, P. Brianceau, and K. Hassan, "Chapter one - building blocks of silicon photonics," in *Future Directions in Silicon Photonics* (S. Lourdudoss, J. E. Bowers, and C. Jagadish, eds.), vol. 101 of *Semiconductors and Semimetals*, pp. 1–41, Elsevier, 2019.

[16] Z. Wang, Y. Cai, H. Jiang, Z. Tian, and Z. Di, "Graphene-based silicon photonic devices for optical interconnects," *Advanced Functional Materials*, vol. 34, no. 9, p. 2307461, 2024.

[17] T. Yin, R. Cohen, M. M. Morse, G. Sarid, Y. Chetrit, D. Rubin, and M. J. Paniccia, "31 GHz Ge *n-i-p* waveguide photodetectors on silicon-on-insulator substrate," *Opt. Express*, vol. 15, pp. 13965–13971, Oct. 2007.

[18] Y. Zhang, S. Yang, Y. Yang, M. Gould, N. Ophir, A. E.-J. Lim, G.-Q. Lo, P. Magill, K. Bergman, T. Baehr-Jones, *et al.*, "A high-responsivity photodetector absent metal-germanium direct contact," *Opt. Express*, vol. 22, pp. 11367–11375, May 2014.

[19] A. Palmieri, M. Vallone, M. Calciati, A. Tibaldi, F. Bertazzi, G. Ghione, and M. Goano, "Heterostructure modeling considerations for Ge-on-Si waveguide photodetectors," *Opt. Quantum Electron.*, vol. 50, p. 71, Feb. 2018.

[20] Synopsys, Inc., Optical Solutions Group, Ossining, NY, *RSoft FullWAVE User Guide, v2023.12*, 2023.

[21] Synopsys, Inc., Mountain View, CA, *Sentaurus Device User Guide. Version V-2023.12*, Dec. 2023.

[22] Y. Okuto and C. R. Crowell, "Threshold energy effect on avalanche breakdown voltage in semiconductor junctions," *Solid-State Electron.*, vol. 18, pp. 161–168, Feb. 1975.

[23] G. Masetti, M. Severi, and S. Solmi, "Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon," *IEEE Trans. Electron Devices*, vol. 30, no. 7, pp. 764–769, 1983.

[24] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," *IEEE Trans. Electron Devices*, vol. 22, pp. 1045–1047, Aug. 1975.

[25] Synopsys, Inc., Optical Solutions Group, Ossining, NY, *RSoft BeamPROP User Guide, v2023.12*, 2023.

[26] K. Yee, "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagation*, vol. 14, no. 3, pp. 302–307, 1966.

[27] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comp. Phys.*, vol. 114, no. 2, pp. 185–200, 1994.

[28] The MathWorks Inc., "rationalfit." `https://it.mathworks.com/help/rf/ref/rationalfit.html`, 2024.

[29] J.-L. Dellis, "Zfit." `https://www.mathworks.com/matlabcentral/fileexchange/19460-zfit`, 2024.

[30] G. Ghione, *Dispositivi per la microelettronica*. Milano: McGraw-Hill, 1998.

[31] A. Bansal, B. C. Paul, and K. Roy, "An analytical fringe capacitance model for interconnects using conformal mapping," *IEEE Trans. Computer-Aided Design*, vol. 25, pp. 2765–2774, Dec. 2006.

[32] M. G. C. Alasio, P. Franco, A. Tibaldi, F. Bertazzi, S. Namnabat, D. Adams, P. Gothoskar, G. Masini, F. Forghieri, G. Ghione, and M. Goano, "Optical power screening effects in Ge-on-Si vertical *pin* photodetectors," in *Proceedings of SIE 2022. 53rd Annual Meeting of the Italian Electronics Society* (G. Cocorullo, F. Crupi, and E. Limiti, eds.), vol. 1005 of *Lecture Notes on Electrical Engineering*, pp. 155–159, Cham, Switzerland: Springer Nature Switzerland AG, 2023.

[33] J. J. Morikuni and S. M. Kang, "An analysis of inductive peaking in high-frequency amplifiers," in *1992 IEEE International Symposium on Circuits and Systems*, vol. 6, pp. 2848–2851, 1992.

[34] S. Shekhar, J. Walling, and D. Allstot, "Bandwidth extension techniques for CMOS amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 2424–2439, Nov. 2006.

[35] M. Gould, T. Baehr-Jones, R. Ding, and M. Hochberg, "Bandwidth enhancement of waveguide-coupled photodetectors with inductive gain peaking," *Opt. Express*, vol. 20, pp. 7101–7111, Mar. 2012.

[36] A. Novack, M. Gould, Y. Yang, Z. Xuan, M. Streshinsky, Y. Liu, G. Capellini, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "Germanium photodetector with 60 GHz bandwidth using inductive gain peaking," *Opt. Express*, vol. 21, Nov. 2013.

[37] B. Razavi, *RF Microelectronics (2nd Edition) (Prentice Hall Communications Engineering and Emerging Technologies Series)*. USA: Prentice Hall Press, 2nd ed., 2011.

[38] A. Meaamar, C. C. Boon, K. S. Yeo, and M. A. Do, "A wideband low power low-noise amplifier in CMOS technology," *IEEE Trans. Circuits Syst. I*, vol. 57, pp. 773–782, Dec. 2010.

[39] R. Jaeger and T. Blalock, *Microelettronica*. Collana di istruzione scientifica, McGraw-Hill Companies, 2009.

[40] FiberLabs Inc., "Optical Communication Band." `https://www.fiberlabs.com/glossary/optical-communication-band/`, 2023.

[41] L. Chrostowski and M. Hochberg, *Silicon Photonics Design: From Devices to Systems*. Cambridge University Press, 2015.