DISSERTAÇÃO DE MESTRADO Nº 971

EXTENDED ELECTRONIC AND OPTICAL SOLUTIONS FOR SENSITIVE INFRARED PHOTODETECTION

Pablo Nunes Agra Belmonte

DATA DA DEFESA: 20/02/2017

Universidade Federal de Minas Gerais

Escola de Engenharia

Programa de Pós-Graduação em Engenharia Elétrica

EXTENDED ELECTRONIC AND OPTICAL SOLUTIONS FOR SENSITIVE INFRARED PHOTODETECTION

Pablo Nunes Agra Belmonte

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do Título de Mestre em Engenharia Elétrica.

Orientador: Prof. Davies William de Lima Monteiro

Belo Horizonte - MG

Fevereiro de 2017

Belmonte, Pablo Nunes Agra.
B451e Extended electronic and optical solutions for sensitive infrared photodetection [manuscrito] / Pablo Nunes Agra Belmonte. – 2017.
146 f., enc.: il.

Orientador: Davies William de Lima Monteiro.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Apêndices: f. 142-146.

Bibliografia: f. 135-141.

1. Engenharia elétrica - Teses. 2. Espectro infravermelho - Teses. 3. Silício - Teses. I. Monteiro, Davies William de Lima. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.

CDU: 621.3(043)

"Extended Electronic and Optical Solutions for Sensitive Infrared Photodetection"

Pablo Nunes Agra Belmonte

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Mestre em Engenharia Elétrica.

Aprovada em 20 de fevereiro de 2017.

Por:

Prof. Dr. Davies William de Lima Monteiro
DEE (UFMG) - Orientador

Prof. Dr. Paulo Fernando Seixas
DELT (UFMG)

Prof. Dr. Wagner Nunes Rodrigues
DFIS (UFMG)

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PABLO NUNES AGRA BELMONTE

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

EXTENDED ELECTRONIC AND OPTICAL SOLUTIONS FOR SENSITIVE INFRARED PHOTODETECTION

BELO HORIZONTE - MG

FEVEREIRO DE 2017

PABLO NUNES AGRA BELMONTE

EXTENDED ELECTRONIC AND OPTICAL SOLUTIONS FOR SENSITIVE INFRARED PHOTODETECTION

Dissertação de Mestrado submetida à Banca Examinadora pelo colegiado ao Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como parte dos requisitos para a obtenção do Título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Davies William De Lima Monteiro

BELO HORIZONTE - MG

FEVEREIRO DE 2017

This work is dedicated to my parents Silvana and João, and to Maisa,
who never failed to inspire and support me every single day of my life.

ACKNOWLEDGEMENTS

I once heard the Buddhist Mantra: "*Shiken Haramitsu Daikoumyo*"; and so far, it became my motto. Among all translations, the one I find the best is: "*Every experience is sacred, and may be the key to the universe we are seeking*". This master degree was one of this enriching experiences, that have opened and will open many possibilities. Still, any experience may only occur if it has been preceded by an opportunity.

Opportunity to live, to learn something new; every time, you have to grab these opportunities. But most often, they are given to you by someone. Therefore, I feel myself obliged to thank those who have, during this journey, given me the chance to become who I am today. First of all, I would like to express my deep gratitude to my advisor, Professor Dr. Davies William de Lima Monteiro, for his always contagious passion for teaching, helping and also learning together. His trust on me have continuously pushed me forward. I am very lucky to have had the chance to assimilate at least a small share of his vast wisdom; not only regarding the technical aspects of the research, but also concerning his always righteous conduct in teaching and mentoring. I find inspiring his willingness to find new methods of teaching, trying to renew a once boring and standard class into an engaging environment, and not falling back after failed experiences.

But most of all, I am very grateful to Davies for creating the *OptMA*^{lab} group, setting over the years a non-hierarchical and collaborating mindset. Of course, I cannot forget Professor Dr. Frank Sill Torres and Professor Dr. Luciana Pedrosa Salles, who also share the same mindset, and never hesitate in helping. The professors, and their complementary knowledge, are the pillars of the *OptMA*^{lab}. I am very lucky to be part of this group, and feel more and more proud as I see how it has grown and matured since I have joined it, more than five years ago. The relaxed atmosphere at the lab, the cooperativism of its members, and their desire to learn are the key factors that turns it into a self-edifying environment and the best place one can work in!

Also, I thank Professor Dr. Paddy James French and Dr. Gregory Pandraud from TU-Delft (Delft University of Technology), who have mentored and immensely helped me with a huge part of the work presented in this thesis.

Finally, but not least important, I would like to thank my family, from the bottom of my heart: my parents Silvana and João; and (the "recently" added new member) Maisa. They are the main responsible for who am I today, always supporting me with whatever I needed along this journey, with (a lot of) patience. Without them, it would be impossible to reach the end.

"If you think we can't change the world,
it just means you're not one of those that will"
Jacque Fresco - The Venus Project

RESUMO

Esse trabalho tem como objetivo analisar as duas partes principais de um sistema de detecção no infravermelho, que são as partes ópticas e a eletrônica de leitura. Além disso, um exemplo idealizado de um sistema de detecção de CO_2 é proposto, de forma a ilustrar como as diferentes partes devem ser integradas em uma aplicação prática. O propósito da eletrônica de leitura é fazer a interface com o fotodetector, para ler o seu sinal de corrente. No entanto, já que fotodetectores para o infravermelho apresentam baixa resistência paralela, alta corrente de escuro e ruído, não é uma tarefa trivial para o circuito de leitura conseguir extrair informação útil do detector. Logo, o emprego de circuitos especiais capazes de lidar com altas correntes de entrada, e também de controlar rigorosamente a tensão de polarização no detector, se faz necessário nesse contexto. Além do mais, circuitos de leitura para o infravermelho, geralmente precisam diminuir a sensibilidade do sistema de detecção, de forma que seja possível ler adequadamente altas correntes. Essa abordagem, no entanto, normalmente diminui a relação sinal ruído (SNR), especialmente para altas correntes. No entanto, somente uma parte dessa corrente está relacionada com o sinal útil, e a diminuição do SNR pode mascarar esse pequeno sinal. Portanto, apresentamos nesse trabalho um circuito denominado de pixel ricochete, que apresenta uma grande capacidade de armazenamento de carga necessária para acomodar altas correntes, e ainda com uma alta sensibilidade, sendo capaz de detectar pequenas variações em uma corrente relativamente alta. De fato, o SNR não pode aumentar além do SNR do fotodetector, que impõe o limite máximo. Porém, quanto maior a sensibilidade, mais perto o SNR do sistema fica daquele máximo imposto pelo detector. Além disso, uma versão especial do pixel ricochete é apresentada: o pixel ricochete cascode. Esse circuito apresenta um desempenho ainda melhor, sendo apropriado para aplicações no infravermelho, principalmente por causa de sua habilidade de manter uma tensão de polarização sobre o detector ainda mais estável.

A parte óptica do sistema é composta sobretudo das microlentes, que aumenta o SNR do sistema ao amplificar o sinal óptico de entrada. Essas lentes precisam ser transparentes no espectro infravermelho, e o silício demonstra ser uma boa escolha, especialmente devido ao seu baixo custo, baixa densidade, e relativamente alto índice de refração. Portanto, é um material apropriado para lentes com distância focal curta, possibilitando a criação de sistemas compactos e leves. As microlentes de silício foram fabricadas usando uma técnica apresentada por de Lima Monteiro et al., 2003, que é mais barata do que os métodos convencionais. As amostras microfabricadas também foram caracterizadas, adquirindo informação de rugosidade e da vergência das lentes. Além disso, um experimento foi montado para a medição dos pontos focais das amostras, mostrando que os wafers de silício microusinados com o processo mencionado se comportam de fato como lentes convexas e convergentes. No entanto, silício é um material que apresenta alta reflectância, e se faz necessária a deposição de camadas antirreflexo sobre a sua superfície, para que a transmitância seja maximizada nos comprimentos de onda alvo.

Para tal, usamos a técnica *LPCVD deposited low stress Si-rich nitride* para depositar camadas de SiN (Nitreto de Silício) de várias espessuras, e medimos suas transmitâncias. Perfis de transmitância muito bons foram encontrados em torno de 4.2 μm (um dos comprimentos de onda em que CO_2 absorve), apesar de que as constantes ópticas usadas para as simulações dos projetos com SiN não eram adequadas. Por este motivo, um novo conjunto de constantes ópticas foi extraído das medições, permitindo que futuros projetos com SiN otimizados para outros comprimentos de onda (dentro de uma faixa do infravermelho) possam ser simulados de forma confiável.

Ao final, ambas as partes do sistema foram integradas visando uma aplicação idealizada para um sistema de detecção de CO_2. Um modelo que descreve o espectro de absorção de CO_2 foi implementado para emular o sistema real. Com isso, pudemos concluir que o uso das microlentes de silício mais a camada antirreflexo, e o pixel ricochete podem potencialmente trazer o SNR do sistema para perto do SNR do detector, para toda a faixa de correntes de entrada, uma vez que a sensibilidade é mantida bem alta. Para esse sistema idealizado, poderiam ser detectadas variações muito pequenas de concentração de CO_2, exemplificando o desempenho realizável com tal sistema.

Palavras-chave: Infravermelho, Circuito de Leitura, Alta Sensibilidade, Microlentes de Silício, Camada Antirreflexo.

ABSTRACT

This work aims to address the two main parts of an infrared photodetection system, which are the optics and the readout electronics. Also, an idealized example of a CO_2 gas detection system is presented, in order to illustrate how the different parts should be coupled together in a practical application. The purpose of the readout electronics is to interface with a photodetector, reading out its input signal. However, since an infrared photodetector features low shunt resistance, large dark current, and large noise, it is not a trivial task for the readout electronics to properly extract signal information from the detector. Therefore, especial circuits able to handle large input currents, and capable of tightly control the voltage bias across the detector are desired in this context. In addition, generally infrared readout circuits need to decrease the sensitivity of the detection system, so that larger currents can be properly read. This approach, however, usually decreases the Signal to Noise Ratio (SNR), especially for larger currents. Nonetheless, only a small portion of this current is related to the useful signal, and decreasing the SNR can mask this small current signal. Therefore, we present in this work a circuit, coined as the bouncing pixel, that presents a large charge storage capacity necessary to accommodate large currents, and yet with a very high sensitivity, being able to sense small differences in relatively high input currents. In fact, the SNR cannot increase above the SNR of the photodetector, that imposes the maximum limit. However, the larger the sensitivity, the closer the overall system SNR is to that of the detector. Also, a special version of the bouncing pixel is presented: the cascoded bouncing pixel. This circuit presents an even better performance, being suitable for infrared applications, mainly due to its ability of maintaining an even more stable voltage bias across the detector.

The optical part of the system is comprised mainly by microlenses, that enhance the SNR of the system by strengthening the input optical signal. These lenses need to be transparent at the infrared range, and silicon proves to be a good choice, due to its low cost, low density, and relatively high refraction index. Therefore, it is a material suitable for lenses with very short focal length, allowing compact and light weighted systems. The silicon microlenses were fabricated using a process technique presented by de Lima Monteiro et al., 2003, that is less expensive than the conventional ones. The micromachined samples were also characterized, acquiring information of roughness and vergence of the lenses. In addition, a setup was assembled in order to measure the focal points of the samples, showing that the silicon wafers micromachined with the presented process actually behave as convex verging lenses. However, silicon features a high surface reflectance, and an anti-reflective coating needs to be deposited over the silicon surface, in order to maximize transmittance at the desired wavelength range. We have used the *LPCVD deposited low stress Si rich nitride* technique to deposit SiN layers of various thicknesses, and measured their transmittances. Very good transmittance profiles were found near 4.2 μm (one of the wavelengths at which CO_2 absorbs), although the optical constants dataset used for the SiN design simulations were not appropriate. Therefore, an updated fitting dataset were

extracted from measurements, allowing future designs with SiN aiming at different wavelengths (within a given range in the infrared) to be reliably simulated.

At the end, both optical and electronic parts are coupled together in the idealized application of a CO_2 gas detection system. A model was implemented describing the wavelength dependent CO_2 gas absorption to emulate the real application. We were able to conclude that the use of the silicon microlenses plus the anti-reflective coating and the bouncing pixel can potentially bring the SNR of the system closer to that of the photodetector, for the whole input range of currents, since the sensitivity is kept very high. For the idealized system, very small changes in the CO_2 gas concentration could be detected, exemplifying the achievable performance of such system.

Keywords: Infrared, Readout Circuit, High Sensitivity, Silicon Microlenses, Anti-Reflective Coating.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

IR	Infrared
PD	Photodetector/Photodiode
SNR	Signal-to-Noise Ratio $[dB]$
SNR_{in}	Input SNR $[dB]$
SNR_{out}	Output SNR $[dB]$
FF	Fill-Factor $[\%]$
ARC	Anti-Reflective Coating
S	Sensitivity $[mV/\mu A]$
ΔT	Integration Time $[\mu s]$
DR	Dynamic Range $[dB]$
HDR	High Dynamic Range
AMS	Analog Mixed-Signal Designer Simulator
λ	Wavelength $[\mu m;\ nm]$
T	Absolute Temperature $[K]$
NIR	Near Infrared
$SWIR$	Short Wavelength Infrared
$MWIR$	Mid Wavelength Infrared
$LWIR$	Long Wavelength Infrared
FIR	Far Infrared
E_g	Energy Bandgap $[eV]$
k_B	Boltzmann Constant $[J/K;\ \ eV/K]$
q	Electron Charge $[C]$
V_T	Thermal Voltage $[V]$
R_0	Junction Shunt Resistance $[\Omega]$
R_s	Series Resistance $[\Omega]$
R_{in}	Input Resistance $[\Omega]$
Φ_0	Photon Flux Density $[(number\ of\ photons)/(m^2.s)]$
η	Quantum Efficiency $[\%]$
A_{det}	Photosensitive Area of the Detector $[m^2]$
B	Equivalent Noise Bandwidth $[Hz]$
I_{dark}	Dark Current $[A]$
$BLIP$	Background-Limited Performance
FOV	Field of View $[\ ^\circ\ ;\ rad]$
T_B	Background/Environment Temperature $[K]$
NEP	Noise-Equivalent Power $[W]$
$Jones$	Specific Detectivity unit $[cm.\sqrt{Hz}/W]$
D^*	Specific Detectivity $[Jones]$
V_{DET}	Voltage bias across the detector $[V]$
ADC	Analog-to-Digital-Converter
δV_{min}	Voltage Resolution (defined by the ADC) $[V]$
SN	Sense Node
V_{DD}	Maximum operational voltage of the circuit $[V]$
GND	Ground reference voltage (defined as $0\ V$)
SFD	Source-Follower Per Detector
APS	Active Pixel Sensor
DI	Direct Injection
BDI	Buffered Direct Injection
GMI	Gate Modulation Input
$CTIA$	Capacitive Transimpedance Amplifier
CMI	Current Mirroring Integration
MOS	Metal-Oxide-Semiconductor
$MOSFET$	MOS Field Effect Transistor
$CMOS$	Complementary MOS
$PMOS$	p-channel MOS transistor
$NMOS$	n-channel MOS transistor
g_m	transistors transconductance gain
$f/\#$	f-number $[1]$
f	focal length $[m]$
τ_A	Atmospheric transmission $[\%]$
τ_O	Optical system transmission $[\%]$
n	Refractive Index
k	Extinction Coefficient
RIE	Reactive Ion Etching
$LPCVD$	Low-Pressure Chemical Vapor Deposition
$PECVD$	Plasma-Enhanced Chemical Vapor Deposition
ALD	Atomic Layer Deposition
$INCT\text{-}DISSE$	Brazilian National Institute of Science and Technology for Semiconductor Nanodevices
$FTIR$	Fourier Transform Infrared
PSD	Power Spectral Density
ε	Injection efficiency $[\%]$
σ/σ_V	RMS noise current/voltage

SUMMARY

1. INTRODUCTION

One of the most deeply spread and significant technologies, that profoundly affected our daily lives in recent modern history, are the optical detectors, also known as photodetectors, light sensors, or optoelectronic sensors. These devices are widely used in many applications, ranging from motion detectors to photocells, from cameras to gas sensors [1].

This kind of sensor usually transduces an input optical signal into an output electrical one; being the output either a charge, a voltage or a current signal. In order to extract any information from these sensors, generally, they are coupled along with a read-out circuit that measures the electrical signal, and therefore, has an impact on detection quality [1]. Read-out circuit topologies vary widely, mainly due to the kind of light sensor it is interfacing with. As the very name states, light sensors are sensitive to light; either visible or *invisible*. And, when it comes to invisible light, a really important spectral range is the *infrared*, for it offers a whole new way of perceiving our surroundings, and hence, many new applications are made possible. Nonetheless, detecting infrared light is not straightforward, since infrared sensors usually have several working constraints, mainly high background noise, which is usually much larger than the phototransduced signal itself.

Read-out circuit topologies have been proposed in order to work around the high background noise issue [2], [3]. This master thesis proposes a pixel circuit architecture that prevents signal saturation in infrared signal detection, or any other system that features a very high background signal or an intense signal itself. The following sections of this introductory chapter elaborate on the aforementioned problem and proposed solution, starting by section 1.1 with a brief explanation of how a read-out circuit works, how it interfaces with its detector, and how to measure its performance. Then, section 1.2 presents a way of improving the signal detection by using *microlenses*. Motivation and goals of this thesis are discussed in section 1.3. Finally, section 1.4 describes the content of the subsequent chapters.

1.1 The read-out circuit and detection noise

Although most of the literature on infrared detection systems defines the read-out circuit as *ROIC* (Read-Out Integrated Circuit) [4], [5], we will herein use the nomenclature *pixel circuit*, a far more conventional terminology, mainly due to the imaging systems marketing [1], [6], [7], with the acronym of *MPix* (Mega Pixels) used to indicate the number of pixels in a photograph camera.

A *pixel* is composed by two main parts: the photodetector (PD), and the pixel circuit itself. The first is responsible for transducing an impinging light beam into a quantifiable electrical signal; usually a current. The latter measures this current through a series of electronic circuitry, accounting for filtering, amplifying, and interfacing the transduced signal to an output processor, where data can be interpreted. Usually, the pixel circuit works as an integrator, producing an

output voltage whose slope over time is proportional to the measured current, as shown in Figure 1. After a given time interval, we can read the voltage signal, which is also proportional to the light intensity over the photodetector.

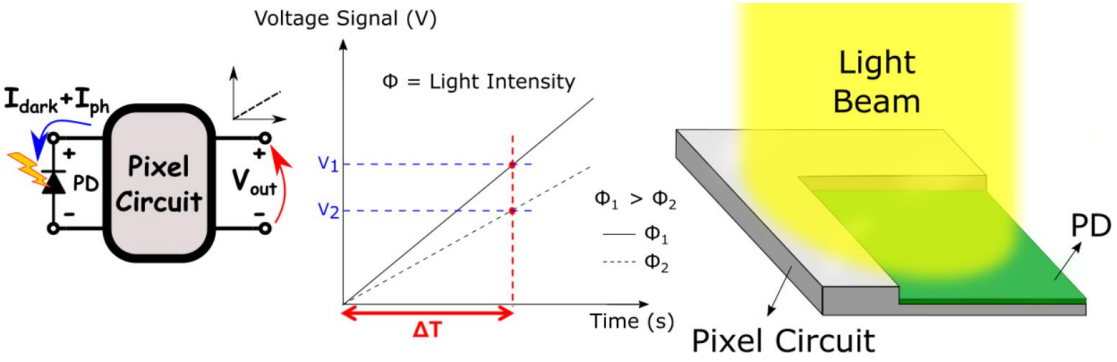

Figure 1: Left: Schematic of a Pixel Circuit interfacing with a PD. Center: Output voltage proportional to the light intensity. Right: A Pixel comprising the photodiode and read-out circuit.

When light is focused on the sensor, it generates what is called the *photocurrent* I_{ph} (or *photogenerated* current), which is proportional to the impinging light flux, namely the number of *photons* arriving at the detector per unit time [1]. The more light it gathers, the larger is the sensor photocurrent. If no light comes to the detector, however, it still produces a non-zero amount of current, due to thermal excitation, called the *dark current* I_{dark} [4], [5]. The sum of both dark and photogenerated currents yields the total current through an optical detector, which in turn is measured by the pixel circuit.

Ideally, dark current should be zero for optimal pixel operation, and it arises with an associated noise, called *dark noise*. For many applications and their respective optical detectors, dark current can be neglected. However, usual infrared light sensors feature very high dark current, which is typically much larger than the target photogenerated current signal [8], [9]. Furthermore, for infrared detectors, background irradiation may produce an undesirable amount of photogenerated current, and consequently, an associated background noise.

A figure of merit quantifying how strong a signal is in respect to its noise floor, is the *Signal-to-Noise Ratio* (SNR), defined as the optical signal power divided by the *Noise Equivalent Power* (*NEP*) [1], [5], as will be explained in chapter 2.

1.2 Optics

The ratio between the detector area and the total pixel area (including both pixel circuit and sensor) is generally much lower than 100%. This relation is called the *fill-factor* (*FF*), which is preferred to be as close to unity as possible, since the bigger the detector area, the more light it gathers, and hence the stronger the photogenerated current signal. The bottom line is that a higher FF improves the SNR and, a technique to counteract a low FF is to deploy microlenses on top of the pixel.

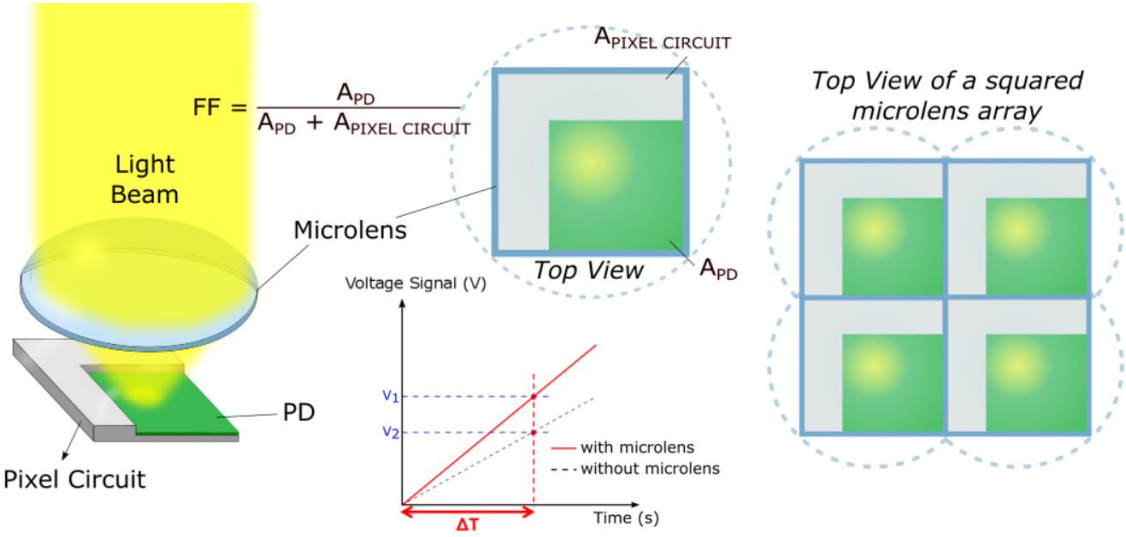

Figure 2: The FF (ratio between PD area and total pixel area), and effect of microlenses in signal strength. The top view of a microlens array is also shown at the right.

Microlenses are literally small lenses that are placed in front of the photodetector, with the purpose of focusing incident light precisely over the sensor photosensitive area, as shown in Figure 2. Since the pixel circuitry is often shielded, light impinging over it will be reflected, and consequently, lost. Also, if the circuit is not shielded, light will affect its performance. Therefore, using microlenses is a must for systems with low FF and relatively weak signals.

To manufacture lenses, there are dozens of available materials in the infrared spectral range, however, only a few are commonly used for optical purposes, like silicon, germanium, ZnSe, and fused silica [5], [10]. Among them, silicon exhibits a number of advantages, mainly its availability in affordable high-quality wafers, low density, mechanical robustness and relatively high *refractive index* (*n*), allowing light-weighted compact optical systems to be built [5]. Although silicon is suitable for infrared, it features high surface reflectance in that spectral range, being that its major disadvantage [11]. Fortunately, we can enhance silicon surface transparency by using an *Anti-Reflective Coating* (*ARC*), which consists of one or more layers of specific materials and thicknesses deposited onto the silicon surface [9], [10]. Hence, for infrared detection systems, silicon microlenses plus its suitable ARC will also be encompassed by this thesis.

1.3 Motivation and goals

The Infrared (IR) spectra is subdivided in five different ranges, as will be explained in details in section 2.1. Some of these ranges already find many applications commercially. However, two of them feature military applications that restrict both the commercial availability of products for civilians and the research and development of such technology in this country. Examples of applications in theses ranges are free-space communication (FSO), gas analysis, radiation thermometers, thermal imaging, remote sensing, spectrophotometry, flame monitors, fire detectors, optical power meters, and many others. Therefore, a national institute named

INCT-DISSE (Brazilian National Institute of Science and Technology for Semiconductor Nanodevices) was created in 2008 in order to promote the development of such technology in this country [12]. This work is a result from a demand of this institute and the collaboration between the Laboratory for Optronics and Microtechnology Applications (OptMA^{lab}) at the Federal University of Minas Gerais (UFMG) and its institutional partners.

Regarding the technology for an infrared optical detection system, as discussed in the previous sections, we can summarize the main limitations by its detector relatively high thermal dark current and low FF, and by undesirable high background irradiation, all of them rendering lower SNR. Furthermore, since dark and background currents might be much larger than the useful signal, it is important to be able to discriminate small differences between two adjacent relatively strong current values.

In order to distinguish small differences among signals, the pixel circuit should have a high sensitivity, i.e., the same input current signal variation should yield the largest possible output voltage variation. This way, even the smallest current signal differences, compared to the dark current and background level, could be discriminated by measurable output voltages. To illustrate, Figure 3 shows two voltage signals integrated by a linear pixel circuit, with two different sensitivities (S_1 and S_2; with $S_1 < S_2$). We can see that for the same input current difference (ΔI), the resulting output voltage difference (ΔV) is larger when the sensitivity is higher, that is, $\Delta V_1 < \Delta V_2$ when $S_1 < S_2$. The output voltages are read after a time interval called the integration time (ΔT).

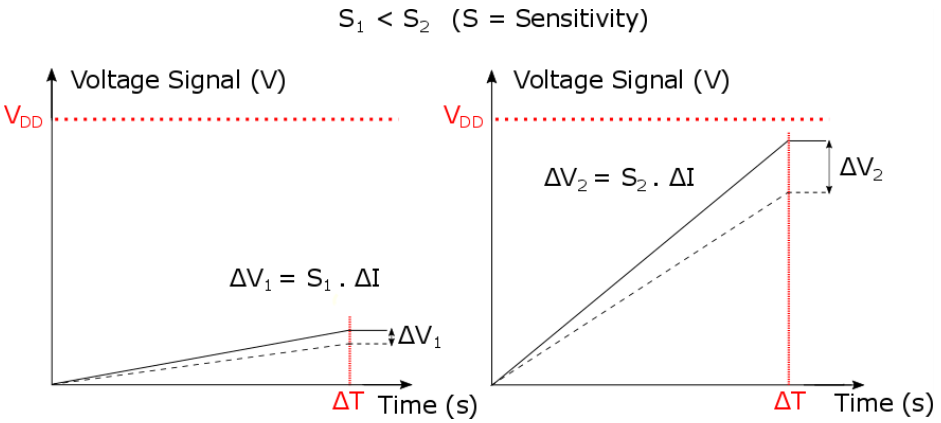

Figure 3: Different values of sensitivity and its effect on the final integrated voltage.

However, a linear pixel circuit with higher sensitivity is able to measure a smaller extent of input current signals, provided a constant output voltage range, that is, the maximal operational voltage of the circuit, shown as V_{DD} in Figure 3. Any current that would be transduced to a voltage higher than V_{DD}, within ΔT, is said to be saturated, as in Figure 4, and information about this signal is lost. Input currents I_1, I_2 and I_3 result in the proportional voltages V_1, V_2 and V_3, respectively. Note, however, that the input current I_4 results in an output voltage V_4 equal to V_{DD}. Therefore, it is impossible to distinguish between any current values larger than the maximum one, in this case, I_3. Also, the higher the sensitivity, the smaller the maximum current that can be integrated.

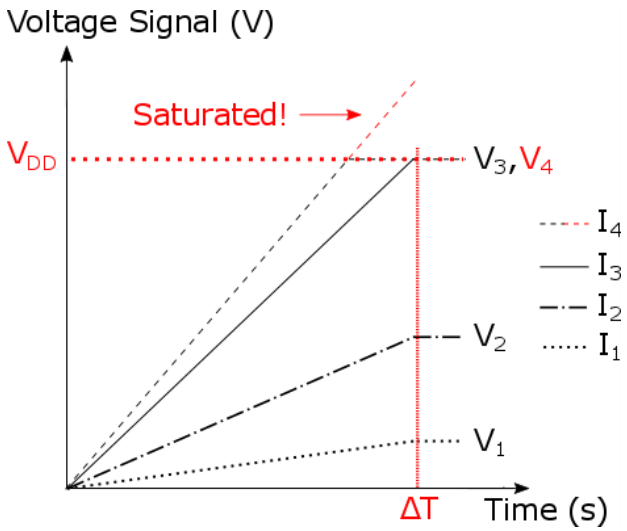

Figure 4: Pixel circuit voltage saturation and maximum readable current.

Roughly, the range from noise level to the highest measurable current is called *Dynamic Range* (*DR*), and pixel circuits featuring *High Dynamic Range* (HDR) are, therefore, desirable in this context. The formal definition of the DR will be presented with more details in section 2.8.2.

This work aims to address high dark and background currents, low FF, and low DR, proposing a measuring system featuring higher SNR and sensitivity. The proposed system is composed of a novel pixel circuit architecture, coined as *bouncing pixel*[1], and a properly designed silicon microlens covered by an ARC. At the end, an emulated application example will be presented in order to illustrate the importance and practical function of each component.

That way, SNR is boosted in two ends: firstly, with the bouncing pixel, that features both high sensitivity and HDR, making the SNR at the output as close as possible to the SNR of the photodetector, and by allowing higher currents to be integrated, which in turn renders in a higher maximum SNR; and secondly, silicon microlenses are capable of focusing an infrared beam precisely over the photosensitive area of the sensor, overcoming the low FF restriction, assuming the use of an ARC to mitigate high silicon surface reflection.

The main goal of this work is to prove the herein described system concept for infrared applications. All presented results will be based on an application driven design flow of a CO_2 gas detection system, as will become clear in the following chapters. For that purpose, simulations will be done emulating the bouncing pixel real test conditions. Also, microfabrication and characterization of the microlenses will be presented, followed by its performance analysis.

[1] In Portuguese: *Pixel Ricochete*.

1.4 Thesis roadmap

Chapter 2 of this work devotes itself to present the background theory necessary to understand both the optical and electronic parts of the proposed detection system, explaining the basic concepts, figures of merit, and challenges regarding IR photodetection and signal transducing. Chapter 3 elaborates on the proposed solutions for the mentioned problems, showing the fabrication process for the silicon microlenses, the ARC design, and the proposed *bouncing* pixel. Chapter 4 presents the results for the microlenses fabrication and characterization, the measurement of the designed ARC, the analysis of the bouncing pixel through simulations, ending with an emulated application of a CO_2 gas detection system, as an example. Lastly, Chapter 5 presets the final conclusion.

2. THEORY

Last chapter briefly introduced the main problems being addressed in this work, and in order to understand their respective solutions, a sound background theory must be presented. This chapter explains the main features that allow us to acknowledge the proposed system as a suitable solution for infrared systems.

Section 2.1 starts by presenting the optical spectrum and, more specifically, by subdividing the infrared into its subclasses and respective applications, followed by how light is absorbed by the atmosphere. Then, section 2.2 elaborates on the photodetection mechanism, explaining how a simple optical detector works, based on *p-n junctions*. Next, section 2.3 focuses on thermally excited carriers, and how light is absorbed by different materials. Section 2.4 elucidates on photodetection in the infrared, listing commonly used light sensors at this spectral range. Further, section 2.5 devotes itself to the equivalent circuit of photodiodes, and noise analysis, necessary to understand dark and background current effects on SNR. Next, in section 2.6, the main figures of merit for pixel circuits are presented. In section 2.7, the conventional pixel circuit topologies are presented, and operation explained. Afterwards, section 2.8 defines HDR, explains why HDR operation is desired, and presents the different types of HDR sensors. Then, section 2.9 elaborates on infrared optics, defining optical gain, and why silicon microlenses plus a single-layer ARC can be used to enhance SNR.

2.1 Optical Spectrum and black-body radiation

Within the electromagnetic spectrum, there is a region of interest called the optical spectrum, shown in Figure 5. Although it accounts for only a small share of the electromagnetic scale, photons in the optical region may have wavelengths (λ) varying from 0.01 µm up to 1000 µm. And, inside that range, there are three major important spectral regions: *Ultraviolet* (from 0.01 µm to 0.4 µm); *Visible* (from 0.4 µm to 0.7 µm), where human eyes can detect light; and *Infrared* (from 0.7 µm to 1000 µm) [13], [14].

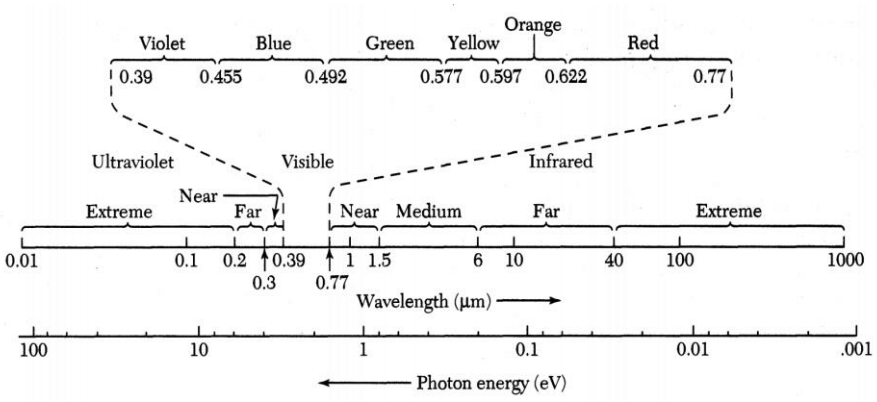

Figure 5: The optical spectrum comprising UV, visible and IR.
Source: [14].

The same way that visible light has its color bands – like red, green or blue – infrared (*IR*) has its own subdivision scheme, as follows: *Near-infrared* (*NIR*), *Short-wavelength-infrared* (*SWIR*), *Mid-wavelength-infrared* (*MWIR*), *Long-wavelength-infrared* (*LWIR*), and *Far-infrared* (*FIR*). Table 1 groups each of the aforementioned regions and their respective wavelengths range.

Table 1: Infrared spectral regions (wavelengths, and 'color' temperatures [peak-wavelengths]).

Spectrum subdivision	Abbreviation	Wavelength	'Color' temperatures (K)	'Color' temperatures (°C)
Near-Infrared	NIR	0.75 – 1.4 µm	3867 / 2071 K	3594 / 1798
Short-wavelength-infrared	SWIR	1.4 – 3 µm	2071 / 967 K	1798 / 694
Mid-wavelength-infrared	MWIR	3 – 8 µm	967 / 363 K	694 / 90
Long-wavelength-infrared	LWIR	8 – 15 µm	363 / 193 K	90 / – 80
Far- infrared	FIR	15 – 1000 µm	193 / 2.9 K	– 80 / – 270

Potential applications to each infrared band can be indicated by the analysis of blackbody radiation, since many objects with finite temperature can have their emitted radiation spectrum approximated to that of a blackbody. The main feature of these spectra are their peak-wavelengths (λ_p), that is, the wavelength at which the emitted radiation is the highest, defined by the objects absolute temperature *T* in Kelvin [5], [9], [10]. Figure 6 shows the approximated spectra of the Sun, a "hot object" at 800 K, and the human-body, plotted in the logarithmic scale.

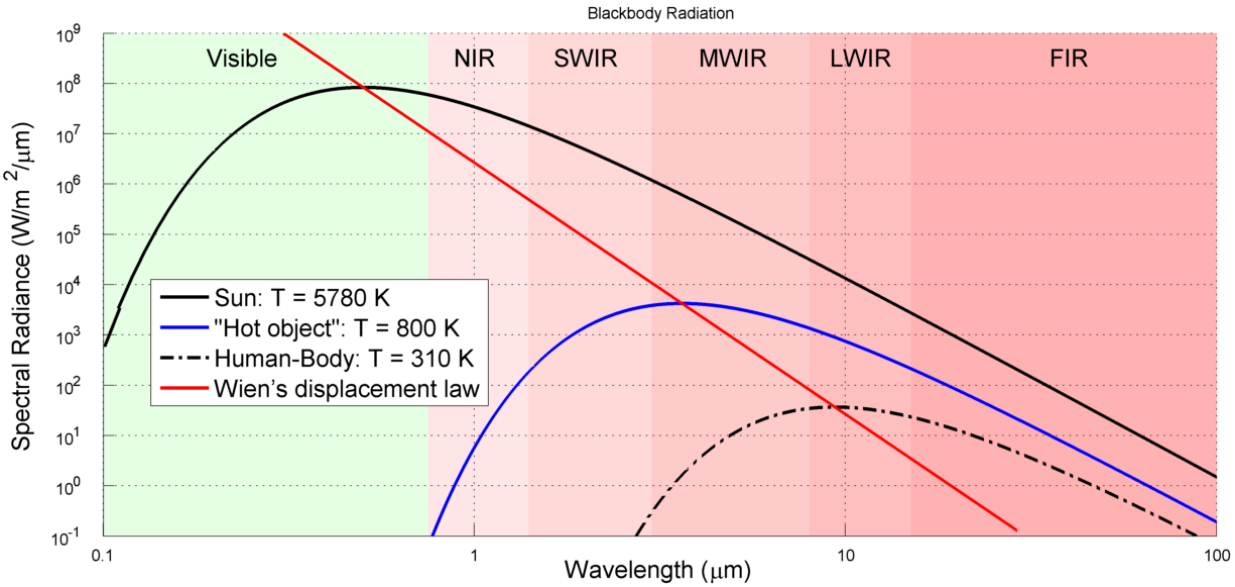

Figure 6: Blackbody spectra in logarithmic scale: the Sun @5780 K, an object @800 K, and the human-body @310 K; Also, the *Wien's displacement law* is presented as a red line.

We can notice that the higher the temperature, the shorter the peak-wavelength. This inverse relationship is given by *Wien's displacement law* [5], [9], [10], also shown in Figure 6:

$$\lambda_p = \frac{2900}{T[K]} \quad [\mu m] \qquad (1)$$

Table 1 also shows the temperatures of blackbodies, whose maximum spectral radiances match the given wavelengths. Using Wien's law, we can estimate the peak-wavelength of the Sun as $\lambda_{p\,Sun} \cong 0.50\ \mu m\ (@5780\ K)$. Although its peak-wavelength is inside the visible range, most of its radiation is comprised by IR, as can be seen in Figure 6 by its long right-hand side tail, providing good background illumination for many applications at NIR and SWIR. IR thermography of high temperature objects is preferred, however, to be performed using MWIR [15]. On the other hand, thermal detection of near-room temperature objects yields its best results at LWIR [15]. For instance, from the Wien's law, the human-body has a peak-wavelength of $\lambda_{p\,Human-Body} \cong 9.36\ \mu m\ (@310\ K)$, which falls on the LWIR region.

In conclusion, IR light can be derived from either direct emission (usually MWIR and LWIR), or reflected by an object (generally NIR and SWIR). To detect reflected or thermally emitted light, it must propagate through a media – generally air – from its source towards a sensor. Radiation is, however, attenuated by both scattering and absorption processes. Rayleigh scattering [16] shows strong dependence on wavelength (λ^{-4}) for gas molecules, especially for smaller wavelengths. However, wavelengths longer than 2 µm exhibit negligible scattering, which is why IR can be used to see through smoke or mist particles [5]. On the other hand, rain, fog and suspended particles are bigger than gas molecules, and scatter IR light to a degree similar of visible [5]. In addition to scattering, atmosphere gas molecules like water vapor, carbon dioxide, and oxygen play a great role in absorbed radiation throughout the IR spectra, confining near-transparent transmission windows to the following ranges: 3-5 µm (@MWIR) and 8-14 µm (@LWIR) [5], [9], as shown in Figure 7.

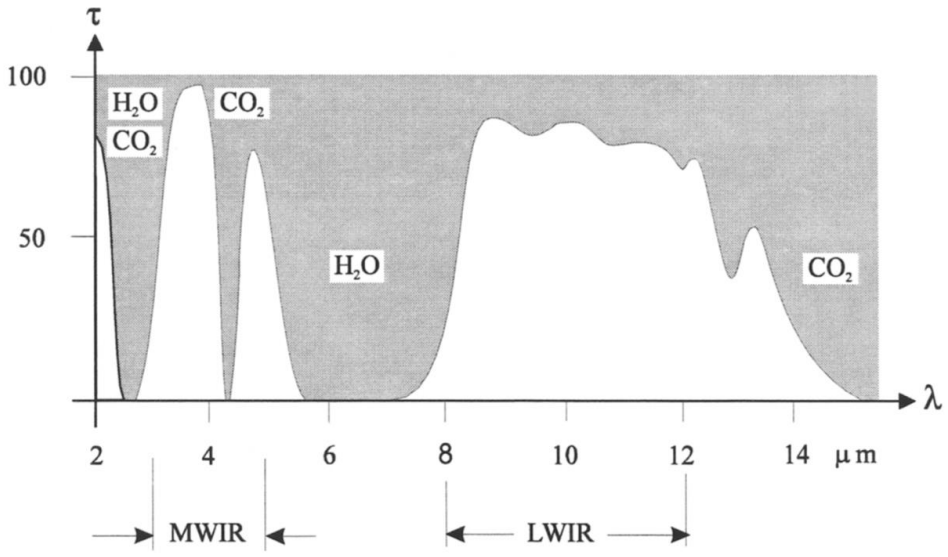

Figure 7: Atmosphere (air) transmission (τ) at IR spectrum and 'transparent windows'.
Source: [9]

The LWIR window is suitable for applications like thermal imaging, once it has higher sensitivity to objects at ambient temperature (λ_p inside LWIR) and better penetration through smoke or mist (longer λ). Nonetheless, the MWIR window generally yields a cheaper system, mainly due to smaller optics and not-so-harsh sensor cooling requirements, compared to LWIR sensors [4], [5], being suitable when contrast at higher temperatures is needed.

Beyond that, if one does not want to see through a medium, but rather identify its composition, or quantify a gas concentration for example, gases optical absorption can yield

plenty of useful information. There are two main types of gas sensors, the open-path, and point detector [15], [17], both of which rely on the amount of radiation absorbed through a gas volume, from an IR source towards an IR sensor, at specific wavelengths. They can be used to detect the leakage of flammable gases, hydrocarbons, carbon dioxide, and alike.

Summarizing, IR can be used for many applications, since many physical phenomena take place in this rather wide spectral range. An IR detection system can be designed based on absorbance, transmittance or reflectance of a given material in this spectrum. Also, it can be conceived on the basis of the temperature or radiated emission of the target object, or by the amount of available background illumination.

2.2 Photodetection mechanism: The p-n Junction

One of the mechanisms to sense IR light is called *photodetection*, which is the operating principle of the *junction* or *photonic sensors*. These are generally formed by semiconductor materials and have one or more junctions between *p-type* and *n-type* regions, which are, respectively, semiconductors with excess of holes and electrons [13], [14]. The simplest device is called a *p-n junction*, or just *photodiode*, schematized in Figure 8. Excess carriers, i.e., electrons in n-type region, or holes in the p-type, are called *majority carriers*; whereas the *minority carriers* account for the respective carriers in lesser concentrations.

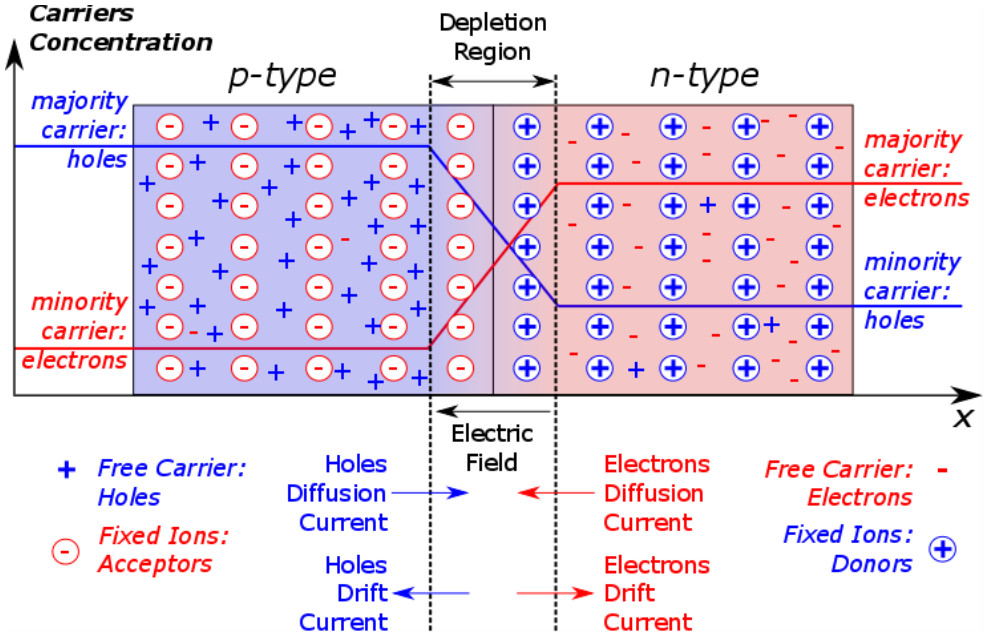

Figure 8: P-N junction diagram in equilibrium.

As can be shown in Figure 8, around the junction there is a *depletion region*, which has no standing free carriers, and has a *built-in electric field*. Whenever a minority carrier, either an electron or a hole, wandering around the neutral region of the p-type or n-type, respectively, comes too close to the depletion region, then the built-in electric field pushes this carrier to the other side, resulting in a *drift current*. In addition, there is always a diffusion tendency pushing the majority carriers to the other region, due to gradient concentration, generating a *diffusion*

current [13], [14]. In equilibrium, both drift and diffusion currents are equal, cancelling each other, and producing a zero net current across the junction. However, if light is focused over the junction, equilibrium may be disturbed.

If a photon is absorbed by the material, it creates an electron-hole pair, say, at the p-type region; and if close enough, the minority carrier (electrons) will diffuse towards the junction being swept by the electric field across the depletion region, reaching the other side. This way, an excess carrier will be created in both *n* and *p* regions, due to the incident photon, forward biasing the photodiode with an *open-circuit photovoltage* [13], [18]. If both terminals of the diode are shorted together, a *short-circuit* photocurrent will flow through the diode. This can be seen on the photodiode I-V curve in Figure 9. As a corollary, the more incident photons, the bigger the photocurrent.

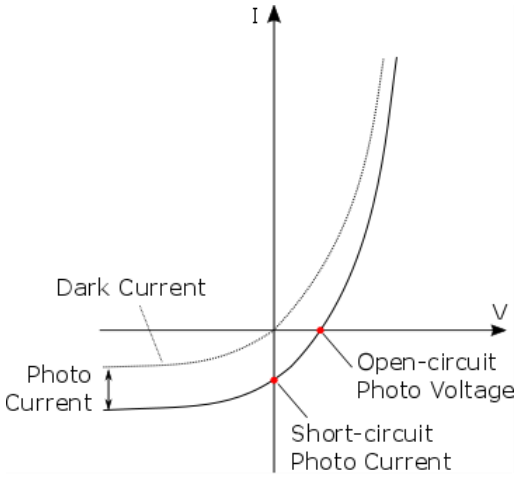

Figure 9: Photodiode I-V characteristic in the dark and under light.
Source: Adapted from [18].

Nonetheless, not all photons generate electron-hole pairs, and not every created minority carrier makes their way through the depletion region, being collected by the diode terminals. The ratio between the number of collected electrons to the number of incident photons is called *quantum efficiency* (η), which is often lower than 100%, depending strongly on wavelength and structural parameters, like junction depth, size of depletion region, and surface reflectance [13], [18]. Figure 10 shows a photodiode diagram and its structural parameters.

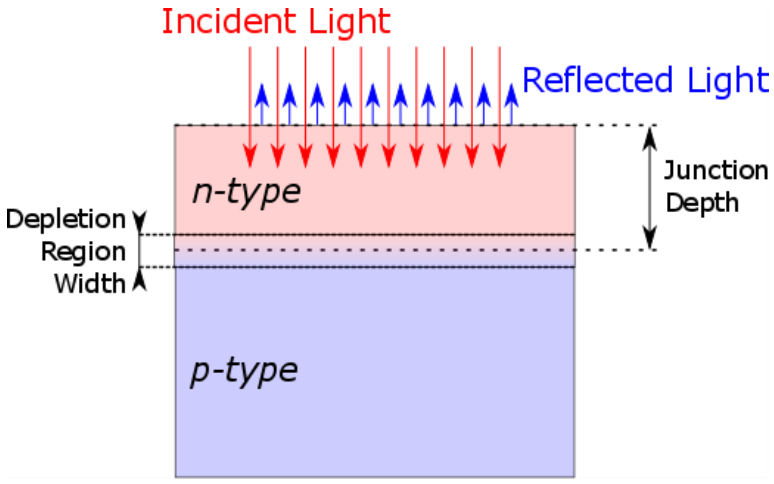

Figure 10: Photodiode cross-section.

2.3 Thermally excited charge carriers and light absorption

Every semiconductor material has a different energy band diagram structure, and it is known from quantum mechanics, that all of them have a forbidden energy gap between their valence and conduction bands [13], [14], as illustrated by Figure 11. At absolute zero temperature, all valence electrons in a semiconductor are participating in chemical bonds with adjacent atoms, and are not available for conduction. In other words, all electrons are in the valence band, and none at the conduction band.

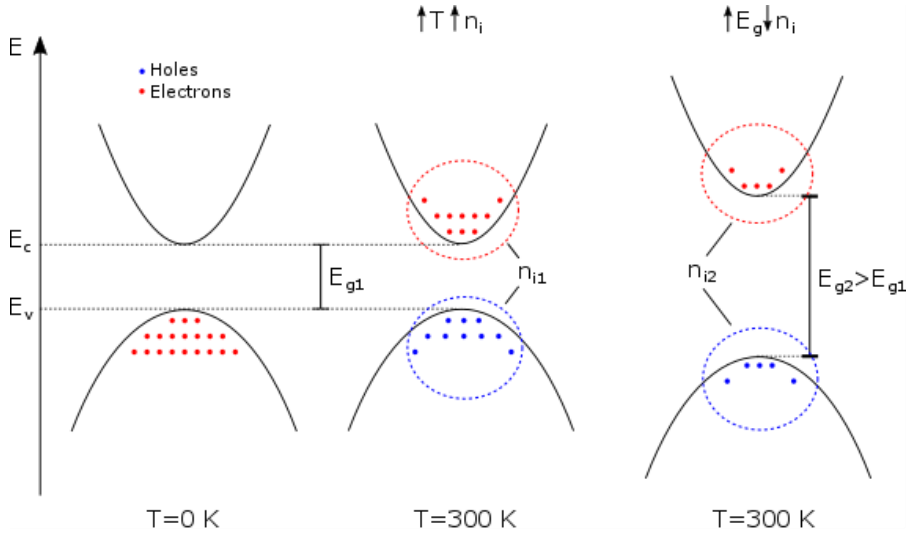

Figure 11: A semiconductor band diagram (for direct-bandgap materials), illustrating the effect of temperature and of the energy gap on the intrinsic-carrier concentration.

However, if temperature rises, energy is imparted to electrons, such that some of them start to have enough energy to "jump" to the conduction band, leaving a hole behind in the valence band. The density of thermally excited carriers is called the *intrinsic carrier concentration* (n_i), and depends on the absolute temperature (T) of the material and on its energy bandgap (E_g) in *eV* (*electron-Volt*). The term E_g is the energy difference between the top of the valence band and the bottom of the conduction band, and is a parameter of the material. The higher the temperature, the higher the electron energy, and therefore the larger the n_i. Also, for materials with narrower bandgaps, less energy is needed for an electron to "jump" to the conduction band, resulting in larger n_i, if compared to a material with higher E_g at the same temperature [13], [14].

The process of an electron "jumping" from the valence to the conduction band, and leaving a hole behind, can also be interpreted as the generation of an electron-hole pair, since now both types of free carriers are available for conduction. Exactly as thermal excitation, when light is absorbed by a semiconductor material, it generates electron-hole pairs. The only difference is the cause of the excitation trigger: light instead of temperature. Furthermore, in order for a photon to be absorbed, its energy must be greater than or equal to E_g, being the cut-off-wavelength (λ_{cut}) for a semiconductor material defined as [13], [14]:

$$\lambda_{cut} = \frac{h.c}{E_g} = \frac{1.24}{E_g[eV]} \quad [\mu m] \qquad (2)$$

where h is the Planck constant, and c is the speed of light.

Wavelengths shorter than λ_{cut} are associated with higher energies, and are absorbed at shallower depths by the material; whereas any wavelength longer than λ_{cut} will render an energy smaller than E_g, and therefore, will not be absorbed. The y-axis absorption coefficient (α) in Figure 12 relates to how strong light is absorbed in a specific material. As can be seen, absorption diminishes with increasing wavelength, until it reaches λ_{cut}, where the absorption coefficient is negligible [13], [14]. On the other hand, for very short wavelengths, the electron-hole pairs are generated very close to the material surface, where there is a high density of dangling bonds, which causes the minority free carriers to recombine. This is called surface recombination, that makes the photogenerated excess carriers to be annihilated. Since surface recombination is typically high, regarding photodiodes, the incident photons with smaller wavelengths are absorbed, but most of them cannot penetrate deep enough in the material – near the depletion region – to be collected and contribute to the photocurrent.

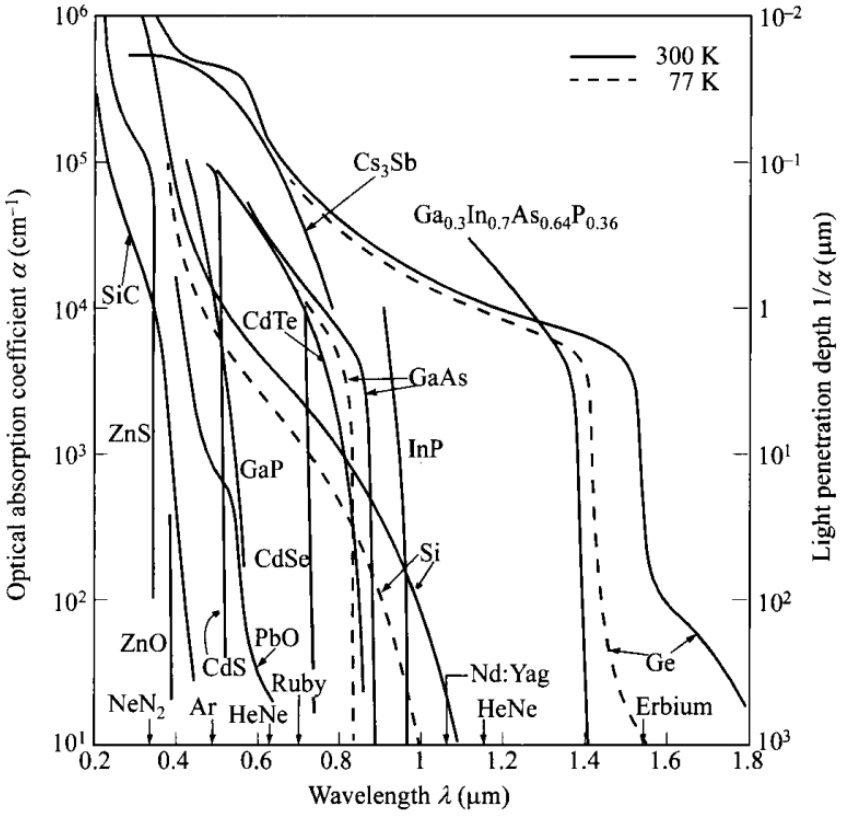

Figure 12: Some semiconductor absorption (α) vs. wavelength.
Source: [14].

Mathematically, an incident photon flux density Φ_0 impinging over a semiconductor surface, decays exponentially with the distance traveled across the material:

$$\Phi(x) = \Phi_0 . e^{-\alpha(\lambda).x} \quad \left[\frac{number\ of\ photons}{m^2 s}\right] \qquad (3)$$

where $\Phi(x)$ is the remaining photon flux at a distance x (considering $x = 0$ at the surface), and $\alpha(\lambda)$ is the wavelength-dependent absorption coefficient from Figure 12. A piece of semiconductor with finite thickness and made of a material with very low absorption coefficient

is considered to be transparent. Silicon, for instance, has $E_g = 1.12\,eV$, and according to Equation (2), for wavelengths longer than $\lambda_{cut} = 1.1\,\mu m$, which corresponds to most of the IR spectrum, it is a transparent material. As a result, silicon cannot be used to detect IR, except for a small portion at NIR, and materials with narrower bandgaps ought to be used [13], [14].

From Figure 12, we could choose germanium as a better infrared detector material, due to its narrower bandgap (longer cut-off-wavelength). Though it detects all NIR, it is still sensitive to only a small portion of SWIR, limiting the number of potential applications for germanium detectors. In addition, the fact that germanium has a narrower E_g also renders a larger n_i, that is, larger thermally excited carrier concentration. In a photodiode, both light and thermal excitation may contribute to the amount of total generated current. However, current created from thermal sources are undesirable, meaning more noise to the device; whereas photocurrent generated from light is generally correlated to a useful signal.

Since IR detectors must have a narrower bandgap, thermally excited carriers cannot be neglected, rendering undesirable dark current and dark noise, sometimes high enough to mask useful photogenerated signals, yielding smaller SNR. A usual way to enhance SNR is by cooling the device [5], [9], [10]. This way, the temperature of the sensor is reduced, resulting in a much lower n_i, and hence bringing the photogenerated signal to a level higher than that of dark noise.

2.4 Infrared photodetectors

There are many available devices capable of detecting IR radiation, and they are classified in two main groups: the *thermal* and *photonic* types [5], [8], [18]. Thermal sensors convert IR radiation indirectly to electrical signals, each by means of a different temperature-dependent measurable property. Thermocouples, thermopiles and bolometers are examples of usual thermal detectors. The mechanism of resistive bolometers, for instance, relies on temperature-dependent resistivity change. Figure 13 illustrates a typical thermal sensor structure, composed of a sensitive element (the IR absorber), a heatsink with constant temperature (the substrate), and a thermometer, which can be any of the detector mechanisms aforementioned [4], [5], [10], [18], [19]. The IR radiation causes a temperature change in the sensitive element, that is measured. The supporting legs are generally thin and long, providing good thermal insulation between the sensitive element and the substrate, necessary for an accurate temperature change measurement. However, for every new measurement, the sensitive element must be cooled down to the temperature of the substrate, which is a slow process. Therefore, it is desirable that the heat capacity of the sensitive element is kept low enough, so that the response time of the sensor is not too large.

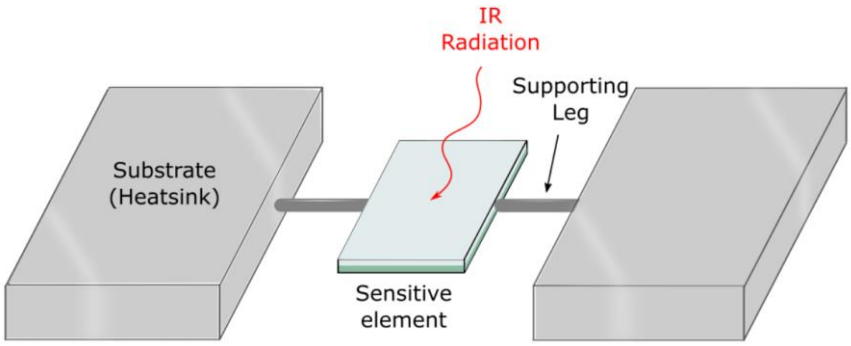

Figure 13: Thermal sensor schematic.
Source: Adapted from [18].

On the other hand, photonic devices work according to the same principle described earlier for p-n junctions, i.e., absorbing photons and generating electron-hole pairs. One of the biggest differences between photonic and thermal sensors is that the photonic type generally needs costly cooling systems, while thermal type works at room temperature [8]. Beyond that, photonic detectors have a sensitivity highly dependent on the wavelength, whereas thermal ones detect IR radiation in a broad-wavelength bandwidth [8]. However, photonic sensors are much faster and much more responsive than their thermal counterparts [4], [5], [10], [18]. From a market point of view, thermal detectors generally are high volume and low performance, but low cost; and photonic sensors are high cost, but low volume and high performance [8]. Also, photonic sensors offer the capability of building larger arrays, and can be readily integrated with readout electronics, generally by using a *flip-chip* technology [4], [5]. Although thermal detectors offer interesting applications, our main concern in this work are the photonic sensors, once we are more interested in their higher performance.

Among photonic detectors, current predominant technology is based on *MCT* (Mercury-Cadmium-Telluride), or HgCdTe. While HgTe is a semimetal with 0 eV bandgap, CdTe is a semiconductor with 1.5 eV bandgap. MCT is made by the mixture of HgTe and CdTe. A more exact composition of the ternary compound is $Hg_{1-x}Cd_xTe$, where *X* and *1-X* are the mole fractions of Hg and Cd in the compound, respectively. By changing their relative mole fractions, it is possible to alter the bandgap of the material. Therefore, it becomes a tunable bandgap semiconductor, capable, in practice, of detecting IR in a spectral range from about 1 µm to 25 µm, depending on the amount of Cd and Hg in the compound [20]. However, in practical terms MCT has several mass production issues, including low uniformity and yield, and high toxicity [5], claiming for new material systems.

Many other material combinations have been developed, like InGaAs, InSb, InAsSb, HgZnTe, HgMnTe, among others. [4], [5], [8], [10], [18]. Emerging technologies have also been proposed and are commercially available, as the *QWIP* (*Quantum well IR photodetector*), which offers the possibility of being implemented in chemically stable large bandgap materials, relying on intrasubband photon transitions, instead of interband [4], [5], as illustrated in Figure 14. It is a lower cost technology compared to MCT and offers better material uniformity, but features lower quantum efficiency and demands lower operating temperatures, besides depending on non-zero bias to work [8]. Although an interesting solution, QWIPs are not based on p-n junction operation, and will not be modelled in this work.

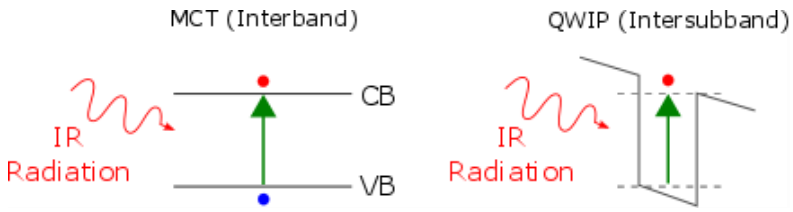

Figure 14: MCT and QWIP energy band transitions.
Source: Adapted from [8].

Another emerging technology for IR light detection are the *Quantum-dot Infrared Photodetectors (QDIPs)*, that compared to QWIPs offer a series of advantages, based on the three-dimensional confinement of carriers. QDIPs have a significant sensitivity to normal incidence of light, and longer lifetime of carriers, leading to lower dark currents [8]. Many works have been done in this direction, including in the INCT-DISSE [12], [21]–[24].

2.5 Photodetector equivalent circuit

As discussed in Chapter 1, in order to read out any useful information from a photodetector, it must be interfaced along with a pixel circuit. To predict performance of both photodetector and pixel circuit operating together, before they are actually fabricated, integrated circuit simulation offers a good option to foresee how the detection system will behave.

A pixel circuit is often composed by a number of integrated electronic components, mainly transistors, which can be easily simulated with highly reliable models [25]. The photodetector, however, usually is a single device, and its behavior under incident light conditions have to be carefully modelled. In the context of this thesis, a photodiode will be modelled and used in further simulations. In fact, the readout circuit to be presented in this thesis – the bouncing pixel – was designed to work with other types of IR detectors as well, like a PIN diode or a QWIP and QDIP. However, since in future work it will be tested using a regular photodiode fabricated in a standard CMOS technology in the same chip as the pixel circuit, then simulations – including during the application example – will be performed with a photodiode model. For that purpose, the used equivalent circuit of the photodiode will be explained in this section.

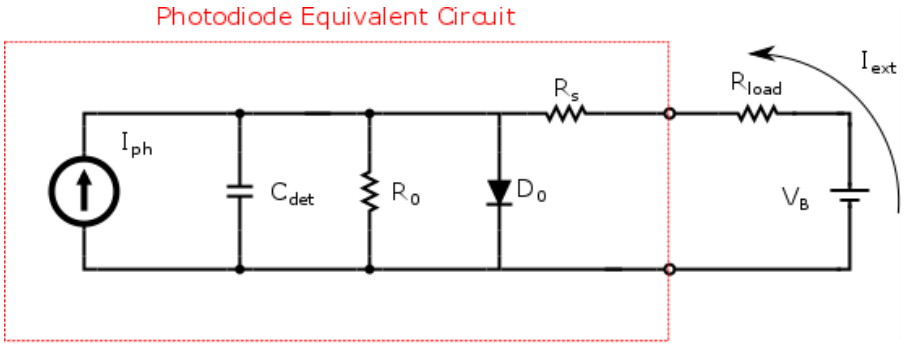

Figure 15: Photodiode equivalent circuit in photoconductive mode.

Figure 15 shows the aforementioned equivalent circuit. The current source I_{ph} represents the photogenerated current, C_{det} the parasitic capacitance, D_0 models an ideal diode, R_s the

photodiode series resistance, R_0 the junction shunt resistance, R_{load} the load resistance, and V_B the external bias voltage. The collected current at the photodiode terminals (I_{ext}) is given by the Shockley equation [1]:

$$I_{ext} = -I_{ph} + I_s\left(e^{\frac{V_B - I_{ext}(R_s + R_{load})}{V_T}} - 1\right) + \frac{V_B - I_{ext}(R_s + R_{load})}{R_0} \quad [A] \qquad (4)$$

where I_s is the reverse saturation current of the diode, and $V_T = k_B T/q$ the thermal voltage, being k_B the Boltzmann constant, q the electron charge, and T the junction absolute temperature. Generally, the p-n junction shunt resistance R_0 is large enough ($\sim 100\ M\Omega$), so that third term in equation (4) can be neglected. Moreover, voltage drop across R_s and R_{load} is small, at least when interfacing with a pixel circuit with zero bias. In this case, the photodetector will only see the input resistance of the circuit ($R_{load} = R_{in}$), that is generally designed to be as low as possible, as will be discussed in sections 2.6 and 2.7. Even when an external bias V_B is applied, it might appear a parasitic resistance from the source (R_{source}), which also is very small, resulting in a total load resistance $R_{load} = R_{source} + R_{in}$ very small as well. In addition, the series resistance of the photodetector can be as low as $10\ m\Omega$. Finally, since current and resistances are small, their product $[I_{ext}(R_s + R_{load})]$ will also be small, and equation (4) can be simplified to:

$$I_{ext} = -I_{ph} + I_s\left(e^{\frac{V_B}{V_T}} - 1\right) \quad [A] \qquad (5)$$

Equation (5) describes the photodiode I-V curve as shown in Figure 9. The first term in this equation describes the photocurrent and can be expanded as [13], [14]:

$$I_{ph} = \eta q A_{det} \Phi_0 \quad [A] \qquad (6)$$

where A_{det} is the photodetector area, η the quantum efficiency, and Φ_0 the photon flux density. Note that equation (6) is valid for a monochromatic light source, for instance, a laser. If a polychromatic source is used instead, then we should consider the spectral components, that is: $I_{ph[\lambda]} = \eta(\lambda) q A_{det} \Phi_{0[\lambda]}$; where $I_{ph[\lambda]}$ is the spectral photocurrent, $\eta(\lambda)$ is the wavelength-dependent quantum efficiency, and $\Phi_{0[\lambda]}$ is the spectral incident photon flux density. The total amount of photocurrent in the detector is given by the integral of $I_{ph[\lambda]}$ for the whole spectrum of the incident light source ($\Delta\lambda = \lambda_2 - \lambda_1$), that is: $I_{ph} = \int_{\lambda_1}^{\lambda_2} I_{ph[\lambda]}\, d\lambda$.

From equation (6), the photocurrent increases with either the area A_{det}, the quantum efficiency η, or the incident photon flux density Φ_0. For a given device and spectral range, A_{det} and η are fixed, and thus I_{ph} linearly scales with Φ_0, and measuring the external current I_{ext} of the photodiode gives information of the amount of light signal impinging over the detector.

However, as in all real devices, beyond signal, there is always noise. Depending on device or operating condition, noise can even overcome signal. The main noise sources in photodetectors are *quantum and photodetector shot noise*, *optical background noise, and photodetector dark current noise* [26]. For noise analysis through simulation, the complete noise equivalent circuit is shown in Figure 16, where $\sigma_{thermal}$ represents the *thermal noise*, and σ_{shot} the *total shot noise*.

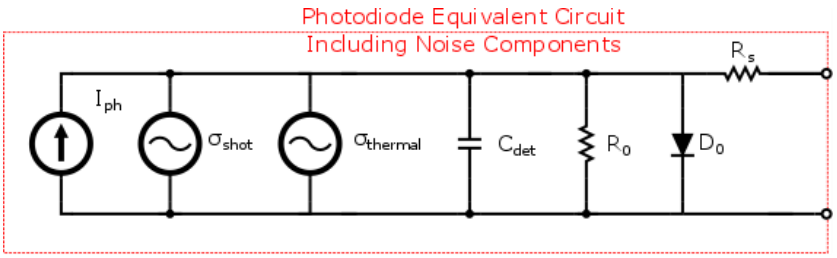

Figure 16: Noise equivalent circuit.

Thermal noise arises from parasitic resistances shown in Figure 16, i.e., either the shunt resistance R_0 or the series resistance R_s. The term $\sigma_{thermal}$ does not take R_{load} into account, since it considers only noise originating from the photodiode. The noise components from other devices will also be taken into consideration, as will become clear in section 2.6. Although resistors exhibit charge neutrality, on the microscopic level free carriers are in a persistent *Brownian motion*, resulting on random local voltage fluctuations in series with the resistance, or equivalently, random currents in parallel with it [26]. Thermal noise is also known as *Johnson noise* [13], and depends on temperature, the equivalent resistance R_{eq}, and the observation bandwidth B:

$$\sigma_{thermal} = \sqrt{\frac{4kT}{R_{eq}}B} \quad [A] \qquad (7)$$

$$\sigma_{V-thermal} = \sqrt{4kTR_{eq}B} \quad [V] \qquad (8)$$

Equations (7) and (8) are equivalent, describing the *RMS* (*root-mean-square*) current and voltage, respectively, being the first used for parallel, and the second for series resistances. Since R_s is considered to be small enough, then its thermal voltage component ($\sigma_{V-thermal}$), which should be in series with R_s, was ignored in the circuit of Figure 16. Therefore, the main thermal noise component is due to the parallel shunt resistance ($R_{eq} = R_0$). The derivation of the RMS noise voltage and current are obtained considering a white-noise *PSD* (*power spectral density*) along the bandwidth B within which noise is observed [26]. The latter is called as the *Equivalent Noise Bandwidth*, that depends on the cut-off ($or\ f_{3dB}$) frequency of the pixel circuit. If the pixel circuit has a lowpass-like transfer function $H(f)$, with a maximum of $H(0)$, then the Equivalent Noise Bandwidth is defined as the rectangle with maximum $H(0)^2$ and width B, such that it has the same area as the area under $|H(f)|^2$. For a circuit with just one pole, the Equivalent Noise Bandwidth is equal to $B = \frac{\pi}{2}.f_{3dB}$.

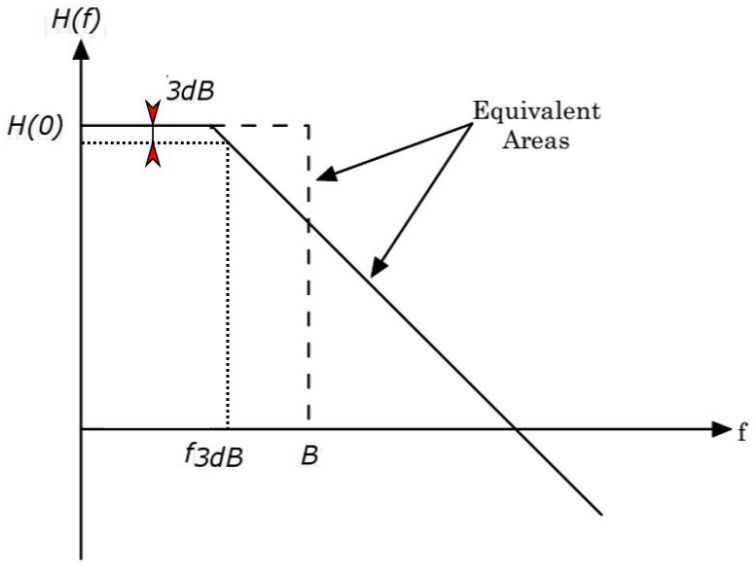

Figure 17: Equivalent Noise Bandwidth.
Source: Adapted from [27].

Current flowing either in the photodetector or pixel circuitry is not smooth nor continuous. It is the summation of individual charge carrier movement pulses, whose average flux through a potential barrier yields the mean current flow. Crossing a potential barrier means whenever a charge carrier goes from a point with potential A to a point with potential B, where $A \neq B$. It can represent the built-in potential in the depletion region of the photodiode, or the potential difference between the terminals of a transistor (drain and source). The number of carriers crossing this potential is random and follows a *Poisson distribution* [26]. If a sufficiently large amount of carriers crosses the barrier, which is usually the case, it can be approximated by a *Gaussian distribution*, whose mean represents the average current flow, and the variance represents the white-noise PSD (in [A²/Hz] units) [26]. The total noise associated with current flowing through a potential barrier is called shot noise, given by equation (9):

$$\sigma_{shot} = \sqrt{2qI_{DC}B} \quad [A] \tag{9}$$

where I_{DC} is the average *direct-current* (*DC*), which in the context of this thesis represents any current passing through the modelled photodetector, and is modelled with a current source in parallel with the noisy device. For instance, a noisy diode is replaced by an ideal noise-free diode in parallel with a current source, whose value equals the total RMS shot noise.

Shot noise arises from five independent sources: due to signal photocurrent, quantum photodetection, incident photon flux, optical background photocurrent, and dark current. Shot noise arisen from electronic photocurrent flowing through the photodiode, random fluctuation on the incident photon flux, and random conversion of photons into electrons, are generally undistinguishable, and the first three sources are modelled as the photocurrent shot noise. Third source is the photogenerated current as a result of ambient radiation unrelated to the signal. And the last source is the dark current, which depends on temperature and voltage bias [26], described by equation (10):

$$I_{dark} = I_s\left(e^{\frac{V_B}{V_T}} - 1\right) \quad [A] \tag{10}$$

which is basically equation (5) without light influence. The dark current is the leakage current when the detector is under bias ($V_B \neq 0$), but not exposed to light. It has components due to *generation-recombination* in the depletion region and tunneling between conduction and valence bands [13], [14]. For long wavelength IR detectors, dark current can be very high, and they must be physically cooled, sometimes to cryogenic temperatures such as the liquid-nitrogen 77 K, or lower in Helium closed-loop cooling systems. Even if the detector is cooled, shot noise is never reduced to zero, since beyond the dark noise there is always a background noise. For a given operating condition and environment, the background radiation is constant, and it limits the maximum performance that the detector can achieve. This is called the background-limited performance (the *BLIP* condition). Background spectral radiation is determined by the environment temperature T_B and the object *field of view* (*FOV*), i.e., the angular extent observable by the detector through its optics [28]. For instance, for $T_B = 300\ K$ and f/2 FOV, for a typical MCT detector, BLIP condition is achievable in the 3-5 µm wavelength range (MWIR) with detector cooling down to 200 K or less, whereas for the 8-14 µm range (LWIR) operating temperatures are cryogenic, around 100 K, to achieve BLIP performance [5]. The difference between MWIR and LWIR detectors operating temperatures, for BLIP condition, are due to their different amount of irradiation for the same background temperature T_B. Moreover, to avoid self-irradiation, the detector housing must also be cooled to the same temperature.

Total shot noise in a photodiode is then given by the influence of all sources aforementioned. Rewriting equation (9) we can calculate the total RMS shot noise:

$$\sigma_{shot} = \sqrt{2q(I_{ph} + I_B + I_{dark})B} \quad [A] \qquad (11)$$

where I_{ph} is the signal photocurrent, I_B the background photogenerated current (uncorrelated to the signal), and I_{dark} the dark current. The bandwidth B is defined by the same frequency as in equations (7) and (8). If a noise source has no correlation with the other, which is the case for thermal and shot noise, we can derive the total RMS noise current (σ_{noise}) as the sum of their variances [26], or:

$$\sigma_{noise} = \sqrt{\sigma_{shot}^2 + \sigma_{thermal}^2} = \sqrt{2B\left(q\left(I_{ph} + I_B + I_{dark}\right) + \frac{2k_BT}{R_{eq}}\right)} \quad [A] \qquad (12)$$

A common figure of merit of a photodetector, or a detection system, is the *Signal-to-Noise Ratio* (SNR), defined as the ratio of signal power and noise power [1], [4], [5]. From equations (6) and (12) we can find the SNR for a photon with wavelength λ as:

$$SNR = \frac{I_{ph}^2}{\sigma_{noise}^2} = \frac{\left(\eta q P_{opt}\frac{\lambda}{hc}\right)^2}{2.B\left(q\left(I_{ph} + I_B + I_{dark}\right) + \frac{2k_BT}{R_{eq}}\right)} \quad [1] \qquad (13)$$

where η is the quantum efficiency, P_{opt} the total impinging optical power (in [W] units), considering a monochromatic light source, *h* the Planck constant, and *c* is the speed of light. The term $A_{det}\Phi_0$ (the product of the photosensitive area and the photon flux density) in equation (6), that represents the amount of incident photons per second, was rewritten as $P_{opt}\lambda/(hc)$, i.e., the total optical incident power, a measurable unit, divided by the energy of a photon given by hc/λ. If a polychromatic source is used, then the photocurrent should be calculated as mentioned earlier in this chapter, as: $I_{ph} = \int^{\Delta\lambda} I_{ph[\lambda]}\, d\lambda$; where $\Delta\lambda$ is the spectrum of the incident light source, and $I_{ph[\lambda]}$ is the spectral photocurrent. The background current is calculated by the same way, with its respective spectral current $I_{B[\lambda]}(T_B, FOV)$, which in turn depends on the background temperature and the optics field-of-view. Finally, the spectral dark current $I_{dark[\lambda]}(T, V_{DET})$ depends on the detector temperature and on the voltage bias across the detector (V_{DET}).

From equation (13) we can find the *Noise-Equivalent Power* (*NEP*), which is the amount of impinging optical power that yields $SNR = 1$, and neglecting shot noise derived from photocurrent signal (in comparison with background and dark current):

$$NEP = \left(\frac{hc}{\eta\lambda q}\right)\sqrt{2.B\left(q(I_B + I_{dark}) + \frac{2k_BT}{R_{eq}}\right)} \quad [W] \qquad (14)$$

NEP is a common measure of the noise floor, and decreases with the temperature of the photodetector, and for larger equivalent shunt resistances R_{eq}, until it saturates to a constant value limited by background and dark current shot noise [13], [14]. The reciprocal of NEP is defined as the detectivity *D*, which correlates with photodetector performance. However, as NEP usually scales with area A_{det} and bandwidth B of the detector, to compare different detectors performances, the most usual figure of merit is the specific detectivity *D** (D-star) given as follows by equation (15):

$$D^* = \frac{\sqrt{A_{det}.B}}{NEP} \quad [Jones] \qquad (15)$$

where the unit $[Jones]$ denotes for $\left[\frac{cm\sqrt{Hz}}{W}\right]$.

2.6 Read-out pixel circuit requirements and figures of merit

The previous section described how to predict the behavior of a general photodetector through simulations, and their deleterious noise components. This section will describe the main desirable characteristics of a pixel circuit, which interfaces with its respective photodetector in order to read out the photogenerated signal. There are many pixel circuit topologies, and besides being responsible for signal reading, they also can improve detector performance by determining some of their operating conditions, for instance, attenuating the influence of noise [29], [30]. Each application demands its own requirements for the used pixel circuit, which may fall into any combination of the following categories: *charge storage capacity, sensitivity or*

charge-to-voltage gain, integration time, noise, dynamic range, and *detector bias stability* [29], [30]. Each of these figures of merit will be elaborated in the present section.

The first desirable feature is a large *charge storage capacity*, meaning that a great amount of charge can be stored in the pixel circuit to handle very bright signals and backgrounds, and large dark currents without saturating the pixel maximum capacity. When signal saturates, information is lost, and therefore a large capacitance is usually employed for that matter, whose drawback is a lower *sensitivity* and increased circuit area [29]. To interface a single photodetector, large areas are not an issue, however, for linear or two-dimensional arrays, a large pixel area is certainly a constraint.

The above-mentioned *sensitivity* is a figure of merit of a photodetection system that describes how close two values of light intensity can be, so that they are still distinguishable from each other [6]. For a voltage-mode pixel circuit, i.e., that converts a photocurrent I_{ph} into an output voltage V_o, the sensitivity S is defined as the derivative of the pixel circuit V_o with respect to the input I_{ph} (considering no dark current), as in equation (16) [6]:

$$S = \frac{\partial V_o}{\partial I_{ph}} \quad \left[\frac{V}{A}\right] \tag{16}$$

For a given system, disregarding noise effects, the read-out maximum resolution will be bounded to the lowest readable voltage δV_{min}, that is, the voltage resolution of the system, which in turn is limited by the systems *ADC* (*analog-to-digital-converter*) number of bits. From equation (16), for a given voltage difference δV_{min}, the larger the sensitivity, the lower the difference between two adjacent photocurrents readable by the pixel circuit, i.e., the higher the current resolution. For a pixel circuit operating in linear mode as a simple integrator, with the photocurrent as a constant current source being integrated in a capacitor with capacitance C_{int}, i.e. with $V_o = (I_{ph}\Delta T)/C_{int}$, equation (16) can be rewritten as [31]:

$$S = \frac{\partial}{\partial I_{ph}}\left(\frac{I_{ph}\Delta T}{C_{int}}\right) = \frac{\Delta T}{C_{int}} \quad \left[\frac{V}{A}\right] \tag{17}$$

where ΔT is the *integration time*, that is, the time interval in which the current is being integrated by the pixel circuit. For low signal applications, the tendency is to operate with larger ΔT, in order to increase sensitivity. On the other hand, for longer IR wavelengths (with high backgrounds of $\sim 10^{18}\,\frac{photons}{cm^2 s}$) or large dark currents, to prevent signal from saturation, ΔT must be kept small, due to finite C_{int} and operating voltages, rendering in lower sensitivities [29]. Another figure of merit is the *photoelectron charge-to-voltage conversion gain* $\left(G_{QV}\right)$, given in volts per electron $[V/e-]$. It is analogous to the sensitivity, and is defined as how much in volts is produced at a circuit output, for a given number of input electrons at the photodetector. For instance, for a given ΔT and I_{ph}, the number of electrons $(\#e^-)$ at the photodetector is given by $\#e^- = \left(\Delta T . I_{ph}\right)/q$, where q is the electron charge. Therefore, G_{QV} can be represented by equation (18), as follows:

$$G_{QV} = \frac{\Delta V_o}{\#e^-} = \frac{\left(\frac{I_{ph}\Delta T}{C_{int}}\right)}{\left(\frac{\Delta T . I_{ph}}{q}\right)} = \frac{q}{C_{int}} \quad \left[\frac{V}{e^-}\right] \tag{18}$$

With smaller capacitances, higher values of G_{QV} can be achieved, which is desirable to rise output signals above noise floor (NEP).

Although pixel circuitry may improve signal quality, it can also add noise to the system. Specially in applications with low background illumination $\left(10^6 - 10^{12}\,\frac{photons}{cm^2 s}\right)$, where background noise is not as issue, electronics noise can dominate, and must be kept as low as possible to maintain performance as close as possible to the BLIP condition [29]. While a larger ΔT may increase sensitivity, it may also introduce *1/f noise*, called *flicker* or *excess noise* [29]. Flicker noise is present in all kinds of electronic devices, and its main currently acceptable explanation is based on the fluctuation of the mobility of charge carriers [32]. The PSD expression for transistors can be rather complex, but mainly dependent on $1/f$, where f is the operating frequency within the equivalent noise bandwidth of the circuit B, and its RMS voltage $\sigma_{V-flicker}$ is more accurately evaluated through semi-empirical built-in models in integrated circuit simulators [25].

Another major noise source is the temporal *reset noise* [29]. This is introduced whenever a switch is used to reset a voltage on a capacitor, which is usually the case for photodetection systems. The reset transistor – that acts as a switch – when conducting, has a non-zero resistance, and therefore a thermal noise associated with it [33]. The noisy voltage that appears in series with the reset resistance gives rise to an uncertainty on the capacitor voltage, and its RMS noise value is given by equation (19). Reset noise is also called "*kTC*" noise.

$$\sigma_{V-kTC} = \sqrt{\frac{k_B T}{C_{int}}} \quad [V] \qquad (19)$$

where T is the circuits absolute temperature, k_B the Boltzmann constant, and C_{int} the integration capacitance.

Another extremely important figure of merit of a photodetection system is the *Dynamic Range* (*DR*). As briefly described in section 1.3, the DR is defined as the ratio between the maximum signal $\left(I_{ph\_max}\right)$ and the minimum RMS noise floor $\left(I_{ph\_min}\right)$ [29]. In other words, the range in which all signal values can be measured and resolved by the pixel circuit. Then, equation (20) defines de DR in decibels (*dB*):

$$DR = 20 * \log\left(\frac{I_{ph\_max}}{I_{ph\_min}}\right) \quad [dB] \qquad (20)$$

The main limiting factor for *DR* is the pixel finite charge storage capacity, and typical values vary between $70 - 80\ dB$ [29]. Dark and background currents are integrated together with the photocurrent signal, and if they are very high, the DR of the system has to be increased, in order to accommodate all integrated charge. Even if charge storage capacity is not an issue, the DR of a system is often limited by its ADC number of bits. For instance, disregarding quantization noise, a 16-bit ADC intrinsically renders a maximum dynamic range of $DR = 20 * \log(2^{16}) \cong 96\ dB$ [29]. It is important to note, that the *quantization noise* arisen from the process of digitalizing the analog signals are not going to be modelled in this thesis, since the main focus is to analyze

the pixel circuit. Any noise that would originate from the ADC would have to be included in the electronic noise component.

Furthermore, when interfacing a detector with its pixel circuit, the best case is when all current produced in the photodetector is transferred to the pixel circuit. However, this is not the usual case, since part of the photogenerated current will be partially lost internally in the detector shunt resistance. Equation (21) defines the percentage of current that is fed from the detector into the pixel circuit, for zero bias operation, called *injection efficiency* (ε):

$$\varepsilon = \frac{I_{ext}}{I_{ph}} = \frac{R_0}{R_0 + R_{in}} \quad [1] \qquad (21)$$

where R_{in} is the input resistance of the pixel circuit, and R_0 is the detectors dynamic resistance (R_d) at zero bias. The dynamic resistance is defined as the inverse of the I-V curve derivative, i.e., $R_d = (\partial I/\partial V)^{-1}$ [4], [5]. Equation (21) is valid only if the bias across the photodetector (V_{DET}) is maintained constant, since the dynamic resistance varies depending on the voltage, even when reverse biased, as can be seen in Figure 9 (an example of I-V curve). The higher the reverse voltage, then the larger the dynamic resistance. The ideal case is when $R_d \to \infty$, where the injection efficiency will be $\varepsilon \to 1$. Equation (21) can be derived based on Figure 18 current divider, which shows the photodetector simplified equivalent circuit.

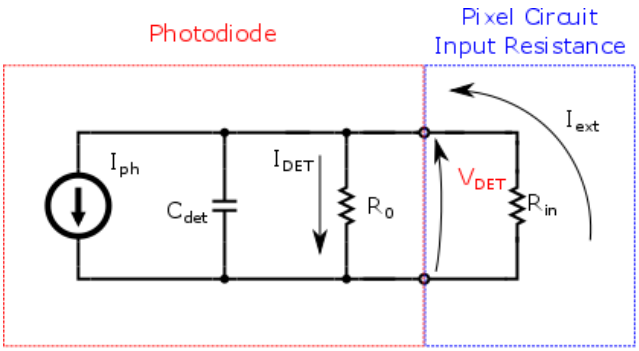

Figure 18: Injection efficiency illustration.

The injection efficiency is an important parameter, and should be as close to the unity as possible. It is a measure of *impedance bridging*, that indicates how faithfully a signal is transferred from the input to the output. For the case of current signals, the output impedance (of the photodetector) should be as high as possible, while the input impedance (of the pixel circuit) should be ideally zero [34], [35]. It is not to be confused with *impedance matching*, where the aim is to transfer the maximum amount of power.

Appropriate pixel circuits can be used in order to achieve $\varepsilon \cong 1$, mainly by providing an almost zero input resistance R_{in}, or by making it much lower than detectors resistance R_0 [5]. For IR detectors aimed at low background illuminations $\left(10^6 - 10^{12}\,\frac{photons}{cm^2 s}\right)$, the shunt resistance should be high enough to yield a good injection efficiency. Special pixel circuits – with very low R_{in} – are of particular interest for LWIR detectors, whose R_0 are usually smaller, and for very high performance applications [5].

In addition, the highest dynamic resistance R_d values occur when the photodetector is not at forward voltage. However, when operating at very small bias, ideally zero, dark current is the

smallest ($I_{dark} \cong 0$), noise is the closest to BLIP ($\sigma_{noise} \cong \sqrt{4k_B TB/R_0}$), and power consumption and responsivity are optimized [5], [29], [36]. Therefore, *detector bias stability* is of paramount importance, and voltage bias V_{DET} across the detector should be as close to zero as possible, for detectors based on p-n junctions.

Lastly, although it is not a figure of merit of pixel circuits, the R_0A *product* is worth mentioning as it is a widely-employed measure of the susceptibility of the photodetector to the deleterious effects caused by dark current and injection efficiency [29], where A is the detector optical area, and R_0 is the detector dynamic resistance, at zero bias. In a photodiode, if the area increases, the resistance decreases likewise, then the R_0A product should be constant, being an intrinsic material property [4], [5], [37]–[39]. The R_0A product is related to dark current, and decreases for longer cut-off-wavelength λ_{cut} materials, causing reduction in injection efficiency (ε), which introduces non-linearity. Once more, especially for LWIR detectors, a tight control of the detector bias might alleviate the problem.

2.7 Conventional IR pixel circuit topologies

This section will present conventional pixel circuits used to read out signals of IR detectors. Also, a comparison between them regarding the earlier described IR figures of merit will be presented – namely: charge storage capacity, sensitivity or charge-to-voltage gain, integration time, noise, dynamic range, and detector bias stability. All pixel circuits described here are voltage-mode, meaning they convert the photodetector current I_{ph} into an output voltage V_o, for a given integration time ΔT, presenting a specific charge-to-voltage gain G_{QV}. In addition, all pixel circuit topologies to be presented have a selection transistor (M_{sel}). These transistors act as a switch, and control whether the integrated voltage is transferred to the output or not. This control is performed through the signal SELECT. The M_{sel} transistors are extremely important when the pixels are operating in an array, since the SELECT signal on the gate of each M_{sel} is able to individually choose a pixel to feed its output voltage to the output bus – where all other pixels are also connected.

The first pixel circuit is the *Source-Follower Per Detector* (*SFD*) [4], [5], [29]. This circuit is widely known by the image-sensor community, working with the visible spectrum, by a different name: *Active Pixel Sensor* (*APS*) [40]. Its unit cell is composed by only three transistors, as shown in Figure 19. Its operation is quite simple: first, the reset transistor M_{rst} is on, setting the sense node (SN) voltage to a reference value, usually V_{DD}, if M_{rst} is a PMOS; then, when M_{rst} is off, the integration capacitance starts to discharge until the end of the integration time, when the SN output voltage is read through a source-follower transistor M_{SF}, at the output bus. As the photocurrent discharges the photodetector capacitance (C_{det}), the voltage at the SN changes, varying the detector bias, and rendering no bias stability.

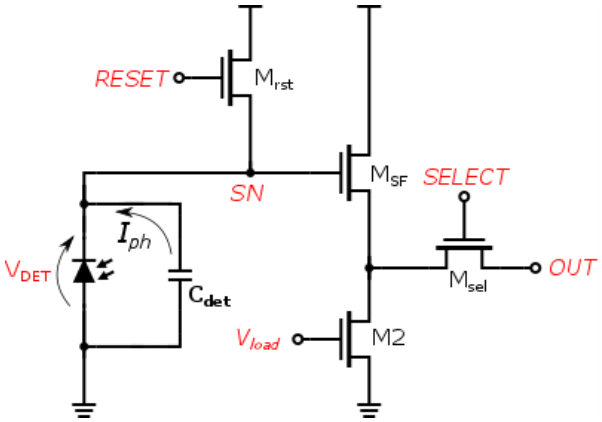

Figure 19: Schematic of the SFD.
Source: Adapted from [29].

In this circuit, M_{SF} and M_2 operate as a source-follower and a load transistor, respectively. The integration capacitance (C_{int}) for the SFD is given by the sum of C_{det}, the gate input capacitance of M_{SF}, and the source/drain capacitance of M_{rst}. For this circuit, the G_{QV} is slightly different than that presented in equation (18), since the output voltage variation is given by $\Delta V_o = A_{SF}.\Delta V_{SN}$, where A_{SF} is the source-follower gain ($A_{SF} < 1$), and ΔV_{SN} is the integrated voltage at the SN node. Therefore, the gain can be derived as $G_{QV} = A_{SF}(q/C_{int})$. Although increased storage capacity is desired in order to be able to transduce high current values without saturation, it reduces the gain, and for large values of C_{det}, the capacitance presents a significant voltage-dependent behavior, yielding a non-linear response gain since there is no bias stability. Furthermore, the G_{QV} of the SFD is small, because the integration parasitic capacitances cannot be made arbitrarily small. In addition, its low G_{QV} makes it more difficult to raise the output signal above the system noise floor in comparison to the circuits presented below, mainly due to its high 1/f noise component [4], [5], [29]. Therefore, the SFD should operate at higher frequencies, resulting, however, in smaller sensitivities due to the smaller integration time ΔT.

The second pixel circuit is called *Direct Injection* (*DI*) [4], [5], [29], [36], composed by an injection transistor M_i, an integration capacitor C_{int}, and a reset transistor M_{rst}, as shown in Figure 20. First, M_{rst} presets the voltage at the capacitor, and during the integration time, the photocurrent is integrated into C_{int} through the injection transistor M_i. Its G_{QV} is given by $G_{QV} = \varepsilon.q/C_{int}$, which is almost the same as equation (18), if $\varepsilon \cong 1$. The input resistance of the DI is given by $R_{in\_DI} = 1/g_{mi}$, where g_{mi} is the transconductance of the injection transistor M_i. The R_{in_DI} can be rather large compared to R_0, especially for low background illuminations, where currents are small, resulting in a small g_{mi} as well, and for LWIR detectors whose R_0 is small. In these cases, the DI features low injection efficiency, since $\varepsilon = \frac{R_0}{R_0+(1/g_{mi})}$, and R_0 is comparable to R_{in_DI}, offering a parasitic path for the photocurrent, instead of the pixel circuit. Beyond that, for LWIR detectors and very low background illuminations, the current flowing through M_i can be small, biasing it in the weak inversion regime, resulting in an even smaller transconductance g_{mi}. Therefore, DI is not suitable for detectors with small R_0 and ultra-low backgrounds $\left(< 10^6 \frac{photons}{cm^2 s}\right)$ [29]. Moreover, bias control is very poor for the DI pixel circuit, mainly because it is photocurrent dependent.

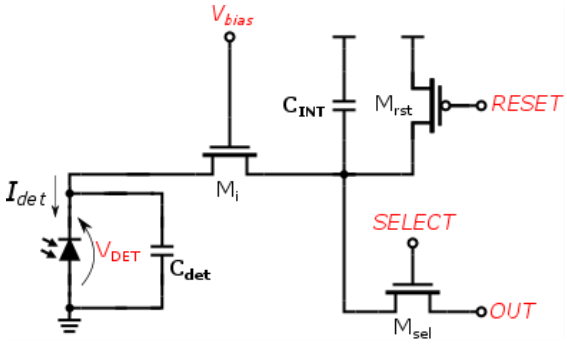

Figure 20: Schematic of the DI.
Source: Adapted from [29].

Another commonly used circuit is the *Buffered Direct Injection* (*BDI*), which is based on the aforementioned DI circuit. The only difference is the presence of an operational amplifier between the photodiode and the gate of the injection transistor M_i, as shown in Figure 21. The gain G_{QV} is the same as for the DI circuit, i.e. $G_{QV} = \varepsilon . q/C_{int}$. However, the introduction of the amplifier reduces the input resistance of M_i by a factor of $(1 + A_o)$, where A_o is the operational amplifier open-loop gain. This way, the injection efficiency is increased dramatically: $\varepsilon = \frac{R_0}{R_0 + 1/(g_{mi}(1+A_o))}$ [29], [41].

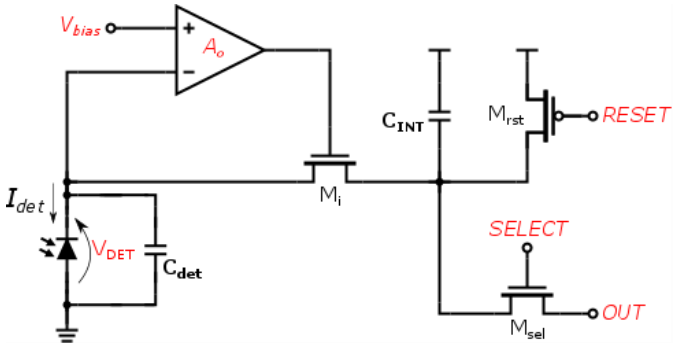

Figure 21: Schematic of the BDI.
Source: Adapted from [29].

Despite its higher injection efficiency, which solves the problem for detectors with smaller R_0 values, it features the same limitation as the DI circuit, regarding low backgrounds $\left(10^6 - 10^{12}\,\frac{photons}{cm^2 s}\right)$. This is because for very low drain currents, g_{mi} becomes smaller, reducing ε and making it dependent on the photocurrent. One main advantage of the BDI circuit is that it provides a feedback with the amplifier, and has very precise bias control, allowing it to provide zero voltage across the detector with very small error (equal to the amplifier input offset).

The next pixel circuit is the *Gate Modulation Input* (*GMI*) [4], [5], [29], [36], disposed in a current-mirror configuration. As shown in Figure 22, transistor M_i is connected as a load (gate and drain short-circuited) and modulates the gate voltage of the output transistor M_o, depending on the detector photocurrent. The output drain current at M_o is then discharged in the capacitor C_{int}, during the integration time. The input resistance of GMI is the same as for

the DI circuit, yielding the same injection efficiency: $\varepsilon = \frac{R_0}{R_0+(1/g_{mi})}$; and the current flowing through M_i is $I_{Mi} = \varepsilon . I_{ph}$. If $V_{bias} = GND$ (the ground reference voltage of the circuit), and both M_i and M_o are equally sized, the currents flowing through them are equal. However, with this configuration it is possible to set a large current gain $\beta = I_{Mo}/I_{Mi} = g_{mo}/g_{mi}$, where g_{mi} and g_{mo} are the transconductances of transistors M_i and M_o, respectively.

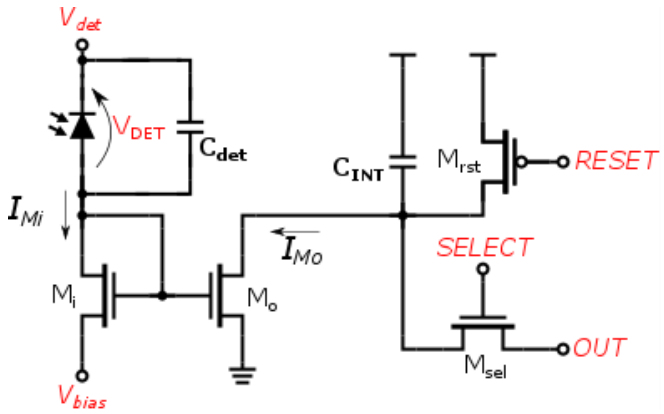

Figure 22: Schematic of the GMI.
Source: Adapted from [29].

From the general case of equation (18), we can find the GMI's charge-to-voltage gain:

$$G_{QV} = \frac{\Delta V_o}{\#e^-} = \frac{\left(\frac{I_{Mo}\Delta T}{C_{int}}\right)}{\left(\frac{\Delta T . I_{ph}}{q}\right)} = \frac{I_{Mo}}{I_{ph}}\frac{q}{C_{int}} = \frac{\beta.(\varepsilon . I_{ph})}{I_{ph}}\frac{q}{C_{int}} = \left(\frac{g_{mo}}{g_{mi}+1/R_0}\right)\left(\frac{q}{C_{int}}\right) \quad \left[\frac{V}{e^-}\right] \qquad (22)$$

where the voltage variation at the output (ΔV_o) is given by the DC current I_{Mo} being integrated at C_{int}, for an integration time ΔT. If current gain (β) is high, equation (22) shows that very large G_{QV} can be obtained, yielding very high detection sensitivity and reduced input-referred noise levels [29], even if large integration capacitances are employed, in order to increase charge storage capacity. Moreover, the larger the current through M_i, the smaller the input resistance, resulting in a larger injection efficiency ε. However, the current gain will decrease likewise, and therefore, the G_{QV} for this circuit self-adjusts depending on the background flux, being suitable for background (and dark current) suppression. Also, for higher backgrounds, the voltage bias (V_{bias}) at the source of M_i can be extended so that it compensates for any further increase in the gate-to-source voltage of M_i, due to higher currents, keeping the current gain steady. In addition, this background suppression feature can be used to achieve higher dynamic ranges (DR), at the cost of non-linearity, since V_{bias} can be used to control the current gain, and therefore the sensitivity [29]. Even so, a major drawback of GMI is that it cannot provide stable zero bias for the photodetector, and has relatively large input resistance $(R_{in\_GMI} = 1/g_{mi})$, featuring lesser injection efficiency.

The following circuit is one of the most widely employed read-out pixel circuits for infrared: the *Capacitive Transimpedance Amplifier* (*CTIA*) [4], [5], [29], [36]. Figure 23 shows the CTIA schematic, which is mainly composed by a reset transistor M_{rst}, an operational amplifier with open loop gain A_o, and the integration capacitance C_{int}. Basically, the reset transistor M_{rst} is used to discharge the capacitor before the start of the integration time. Next, the current from

the photodetector is injected into the circuit, and integrated in C_{int} during ΔT. The pixel circuit (not the amplifier) has a very low input impedance, and the ability to provide stable bias across the detector. Therefore, for a high enough gain A_o, current flowing through the parasitic shunt resistance R_0 is negligible (if $V_{bias} = 0\ V$), yielding $\varepsilon \cong 1$ and a very linear response.

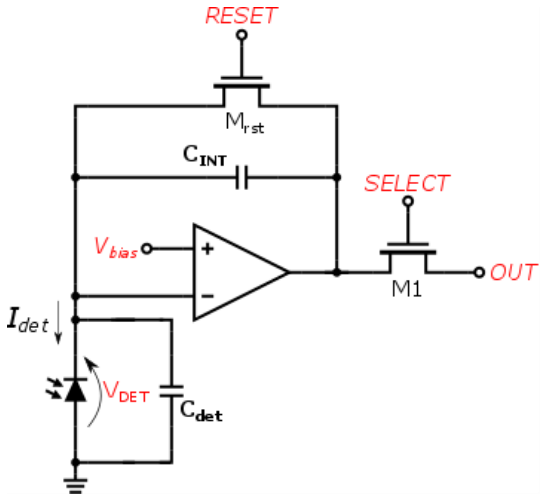

Figure 23: Schematic of the CTIA.
Source: Adapted from [29].

The charge-to-voltage gain for the CTIA is: $G_{QV} = q/\left(C_{int} + \frac{C_{int}+C_{det}}{A_o}\right)$ [29]. However, since virtually zero bias can be delivered to the photodetector, the CTIA prevents almost entirely that charge is integrated in the detector capacitance C_{det}. Consequently, for sufficiently large A_o, G_{QV} can be approximated to equation (18). Another advantage of the CTIA is that its integration capacitance can be made arbitrarily small, yielding large sensitivity and low noise performance, allowing it to operate at very low backgrounds $\left(< 10^6 \frac{photons}{cm^2 s}\right)$. In addition, it works fairly well for high backgrounds $\left(> 10^{18} \frac{photons}{cm^2 s}\right)$, being limited by the charge storage capacity, i.e., the integration capacitance C_{int}, and the maximum operating voltage of the circuit V_{DD}. Clearly, if it features a higher sensitivity, the charge storage capacity – and hence the DR – will be restricted. Main disadvantages of the CTIA are the large circuit (due to the operational amplifier), which renders a higher power consumption, the additional noise introduced by the amplifier, and the absence of current gain, like in the GMI.

Finally, another very important pixel circuit is the *Current Mirroring Integration* (*CMI*) [36], [42]–[45]. Figure 24 shows the CMI circuit, composed by four transistors, and an integration capacitor C_{int}. Current flowing through both M_{N1} and M_{P1} is $I_1 = \varepsilon . I_{ph}$. Since both PMOS transistors (M_{P1}, M_{P2}) are equally sized and have the same source-to-gate voltages, they form a unity current mirror, being the current through M_{P2}, and hence at M_{N2}, given by $I_2 \cong I_1$. These currents are not exactly the same due to possible mismatches between the transistors in the current mirrors, and due to a possible non-zero voltage across the photodiode (V_{DET}). Provided that $I_2 \cong I_1$, then the NMOS transistors (M_{N1}, M_{N2}) will also form a unity current mirror, since their currents and sizes are the same, resulting in equal gate-to-source voltages – respectively V_{GSN1} and V_{GSN2}. Since the source of M_{N2} is connected to GND, then $V_{GSN2} = V_N$. In addition, $V_{GSN1} = V_N - V_{DET}$. However, since $V_{GSN1} \cong V_{GSN2}$, therefore $V_{GSN1} \cong V_N$, and the bias across the detector V_{DET} will be stable and close to 0 V. Transistor M_{P3} copies the input current to the output node, yielding a current of $I_{out} \cong \beta . I_1 = \beta . \varepsilon . I_{ph}$, where β is the current gain between

M_{P3} and the double-current mirror, that can be written as $\beta = g_{mp3}/g_{mn1}$. During ΔT, the output current is integrated into C_{int}.

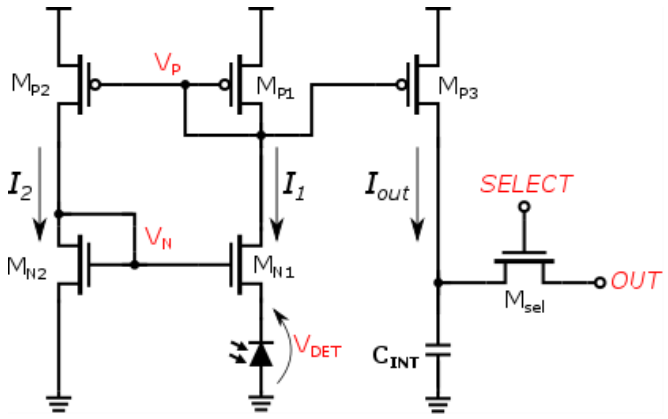

Figure 24: Schematic of the CMI.
Source: Adapted from [29].

In fact, as mentioned earlier, because of threshold voltage mismatches between transistors, current mirroring is not perfectly equal, and V_{DET} will be stable, but not at exactly zero volts [36], [42], [46]:

$$V_{DET} = \frac{K_p}{K_n}\Delta V_{tp} + \Delta V_{tn} \quad [V] \qquad (23)$$

where K_p and K_n are, respectively, PMOS and NMOS geometry constants, dependent on the width and length of the transistors, ΔV_{tp} is the threshold mismatch between the PMOS transistors, and ΔV_{tn} between the NMOS'. In addition, to calculate the injection efficiency, one has to know the input resistance of the CMI circuit, given as [36], [42], [43], [46]:

$$R_{in\_CMI} = \left(1 - \frac{g_{mn1}.g_{mp2}}{g_{mn2}.g_{mp1}}\right).\frac{1}{g_{mn1}} = \frac{\Delta g_m}{g_{mn1}} \quad [\Omega] \qquad (24)$$

where g_{mn1}, g_{mn2}, g_{mp1} and g_{mp2} are, respectively, the transconductances of transistors M_{N1}, M_{n2}, M_{P1} and M_{P2}. The dimensionless term Δg_m is defined as the *equivalent transconductance mismatching*, being zero for perfectly matched transistors. From equations (21) and (24), the injection efficiency can be calculated as $\varepsilon = \frac{R_0}{R_0 + (\Delta g_m/g_{mn1})}$. Similarly to the GMI circuit, as in equation (22), the charge-to-voltage gain is given by:

$$G_{QV} = \frac{\Delta V_o}{\#e^-} = \frac{I_{out}}{I_{ph}}\frac{q}{C_{int}} = \frac{\beta.(\varepsilon.I_{ph})}{I_{ph}}\frac{q}{C_{int}} = \left(\frac{R_0.g_{mp3}}{R_0.g_{mn1} + \Delta g_m}\right)\left(\frac{q}{C_{int}}\right) \quad \left[\frac{V}{e^-}\right] \qquad (25)$$

in which the current gain is considered, for simplicity, to be given by $\beta = g_{mp3}/g_{mn1}$, being g_{mp3} the transconductance of the output transistor M_{P3}.

For the ideal case wherein there is no threshold and transconductance mismatching, from equations (23), (24) and (25), we can find that $V_{DET} \to 0\ V$, $\Delta g_m \to 0$, $R_{in\_CMI} \to 0\ \Omega$, $\varepsilon \to 1$, and $G_{QV} \to \beta.q/C_{int}$. All these results are major advantages of the CMI circuit for IR

photodetection, and they do not differ very much from reality, since the CMI can be designed with transistors physically very close to each other, greatly reducing their relative threshold and geometry mismatching [43], [47]. Consequently, the real injection efficiency will be very close to 100 %, even if R_0 is very small, which is the case for LWIR detectors at cryogenic temperatures [4], [5] – and also for MWIR at higher temperatures (150-200 K). Furthermore, the CMI features very stable bias for a wide range of photocurrents, and although V_{DET} is not perfectly zero, it is close enough to benefit from all advantages discussed at the end of section 2.6. Moreover, the possibility of including a current gain β allows to increase the G_{QV} in a great extent without having to reduce the capacitance. Bottom line is that the CMI can feature both high sensitivity and charge storage capacity, with low noise performance. Like the CTIA, the CMI works well for high backgrounds and dark currents, being limited by C_{int} and maximum voltage V_{DD}.

Table 2: Summary of IR pixel circuits.

Type of Sensor	Acronym	$G_{QV}\ \left[\frac{V}{e^-}\right]$	$R_{in}\ \ [\Omega]$	β	V_{DET}
Source-Follower Per Detector	*SFD*	$A_{SF}(q/C_{int})$	---	---	*Unstable*
Direct Injection	*DI*	$\varepsilon.\frac{q}{C_{int}}$	$\frac{1}{g_{mi}}$	---	*Unstable*
Buffered Direct Injection	*BDI*	$\varepsilon.\frac{q}{C_{int}}$	$\frac{1}{g_{mi}(1+A_0)}$	---	*Stable*
Gate Modulation Input	*GMI*	$\left(\frac{g_{mo}}{g_{mi}+1/R_0}\right)\left(\frac{q}{C_{int}}\right)$	$\frac{1}{g_{mi}}$	$\frac{g_{mo}}{g_{mi}}$	*Unstable*
Capacitive Transimpedance Amplifier	*CTIA*	$\frac{q}{\left(C_{int}+\frac{C_{int}+C_{det}}{A_o}\right)}$	*Ideally zero, depending on input impedance of the amplifier*	---	*Stable*
Current Mirroring Integration	*CMI*	$\left(\frac{R_0.g_{mp3}}{R_0.g_{mn1}+\Delta g_m}\right)\left(\frac{q}{C_{int}}\right)$	$\left(1-\frac{g_{mn1}.g_{mp2}}{g_{mn2}.g_{mp1}}\right).\frac{1}{g_{mn1}}$	$\frac{g_{mp3}}{g_{mn1}}$	*Stable*

Table 2 presents a summary of all IR pixel circuits. After the presentation of all circuits above, we can find that there is generally a trade-off between the performance figures of merit. For instance, if the charge sensitivity (G_{QV}) is high, it usually implies in limited charge storage capacity and lower DR (except the GMI and CMI). The same pattern can be seen for: low noise performance vs. charge storage capacity; circuit complexity (e.g. due to amplifier) vs. noise, circuit size and power consumption. Also, for newer integrated-circuit technology processes, circuits can be made much smaller, fitting more complexity in smaller areas; yet, they feature smaller maximal operational voltage V_{DD}, limiting the DR. Choosing the appropriate circuit requires knowing what kind of detector it is going to interface with, the chip process technology, and the application requirements. Generally, low background applications in the SWIR and MWIR ranges require either low readout electronics and photodetector noise to keep operation as close to BLIP as possible. In contrast, the LWIR range features high background applications, demanding a tight bias control – to both minimize noise and to reduce dark current – and large charge storage capacity, and a higher DR [29].

2.8 High Dynamic Range Pixel Circuits

2.8.1 Motivation

As mentioned at the end of last section, some applications require a larger storage capacity, mainly to achieve higher dynamic ranges, i.e., to be able to increase the extent of signals that can be read by the pixel circuit. This is an obvious requirement for LWIR detectors, since the background photon flux is high for ambient temperature (T_B=300 K), and thereafter its related background current. Additionally, LWIR detectors have longer cut-off-wavelength, or equivalently, smaller bandgaps E_g, as described by equation (2), being more susceptible to thermal excitation, and therefore featuring higher dark currents. SWIR and MWIR detectors also feature large dark currents, and may need to operate with pixel circuits with a large storage capacity. Cooling the detector can reduce dark current, since less thermal energy is available, which in turn reduces noise [5]. However, there is a lower temperature limit, below which noise is not reduced any further, and is called the BLIP temperature (T_{BLIP}). Achieving T_{BLIP} has the objective of reducing the detector noise to the same level as that of the optical background noise, which is the highest performance of a detector at a given ambient temperature T_B, and detector angular FOV. If the operating temperature is higher than T_{BLIP}, the effect is that both shot and thermal noise of the detector – due to larger dark current and temperature itself – will be larger than the background photocurrent shot noise. However, it does not mean that the detector will not work properly, but only that its specific detectivity (D*) will be lower than its possible maximum value. The effect of increasing temperature can alter not only the maximum D*, but also the wavelength at which its peak occurs. Generally, as temperature decreases, the peak-detectivity will be shifted to longer wavelengths [9]. For higher temperatures, the decrease in D* can also be seen as an increase in the noise floor of the system (NEP), that renders in smaller SNR and DR.

Systems that can accommodate a very wide swing of input signals are labeled as High Dynamic Range (HDR) systems. A very attractive feature of HDR systems is their ability to present higher values of SNR. Equation (13) – which describes the photodetector SNR formula – illustrates this advantage, observing that the numerator of the ratio increases faster than its denominator, regarding the input signal I_{ph}, meaning that the higher the photocurrent, the larger the SNR. This is mainly because the RMS noise current σ_{noise} varies as the square root of the photocurrent, as can be noted from equation (12). We can rewrite equation (13) of the SNR, but now accounting for electronics and reset noise, and considering the output of a voltage-mode pixel circuit. The following SNR calculation takes into account the output-referred noise [25], [48], i.e., SNR will be the ratio between the output signal, and the total output noise (SNR_{out}):

$$SNR_{out} = \frac{V_{signal}^2}{\sigma_{V-noise}^2} = \frac{\left(\frac{I_{ph}.\Delta T}{q}.G_{QV}\right)^2}{\sigma_{V-electronics}^2 + \sigma_{V-kTC}^2 + \left(\frac{\sigma_{noise}.\Delta T}{q}.G_{QV}\right)^2} \quad [1] \qquad (26)$$

where V_{signal} is calculated based on equation (18), $\sigma_{V-noise}$ is the total noise voltage at the output of the pixel circuit, $\sigma_{V-electronics}$ is the total RMS noise voltage due to the pixel circuit, σ_{V-kTC} is the reset noise voltage given by equation (19), and last term of the denominator accounts for the equivalent detector RMS noise current σ_{noise}, given by equation (12), integrated through the pixel circuit. Since none of these noise sources are correlated, their variances can be summed. Expanding σ_{noise} and σ_{V-kTC} yields:

$$SNR_{out} = \frac{I_{ph}^2 \left(\frac{\Delta T}{q}.G_{QV}\right)^2}{\sigma_{V-electronics}^2 + \frac{kT}{C_{int}} + \left(\frac{\Delta T}{q}.G_{QV}\right)^2 . \left(2q(I_{ph} + I_B + I_{dark})B + \frac{4kT}{R_{eq}}B\right)} \quad [1] \qquad (27)$$

It is clear from equation (27) that the higher the photocurrent I_{ph}, the higher the SNR, even taking the pixel circuit noise into account. Whereas HDR topologies have a large charge storage capacity, then HDR pixel circuits render higher maximum output SNR values, since larger photocurrents can be integrated.

Before presenting the HDR pixel circuit topologies, one last remark needs to be made, regarding noise and SNR. First, equation (13) describes the SNR of the photodetector alone. When it is attached to a pixel circuit, the resulting SNR changes, as shown in equation (27). To measure how much it changes, i.e., the influence of the addition of the pixel circuit, the used figure of merit is the *Noise Factor* (*F*) [25], [49], defined as the ratio between the input and the output signal-to-noise ratios:

$$F = \frac{SNR_{in}}{SNR_{out}} \quad [1] \qquad (28)$$

where SNR_{in} is the pixel-circuit input SNR (of the photodetector alone). If the pixel circuit does not degrade noise performance, F will be close the one. However, very often F is larger than one, since more noise is introduced from the circuitry. This additional noise voltage at the output $\sigma_{V-pixel}$ is represented by the first two terms of the denominator of equation (27), or:

$$\sigma_{V-pixel} = \sqrt{\left(\sigma_{V-electronics}^2 + \frac{kT}{C_{int}}\right)} \quad [V] \qquad (29)$$

In fact, the SNR of the input of the pixel circuit (SNR_{in}) is given by equation (13), and combining it with equations (26), (28) and (29), we can find the noise factor:

$$F = \frac{SNR_{in}}{SNR_{out}} = \frac{\left(\frac{I_{ph}^2}{\sigma_{noise}^2}\right)}{\left(\frac{I_{ph}^2 \left(\frac{\Delta T}{q}.G_{QV}\right)^2}{\sigma_{V-pixel}^2 + \sigma_{noise}^2 \left(\frac{\Delta T}{q}.G_{QV}\right)^2}\right)} = \frac{\sigma_{V-pixel}^2 + \sigma_{noise}^2 \left(\frac{\Delta T}{q}.G_{QV}\right)^2}{\sigma_{noise}^2 \left(\frac{\Delta T}{q}.G_{QV}\right)^2} \quad [1] \qquad (30)$$

Additionally, in regard to all pixel circuits discussed in the previous section, and recalling equation (18) – of the G_{QV} –, we can find that a more general expression for the charge-to-voltage gain can be given as follows:

$$G_{QV} = \frac{\Delta V_o}{\#e^-} = \beta . \varepsilon . \frac{q}{C_{int}} \quad \left[\frac{V}{e^-}\right] \qquad (31)$$

where β is the current gain and ε the injection efficiency. Joining and regrouping equations (30) and (31):

$$F = \frac{\sigma_{V-pixel}^2 + \beta^2 . \left(\sigma_{noise} . \varepsilon . \frac{\Delta T}{C_{int}}\right)^2}{\beta^2 . \left(\sigma_{noise} . \varepsilon . \frac{\Delta T}{C_{int}}\right)^2} = \frac{\sigma_{V-pixel}^2 + \beta^2 . \sigma_{V-noise-PD}^2}{\beta^2 . \sigma_{V-noise-PD}^2} = 1 + \left(\frac{\sigma_{V-pixel}}{\beta . \sigma_{V-noise-PD}}\right)^2 \quad [1] \qquad (32)$$

in which $\sigma_{V-noise-PD}$ is the integrated noise voltage only due to the RMS noise current of the photodetector (σ_{noise}), considering no current gain, i.e., with $\beta = 1$. Clearly, if $\sigma_{V-noise-PD}$ is much higher than $\sigma_{V-pixel}$, then $F \to 1$. We can play with various circuits parameters in order to make $\sigma_{V-noise-PD} \gg \sigma_{V-pixel}$, like assuring a high injection efficiency, or by increasing the integration time, or even decreasing the integration capacitance. However, we generally cannot increase ε above unity, larger ΔT may slow down system operation, and smaller C_{int} may render limited DR and charge-storage-capacity. The only factor that does not appear to show substantial drawback is the increase of β.

Increasing the current gain, we can see that the impact of noise introduced by the pixel circuit is substantially reduced, and it can be considered as a noiseless component. That could also be concluded if we have considered the input-referred noise. Although noise is measured at the output of the pixel circuit (*output-referred noise*), it can be referred back to its input, and then compared to the input stage noise [25], [49]. If the circuit has a sensitivity *S*, that is equivalent to a DC transimpedance gain, in [V/A], then the input-referred noise PSD $\left(\sigma_{input-noise}^2\right)$ is simply given by the output-referred noise PSD $\left(\sigma_{V-output-noise}^2\right)$ divided by the squared gain (S^2), or:

$$\sigma_{input-noise}^2 = \frac{\sigma_{V-output-noise}^2}{S^2} \quad [A^2] \qquad (33)$$

in which $S = \left(\frac{\Delta T}{q} . G_{QV}\right) = \left(\beta . \varepsilon . \frac{\Delta T}{C_{int}}\right)$. This way, if we apply equation (33) to the total noise voltage at the output of the pixel circuit $\sigma_{V-noise}$, given by the term at the denominator of SNR_{out}, from equation (26) – also given by the numerator of equation (30) –, then:

$$\sigma_{input-noise}^2 = \frac{\sigma_{V-noise}^2}{S^2} = \frac{\sigma_{V-pixel}^2 + \sigma_{noise}^2 \left(\frac{\Delta T}{q} . G_{QV}\right)^2}{\left(\frac{\Delta T}{q} . G_{QV}\right)^2} = \frac{\sigma_{V-pixel}^2}{\left(\frac{\Delta T}{q} . G_{QV}\right)^2} + \sigma_{noise}^2 \quad [A^2] \qquad (34)$$

The input-referred noise is used only to get a clue of how the pixel circuit would corrupt the input signal. We can realize from equation (34) that, besides the noise term from the photodetector $\left(\sigma_{noise}^2\right)$, an extra term is added, comprising the "fictitious" pixel circuit input-referred noise. Once more, we can find that this term is drastically reduced if the circuit presents a higher current gain β, and therefore, a higher sensitivity S – or transimpedance gain.

However, increasing S in any way will boost both signal and noise (keeping SNR), meaning that more charge will be stored in the pixel circuit. The bottom line is that, in order to mitigate noise influence from the pixel circuit, a high sensitivity gain has to be employed, and the best way to do that is by increasing the current gain β, so that the integration capacitance can still be maintained large enough to achieve a higher DR, necessary to accommodate all charge coming from an IR detector. That means not only noise, but also dark and background currents, along with signal itself.

2.8.2 Dynamic Range Definition

As introduced in section 2.6, the DR quantifies the ability of a sensor to satisfactorily detect a given range of input signal intensities. The widest the intensity range of input signals, the higher the DR. It is defined as the ratio of the largest nonsaturating input signal to the smallest detectable one [50], as shown by equation (20) – a logarithmic ratio between maximum and minimum signals. The largest input signal $\left(I_{ph\_max}\right)$ is given by:

$$I_{ph\_max} = I_{sat} - (I_B + I_{dark}) = \frac{V_{DD}}{\Delta T}.\frac{q}{G_{QV}} - (I_B + I_{dark}) \quad [A] \qquad (35)$$

where I_{sat} is the maximum current at the detector that can be integrated by the pixel circuit, or equivalently, considering equation (18), $I_{sat} = \frac{V_{DD}}{\Delta T}.\frac{q}{G_{QV}} = \frac{V_{DD}}{S}$, in which V_{DD} is the maximal operational voltage and S is the sensitivity (or transimpedance gain); I_B and I_{dark} represent, respectively, the background and dark currents. On the other hand, the minimum detectable current $\left(I_{ph\_min}\right)$ is given by the standard deviation of the input-referred noise PSD of equation (34), but in dark conditions (without the I_{ph} term) [50]:

$$I_{ph\_min} = \sqrt{\frac{\left(\sigma^2_{V-electronics} + \frac{kT}{C_{int}}\right)}{\left(\frac{\Delta T}{q}.G_{QV}\right)^2} + \left(2q(I_B + I_{dark})B + \frac{4kT}{R_{eq}}B\right)} \quad [A] \qquad (36)$$

Therefore, with equations (20), (35) and (36), and considering the ratio (I_{ph_max}/I_{ph_min}), we can write the DR definition as:

$$DR = 20 * \log\left\{\frac{\frac{V_{DD}}{\Delta T}.\frac{q}{G_{QV}} - (I_B + I_{dark})}{\sqrt{\frac{\left(\sigma^2_{V-electronics} + \frac{kT}{C_{int}}\right)}{\left(\frac{\Delta T}{q}.G_{QV}\right)^2} + \left(2q(I_B + I_{dark})B + \frac{4kT}{R_{eq}}B\right)}}\right\} \quad [dB] \qquad (37)$$

2.8.3 Types of HDR Sensors

Equation (37) provides insights on how to enhance the DR of a specific system. In general terms, by either increasing I_{ph_max} or decreasing I_{ph_min}, the DR can be extended. Generally, HDR techniques only increase I_{ph_max}. However, in the same way as for the SNR, I_{ph_min} can be decreased if noise introduced from the pixel circuit is mitigated, which is done mainly by increasing the charge sensitivity G_{QV}. Instead, to extend I_{ph_max}, we can either increase V_{DD}, or decrease the sensitivity gain S to "slow down" the rate at which the response reaches V_{DD}. However, this last approach of reducing the sensitivity may decrease the SNR [6], [50]. In addition, increase the V_{DD} may not be possible, since it is restricted to the maximum voltage a given CMOS technology can handle [25]. Recalling equation (17), but now for the general case, the sensitivity for linear voltage-mode pixels is given by:

$$S = \frac{\partial V_o}{\partial I_{ph}} = \frac{\Delta T}{q}.G_{QV} = \beta.\varepsilon.\frac{\Delta T}{C_{int}} \quad \left[\frac{V}{A}\right] \qquad (38)$$

With equation (38) in mind, we will briefly present the main pixel circuits for HDR operation, that are usually divided into seven different categories [6].

The *Clipping Sensors* use Capacitance Adjustment [51], [52], which combines the response of different capacitances. This means that, in order to achieve HDR operation, the pixel circuit transimpedance needs to change during the integration time. This renders in a nonlinear response, and a more complex control of the pixel and signal processing [50]. In addition, it features more kTC noise, lesser SNR and sensitivity for higher illuminations, when the extra capacitance is added [6]. When capacitances are switched, SNR dips may occur, i.e., the SNR falls abruptly, and starts to increase again with photocurrent [6], [50]. This occurs mainly because in the calculation of SNR_{out} in equation (27), if we consider $B = \frac{1}{2.\Delta T}$, we can find that the SNR is proportional to ΔT, and the different capacitances are integrated during integration times that are fractions of ΔT.

The *Companding Sensors*, or *Logarithmic Sensors*, compress the pixel response in a logarithmic fashion. This results in a very wide range of photocurrents that can be integrated in a rather small voltage swing. In practice, it means that the pixel circuit transimpedance sinks proportionally to the input photocurrent, yielding a continuously smaller sensitivity and limited SNR [6]. This circuit has a very simple control and high DR, but with signal nonlinearity and very poor sensitivity at high illumination levels [53].

The *Multimode Sensors* operate in both logarithmic and linear modes [54], [55], achieving very high DR. For lower photocurrent levels it operates at the linear mode, keeping its characteristic higher sensitivity. For higher photocurrents, however, it operates at the logarithmic mode, allowing it to transduce higher currents, but also rendering very poor sensitivity. When swapping from one mode to another, it also presents a SNR dip. It demands a more sophisticated pixel circuit control, along with nonlinearity, and non-trivial signal processing.

Frequency-based Sensors operate converting light intensity in frequency pulses [56], [57]. The higher the light intensity, the higher its frequency. They feature linear response, while

achieving HDR with constant SNR for the whole DR – a very good feature. However, the maximum SNR can only be achieved at the expense of limiting the sensor ability to detect high photocurrents [6]. On the other hand, since they detect only saturation events, their capability to sense low photocurrents is only possible for lower SNR values.

Time to First-Spike (*TTS*) *Sensors* [58], [59] count the time required for the pixel circuit to saturate. Information of both saturation voltage and time-to-saturation is recorded and used to reconstruct the output signal, allowing very high DR values. Still, they suffer under steadily lower sensitivities for higher photocurrents, nonlinearity, and constructive limitations on smallest and largest readable signals [50]. Further, they need very precise time controlling and high resolution of the saturation time representation [6].

Both categories *Global-control-over-the-integration-time Sensors* [60], [61] and *Autonomous-control-over-the-integration-time Sensors* [62], [63] implement multiple capture algorithms. This algorithm exposes the photodetector during at least two different integration times, and read their response. Smaller integration times are required to read higher photocurrents, whereas longer integration times are suitable for smaller currents. All responses are combined and used to reconstruct the HDR signal. However, with smaller ΔT values, as can be concluded from equation (38), sensitivity is smaller. Furthermore, since ΔT is subdivided in two or more integration times, the SNR will present dips (actually looking like a sawtooth curve) [6]. In addition, they require complex signal processing and control, and operate slower, due to the need of multiple exposures. The difference between the two categories are their pixel reset control. If the system is composed of several pixel circuits, the reset control can be global, or inside the own pixel circuit [6].

Table 3 shows a comparison between all kinds of HDR sensors:

Table 3: Summary of HDR types of sensors.

Type of Sensor	Principle	Linear	SNR	Additional info
Clipping	Adjusts the sensitivity by switching capacitances	No	Presents dips when switching capacitances. Lesser SNR for higher illuminations	Complex control, and signal processing
Companding/ Logarithmic	Compresses the response in a logarithmic fashion	No	Limited SNR. Very poor sensitivity at higher illuminations	Small voltage swing. Simple control
Multimode	Operates both in logarithmic and linear modes	No	Presents dips when switching from linear to log	More sophisticated control
Frequency-based	Converts light intensity into frequency pulses	Yes	Constant SNR. Higher SNR at the expense of limiting the maximum photocurrent	Difficulty in detecting low light intensities
***Time to First-Spike* (*TTS*)**	Counts the time to first saturation	No	Limited SNR. Steadily lower sensitivities	Precise control and high resolution on saturation time representation. Limited maximum and minimum currents
Global-/Autonomous - control-over-the-integration-time	Uses multiple capture algorithm, exposing sensor to multiple integration times	No	Presents dips when changing integration time. Smaller sensitivities when decreasing ΔT	Complex control and signal processing. Slower operation due to multiple ΔT

We can summarize from all kinds of HDR sensor concepts, that their main drawbacks are [6], [31], [50]: poor sensitivity, mainly at higher photocurrents; nonlinear response; slow operation (multiple captures required); somewhat sophisticated control; non-trivial signal processing; and the occurrence of SNR dips.

In order to implement any of the abovementioned HDR concepts for IR detection, we should choose one of the pixel circuit topologies mentioned in section 2.7, namely, the SFD, DI, BDI, GMI, CTIA, or CMI. In next chapter, more specifically in section 3.3, we will choose and modify one of these topologies, in order to achieve the best result for HDR operation, and with as less disadvantages as possible.

2.9 Basic Infrared Optics

We have seen so far that the pixel circuit is a very important component of an IR photodetection system. Nevertheless, the systems optics are also equally important, and deserve to be revised. In this section, basic optical concepts are going to be shown to demonstrate how optics play a crucial role in the overall system performance.

By performance, we mean primarily the signal-to-noise ratio, discussed in previous sections. Beyond the definitions shown before, another usual way of expressing the SNR is in the *dB* units. If we take equation (13), for instance, where $SNR = I_{ph}^2/\sigma_{noise}^2$, this could be expressed in *dB* as $SNR_{dB} = 10 * \log(I_{ph}^2/\sigma_{noise}^2)$, or $SNR_{dB} = 20 * \log(I_{ph}/\sigma_{noise})$. Equivalently, that could be expressed using power units, instead of current or voltage: $SNR_{dB} = 10 * \log(P_{opt}/P_{noise})$, where P_{opt} is the signal power, and P_{noise} the noise power. The conclusion is that it does not matter how the "pure" ratio is calculated, as in equation (13), since it is always normalized when converting to *dB*, in the logarithmic scale.

We have discussed in section 2.8.1, that if we employ a large transimpedance gain, that is, a high sensitivity S, the pixel circuit responsible for signal readout would have a noise factor $F \rightarrow 1$, and in this case the output SNR value would be very close to the input one. This means that the final SNR of the system would be almost the same as the SNR of the photodetector, given by equation (13). Therefore, the maximum possible SNR is dictated by the photodetector, and can be rewritten in optical power units [9], as:

$$SNR_{in} = \frac{P_{opt}}{NEP} \quad [1] \qquad (39)$$

where NEP is the noise equivalent power, shown by equation (14). We know that, for a specific optical system with *focal length f*, and *optical aperture D*, we can write the optical power of the incident signal (P_{opt}) at the surface of the photodetector as [9]:

$$P_{opt} = \frac{A_{det}}{4.\left(\frac{f}{D}\right)^2}.E \quad [W] \qquad (40)$$

in which A_{det} is the sensitive area of the photodetector, and E is the amount of irradiated signal power that reaches the detector, namely, the signal power density in $[W/m^2]$ units. This is valid when the source of the signal is at infinity (at least far away) with respect to the detector. If we consider a squared detector with length and width equal to L, then:

$$P_{opt} = \frac{L^2}{4.(f/\#)^2}.E \qquad [W] \tag{41}$$

where we have replaced (f/D) by $(f/\#)$, being this term called the *f-number* of the lens system [9]. Recalling equation (15), containing the relation between NEP and the specific detectivity D*, and from equations (39) and (41), the SNR_{in} can be expressed as follows [9]:

$$SNR_{in} = \frac{D^*.L}{4.(f/\#)^2.\sqrt{B}}.E \quad [1] \tag{42}$$

where B is the Equivalent Noise Bandwidth, and we have used $A_{det} = L^2$ in equation (15). Expression (42) is rather simple, but demonstrates the strong relation between the SNR of the final system and its optics. Smaller $f/\#$ objective lenses mean larger SNR, i.e., lenses with short focal lengths – high optical powers – and large diameters. The problem with smaller $f/\#$, however, are the introduction of optical aberrations [9]. This is an issue mainly when an array of detectors is employed, for instance, for image sensors, since the aberrations introduced by the lens system will distort the image of the object being observed. For a single (or a few) detectors, however, it should not be a problem.

If we inspect more closely the variable *E* in equation (42), we could see that this power density actually is just a fraction of the amount of power density emitted by the signal source, namely $E = \alpha_P.E_{source}$, where α_P is a power scale factor ($\alpha_P \leq 1$). Also, these power densities are given by the integration of their respective spectral power densities – $E_{[\lambda]}$ and $E_{source[\lambda]}$ – over the wavelength range at which the detector operates. In fact, if the detector is a photodiode, its specific detectivity D* will also vary with the wavelength, and equation (42) must be integrated over the spectral range to yield the correct SNR. However, for a narrow spectral range $\Delta\lambda$, centered at λ_M, we can calculate the SNR by simply assuming a constant specific detectivity $D^*(\lambda_M)$, and calculating the irradiated power density at the detector by $E = \alpha_P.\left(E_{source[\lambda]}.\Delta\lambda\right)$ [9].

Term α_P can be understood as a transmission power loss factor. This is because the emitted radiation from the source must travel through an optical path until it reaches the detector. In general, it means both the atmospheric transmission (τ_A) and the optical system transmission (τ_O) [9]. Therefore, we can write $\alpha_P = \tau_A.\tau_O$. We can rewrite equation (42) accounting for the transmission factors as:

$$SNR_{in} = \frac{D^*.L}{4.(f/\#)^2.\sqrt{B}}.\tau_A.\tau_O.E_{source} \quad [1] \tag{43}$$

Calculating τ_A is a really complex matter, since it varies with wavelength, altitude, temperature, humidity, atmosphere composition, and so on [9]. It is especially important for IR photodetection in spatial applications, and monitoring of large areas using cameras or

open-path detectors. For other applications that are not too much affected by small variations in τ_A, constant values for the atmospheric transmission can be applied, depending on which atmospheric wavelength window is considered, namely, the MWIR or LWIR window. In contrast, for gas detection systems, τ_A varies depending on the amount of the target gas that is absorbed. In section 4.4, a glimpse of how transmission is calculated will be presented and used for the CO_2 gas detection system. In fact, for this case term E_{source} in equation (43) is considered constant, and is the IR source power density. The actual signal is calculated comparing the response of two gas volumes: one where the amount of the target gas has to be measured; and a control (with a reference gas, e.g. N_2), or a previously calibrated volume (also fed with the reference gas) [17].

From the optics point of view, two terms can be considered in order to improve SNR. First one, we have already observed in this section, which is the $f/\#$ of the lens system. Second one, is the optical transmission τ_O, which can be improved mainly by using *Anti-Reflective Coatings (ARC)* over the lens system.

2.9.1 Optical Gain

The simplest optical system is composed by only a pinhole – a tiny aperture. This solution however limits the amount of light that can be collected by the optical detector by a great extent. A very simple, but effective alternative, is to use an objective lens, which gathers much more light than a pinhole. We have said that a way of boosting SNR is through the $f/\#$, and this can be done by choosing the proper objective lens focal length and diameter. However, we still can provide an optical gain, and henceforth, improve the system SNR by using other kinds of optical elements. These are the *immersion lenses*, *light pipes*, and *field lenses* [9], [64]. They allow the reduction of the detector physical size, rendering more compact systems. Also, they permit a more integrating effect regarding the sensitive area of the detector, i.e., they improve the *fill-factor* (*FF*) – the percentage of the detector area sensitive to light.

First type of lens is the *immersion lens*, shown in Figure 25 . As the very name states, it is an optical element immersed at the rear surface of the detector. The main purpose of this lens is to refract the incoming oblique rays that would not reach the detector, if the lens was not there. The optical gain (OG) in this case, considering air as the medium, is given by $OG = L_{virtual}/L = n$, where $L_{virtual}$ is the linear size the detector should have to gather the same amount of light if there was no lens, *L* is the allowed reduction of the detector size when the lens is employed, and n is the *refractive index* of the lens material. The SNR is multiplied by the optical gain, since we now use $L_{virtual}$ in equation (43), instead of *L*, and hence multiplying the SNR by *n* [9], [64]. As a reference, silicon has a refractive index of about 3.4.

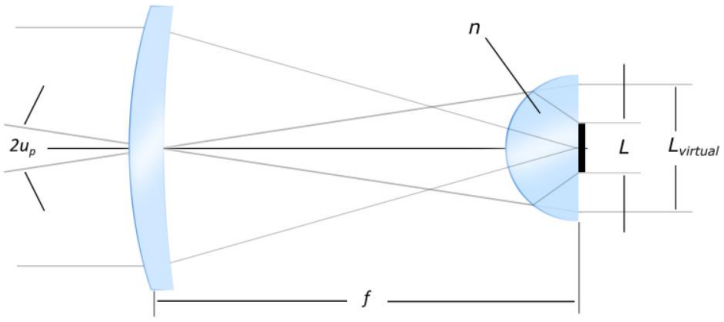

Figure 25: Immersion lens placed at the rear side of the detector.
Source: Adapted from [9].

The exemplified immersion lens just discussed has a hemispherical shape. It may have different formats though, like the hyperhemisphere, which may present an optical gain of n^2 [9], [64].

Light pipes are kind of hollow reflective cones that concentrate the input energy into an area smaller than its opening, where the detector should be placed, like shown in Figure 26. The optical gain is the ratio between the cone opening and the detector area, being limited for the presented geometry, though.

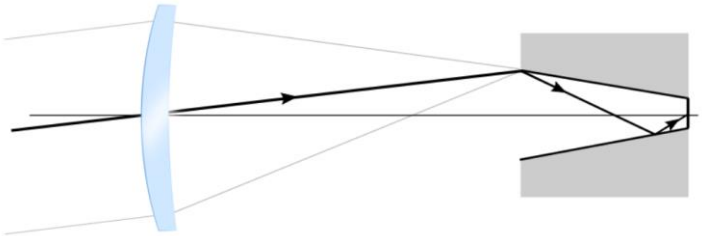

Figure 26: Reflective light pipes placed at the rear side of the detector.
Source: Adapted from [9].

Last kind of optical element is the *field lens*. These lenses may be placed at the focus of the objective, like in Figure 27, or very close to it. If placed at the focus, it adds no optical power to the lens system [9], [65]. The field lens power is calculated to image the *aperture stop* of the objective into the detector. The aperture stop is defined as the physical element that limits the cone of light that can pass through the lens system [9], [66]. It can be an iris, or the edge of some lens, being in this case the edges of the objective lens.

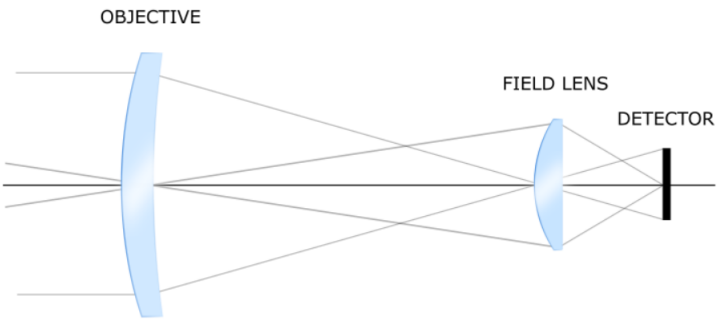

Figure 27: Field lens placed at the focus of the objective.
Source: Adapted from [9].

Basically, the field lens increases the SNR by the ratio between the area of the Field Lens (A_{FL}) and the area of the detector (A_{det}). If we assume a squared field lens, then the optical gain (OG) is given by $OG = A_{FL}/A_{det} = (L_{FL}/L)^2$, where L_{FL} and L are, respectively, the linear size of the field lens and the detector. Then, the effect of the field lens in equation (43) can be seen below:

$$SNR_{in} = \left(\frac{L_{FL}^2}{L}\right)\frac{D^*}{4.(f/\#)^2.\sqrt{B}}.\tau_A.\tau_O.E_{source} \quad [1] \qquad (44)$$

It seems that the field lenses offer a great advantage regarding the SNR, boosting incoming signal. It is clear that the field lens should be larger than detector itself, in order to overcome the low fill-factor issue.

2.9.2 Microlenses

We might have a case in which the size of the detector is very small. We may also have an array of detectors, instead of a single one. In these situations, we shall consider using a type of lens similar to the field lens: the *Microlens*. These are literally small lenses, placed very close to the (also small) detector, with the purpose of enhancing the amount of light it senses. It is particularly interesting for pixels or detectors with very low fill-factor (FF), just as discussed in the last section. In addition, a detector may also have a low FF, in which case it is interesting to use microlenses. The FF of a detector is defined as the ratio of its photosensitive area and its total area, for instance when the photodetector area is partially covered by a metal area.

For the sake of generality, let us consider the case where we have a system with an array of detectors, either a linear array or a matrix. We might employ for this situation a microlens array, with the same format as the detector array. The purpose is that each microlens focuses incoming light over its respective photodetector. This way, the optical gain will be similar to that of the field lens, namely, the ratio between the area of each microlens ($A_{microlens}$) and the sensitive detector area (A_{det}), that is: $OG = A_{microlens}/A_{det}$.

Microlenses are very interesting for applications that need miniaturization, being even more attractive since it reduces the size, and hence the cost of packaging [67]. Building these kind of lenses for the infrared spectra, however, require appropriate materials. They should be transparent and have a high enough refractive index, allowing a more compact system, since the microlens can add power to the optical system. Find a material suitable for IR optics is a real challenge, mainly because there is not really a great amount of transparent materials in these spectra. Among the available ones, there are some hygroscopic, toxic, and other hazardous features that make them impractical for real systems [10].

From the most popular materials, we can name Germanium, ZnSe, and Silicon as good materials for IR optics, among a very few others. Germanium has a very high refractive index (about 4), and a wide transmission band, from 2 µm up to 17 µm [5], [10], being particularly useful for the LWIR range. It is quite brittle, but good to polish, stressing that optical elements need to be polished to show acceptable performance [67]. A main problem of germanium is its large gradient of the refractive index on temperature (dn/dT), meaning that it must be

athermalized. Also, in the LWIR range, for temperatures above $57\ °C$, the absorption coefficient (α) increases [5]. This is the same coefficient as in equation (3).

Another good material is the ZnSe, which has a relatively high n (about 2.4), and transmission window from 0.55 µm until 20 µm. However, it is relatively soft and weak, and more expensive than germanium. It features a lower dn/dT, about 15 % of that of germanium [5], [10].

Among many of the available materials for IR, Silicon is a very low-cost choice, and one with the smallest densities, therefore allowing very light-weighted systems. It has high refractive index (about 3.4), but relatively large dn/dT (about 40 % that of germanium). Both these values are between those presented for ZnSe and Ge. In contrast, it has a narrower transmission window, from 1.2 µm up to 9 µm, being more suitable for MWIR, rather than LWIR [5], [10]. However, it still features acceptable transmittance within the LWIR range, depending on application. And recalling equation (3), although the absorption coefficient may increase for $\lambda > 9\ \mu m$, if the optical component thickness is not too large (reasonable for microlenses), it may transmit a tolerable amount of light that still justifies adding this optical component to improve SNR. In addition, despite the fact that silicon is inexpensive, it is difficult to fabricate through conventional diamond turn method, which could render its manufacturing more expensive [10]. Nevertheless, there are alternative methods for silicon microfabrication that are not expensive, offering a feasible solution [67], [68]. We are going to use the method presented by *de Lima Monteiro et al.*, *2003*, in section 3.1, in order to fabricate the microlenses. Among others, this method also has the advantage of reducing alignment cost, since it results in very precise optical axes of the microfabricated lenses.

From equation (44), we can conclude that boosting SNR with an optical gain will only be advantageous if the introduction of new optical elements does not result in less transmission (τ_O). The aforementioned materials feature a transmission window within the IR spectral range, but not yet 100 %. In fact, although higher refractive indices may be a good feature, allowing small lenses with higher power, it has a negative effect on surface reflection, which may reduce transmission by a great extent. To improve the parameter τ_O in equation (44), anti-reflective coatings (ARC) need to be used.

2.9.3 Anti-Reflective Coatings

We have said in last subsection that the larger the refractive index, the higher the surface reflectance [10]. Indeed, this is an unwanted feature, since a lot of incoming power would be lost in the optical path before reaching the detector. Silicon, for instance, has a surface reflectance of about 46 % in the MWIR range, and germanium has more reflectance due to its higher refractive index.

Adding the ARC will improve transmittance at the desired range of wavelengths. The wider the range of wavelengths within which we want to enhance transmittance, the more complex the ARC. It can be just a single-layer of a specific material, with appropriate thickness, or a stack of several layers, with precise control on their thicknesses. Since we are adding the microlenses as an optical component capable of improving the SNR, which already can increase the cost of

the system, we will focus only on simple anti-reflective coatings, in order to minimize the impact on the overall system cost. By simple, we mean the single-layer ARC with non-exotic materials.

Single-layer ARCs are cheaper, but also will enhance transmittance in a usually narrow wavelength range. This is because single-layer ARCs are optimized to a specific wavelength. The selection of this wavelength is controlled by the thickness of the ARC material. The reflection losses are diminished by adding a thin film layer with a value of refractive index (n_{ARC}) larger than the refractive index of the medium (n_0), and smaller than that of the substrate (n_s), i.e., $n_0 < n_{ARC} < n_s$, optimally given by their geometric mean, as follows [9], [65]:

$$n_{ARC} = \sqrt{n_0 . n_s} \quad [1] \tag{45}$$

In addition, the thin film physical thickness (t_{ARC}) should be given by the following expression:

$$t_{ARC} . n_{ARC} = \rho . \frac{\lambda}{4} \quad [\mu m] \tag{46}$$

where λ is the target wavelength given in [µm], and ρ is an odd integer number. The product of $t_{ARC} . n_{ARC}$ is called the optical thickness.

The principle behind this interference phenomenon is based on the interactions at surface boundaries. Figure 28 illustrates the concept. When a light wave refracts from one material to another, part of this wave is transmitted, and the other is reflected. The transmitted wave is always in phase with the incident wave. If traveling from an optically sparser to an optically denser material, namely from a low-refractive index to a high-index, then the reflected wave is half-wave (λ/2) phase shifted. On the other hand, when traveling from the high-index to a low-index, there is no phase shift in the reflected wave [9]. This happens due to continuity conditions and energy conservation at the boundary.

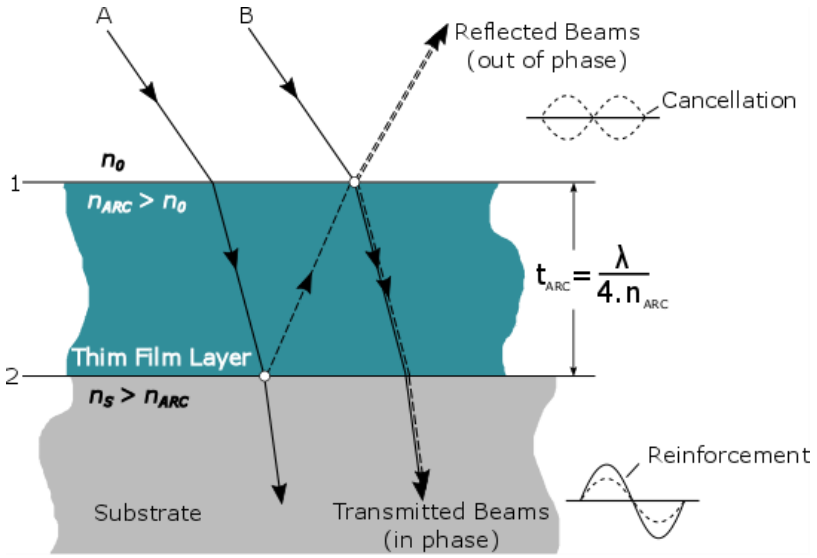

Figure 28: Interference principle behind the ARC.
Source: Adapted from [9].

From Figure 28, when beam A enters surface 1, it has say, phase shift of 0° (in phase). If we choose the optical thickness of the ARC layer to be λ/4, then the transmitted ray will reach surface 2 with λ/4 phase shift, and when reflected will have an additional λ/2 phase shifting, i.e. a total of 3λ/4. This reflected beam will reach surface 1 again with a total phase shift of λ (= 3λ/4 + λ/4), which is the same as a 0° phase shift. Beam A leaves surface 1 at the same point ray B enters this surface. In addition, beam B will also enter surface 1 with 0° – assuming a coherent light source. However, ray B will be reflected by surface 1, and the reflected beam will have a phase shift of λ/2. Both reflected rays (from beam A and B) will leave the ARC material (towards air) at the same point, but out of phase (λ and λ/2, respectively). This causes a destructive interference, being the responsible for diminishing the overall reflectance of the material around wavelengths close to λ. Furthermore, beam A will reflect again at surface 1, eventually reaching surface 2 with a phase shift of λ/4. The transmitted ray B will also reach surface 2 with λ/4 phase shift. Therefore, the summing of this rays cause a constructive interference, which reinforces transmission [9].

In fact, the optical thickness should be any odd multiple of λ/4, as described by term ρ in equation (46). This means that, if we choose the adequate optical thickness, with $\rho = 1$, for $\lambda = 4.2\ \mu m$, then the maximization in transmission will also occur for the wavelengths of $4.2/\rho\ \mu m$, like $1.4\ \mu m\ (= 4.2/3)$, $0.84\ \mu m\ (= 4.2/5)$, $0.6\ \mu m\ (= 4.2/7)$, and so on.

The problem for the ARC consists in finding the appropriate material that can be deposited over the substrate (regarding adherence, mechanical stress, etc.), with a standard technique that allows precise controlling of its thickness. In addition, we must have available the optical properties of the material, namely the refractive index (n) and the extinction coefficient (k) for the wavelength range of interest, in order to simulate and find the optimal thickness. For our purpose, materials that are toxic, radioactive or cause any hazard are excluded, even if their properties fit perfectly the application.

3. TECHNOLOGICAL ASPECTS

In this section the theoretical background presented in chapter 2 will be applied, with the purpose of designing part of an infrared optical system. These parts are the optics, comprising the microlens fabrication and its respective ARC, and the novel idea for HDR sensors. In section 3.1, the technique for microlens fabrication is presented, along with a comparison with conventional methods. Next, in section 3.2, the strategy for the ARC design is shown, listing the standard techniques and materials available at the time of the experiments. Then, section 3.3 presents the proposed pixel circuit for HDR operation, introducing its principle of operation and circuit architecture.

3.1 Single-mask silicon microlens fabrication

We have stated in the previous chapter, that the use of microlenses can be beneficial for the overall system SNR. Also, among the available materials we have chosen silicon as one that presents more advantages than drawbacks. Nevertheless, the most usual method for lens fabrication is by diamond turning, which is impractical for silicon, due to its hardness [10]. In addition, this mechanic fabrication method has a resolution lower than desirable for silicon small lenses, which can range from the *mm* up to the *μm* scale. Therefore, other types of process ought to be used, based on micromachining of silicon.

For silicon microfabrication, the state of the art techniques are: *Thermal Reflow* [69], [70]; *Defocused Lithography Exposure* [71]; and *Grayscale Lithography* [72]. All these methods initially form the desired lens shape on a photoresist – material chemically sensible to light – and then the pattern is transferred to the silicon substrate by employing *RIE* (*Reactive Ion Etching*) [73]. Figure 29 illustrates the microlenses pattern creation and its transfer. The RIE technique needs exact controlling of reactant ion gases proportions, namely SF_6 and O_2, for a correct etching of both silicon and photoresist. This control is crucial for the reproducibility of the lenses.

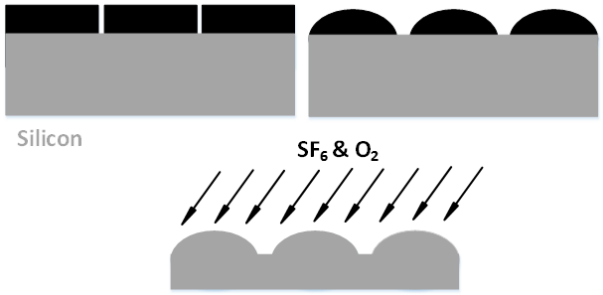

Figure 29: *Photoresist* (black) pattern transfer to *silicon* (gray) using RIE.

All three techniques abovementioned are quite sophisticated and expensive. For instance, to form the desired lens shape on photoresist, the *Thermal Reflow* technique imposes tight

control on resist baking temperature and time; the *Defocused Lithography Exposure* needs rigorous focus requirements, and tightly controlled etching time; and the *Grayscale Lithography* utilizes a very expensive mask type, and arrays are difficult to be obtained.

A technique has been proposed by *de Lima Monteiro et al.*, *2003* [67], [68] that is far less complex than the aforementioned ones. This method is based on KOH anisotropic etching of silicon, with an inexpensive single-mask process. The technique is based on the experimental evidence – first observed by *Kendall et al.*, 1988 [74], [75] – that maskless KOH etching of <100> silicon, starting from an inverted pyramidal pit, results in an almost spherical depression, as illustrated by Figure 30. The initial pyramid is formed by four <111> planes. During etching, <411> planes have an etching rate faster than that of planes <111>, until eventually they form a new pyramid. Further etching will result in a rounded surface due to the subsequent etching of planes <N11> (N > 4), which overtake each other consecutively [68].

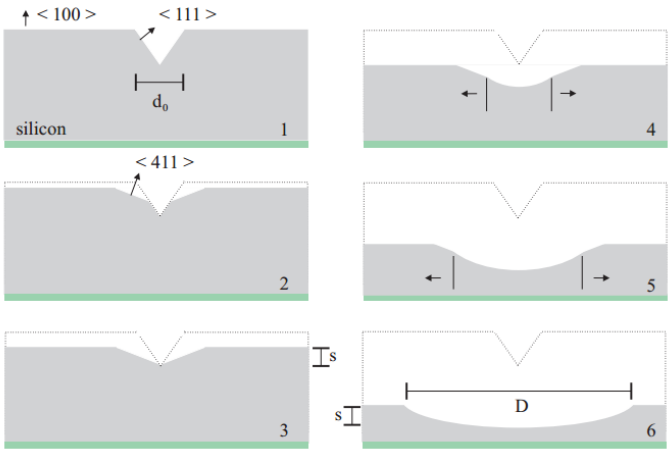

Figure 30: Maskless KOH silicon etching of an inverted pyramidal pit.
Source: [68].

Certainly, etch parameters and initial size of the pyramidal pits will determine the depth and curvature radius of the etched depression. To control the size of the pyramids, we must simply determine the hole diameters in an initial etch-resistant SiO_2 mask. Figure 31 shows a pyramidal pit formed with depth equal to $d_0/\sqrt{2}$, where d_0 is the opening diameter in the initial SiO_2 layer. This layer and hole patterning is molded through standard microelectronics photolithography, with a binary mask. Anisotropic KOH etching of these holes will form the initial pyramids. Afterwards, the oxide is removed, and silicon substrate is etched again with KOH, resulting in the spherical shapes of Figure 30.

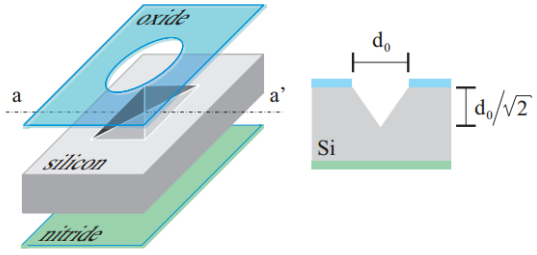

Figure 31: Initial pyramidal pit formation.
Source: [68].

The resulting spherical shapes are undoubtedly concave. However, we aim to microfabricate convex lenses. This can be done through the *ball-stamping* principle [68], which is the superposition of several concave surfaces, as demonstrated by Figure 32. Each initial pyramid has a different pit size, and the bigger it is, so will be the curvature radius of the formed concave depression. With different pits placed very close to each other, the desired convex profile can be shaped by the laterally overlapping spherical surfaces. In fact, any aspherical surface can be approximated using this ball-stamping principle.

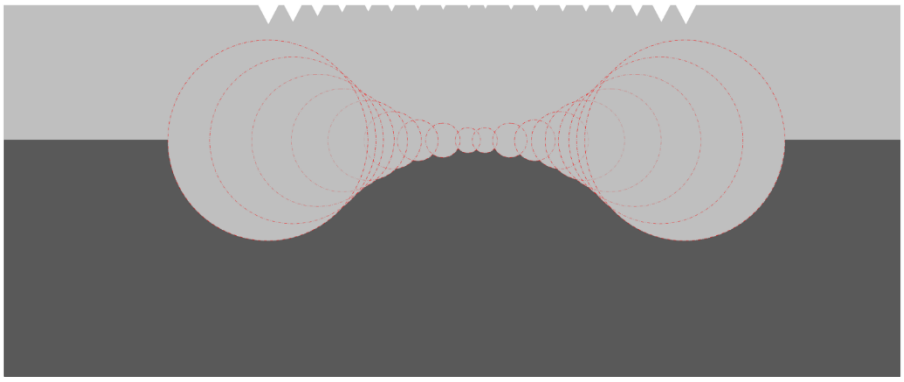

Figure 32: Ball-Stamping principle for convex lens formation: Dark grey represents final silicon structure, and light grey the etched silicon.

A computer program developed by *de Lima Monteiro* generates the initial mask, containing all holes that will be transferred to the oxide layer during photolithography, for a given input surface profile. In this thesis, we have used the pattern of holes shown in Figure 33. This mask was designed for a small 3x3 microlens array, with sagittae equal to approximately 11 µm, and a 1 mm microlens pitch.

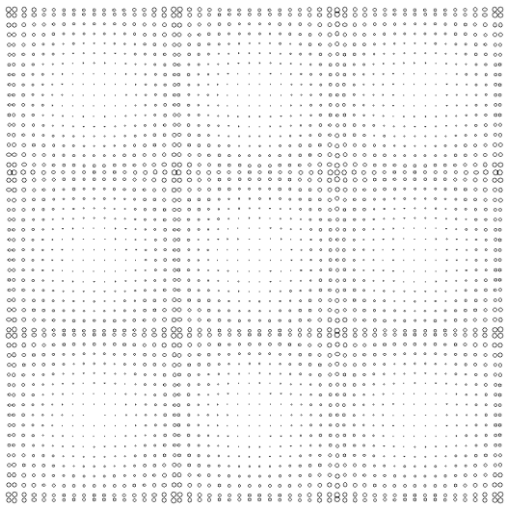

Figure 33: Mask design with pattern of holes for a 3x3 microlens array.

Since it needs just a single-mask, this technique is cheaper than the usual methods for microlens fabrication. It uses KOH aqueous solution, a reliable and standard solution in bulk micromachining of silicon. In addition, it is not needed to form the lens shape first on photoresist, which eliminates the RIE step. This makes the process cheaper and more reproducible, noting that it does not depend on the used photoresist etching ratio – important

for RIE. Furthermore, the fitting error between the final surface and the desired one can be alleviated with a denser initial mask, i.e., with more holes to represent the same surface, being only limited by the photolithographic resolution. Lastly, since there is ample flexibility for different shapes, a 100 % fill-factor microlens array is made possible.

3.2 Single-layer ARC Design

We have stated in last chapter, that silicon has a high surface reflectance, and using anti-reflective coatings is a must. However, it is not easy to find the suitable materials for the ARC, mainly due to lack of information on optical properties for specific fabrication processes. It is relatively easy to access the optical properties of standard thin-film materials for the visible spectrum, but harder for the infrared. Also, optical constants may vary depending on the thin-film deposition technique and the process parameters [19], [76].

These properties are the refractive index (n) and the extinction coefficient (k) [19], [77]. Both of them are wavelength-dependent, and extrapolation of these values from the visible to the infrared might generate unreliable results. Furthermore, there are some restrictions to the $(n_{ARC},\ k_{ARC})$ dataset to be used: the refractive index n_{ARC} should be as close as possible to that calculated by equation (45), considering the medium as air, and substrate as silicon; and the extinction coefficient k_{ARC} should be negligible, so that the material is transparent at the aimed wavelengths. In fact, there is a difference between the extinction coefficient (k), and the absorption coefficient (α) presented in equation (3), being their relation as follows: $\alpha = 4\pi k/\lambda$.

Since our main goal here is to find the simplest ARC design, but yet an effective one, the employed thin-film deposition technique needs to be less expensive, reliable, and the material properties should be well-known. Therefore, standard materials should be used, i.e., common thin-film materials that can be deposited with standard microelectronics techniques. This is extremely important because the wavelength for which transmission is optimized is heavily dependent upon the deposited material thickness (t_{ARC}), as demonstrated by equation (46). Subsequently, to prove that a design works, it must be actually fabricated and measured, and for that matter standard processes are more convenient. The thin-film depositions presented along this thesis were performed in a *Clean Room* environment, at the *Else Kooi Laboratory* – former *Dimes Technology Centre* – in Delft, Netherlands, in 2013.

Table 4: Standard deposition techniques and materials available at Else Kooi Laboratory.

Deposition Technique	Available Materials
LPCVD	SiO_2, SiO, SiC, SiN, Si, poly Si
PECVD	SiO_2, SiO, SiC, SiN
Epitaxial Growth	Si, poly Si, SiGe
ALD	TiO_2, TiN, Al_2O_3, HfO_2

Table 4 shows the limited set of available test materials and techniques accessible at the time of this research, where *LPCVD* stands for *Low-Pressure Chemical Vapor Deposition*, *PECVD* for *Plasma-Enhanced Chemical Vapor Deposition*, and *ALD* for *Atomic Layer Deposition* [73].

Some materials can be deposited using different methods, but as previously observed, their optical properties may diverge.

There is the possibility that one of these materials could fit perfectly the aimed application for infrared. But there are also cases in which just a single material layer simply cannot meet the specifications to work as ARC, with significant transmission at the desired wavelengths. In these cases, a stack of materials needs to be used, usually featuring a *HLH* or *LHL* configuration, where *H* and *L* stand for layers with higher and lower refractive indices, respectively [78], [79]. These multilayer stacks can be interpreted as equivalent metamaterial layers, used to overcome the limitation of not having the appropriate material.

A multilayer approach may result in a more complex design, and therefore resulting in a more expensive final structure, where the error in the deposited thickness of each layer (compared to the designed ones), when combined, may result in a final transmission band very different than the designed one. Nevertheless, a multilayer ARC generally has a wider transmission band, which is not the purpose of this thesis. The main goal is to find the simplest design possible that can be easily manufactured. Yet, regardless of the design approach, before actual fabrication, we must first simulate the proposed coating using proper software. Simulation is vital, since it reduces the cost of the whole process. Fabrication of the test structures is expensive, time consuming, and should be performed only when satisfactory simulation results were achieved.

Beyond the restricted or difficult access to material optical properties at the IR spectrum – and with conventional deposition techniques – there is also lack of data for this wavelength range in existing software. Therefore, to be able to start designing an ARC, the chosen strategy was to identify an appropriate software, list the materials in its database, check which of them are standard ones (contained in Table 4), and verify if their optical properties cover the aimed spectral range. Unfortunately, it is quite difficult to meet all of these constraints.

The selected software is called *TFCalc™* – supplied by Software Spectra, Inc., USA – a very simple to use, and yet powerful program, specially developed for optical calculations [79], [80]. For the purpose of this thesis, the anti-reflective coating is deposited on both sides of the silicon substrate, such as in Figure 34. This is because the incoming IR radiation is considered to originate from the incident medium, whereas the photodetector is placed in the exit medium, behind the substrate.

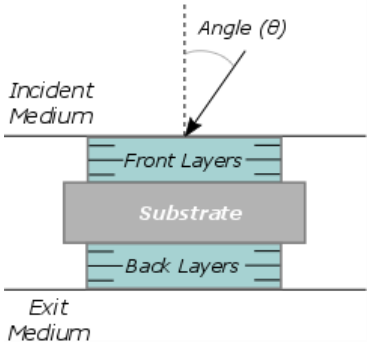

Figure 34: Optical Simulation Setup.
Source: Adapted from [80].

To use TFCalc™, we need to set up the simulation environment, as in Figure 34, setting the silicon substrate thickness (525 µm for a bare 4-inch wafer), incident and exit medium materials (air), and the ARC material, symbolized as the front and back layers. Additionally, the aimed wavelength has to be set. The software runs optimization routines in order to find the best thickness of the ARC material to provide maximum transmission at the particular wavelength.

We have chosen three different target wavelengths for the ARC designs, namely 4.2 µm, 8.2 µm, and 9.3 µm. Enhancing transmission in these wavelengths is suitable for applications like gas analysis, radiation thermometers, thermal imaging, remote sensing, spectrophotometry, flame monitors, fire detectors, optical power meters, among others [81]. In addition, this project was done in the context of the *INCT-DISSE* (*Brazilian National Institute of Science and Technology for Semiconductor Nanodevices*) [12], whose main goal is to develop infrared photodetectors. All the three wavelengths mentioned above are of main interest to this institute.

Although TFCalc™ has a wide range of materials in its database, almost none covered the IR spectrum at these wavelengths, and preliminary simulations shown that none of them yielded satisfactory results. Most likely due to the lack of IR data. In these cases, the software automatically extrapolates the available data points to the wavelengths of interest, which may or may not be a good approximation. Only after actual fabrication and measurement of the designed samples, we can support simulated data. This is a very time consuming and expensive approach.

Fortunately, we could get some different materials optical properties with colleagues at the *Else Kooi Laboratory*, however, without knowing the exact process techniques and conditions. One of the materials presented real promising results: the *SiN*. With just a single-layer of SiN, we could find three different designs that could enhance the transmittance at each target wavelength (4.2 µm, 8.2 µm, and 9.3 µm) to above 90%. Each design was found by setting the target wavelength in TFCalc™, and running an optimization algorithm to maximize transmittance at this wavelength. The optimization tool finds the best thickness of the ARC so that transmittance is the highest. We also have set the reference wavelength as the target wavelength. This reference is used to find the initial thickness guess, based on equations (45) and (46), for $\rho = 1$. Best preliminary designs are summarized in Table 5.

Table 5: Best SiN designs obtained from preliminary simulations.

Target Wavelength	SiN thin-film thickness
4.2 µm	488 nm
8.2 µm	958 nm
9.3 µm	1082 nm

Results shown in Table 5 were based on a no specified thin-film process technique – assumedly one of Table 4 – and optical constants confined to the range from 250 nm up to 6 µm. Therefore, it cannot be totally trusted, at least for $\lambda > 6\,\mu m$. Yet, it is a starting point. From this point onwards, the best strategy would be to test (deposit & measure) the designs listed in Table 5, mainly for 4.2 µm, which falls on the range where there were optical data, with no need for data interpolation.

Next chapter presents simulated and measured data for SiN thin-films over silicon substrate, using the *LPCVD-deposited low-stress Si-rich nitride* technique [19], [82]. Optical transmittance measurements were conducted with *Fourier Transform Infrared* (*FTIR*) spectroscopy [83]. The LPCVD SiN were deposited onto a bare silicon wafer, and substrates previously micromachined with the microlenses shown in previous subsection.

3.3 The Bouncing Pixel

We have covered so far in this chapter the optics portion of the photodetection system. We have presented the silicon microlenses and a non-expensive way to microfabricate them, along with a methodology and standard techniques for the anti-reflective coating design and deposition. The other portion of the system comprises the read-out circuitry, as addressed in previous chapter, where in section 2.8 we have discussed why HDR sensors offer advantages for IR photodetection systems. Common HDR topologies were presented along with their weaknesses – mainly poor sensitivity, and nonlinear response, as presented in Table 3. Furthermore, to implement HDR operation, we should use one of the standard pixel circuit topologies for IR outlined thus far – the SFD, DI, BDI, GMI, CTIA, or CMI – and find a way on how to modify them for HDR operation.

Assuming we were not concerned with HDR operation, a pixel circuit topology that could overcome all limitations of the HDR sensors – enumerated at the end of subsection 2.8.3 – would be a pixel circuit with a linear response, and constant transfer function for the whole DR, i.e., a constant sensitivity, or transimpedance gain (for a voltage-mode pixel). From the topologies of section 2.7, either the BDI, CTIA or CMI topologies could be employed, since their injection efficiency ε can be made close to one, and charge-to-voltage conversion gain – related to the sensitivity – can be held fairly constant. The CMI pixel circuit however, is the only topology that provides both high injection efficiency (due to low input impedance R_{in}) and current gain capability, making the response not only fairly linear, but with a high conversion gain, providing a noise factor F close to one, and good sensitivity.

Besides offering advantages regarding IR photodetection, the CMI also overcomes many of the HDR sensors weaknesses, shown in Table 3. Unquestionably, it still features a very limited DR, especially when higher sensitivities are employed. Therefore, the question that arises is: how do we alter the CMI pixel circuit with the purpose of enhancing the DR, and still maintain minimal negative impact?

From equation (37), we know that increasing V_{DD} leads to a higher DR. If it was possible to increase the voltage swing to a very large extent, we could achieve our goal with the CMI pixel circuit. In practice, however, enhancing the voltage swing while conserving a constant integration time is not normally done as a HDR approach [36], [45], since standard integrated circuit technologies preclude any considerable increase in the – usually low – recommended chip voltage supply [25].

Nevertheless, we can emulate the voltage swing extension. We have envisioned a pixel circuit topology based on the CMI architecture, herein coined as the *Bouncing Pixel*[2], which operates with very high DR and still benefits from almost all advantages of the CMI [31], [84].

3.3.1 Principle of Operation

To help explaining, the case in Figure 4 must be recalled, where the largest signal integrates upwards until it reaches saturation. The idea of the bouncing pixel is to ensure that the output voltage across the capacitor (V_{CAP}) will never saturate. Instead, the signal bounces at a predefined reference upper voltage (V_{max}), and continues to integrate downwards, changing its direction but with the same absolute slope as before, until it eventually reaches a predefined lower reference voltage (V_{min}). Then, the signal re-bounces, and this process repeats itself whenever the signal would saturate within the stipulated voltage-difference boundaries ($\Delta V = V_{max} - V_{min}$).

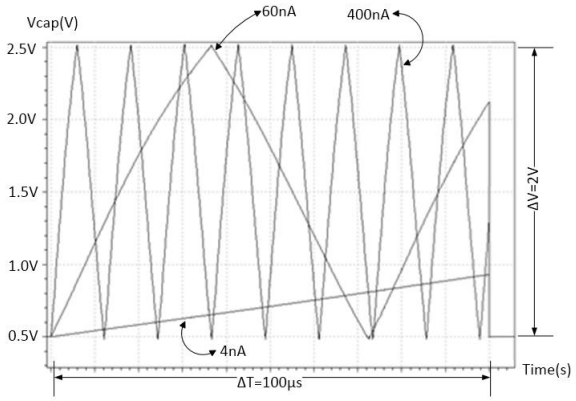

Figure 35: Integrated Output Voltage across the capacitor for three different photocurrents. *Source:* [31].

Figure 35 illustrates the concept, for three different input photocurrents, and a CMOS technology with $V_{DD} = 3.3\ V$ [31]. During simulations presented in chapter 4, a technology with an even smaller voltage will be presented: $V_{DD} = 1.8\ V$. When the photocurrent is $I_1 = 4\ nA$, it integrates as usual, and the output voltage V_{CAP} is read at the end of the integration time. However, for $I_2 = 60\ nA$, a photocurrent whose voltage signal would normally be saturated for this case, within the integration time of 100 µs, the bouncing effect is demonstrated, and information is not lost. Also for a much higher photocurrent, $I_3 = 400\ nA$, signal does not saturate yet. However, to reconstruct the HDR signal, only information about the integrated voltage V_{CAP} does not suffice.

Therefore, the reconstructed signal (V_R) comprises both the integrated voltage across the capacitor V_{CAP}, and the number of bounces (N), i.e. how many times the integration curve changed direction. The first is an analog signal, and is read at the end of the integration time.

[2] In Portuguese: *Pixel Ricochete*.

The latter is a digital signal. Combining both of them, HDR operation can be achieved. For the general case, we have that the reconstructed signal V_R is given by:

$$\begin{cases} V_R = N.\Delta V + (V_{CAP} - V_{min}) \ \ [V]: \quad for\ N\ even \\ V_R = N.\Delta V + (V_{max} - V_{CAP}) \ \ [V]: \quad for\ N\ odd \end{cases} \tag{47}$$

Assume that the time spent by the signal to integrate upwards from V_{min} to V_{max} is given by t_{up}, and downwards from V_{max} to V_{min} by t_{down}. Also, if the slope does not vary along the integration time, meaning that the input photocurrent I_{ph} is constant, then $t_{up} = t_{down} \equiv t_0$.

Similarly, the time spent from the last re-bounce until the end of the integration time can be defined as t_N. If no bounce occurs, then it operates as a regular linear pixel circuit. We can rewrite equation (47), for N even, as:

$$V_R = N.\left(\frac{I_{ph}.t_0}{q}.G_{QV}\right) + \left(\frac{I_{ph}.t_N}{q}.G_{QV}\right) = \left(\frac{I_{ph}}{q}.G_{QV}\right).(N.t_0 + t_N) \ \ [V] \tag{48}$$

where the terms ΔV and $(V_{CAP} - V_{min})$ were rewritten taking equation (18) into account. In addition, we can see that:

$$\Delta T = N.t_0 + t_N \ \ [s] \tag{49}$$

in which ΔT is the only parameter that can be actually changed externally, after fabrication of the pixel circuit, where N, t_0 and t_N change depending on the input signal intensity.

From equations (48) and (49), we can conclude:

$$V_R = \frac{I_{ph}.\Delta T}{q}.G_{QV} \ \ [V] \tag{50}$$

Although we have concluded that for N even, it is also easy to demonstrate the same for N odd. Equation (50) is exactly the same as for a linear pixel circuit. Therefore, besides featuring a high DR, it also offers a linear response for the whole DR, without the need to compress the response, which would render in decreased sensitivity. Moreover, recalling the result of equation (38) for the sensitivity ($S = \beta.\varepsilon.\Delta T/C_{int}$), we can assume that it will also be held constant for the whole DR, if none of its parameters change. Additionally, even with smaller integration capacitances, suitable for high sensitivities, and no current gain, a high DR can be achieved.

A convenient parameter used to calculate how a pixel circuit topology would extend the DR is called the Dynamic Range extension Factor (DRF) [6], and it is usually used to compare between different HDR topologies. The DRF is calculated by the ratio (in *dB*) between the largest measurable photocurrent by the pixel circuit (I_{N_max}), and the largest signal that could be read by the linear pixel if there was no bouncing feature (I_1). In other words, I_{N_max} is the photocurrent corresponding to the last re-bounce that results in the counter saturation, and I_1 corresponds to the current when the first re-bounce occurs. The number of bounces is calculated by a digital counter with b_c bits, and the maximum count is $N_{max} = 2^{b_c} - 1$, where

the sub-index "c" in b_c relates to the counter. The largest measurable value occurs when the counter overflows, namely, when $N \rightarrow 2^{b_c}$. Then, from equation (50):

$$I_{N\_max} = V_{R\_max}\left(\frac{q}{\Delta T.\, G_{QV}}\right) = \left(2^{b_c}.\Delta V\right)\left(\frac{q}{\Delta T.\, G_{QV}}\right) \quad [A] \qquad (51)$$

where the maximum value of V_R is given by ΔV times the number bounces when the counter overflows $\left(2^{b_c}\right)$. Likewise, for the first re-bounce $(N \rightarrow 1)$, we can predict I_1 by:

$$I_1 = \Delta V\left(\frac{q}{\Delta T.\, G_{QV}}\right) \quad [A] \qquad (52)$$

According to equations (51) and (52), we can calculate the DRF:

$$DRF = 20 * \log\left(\frac{I_{N\_max}}{I_1}\right) = 20 * \log\left(2^{b_c}\right) = b_c.\, 20 * \log(2) \cong 6.02 * b_c \quad [db] \qquad (53)$$

from where we can notice that the DRF can be linearly extended depending on the number of bits of the counter. Basically, the bouncing pixel achieves HDR operation by extending the maximum operating voltage (V_{DD}) to a virtual voltage of $V_{R\_max} = 2^{b_c}.\Delta V$.

Lastly, to calculate the final DR of the system one does simply have to sum the conventional linear pixel circuit DR plus the DRF:

$$DR_{final} = DR_0 + DRF \quad [db] \qquad (54)$$

in which DR_0 is the dynamic range of the pixel without bouncing. If we use an ADC of b_a bits (where the sub-index "a" relates to the ADC), then, the maximal DR of the system would be given by:

$$DR_{final\_max} = DR_{0\_max} + DRF \cong 6.02 * (b_a + b_c) \quad [db] \qquad (55)$$

where the ADC number of bits limits the DR_0, since all this range is measured when no bounce has occurred, through the analog signal V_{CAP}. Taking the example of section 2.6, with an ADC of $b_a = 16\ bits$ and a counter of $b_c = 8\ bits$, a maximum dynamic range of $DR_{final} \cong 144\ dB$ could be achieved. This is a much higher DR, compared to the $DR_{0\_max} = 96\ dB$ for a linear pixel without the bouncing feature, and $b_a = 16$ bits. In this case, the counter provided $DRF = 48\ dB$, that could be extended with more bits in the counter. Equation (54) is true when the sum of both dark and background currents $(I_B + I_{dark})$ are small compared to the maximum current signal that can be read by the pixel circuit. This might not be correct for IR detectors, and in these cases the actual DR_{final} has to be calculated through using equation (37). For the bouncing pixel however, equation (37) has to be rewritten accounting for the voltage extension, i.e., using $V_{DD} = 2^{b_c}.\Delta V$, which makes DR_{final} as calculated by equation (54) as a good approximation.

3.3.2 Pixel Circuit Architecture

We have developed a pixel circuit based on the CMI topology, in order to implement the bouncing pixel idea. Figure 36 shows the pixel *unit cell*. Although we have focused so far in the case where a single pixel interfaces with its respective photodetector, it can be arranged in an array fashion, wherein more than one pixel can work together. This array could be a linear array, or a matrix, and its size depends on the application. Therefore, for the sake of generality, we will herein show the case where an array is employed, and in this case, the unit cell is the one that should be replicated, as will soon become clear. For a single pixel operation, we must simply employ a single unit cell.

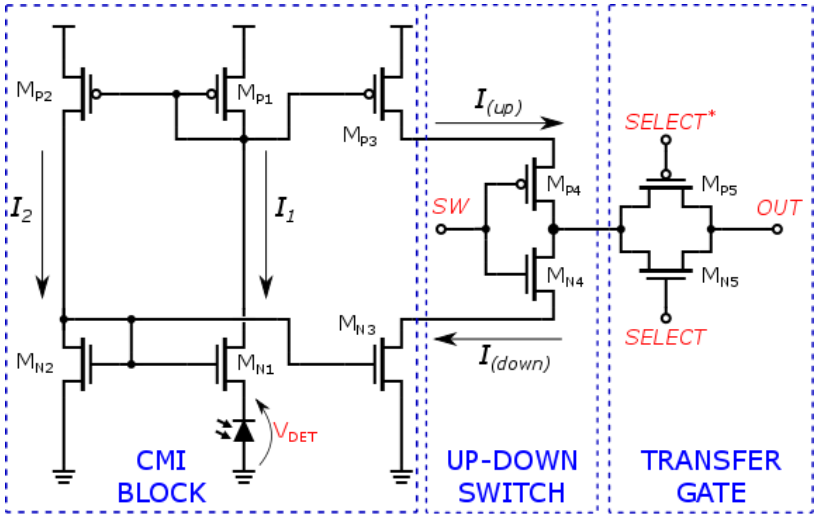

Figure 36: CMI based pixel circuit unit cell.

The unit cell is composed by 10 transistors, and can be subdivided in three different regions: the *CMI block*, the *up-down switch*, and the *Transfer Gate*. The CMI block is almost the same as from Figure 24, but without the integration capacitor, and with the addition of an extra NMOS mirror transistor M_{N3}. It is important to note that transistors M_{P3} and M_{N3} mirror the injected photocurrent $\left(I_1 = I_2 = \varepsilon . I_{ph}\right)$ in a different direction in respect to the next block, respectively, $I_{(up)}$ and $I_{(down)}$. In addition, note that $I_{(up)} = I_{(down)} = \beta . \varepsilon . I_{ph}$. Next, we have the up-down switch, which is built like an inverter and is responsible to choose whether $I_{(up)}$ or $I_{(down)}$ are fed to the output of the unit cell. Signal *SW* is responsible for choosing one of the currents: when SW is logic high, $I_{(down)}$ is selected; and with SW low, $I_{(up)}$ is chosen. Finally, the Transfer Gate (TG) block works as an analog switch, and is important for offering *sample & hold* capability. The TG also permits array applications, in which the signal *SELECT* is used to take the response of a specific pixel circuit in the array. At the end, the generated current at the unit cell output is fed to the output integration capacitor, namely C_{int}, located outside of the unit cell, and shared by pixels in an array.

Figure 37 shows the circuit external to the unit cell, i.e., the circuit shared by pixels in an array. It consists of the *integration capacitor*, two *Comparators*, one *Counter*, one *Shift Register*, and additional logic circuitry. Since the SELECT signal allows only one unit cell to operate at a time, the current integrated in the external capacitor C_{int} is derived only from the selected pixel

during the integration time, while all others are “turned off”. All the process for data acquisition herein described must be repeated for the subsequent pixels.

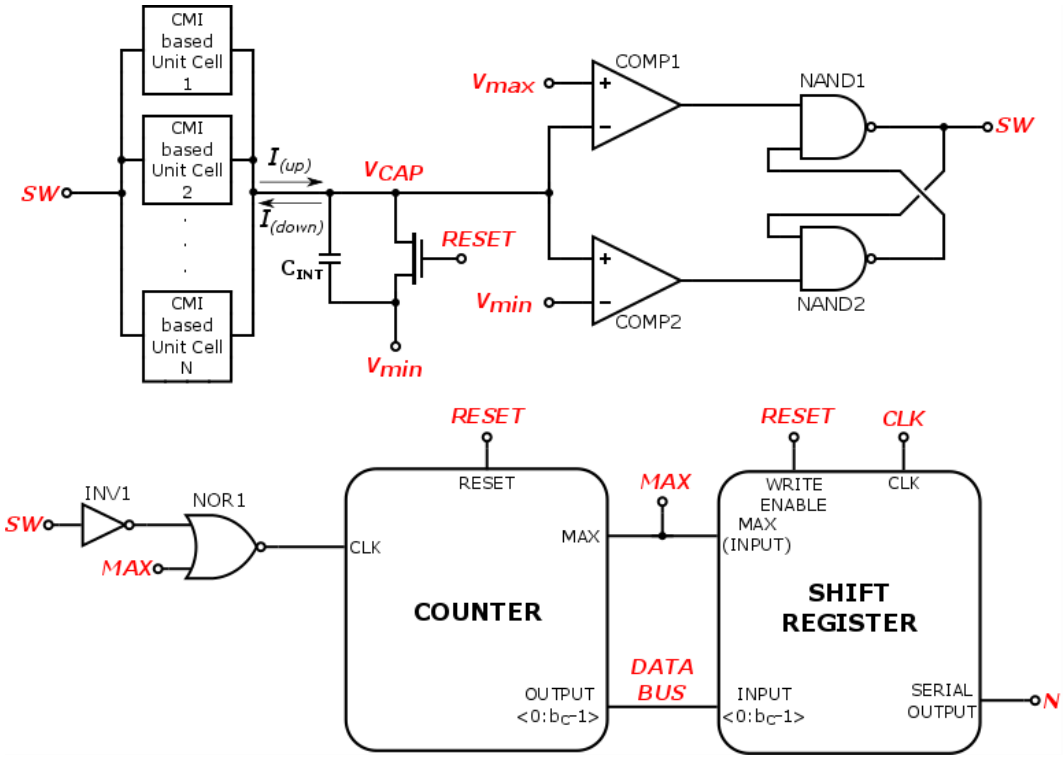

Figure 37: The array shared circuit.

The signal that manages the pixel operation is the RESET. When this signal is logic high, the capacitor C_{int} is short-circuited, discharging any previously accumulated charge, setting the node V_{CAP} to a voltage equals to V_{min} – the starting point for integration – and SW is set to low. When RESET is toggled to logic low, then the integration time starts. At this point, current fed from the unit cell, namely $I_{(\mathrm{up})}$, will start to increase the voltage V_{CAP}, until it reaches V_{max}, just as explained in section 3.3.1, when the first comparator (COMP1) fires a signal that toggles SW from logic low to high. This will change the integration direction, now choosing current $I_{(\mathrm{down})}$ for integration by the up-down switch. Eventually, the second comparator (COMP2) may be triggered, if V_{CAP} reaches V_{min}, toggling SW once again. The latch composed by the NAND1 and NAND2 gates are responsible for the generation of the signal SW, depending on the outputs of the comparators.

The integration time ΔT is defined as the time interval in which RESET is set to low and SELECT is high, because the transfer gate is feeding the unit cell current to the capacitor, which in turn is free to integrate. Figure 38 shows the temporal signal diagram for the bouncing pixel. During ΔT, the Counter digital circuit will count at every edge of the squared signal SW (not only at the rising or falling edge), meaning every time the signal V_{CAP} reached either of the reference voltages, V_{max} or V_{min}. Therefore, at the end of ΔT, the binary value stored in the counter is exactly the desired number of bounces *N*, required for the reconstructed signal V_R in equation (47). A counter with b_c $bits$ can eventually saturate, i.e., when $N \rightarrow 2^{b_c}$, and the circuit must recognize whether it has saturated or not. Otherwise, it would not be possible to distinguish between a very high photocurrent and a very low, since the counter would start counting over again. For that matter, the implemented counter has actually a number of bits equal to $b_c + 1$.

The additional bit, namely MAX, accounts for the counter overflow (if MAX = 1). This signal works like a flag, and tells the counter to stop counting through a small logic at the *clk* input of the Counter. Note that this logic could be integrated in the counter itself, but was put outside, in diagram of Figure 37 for clarity.

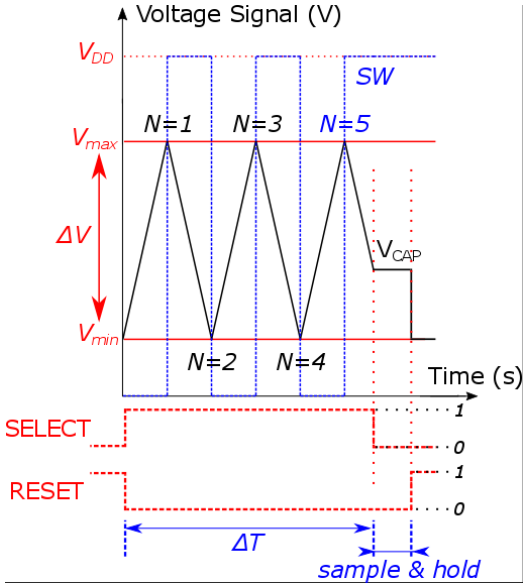

Figure 38: Temporal bouncing pixel signal diagram.

Now, we have two choices concerning the Shift Register block: *A)* We could implement a complete system in a single CMOS chip, that is, all signal processing encompassing the pixel circuits themselves, the ADC blocks, processing unit, memory, etc. Then, in this case the Shift Register would not be necessary, since the output data bus from the Counter could be directly accessed by the internal processing unit. Or option *B)*, where signal processing would be done outside of the chip where the pixel circuit was implemented, using for instance a Microcontroller, a DSP or a FPGA, which is the most general case. For that situation, the Shift Register becomes necessary because it simplifies the interface of the pixel circuit with the external units, since only one signal would be necessary, namely the serial output of the Shift Register. Hereafter, we have chosen the second option, as shown in Figure 37.

As the very name states, the Shift Register shifts its input bits one by one in a stream fashion. When RESET toggles from low to high, the Shift Register writes out the counted data bus value to its internal registers. Right away, it starts to shift the stored value at every rising edge of its input clock (*CLK*). This clock must be provided by the external control unit, and is different from the internal SW signal. At the beginning of the streaming, a start signal is sent, composed of two bits 10_2 (base-2). This start signal is used to create a read protocol between the chip and its control unit. The Shift Register has the same number of bits that the Counter does $(b_c + 1)$. After the start signal, the following $b_c + 1\ bits$ are streamed out, beginning with MAX (the overflow bit), and ending with the *Least Significant Bit* (*LSB*). For instance, for $b_c = 8$, and if the serial output is the binary value 10011111100_2, then, the first two (10_2) would be the start signal, the third (0_2) the MAX bit, and the following eight (11111100_2) could be translated as a decimal 252_{10} (base-10). Lastly, it is important to state that the RESET must be kept at logic high at least longer than $(b_c + 3)$ clock cycles. This communication protocol was used for simplicity, but in a future work, we can implement a standard protocol, like the SPI, or the I^2C.

4. SIMULATIONS AND RESULTS

Last chapter was devoted to present solutions for the problems discussed in chapter 2. We have presented the microlenses as an option for compact systems, using silicon as substrate, a material suitable for the infrared. Also, a non-expensive technique for silicon micromachining – compared to the usual microfabrication processes – was selected and explained. However, since silicon has a high reflectance, it is supposed to have an anti-reflective coating on its surface, optimized for the desired wavelengths. This way, we have proposed to use standard materials and deposition techniques for the ARC layer, and the simulation method for its optimization was introduced. Further, a pixel circuit for HDR operation was presented – appropriate for the infrared with high background and dark currents – along with its principle of operation and architecture.

This chapter aims at bringing all parts of the system together, for a specific application in the infrared spectra. Following an application driven design flow facilitates the elucidation concerning the proposed solutions. Therefore, we have chosen a CO_2 gas detection system, with absorption wavelength at 4.2 µm, as the target application. This is one of the wavelengths mentioned in last chapter, which corresponds to the detectors developed within the INCT-DISSE network [12]. The microlenses were fabricated and characterized, while the ARC layer was optimized for this wavelength. A commercial detector was chosen, and its electrical characteristics were used for the simulation of the bouncing pixel. A CO_2 gas detection system can be used in oil refineries, in exhaust pipe monitoring, for petrochemical industry, for combustion efficiency monitoring, for generators and motors monitoring, among others. The microfabrication of the microlenses and all thin-film depositions presented in this work were performed in a Clean Room environment, at the Else Kooi Laboratory – former Dimes Technology Centre – in Delft, Netherlands.

4.1 Microlenses fabrication and characterization

We have stated in previous chapter that the use of microlenses can be beneficial to improve the SNR of the system, especially when the detectors feature a low fill factor, which is generally the case. We have replicated the pattern mask shown in Figure 33 all over a 4-inch silicon wafer (diameter of about 10 cm). This wafer has a thickness of 525 µm, being thick enough to perform all the needed etchings to form the final patterns, and still present good mechanical stability. Additionally, the silicon wafer starts with a grown silicon oxide layer over its surface, which is useful after the lithography step, as will become clear later. Besides, a silicon nitride layer should be deposited on the backside of the wafer to protect this side from KOH and especially HF etchings.

The mask was designed by *de Lima Monteiro* for a 3x3 microlens array, with sagittae of 11 µm, and pitch of 1 mm. The mask pattern was first transferred to a photoresist (PR) layer deposited all over the wafer, through photolithography. Figure 39 shows microscope images of

the mask pattern of holes. Note that the smallest hole, and the density of holes is purely defined by limitations in the lithographic technology. Therefore, the smoothness of the final lens shape can be improved by using lithography with smaller feature sizes, where smaller and closer apertures can be patterned.

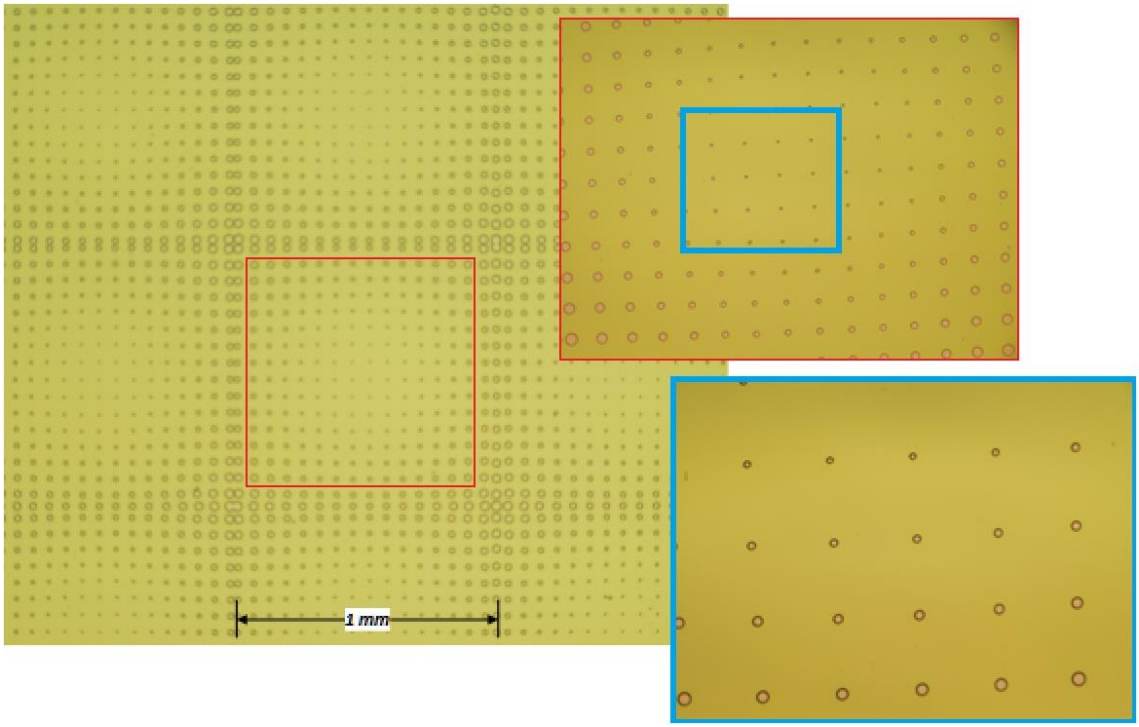

Figure 39: Microscope images of mask pattern on photoresist.
Source: [67].

After the patterning of the initial holes on the photoresist by lithography, a BHF (HF Buffer)[3] is performed in order to remove the oxide where there are apertures in the PR layer. This creates a window that exposes the silicon substrate in the same positions of the initial holes. The substrate is protected elsewhere. Afterwards, the first KOH etching is performed, anisotropically etching the exposed silicon, and forming the inverted pyramidal pits, like presented in section 3.1. The KOH also removes the PR. The time needed for etching is defined by the largest hole, which guarantees that all pyramids are correctly formed. This is due to the KOH anisotropic etching, that is, the etching ratio of the <100> plane is much higher than in the <111> plane (which form the pyramid walls). Therefore, when a pyramid is formed, since the <100> planes are no longer being exposed to KOH, then it automatically acts as etch stop. Figure 40 shows the inverted pyramids on a microscope, at different focal planes. In the clockwise direction, starting from the upper left image, the focal plane deepens ranging from the bases of the pyramids until their upside-down apexes. There is also a rectangular base in Figure 40, instead of a squared one, which did form a “roof-like” shape, rather than a pyramid. This is due to two operalapping holes in the initial mask, or due to underetching of two adjacent apertures that were too close.

[3] A buffered HF solution is commonly used in microfabrication, in order to perform a more controlled etching process [85].

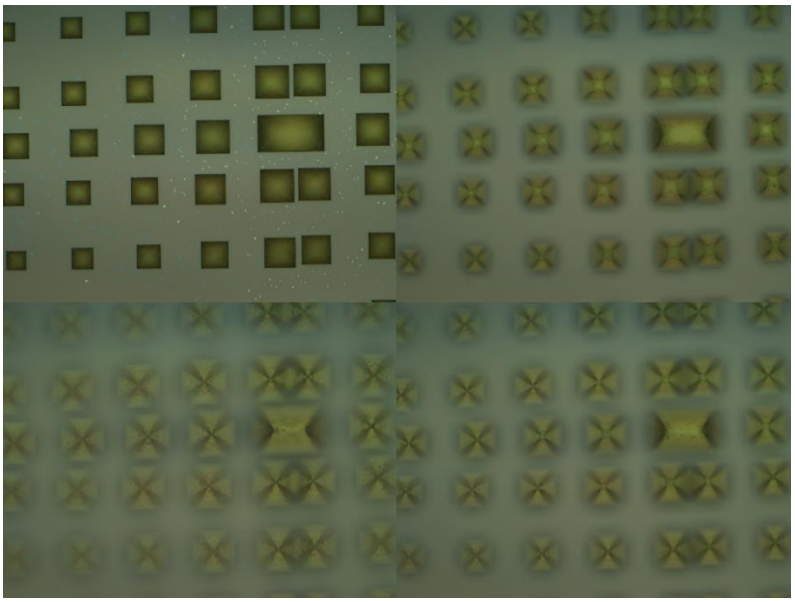

Figure 40: Microscope images of pyramids formation after first KOH etching. Focal plane at different depths.
Source: [67].

Then, the silicon oxide mask is entirely removed with the BHF, exposing all the silicon substrate. Thereafter, a maskless KOH etching was carried out to finish the etching process. This final etching lasts for a time interval of $t_{\mu lenses} = h/R_{<100>}$, where $R_{<100>}$ is the KOH etching ratio in the silicon <100> (normal) plane, and h is the etching depth. This last term represents the desired etching depth to achieve the target microlens shape, and $t_{\mu lenses}$ is calculated based on that. In fact, the KOH etching is performed with an aqueous KOH solution (KOH:H_2O) with specific KOH percentage in weight. Term $R_{<100>}$ depends on temperature and this weight percentage. Therefore, for a given KOH solution, both time and temperature must be controlled to achieve the right etching depth h. Figure 41 shows the final 4-inch silicon wafer produced at DIMES (TU-Delft), with the 3x3 microlenses array mask replicated all over its surface.

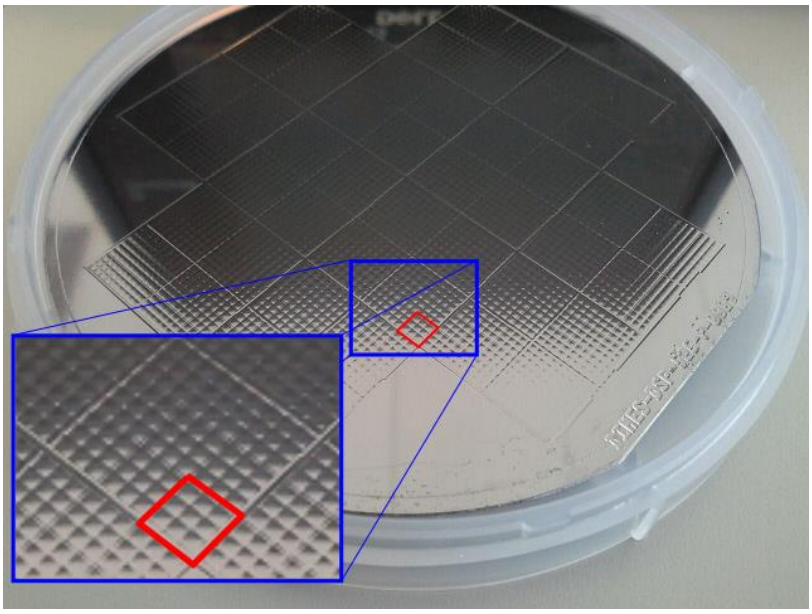

Figure 41: A 4-inch wafer with the microlenses pattern. A highlight is given at the 3x3 pattern.

For this case, the etching depth was designed to be $h = 100\ \mu m$, rendering in a final silicon substrate thickness of $t = 325\ \mu m$, since the backside was not protected from etching, that is, about $100\ \mu m$ were removed from both surfaces from the initial $525\ \mu m$ wafer. From Figure 41, we can notice that the wafer is subdivided into 48 big squares. These are referred as *dies*, and have a width approximately equal to 10 mm for this process. Lithography was performed in a *stepper* machine, exposing each die area subsequently, one die at a time. The exposed area is, therefore, equal to the die area, which is bigger than the original 3x3 microlens pattern. Then, the actual mask has 9 arrays of 3x3 microlenses in order to fill all the die area. Summarizing, each die has a total of 81 microlenses, and all of them have exactly the same pattern.

After microfabrication, the lenses had to be characterized in order to confirm their shapes and functionality as actual lenses. First, lens shape and roughness were measured using a Zygo® Optical Profilometer, and measurements analyses and plots were performed using a software called MetroPro® (from Zygo®). Figure 42 shows a 3D plot of one of the 3x3 microlens arrays on the wafer. Note that the X and Y axes – accounting for the length and width – are in a different scale compared to the Z axis – related to the lens height.

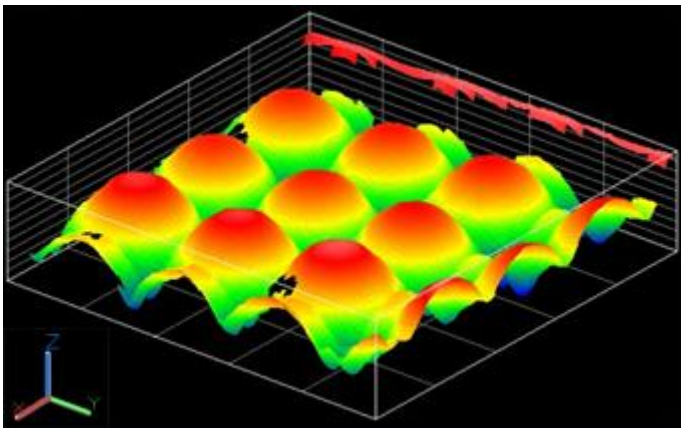

Figure 42: Optical Profilometry: 3D Plot of a 3x3 microlens array.
Source: [67].

As mentioned earlier, there is an intrinsic roughness associated with the fabrication process, and it is important that it does not exceed a minimum acceptable RMS roughness, that results in an appropriate lenslet surface shape, for the given wavelength. In other words, to consider the surface as smooth, its RMS roughness has to meet the following criterion [85]–[87]:

$$RMS < \frac{\lambda\ [\mu m]}{RF.\cos\theta}\quad [\mu m] \qquad (56)$$

where λ is the wavelength, being 4.2 μm for this case, θ is the angle of incidence with respect to the normal, and RF is a value that can be $RF = 8$ for the Rayleigh criterion, or $RF = 32$ for the Fraunhofer – a more stringent – criterion. For the target wavelength, if we consider the most rigorous case in equation (56), namely for the Fraunhofer criterion with normal incidence, then the maximum roughness should be $RMS < 4.2/32 \cong 0.131\ \mu m$. For the Rayleigh criterion it is given by $RMS < 4.2/8 \cong 0.525\ \mu m$.

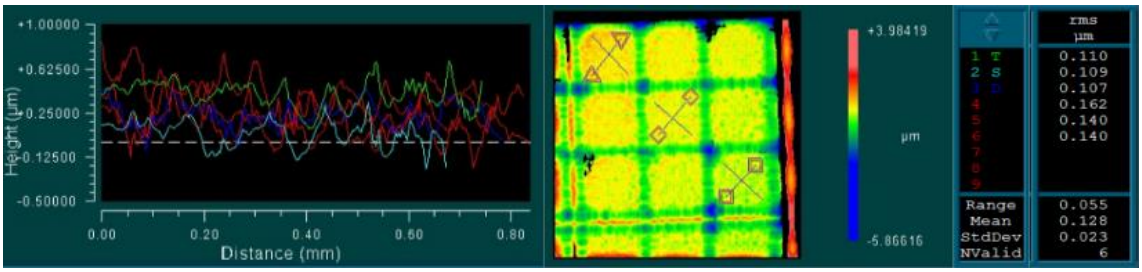

Figure 43: Roughness measurement after high-pass spatial filter.

To perform the RMS roughness measurements, a high-pass spatial filter was applied to the lens topography to remove the low-frequency spherical shapes, and the average roughness was calculated, as shown in Figure 43, resulting in $RMS = 128 \pm 23\,nm$. Although the average meets the most rigorous criterion, the actual roughness still can vary between $105\,nm$ and $151\,nm$, where the highest value violates the Fraunhofer criterion. Nonetheless, with more measurements, it is possible that the standard deviation could decrease, and a more reliable approximation to the mean could be achieved. Yet, it satisfies by far the Rayleigh criterion, and the surface can be considered smooth, by equation (56), for any wavelength satisfying $\lambda > 8 * 151 = 1208\,nm$. Once more, the roughness can be made still smaller if a denser pattern of holes is used in the initial mask of Figure 33, probably satisfying even the Fraunhofer criterion for the target wavelengths.

In contrast to Figure 43, Figure 42 (presented before) does not appear to present any roughness. This is because we have applied a low-pass spatial filter, exactly to eliminate the roughness and accurately calculate the sagittae of the lenses. Figure 44 shows the same filtered surface, and the average sagitta calculation, resulting in a value of $s = 12.8 \pm 0.6\,\mu m$.

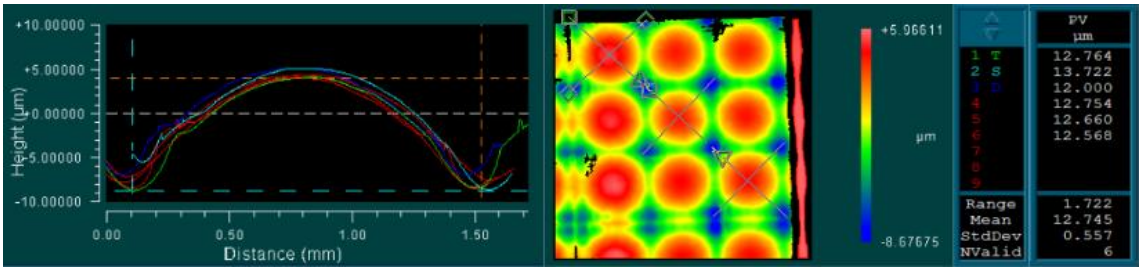

Figure 44: Sagitta measurements after low-pass spatial filter.

The initial mask was designed to have a sagitta of 11 µm. However, the measured one was larger than the desired. The final sagitta depends on the largest hole in the pattern mask of Figure 39 – that was $d_{max} = 31.6\,\mu m$ – and on a parameter m [68]. This parameter is the ratio between the etching ratios of the crystallographic planes <411> and <100> of silicon using KOH, that is: $m = R_{<411>}/R_{<100>}$. The sagitta of 11 µm is considered using $m = 1.4$. However, if the used temperature is around 80 °C, instead of 85 °C, for a watery KOH solution 33% in weight, the parameter increases to $m = 1.688$ [88], [89], resulting in a final sagitta of $s' = 13.0\,\mu m$. That value is much closer to the measured one. Therefore, the inconsistency between the designed and measured value can be explained by the lack of the right apparatus for controlling the temperature. We can conclude that maintaining the temperature steady is of paramount importance during the etch process. Notice, however, that even being sensitive to the

temperature, the process still imposes less variables to be controlled than the RIE process, mentioned in section 3.1.

Knowing the sagitta, in order to calculate the lens power $\left(\Phi_{diopters}\right)$ and focal distance $(f = 1/\Phi_{diopters})$, we have to find the radius of curvature of the lenses (r_C), given by equation (57), as follows:

$$r_C = \frac{s}{2} + \frac{D^2}{8.s} \quad [m] \tag{57}$$

where D is the largest diameter of the microlenses, given by $D = \sqrt{2}\, mm$, since the pitch of the squared array is 1 mm. The radius of curvature can be calculated to be $r_C = 19.6 \pm 0.9\, mm$, using an average deviation, since $r_C(s)$ is not a linear function.

Then, with the *Lens-Maker's Formula*, it is possible to find the lens power:

$$\Phi_{di} = \left(\frac{n_{Si}}{n_{air}} - 1\right)\left(\frac{1000}{r_C\ [mm]}\right) \quad [diopters] \tag{58}$$

in which $n_{Si} \cong 3.43$ and $n_{air} = 1$ are, respectively, the refractive indexes of silicon and air. The lens power is given in diopters: $[diopters] = [m^{-1}]$. Equation (58) is given for a thin lens, that is the case for the microlenses, since their radii of curvature are much larger than the total thickness of the lens along the optical axis, which is the sum of the sagitta and the wafer thickness.

Evaluating equation (58) yields a power of $\Phi_{diopters} = 123.6 \pm 5.4\, diopters$, or equivalently, they present an average focal distance equal to $f = 8.09 \pm 0.35\, mm$. We can see that it presents a very short focal length, suitable for compact systems. Even though the variation in the sagitta may seem large, the actual focal distance average variation is less than 5%. For instance, for image applications, even if a lens is not focusing an object precisely on the image plane, depending on the size of the light spot, an image can still be on focus. The maximum size the spot can have is called the Circle of Confusion (CoC). This value varies a lot in the literature, and depends on the type and size of the sensor, among other factors [90]. The CoC can be translated in a focal distance maximum deviation as $DoF = 2.CoC.f/\#$, where DoF is defined as the depth of focus, and $f/\#$ the f-number of the lens, that is, the ratio between the focal distance and the aperture of the lens. Basically, the DoF measures the maximum deviation the focus can have, such that the image is still in focus. For a typical $CoC = 0.03$, then $DoF = 0.34\, mm$. We can see that the maximum variation in the focus is $\Delta f = 0.70\, mm$, being larger than the DoF. Therefore, for image applications that require strict focus requirements, a new lens with a larger f-number should be designed**.**

Ultimately, in order to verify if the microlenses are functional as actual verging lenses, we have prepared a setup for measurements of the focal points. The setup consisted of a *transparency (or overhead) projector* as the light source, followed by the microlens wafer, an additional lens, and a camera, all disposed in the vertical direction. The projector comes with a high-power halogen lamp, which emits light in a wide spectrum. Also, the overhead projector has a supporting pole, and attached to it, an arm, which holds its projection head, mirror and lens. These last three components were removed, and the projector arm used to hold the

camera that has an ordinary silicon image sensor. Being made of silicon means it detects light up to $\lambda_{cut} \cong 1.1\ \mu m$ – using a bandgap of $E_g \cong 1.12\ eV\ @300\ K$ in equation (2). In other words, it senses light in the visible and in a portion of the NIR spectrum.

The used image sensor was a *NET Foculus FO442SB CCD camera* [91], but without the usual anti-infrared filter of commercial cameras. In fact, an *850 nm infrared long-pass filter* [92] was employed to ensure that no visible light reaches the detector. Figure 45 shows the relative responsivity of the camera (blue dashed lines) and the transmittance of the IR filter (red dots). In addition, by multiplying these curves, we can see the net responsivity of the camera (black solid line), which is rather low at NIR. However, since the light bulb of the overhead projector is quite strong ($410\ W$ of electrical power), a considerable amount of optical power can be delivered to the setup, and using the right settings in the camera (like the shutter speed, gain and exposure) can render in good focal point images.

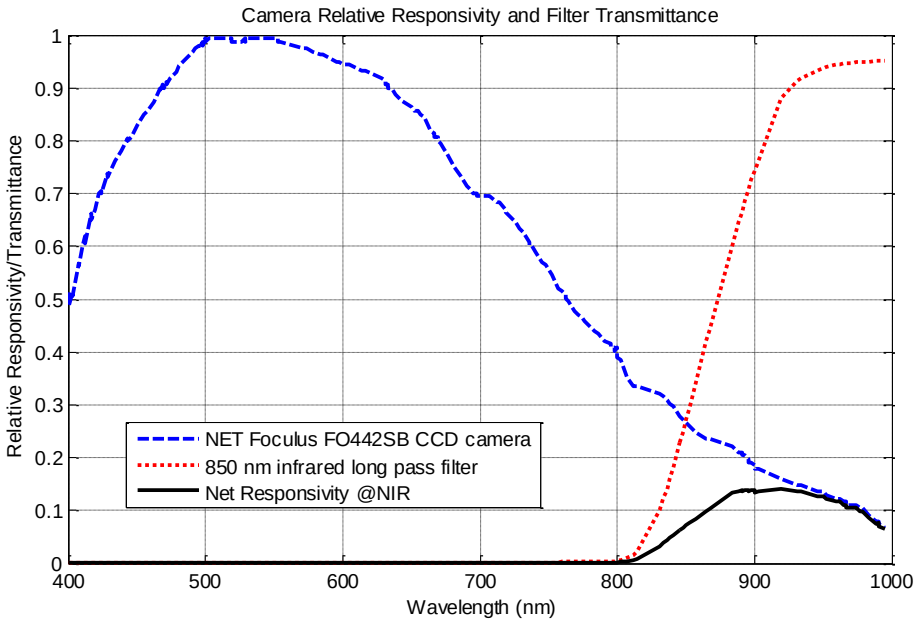

Figure 45: Normalized responsivity of the camera (blue dashed line); transmittance of the IR filter (red dots); and net responsivity (black solid line).
Source: Adapted from [91] *and* [92].

Figure 46 demonstrates the focal points acquired by the setup. At least one 3x3 microlens array is shown and with sharp focal points.

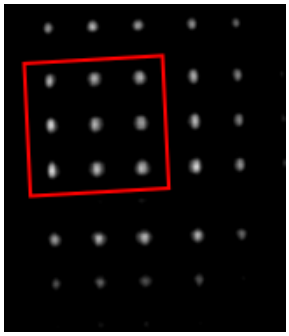

Figure 46: Image of the focal points acquired from the setup, highlighting one 3x3 array.

We may notice that we have used silicon both for photodetection and as an optical component. That seems to be a contradiction at first sight, since silicon should absorb below $\lambda_{cut} \cong 1.1\,\mu m$, and the microlenses would be opaque. However, although it absorbs, the silicon wafer has a finite thickness of $t = 325\,\mu m$. Table 6 presents some values of absorption coefficient for silicon at wavelengths within the range of the camera net responsivity.

Table 6: Some values of the wavelength-dependent absorption coefficient (α) for silicon, and the equivalent penetration depth[4] (δ=1/α)[93].

Wavelengths (λ)	Silicon absorption coefficient (α)	Penetration depth (δ=1/α)
0.80 µm	785.40 cm^{-1}	1.2732 µm
0.85 µm	591.36 cm^{-1}	1.6910 µm
0.90 µm	279.25 cm^{-1}	3.5810 µm
0.93 µm	270.24 cm^{-1}	3.7004 µm
0.94 µm	133.68 cm^{-1}	7.4806 µm
0.95 µm	132.28 cm^{-1}	7.5597 µm
0.96 µm	130.90 cm^{-1}	7.6394 µm
0.98 µm	128.23 cm^{-1}	7.7985 µm

Then, from equation (3), we know the relation between the amount of photon flux impinging on the silicon surface Φ_0, and the amount transmitted $\Phi(t)$ for a wafer with thickness t. As an approximation, consider a light source with a constant spectral intensity all over the wavelength range Δλ: $\Phi_{0\lambda}\ [(number\ of\ photons)/(m^2 s.\mu m)]$, where $\Phi_{0\lambda} = \Phi_0/\Delta\lambda$. Then, we can numerically integrate for the total amount of light passing through the wafer, and measure the ratio T of transmitted light.

$$\Phi(t) = \int_{0.80\ \mu m}^{0.98\ \mu m} \Phi_{0\lambda}.e^{-\alpha.t}d\lambda \ \rightarrow\ T = \frac{\Phi(t)}{\Phi_0} = \frac{1}{\Delta\lambda}\int_{0.80\ \mu m}^{0.98\ \mu m} e^{-\alpha.t}d\lambda \cong 0.3562\ \% \qquad (59)$$

Equation (59) is an approximation, but gives an idea of how much light passes through the silicon wafer. Although $T \cong 0.3562\ \%$ is a rather small fraction, it is not zero. For instance, using equation (59), but for Δλ ranging between $0.80\ \mu m \leq \lambda \leq 0.93\ \mu m$, renders in a much smaller transmittance: $T \cong 0.0053\ \%$, which is considerably closer to zero. With a sufficiently powerful light source (irradiance of $I_{lamp} \cong \frac{410\ W}{0.024\ m^2} = 17066\ \frac{W}{m^2}$), and an appropriate configuration of the camera (longer exposure times), good quality images could be acquired, as in Figure 46. The total power arriving over the camera can be estimated as $P_{camera} \cong I_{lamp}.H.V.L_{pixel}^2.T$, where $H = 1392$ and $V = 1040$ are the number of pixels in the horizontal and vertical directions, respectively, and $L_{pixel} = 6.45\ \mu m$ is the size of each pixel in either direction, and $T \cong 0.3562\ \%$ the transmittance through the wafer. This results in $P_{camera} \cong 3.66\ mW$. Of course, the optical power emitted by the lamp will be smaller than $410\ W$ (the electrical power), but will be strong enough, so that the wafer is sufficiently transparent for the measurements.

[4] *Penetration depth* (δ) is defined as the depth at which approximately 63% of incoming light has been absorbed by the material. For a depth of 5.δ, it is considered that most of the incoming light is absorbed.

Figure 47 shows an example where two wafers with microlenses were placed one above the other, both between the light source and the camera. Even with the two silicon wafers, light could still be detected at the camera. It is worth remembering that the target wavelength is 4.2 μm, for which the transmittance is even higher.

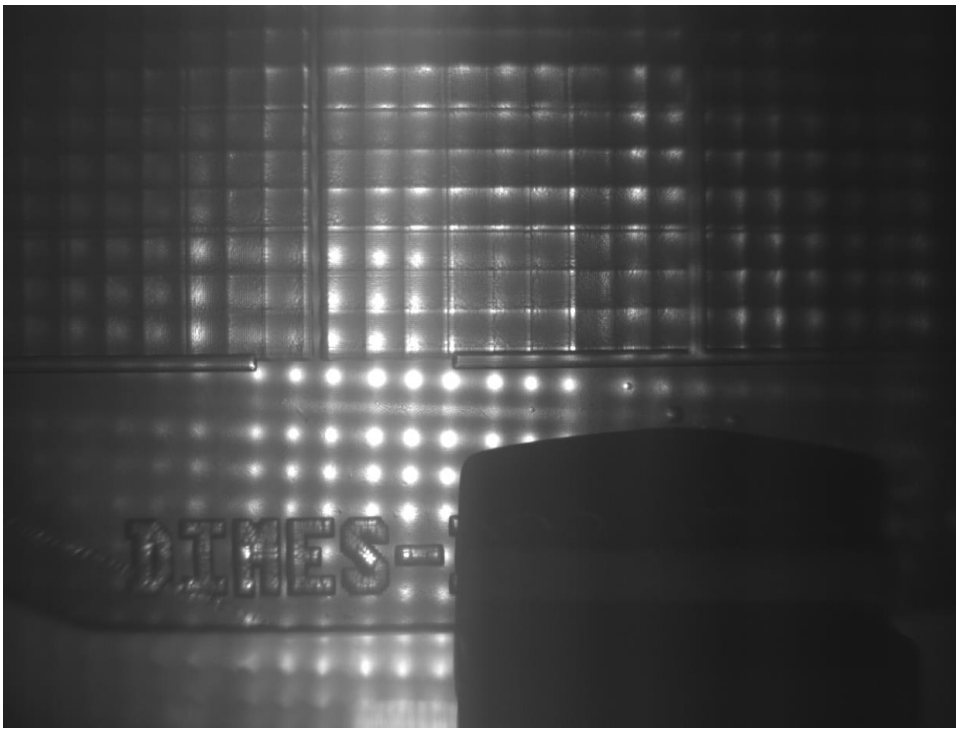

Figure 47: Two wafers placed between the light source and the camera.

In fact, this experiment was solely performed to demonstrate that the microlenses actually work as verging lenses. It was executed with a silicon camera (@NIR) because we did not dispose of an infrared camera capable of sensing @MWIR, i.e., at 4.2 µm and beyond.

4.2 Anti-reflective coating measurements

4.2.1 Preliminary simulations and depositions

In section 3.2, the simulation method was presented, using the software TFCalc™. Also, we have anticipated some simulation results in Table 5. These results from TFCalc™ – with no particular thin-film process technique and a single SiN layer – are shown as follows in Figure 48, Figure 49, and Figure 50, respectively aiming at the wavelengths of 4.2 µm, 8.2 µm, and 9.3 µm:

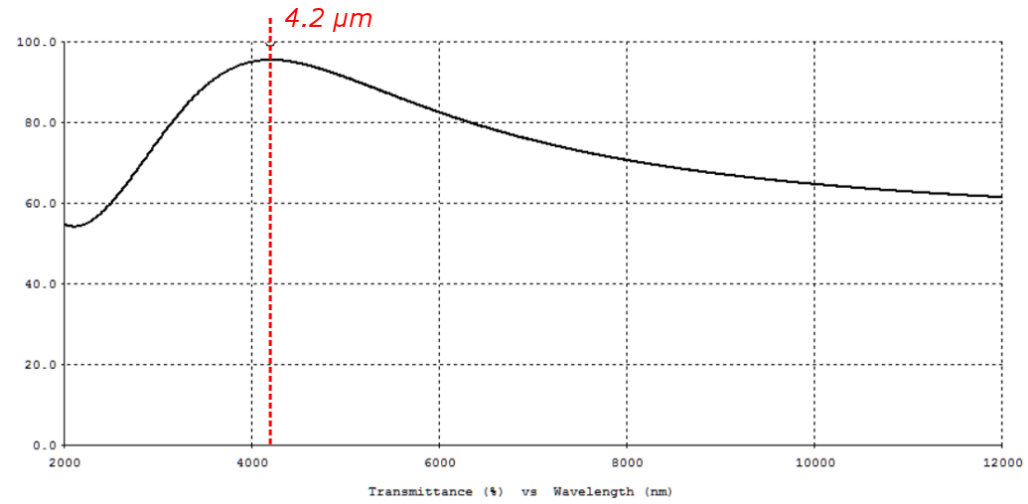

Figure 48: Single-Layer SiN simulation targeting 4.2 µm with 488 nm thickness.

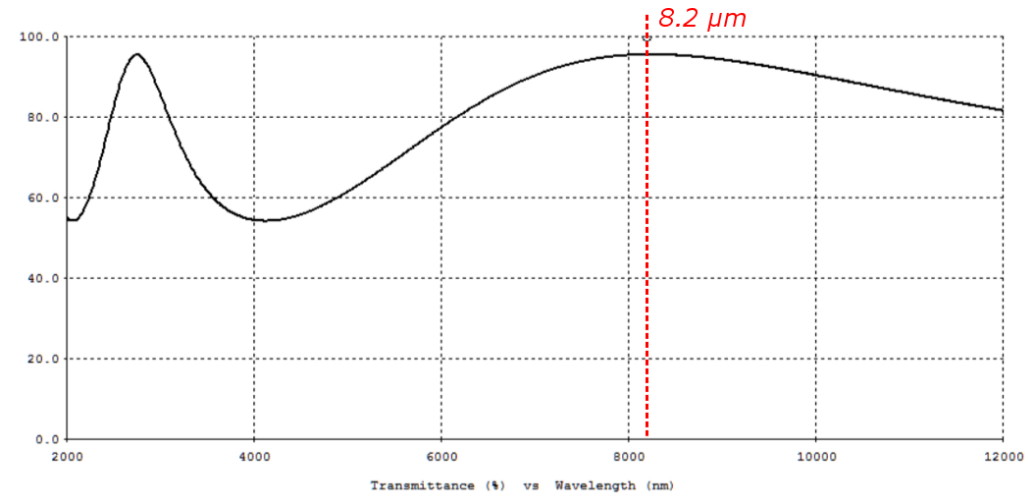

Figure 49: Single-Layer SiN simulation targeting 8.2 µm with 958 nm thickness.

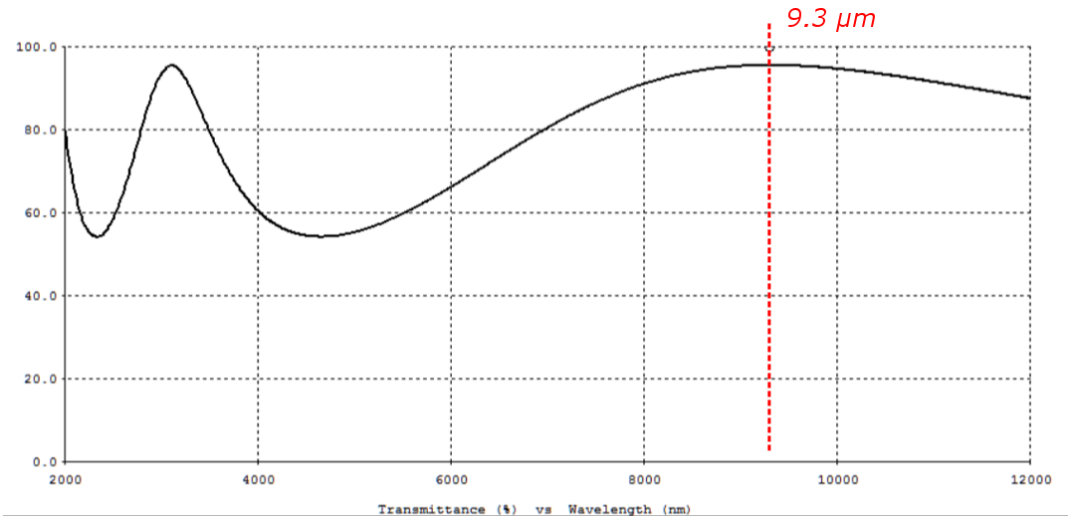

Figure 50: Single-Layer SiN simulation targeting 9.3 µm with 1082 nm thickness.

Clearly, the results were very promising. Therefore, since the optical data used for the simulations ranged up to 6 µm, we have first deposited the designed thin-film layer for 4.2 µm, where there was no need for data interpolation. Also, 4.2 µm is the wavelength of the CO_2 gas detection system that will be presented at the end of this thesis.

For the first depositions, a bare silicon wafer without any KOH etch was used. The SiN was deposited over both front and back surfaces of the wafer and, after first deposition, the actual thickness was measured, yielding 490 nm instead of the designed 488 nm. Measurements of the transmittance were performed using a FTIR spectrometer (*Thermo Scientific Nicolet 5700*).

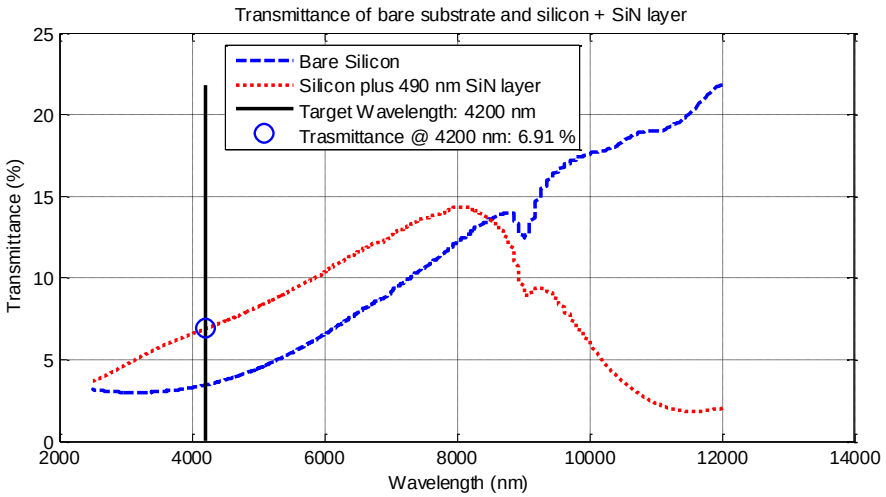

Figure 51: Transmittance through bare Silicon, and through a 490 nm SiN layer deposited over silicon, aiming 4.2 µm. Both silicon wafers had only the front side polished.

It is evident from Figure 51 that the measured transmittance profile does not resemble the designed profile shown in Figure 48. These measurements show the transmittance through the wafer, i.e., considering that the detector would be put behind the silicon substrate, like in Figure 34, and considering air as the background medium. In fact, to see only the effect of the silicon nitride layer on transmittance, we can divide the *silicon + SiN* measurement by the one only with the *bare silicon* substrate.

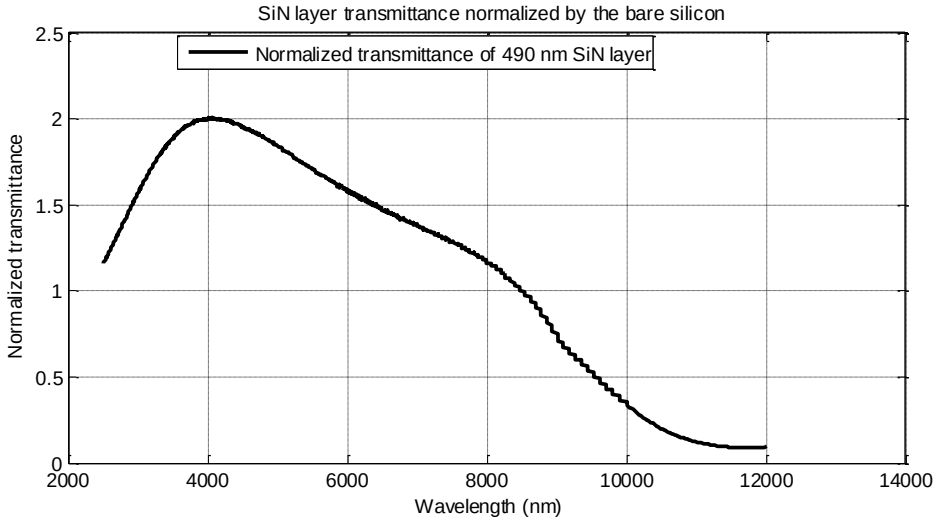

Figure 52: Normalized transmittance of the SiN layer.

Figure 52 shows the normalized transmittance of the SiN layer. This new curve shows how the thin-film layer enhances the overall transmittance, and at a wavelength very close to the designed, namely @4.016 µm, practically doubling the bare silicon transmittance. This new result resembles what is shown in Figure 48, at least near the target wavelength. Although that is a good sign that the nitride layer is acting as an ARC, it still does not present the desired behavior. From the literature, we know that the bare silicon transmittance should be around 55 % for 4.2 µm. In Figure 51, however, the bare silicon transmittance does not exceed 25 %, and is less than 5 % at 4.2 µm.

We have suspected that this behavior could be due to the rough back surface of the sample wafer. Actually, this first measurement was done over a silicon substrate with only the front side polished. The decreasing transmittance trend of the bare silicon curve (blue) for smaller wavelengths, presented in Figure 51, supports this argument. In fact, the whole curve is dominated by scattering, and it is known that larger wavelengths are less scattered, for the same roughness. Therefore, we have replicated the same deposition as before, but now using a *double-side polished wafer* (*DSPW*). The measured thickness of the SiN layer for this new case was 495 nm.

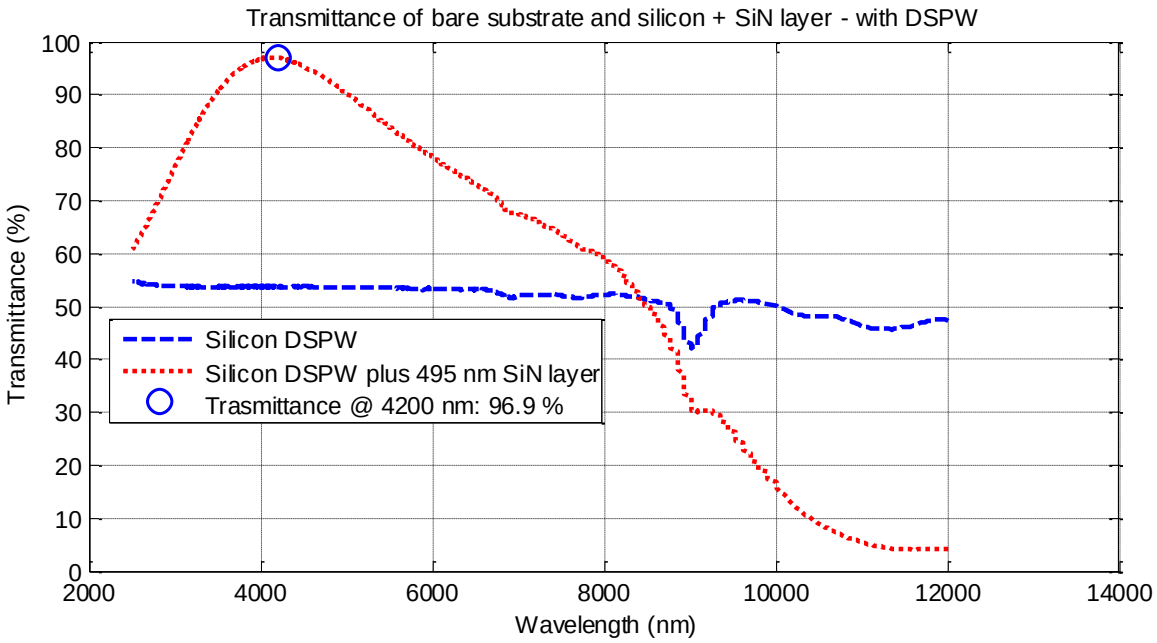

Figure 53: Transmittance through bare Silicon, and through a 495 nm SiN layer deposited over silicon, aiming 4.2 µm, and using DSPW.

This time, the FTIR measurements of transmittance through the wafer present much higher transmittance, without normalizing the curves. The SiN maximum transmittance is 97 % and occurs @4.181 µm. Also, for the target wavelength, transmittance is almost the same: 96.9 %. This result implies that using double-side polished wafer is a must for optical applications, near these wavelengths!

Although the maximum was not reached for the designed wavelength and thickness, it is sufficiently close. Still, comparing Figure 48 and Figure 53 there is a major difference regarding the long "tail" of the curve, for $\lambda > 6\ \mu m$. In practice, SiN absorbs much more than simulated, which suggests that the used optical constants in TFCalc™ were not the right ones. This is mainly due to the lack of optical data for $\lambda > 6\ \mu m$, where the extinction coefficient *k* might be higher

than the extrapolated. In order to better investigate the behavior of SiN at these wavelengths, we have carried out two more depositions, corresponding to the wavelengths of 8.2 µm and 9.3 µm, related to the simulations in Figure 49 and Figure 50, respectively.

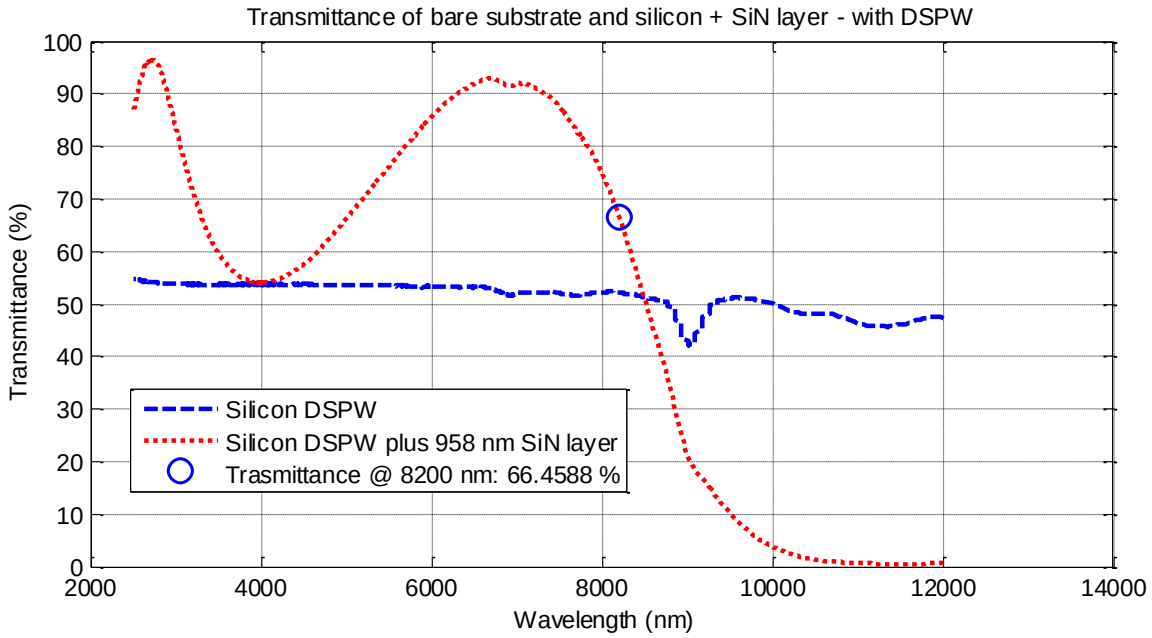

Figure 54: Transmittance through bare Silicon, and through a 958 nm SiN layer deposited over silicon, aiming 8.2 µm, and using DSPW.

Both aimed and measured thickness for Figure 54 were 958 nm. At 8.2 µm, the transmittance was 66.46 %. It shows two maxima, namely @2.71 µm and @6.70 µm, with transmittances of 96.65 % and 92.91 %, respectively. From simulation data of Figure 49, the last peak on the right-side should happen close to 8.2 µm, which is shifted in real data.

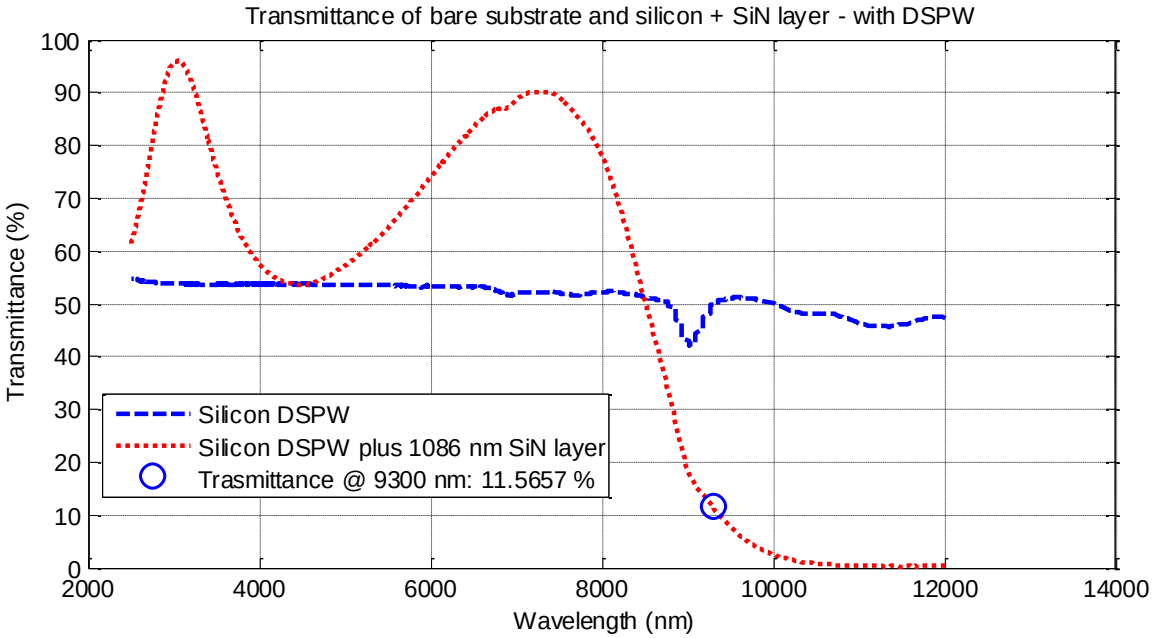

Figure 55: Transmittance through bare Silicon, and through a 1086 nm SiN layer deposited over silicon, aiming 9.3 µm, and using DSPW.

The aimed thickness for Figure 55 was 1082 nm, and the measured one 1086 nm. At 9.3 µm transmittance was 11.57 %. The two maximum transmittances were 96.10 % @3.06 µm and 90.22 % @7.25 µm. Again, from Figure 50 simulation, the maximum at the far right should be closer to 9.3 µm, which did not happen at all.

The last two measurements corroborate the fact that there is a strong absorption component for $\lambda \geq 8\,\mu m$. Therefore, the optical data for the simulation has to be improved in order to reliably predict the behavior of the thin-films at larger wavelengths.

4.2.2 Redesign and new depositions

After starting the simulations and deposition processes, we have identified two papers that had extracted the optical properties at the infrared range of our interest. *Kischkat et al.*, *2012* [77] used *Reactive Sputtering* for SiN thin-film formation, and observed great variance of (n, k) data depending on the different employed reactive gas flow rates. Therefore, there is a great likelihood that their data will not fit the experimental results shown in this thesis. In addition, SiN sputtering was not an available technique at the time of the experiments.

Next, *Wicaksono et al.*, *2009* [19] used the same technique as we had – *LPCVD-deposited low-stress Si-rich nitride*. Consequently, we could have employed data from this paper, which consists of a *Lorentz Oscillator Model* [94], whose fitting constants were available in the article. We have implemented this model, and generated the (n, k) dataset to be used in TFCalc™. In order to compare with the previous depositions, the simulations with the new optical constants used the measured deposited thicknesses, namely 495 nm, 958 nm, and 1086 nm.

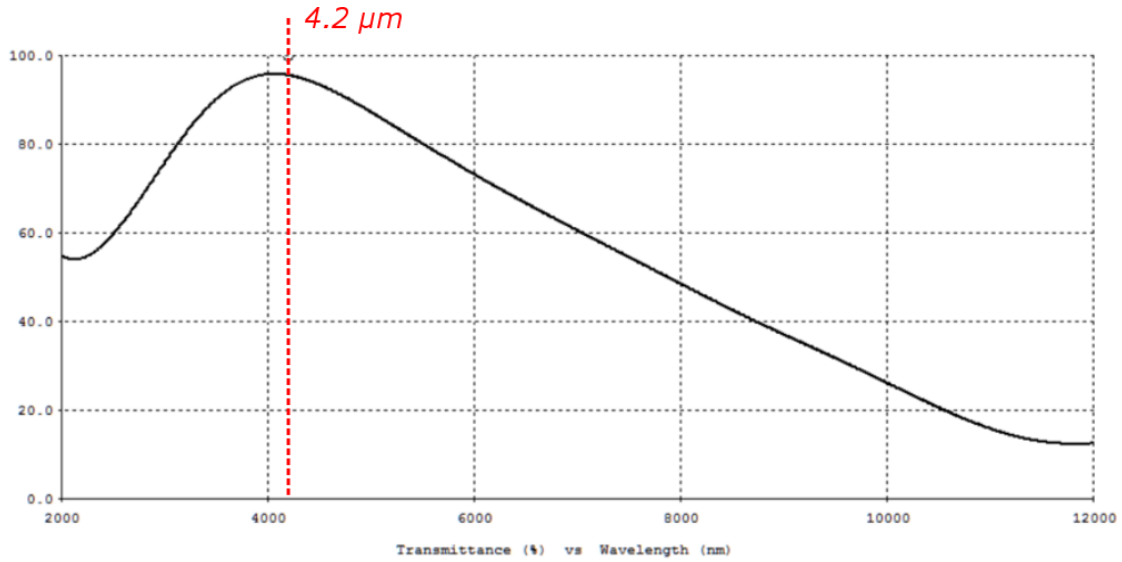

Figure 56: Single-Layer SiN simulation with 495 nm thickness, and new dataset, aiming 4.2 µm.

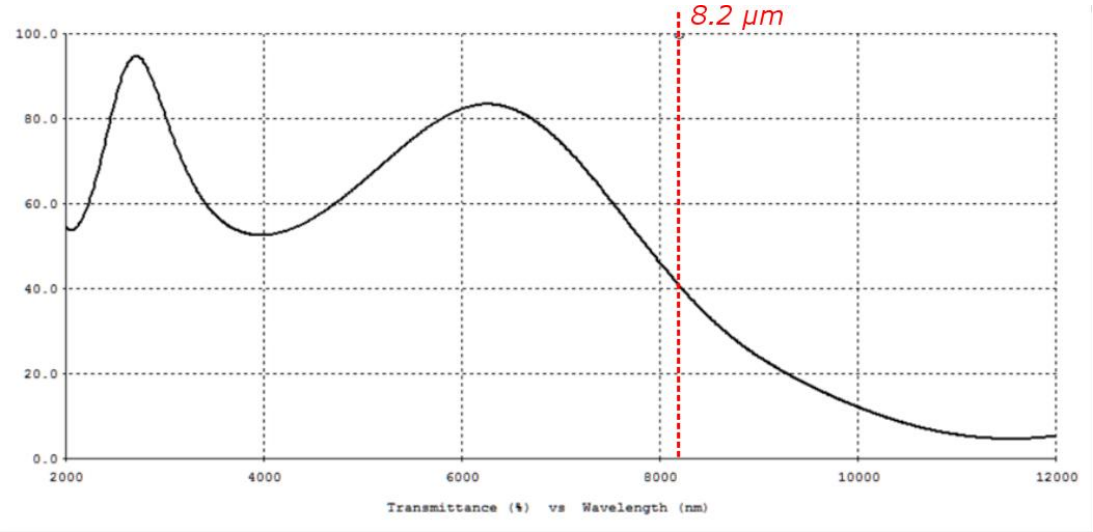

Figure 57: Single-Layer SiN simulation with 958 nm thickness, and new dataset, aiming 8.2 µm.

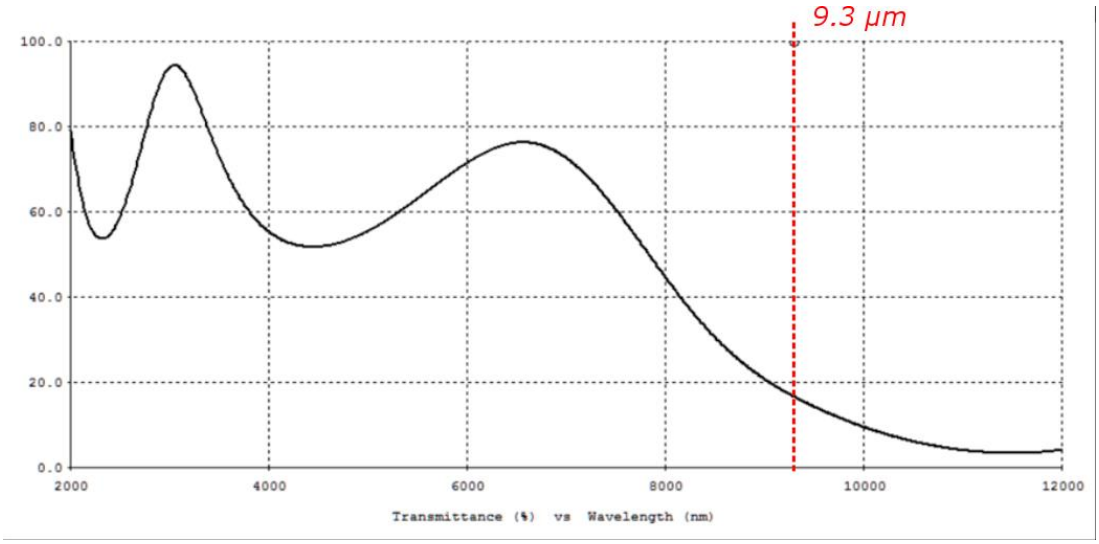

Figure 58: Single-Layer SiN simulation with 1086 nm thickness, and new dataset, aiming 9.3 µm.

The major information from simulations shown in Figure 56, Figure 57 and Figure 58 were summarized in Table 7. That is, the maximum transmittances at the largest wavelengths. For 495 nm, measured and simulated data are fairly close. However, for 958 nm and 1086 nm, neither the maximum transmittances nor the wavelengths at which they occur are well-matched. While for the previous optical constants dataset, used in subsection 4.2.1, simulations did not show any absorption for $\lambda \geq 8\ \mu m$, for the new dataset, it shows too much absorption. In other words, comparing to the measured data, the new dataset probably has a series of extinction coefficients that are higher than the real ones, for the considered wavelength range.

Table 7: Comparison between measured and simulated maximum transmittances for the three different thicknesses.

Thickness	Measured	Simulated
495 nm	97.00 % @4.18 µm	95.80 % @4.06 µm
958 nm	92.91 % @6.70 µm	83.30 % @6.26 µm
1086 nm	90.22 % @7.25 µm	76.20 % @6.56 µm

Since our main application is at 4.2 µm – for the CO_2 gas detection system – and simulations were reasonable near the measured data for this wavelength, then we have redesigned and found the optimized thickness for this wavelength. Figure 59 shows the updated simulation, rendering in a SiN layer of 512 nm, with transmittance of 95.7 %.

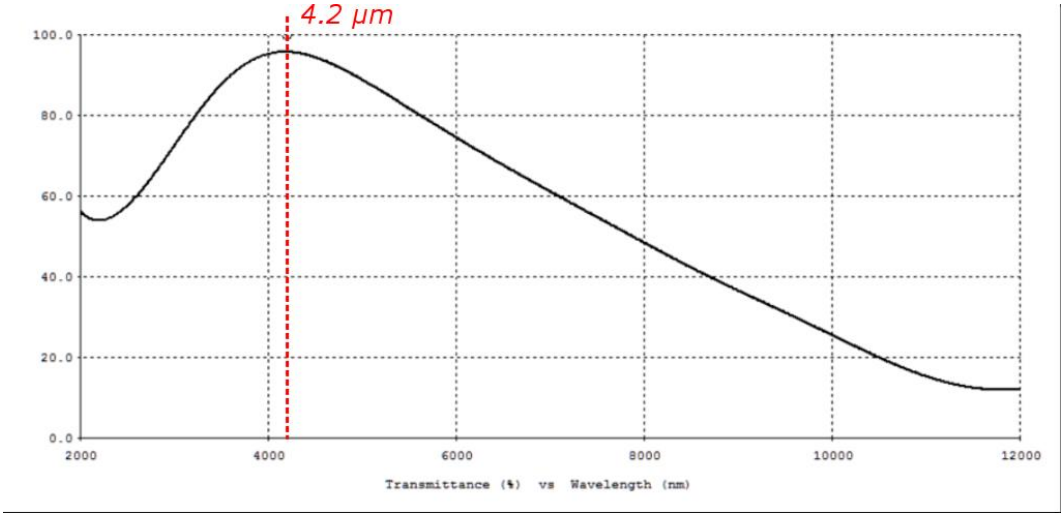

Figure 59: Single-Layer SiN simulation targeting 4.2 µm with new dataset, and 512 nm thickness.

The next step was to deposit this newly designed thin-film layer, but now over a wafer with the microlenses, instead of a bare silicon. The simulation was still not very reliable for wavelengths other than $4.2\ \mu m$, however, since the main idea at the moment is to integrate the microlenses with the ARC that could be used for a CO_2 gas detection system, then the results are sufficient for now. After performing all steps described in section 4.1 to fabricate the microlenses, a wafer just like the one in Figure 41 – a DSPW micromachined with KOH to form the microlenses – was deposited with SiN. Aimed thickness was 512 nm, but measured one was 523 nm. Since this wafer had undergone etching with KOH to produce the microlens profiles, a little bit of roughness was introduced. In section 4.1, the roughness due to the convex lens pattern approximation and KOH were measured, and concluded that is not sufficient to degrade the optical quality by a great extent (by the Rayleigh criterion). The roughness solely due to KOH, however, is still lower than the measured, and therefore might not be a concern.

Table 8 shows a summary of all fabricated and measured wafers. The last samples to be measured are the ones with the fabricated microlenses, with and without SiN. Note that they are also the only ones with a thinner substrate, due to KOH etching of both sides (since the back side was not protected against etching).

Table 8: Summary of all tested wafers.

Sample Description	Wafer Final Thickness (µm)	Wafer description
Bare silicon wafer	525	Front side polished
490 nm SiN layer on silicon wafer	525	Front side polished
Bare silicon wafer	525	Double side polished
495 nm SiN layer on silicon wafer	525	Double side polished
958 nm SiN layer on silicon wafer	525	Double side polished
1086 nm SiN layer on silicon wafer	525	Double side polished
Bare silicon wafer (with microlenses)	325 (due to KOH thining)	Double side polished
523 nm SiN layer on silicon wafer (with microlenses)	325 (due to KOH thining)	Double side polished

Figure 60 exemplifies where the spot positions were aimed at, for the FTIR measurements of the wafer with microlenses. Two different groups of measurements were undertaken: first, measuring both bare silicon and silicon + ARC, with the light beam of the spectrometer aiming at the wafer (in a flat region where there were no lenses). Second, measurements with and without ARC, and the light beam pointing at the lenses.

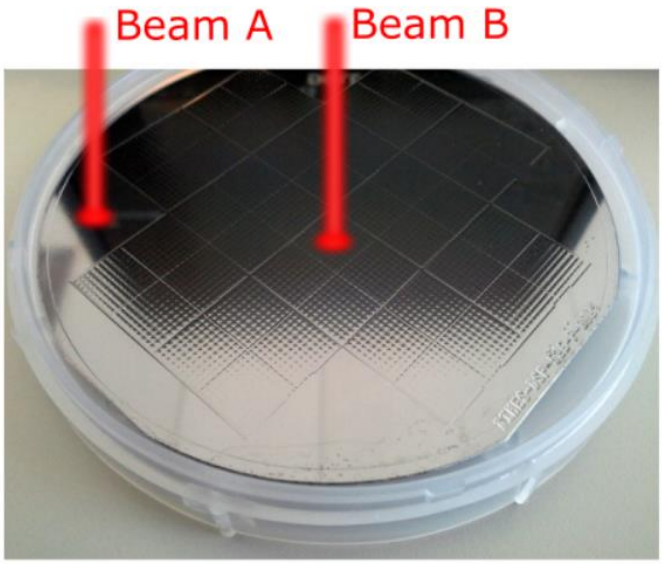

Figure 60: FTIR measurements through flat region of the wafer (beam A) and though the lenses (B).

Figure 61 indicates the FTIR measurements on the microlens wafers with and without SiN. We can see that the transmittances for the wafer with 523 nm SiN layer (at 4.2 µm) were 91.96 % and 71.84 %, aiming the beam at the wafer and at the lenses, respectively. We can notice that for the same wafer, there is a considerable difference when aiming the light beam at the lenses and at the flat region of the wafer. Also, the overall transmittance for the bare silicon wafer aiming at the lenses decreases a lot compared to the beam at the flat region of the wafer.

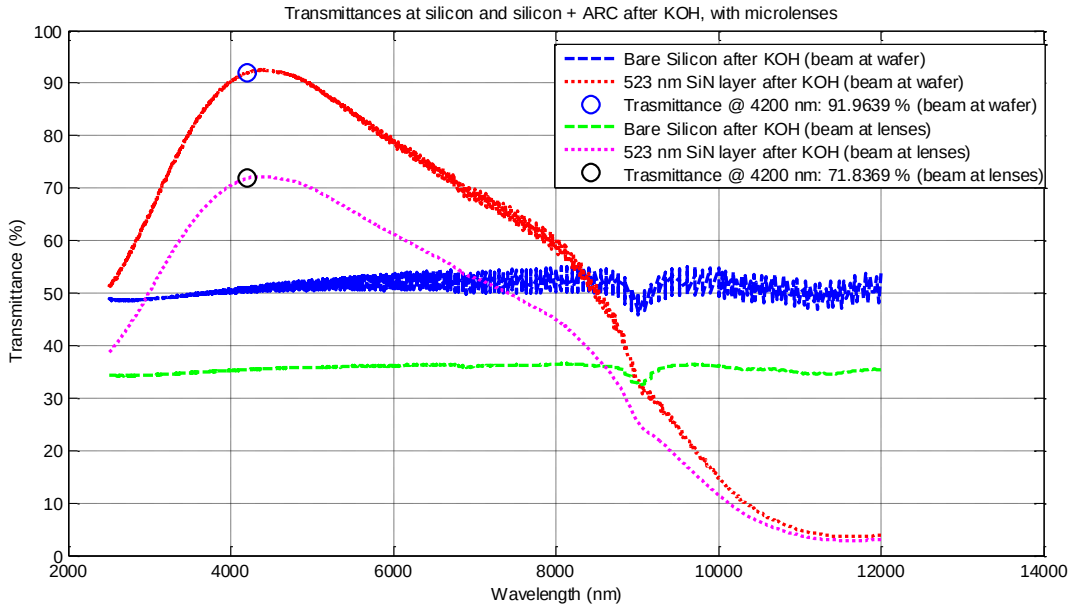

Figure 61: Transmittance measurements for two silicon wafers with microlenses: with and without ARC, and FTIR beam at the flat region of the wafer and at the lenses.

The reason behind this decreasing in transmittance for the beam at the lenses, is probably because the vertical distance between the sample and the FTIR detector was fixed, and much bigger than the focal distance of the microlenses ($> f = 8.092\, mm$). Also, the infrared light beam was larger than the microlenses. The beam had an elliptical shape with minor and major axes of about 3 mm by 10 mm, respectively. Therefore, in this case, the lenses actually diverge light, making only part of the incoming infrared light to be captured by the detector.

Table 9 compiles the maximum transmittances for the SiN wafer and the wavelength at which they occur, for both light beam positions. Also, for the same wavelengths, transmittances through the microlenses silicon wafer without SiN are presented.

Table 9: Maximum transmittance for the SiN layer at different beam positions, and their respective transmittance for the bare silicon wafer.

	With SiN coating	Bare Silicon
Beam through the wafer	92.56 % @4.40 µm	50.40 % @4.40 µm
Beam through the lenses	72.22 % @4.38 µm	35.50 % @4.38 µm

Comparing results from Table 9 and from Figure 53 (for wafers without the microlenses, and therefore, with no KOH etching), although each maximum transmittance occurs at different wavelengths, it is clear that there is a drop in the transmittance for the SiN with beam at the wafer by $97.00 - 92.56 \cong 4.44\,\%$. Additionally, for a DSPW bare silicon without any KOH etching, transmittance is about $55\,\%$, whereas for the wafer after KOH is $\sim 50\,\%$. Consequently, we can conclude that the roughness introduced by the KOH etching reduced the transmittance in no more than 5 %, compared to a DSPW, at least for wavelengths near 4.2 µm.

It is important to notice, however, that a basic step presented in Figure 30, consisting of depositing a protective nitride layer at the back side of the wafer, during the KOH etching process, was NOT done for the fabricated wafers. This nitride layer at the backside surface protects the oxide layer (placed right underneath it) from the HF buffer etching. The oxide layer, in turn, protects the silicon surface from the KOH etching, that introduces roughness, and thins the final sample wafer. Therefore, the final roughness could be diminished, and the transmittance loss reduced if this protective layer was employed.

4.2.3 Transmittance simulations improvement

Although simulation results were fairly satisfactory near 4.2 µm, for $\lambda > 6\,\mu m$ the optical constants dataset still needed to be improved. This is extremely important for future works, since it gives the ability and flexibility to make new designs aiming at different wavelengths, with more reliability. For that purpose, we have used a free software called *RefFIT*, by *Alexey Kuzmenko*, from University of Geneva [95], [96]. This is a really powerful and flexible software, able to get optical data of the analyzed dielectric material through common optical measurement techniques, like reflectivity, transmittance, or ellipsometry. This is done by fitting the optical spectra using models for the dielectric function of the material, with a set of adjustable parameters. There are two major types of models: a) a mathematical formula, based on the Drude-Lorentz (*DL*) model; b) or a variational dielectric function, also called as

"free-shape" [95]. We have chosen the first option for simplicity, since it is more user-friendly, although the second one might generate better results. The formula-defined dielectric function used in RefFIT is given as follows, and is comprised by a sum of Lorentz oscillators:

$$\varepsilon(\omega) = \varepsilon_\infty + \sum_j \frac{\omega_{pj}^2}{\omega_{oj}^2 - \omega^2 - i\gamma_j\omega} \qquad (60)$$

where $\varepsilon(\omega)$ is the complex dielectric function; ε_∞ is the high-frequency dielectric constant contribution of all oscillators; ω_{pj}, ω_{oj} and γ_j are, respectively, the plasma (or longitudinal) frequency, the resonant (or transverse) frequency, and the linewidth (or scattering rate/damping) corresponding to the j^{th} Lorentz oscillator [96], [97]. Equation (60) is used to model the material and fitting of measured optical data.

Data fitting in RefFIT is done by the chi-square (χ^2) minimization technique. Since the minimization function is generally non-linear for optical modeling, then it is indispensable to use the *Levenberg-Marquardt Algorithm* (*LMA*). The LMA is commonly used for non-linear least squares solving, and is implemented in RefFIT. In order to generate a good model, the more experimental data, the better. In this case, simultaneous fitting of several datasets is needed, and RefFIT simply employs a weighted sum of all χ^2 functions to be minimized. For instance, if we have N_D datasets to be fitted at once, then the new chi-square function to be minimized is given by: $\chi^2 = \sum_{i=1}^{N_D} w_i.\chi_i^2$, where χ_i^2 are the individual chi-square terms, and w_i their respective weights. This is a simple way of converting a multi-objective optimization problem into a mono-objective one. From now on, we have assumed $w_i = 1$ for all experimental datasets ($\forall\ i$).

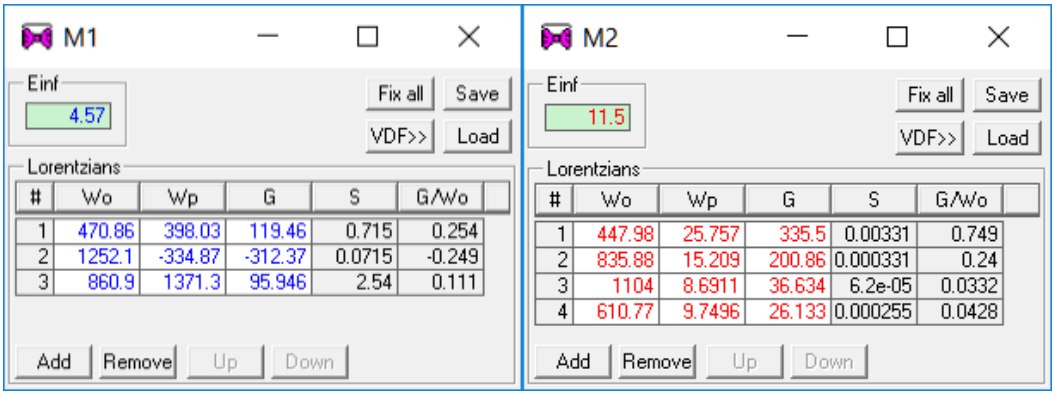

Figure 62: Examples of Lorentz Oscillator models in RefFIT: M1 SiN; M2 Si.
Source: Adapted from [96].

Figure 62 shows two examples of Lorentz oscillator models for dielectric materials – in this case for SiN and Si. Regarding Figure 62 and equation (60), *Einf* refers to ε_∞, *Wo* to ω_{oj}, *Wp* to ω_{pj}, and G to γ_j, being each line a different Lorentzian term (j^{th}). The additional terms S and G/Wo represent, respectively, the *Lyddane-Sachs-Teller* relation, and the normalized damping [96]–[99]. The first can be written as $S = (W_p/W_o)^2 = \varepsilon(0)/\varepsilon$, where $\varepsilon(0)$ is the value of the dielectric function at zero frequency, and the second is a measure of the broadening in each Lorentz oscillator [97].

We can build as many different models, as different materials there are. However, in our practical case, the transmission measurements were performed aiming a light beam through a stack of materials, and not by a single material layer – as modeled in Figure 62. This way, another different special model has to be used, which accounts for situations where there is a multilayer. In order to select this special model, we have to replace the value in the field *Einf* (ε) by the code of the wanted model, which in this case is *"-33"*, as can be seen in Figure 63. Detailed information about this model can be found in the RefFIT manual [96], under the code *"-33"*, and name *"Reflectance/transmittance/ellipsometry of a multilayer"*. The input parameters for this model are listed on Table 10.

Table 10: Special model (code -33) input parameters for a 3-layers' stack.

Line number #	Wo	Wp	G
1	-	Angle of incidence (θ) [degrees]	Polarizer angle (P) [degrees]
2	Model number of layer #1	Thickness of layer #1 [cm]	Layer #1 coherence (0 or 1)
3	Model number of layer #2	Thickness of layer #2 [cm]	Layer #2 coherence (0 or 1)
4	Model number of layer #3	Thickness of layer #3 [cm]	Layer #3 coherence (0 or 1)

On Table 10, the first line is related to the input light characteristics, namely the angle of incidence θ, and the polarizer angle P, both varying from 0° to 90°. Layers 1, 2 and 3 can be configured independently, and with different material models, respectively at lines 2, 3 and 4. Each of these lines refers to another model window that describes a specific material. Columns 1, 2 and 3 represent, respectively, the model number, the thickness in [cm], and the coherence of each layer. If a layer is treated as coherent, then it means that the phase difference between the multiple internal reflected rays is taken into account (just as in Figure 28). Otherwise, for an incoherent layer, this phase difference is disregarded, which is a good approximation for thick layers – mainly substrates. One should use 0 for incoherent layers, and 1 for a coherent one.

A practical example of this multilayer model is shown in Figure 63, along with the corresponding sample cross section. There, both incidence and polarizer angles are 0°. Lines 2 and 4 use Model 1 (SiN), and line 3 use Model 2 (Si), from Figure 62. In other words, a stack of SiN–Si–SiN is formed, just like in our transmittance measurements with the FTIR – as schematized in Figure 34. The thicknesses of the SiN layers are the same ($4.95x10^{-5}\ cm\ =\ 495\ nm$), and the silicon substrate equal to $525\ \mu m$. In addition, the SiN layers are considered to be coherent (thin-film), and the silicon substrate incoherent (thick).

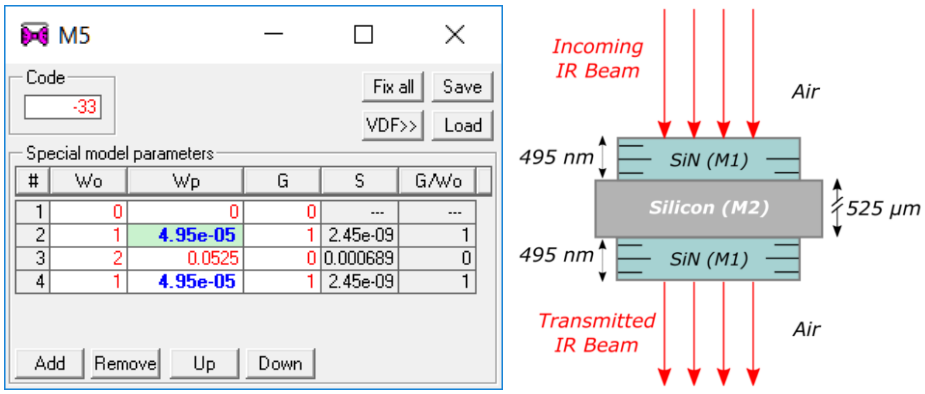

Figure 63: Example of special model (code -33) for multilayer optical fitting, and the corresponding sample cross section.
Source: Adapted from [96].

By using these models, we can import the optical measured data, and try to fit them using RefFIT. Bearing this in mind, we have gone through the following methodology:

1. Imported into RefFIT the measured transmittance data from the FTIR, namely the DSPW silicon substrate, and the stack SiN + Silicon + SiN, shown in Figure 53, Figure 54 and Figure 55.
2. Created two ordinary models, M1 and M2, respectively for the SiN and for the Si substrate. These models were previously shown in Figure 62. We have used 3 Lorentzians for M1 and 4 for M2. Note that the initial values of each parameter of M1 and M2 are not the ones shown in the image, since we do not know them in advance.
3. Then, a special model M6 (code -33) was created to build a "stack of one material". For this case, only M2 (Silicon) was added as a layer, corresponding to the (incoherent) bare silicon DSPW, with thickness of 525 µm, as in Figure 64. This is to emulate the measurement of bare silicon transmittance alone.

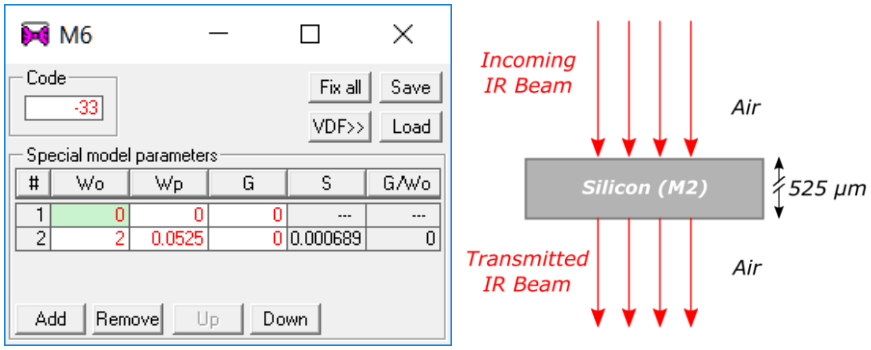

Figure 64: Special model (code -33) for DSPW bare silicon.
Source: Adapted from [96].

4. Then, the chi-square fitting was performed, aiming to find the best M6 model that approximates to the bare silicon DSPW FTIR transmittance measurement. Figure 65 exhibits the RefFIT fitting window for this case. Note that M6 does not have any adjustable parameter (model, thickness, angles and coherence are fixed). However, it refers to the M2 Lorentz oscillator model (for silicon). Parameters inside M2 are the actual variables used by the chi-square minimization.

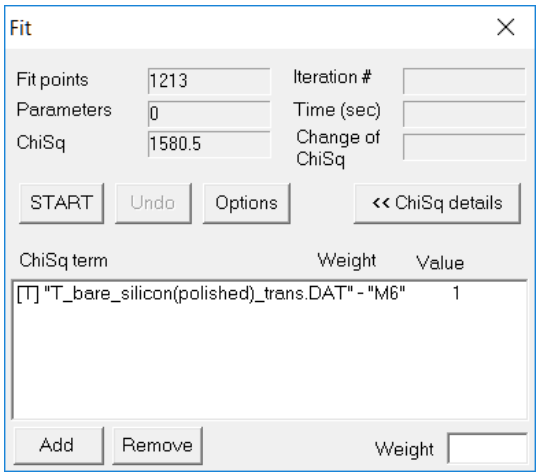

Figure 65: Chi-square fitting between M6 and bare silicon DSPW transmittance.
Source: Adapted from [96].

5. After the χ^2 minimization, a set of parameters was found for M2 that fits best the bare silicon transmittance profile. These parameters are shown in M2 from Figure 62. Also, Figure 66 presents the measured transmittance data (dashed blue line) and fitting model (solid red line). The FTIR spectrometer measures transmittance from 400 cm^{-1} up to 4000 cm^{-1}, where [cm^{-1}] is a wavenumber unit, commonly employed in Physics. Conversion from wavenumber in [cm^{-1}] to wavelength in [nm] is done by: $wavelength = 1x10^7/wavenumber$. This results in a range from 2.5 µm up to 25 µm, shown at the upper plot of Figure 66. However, since our main interest lies at approximately $\lambda \leq 12\,\mu m$, a second plot is shown with a zoom at the wavelengths from 2.5 µm up to 12 µm.

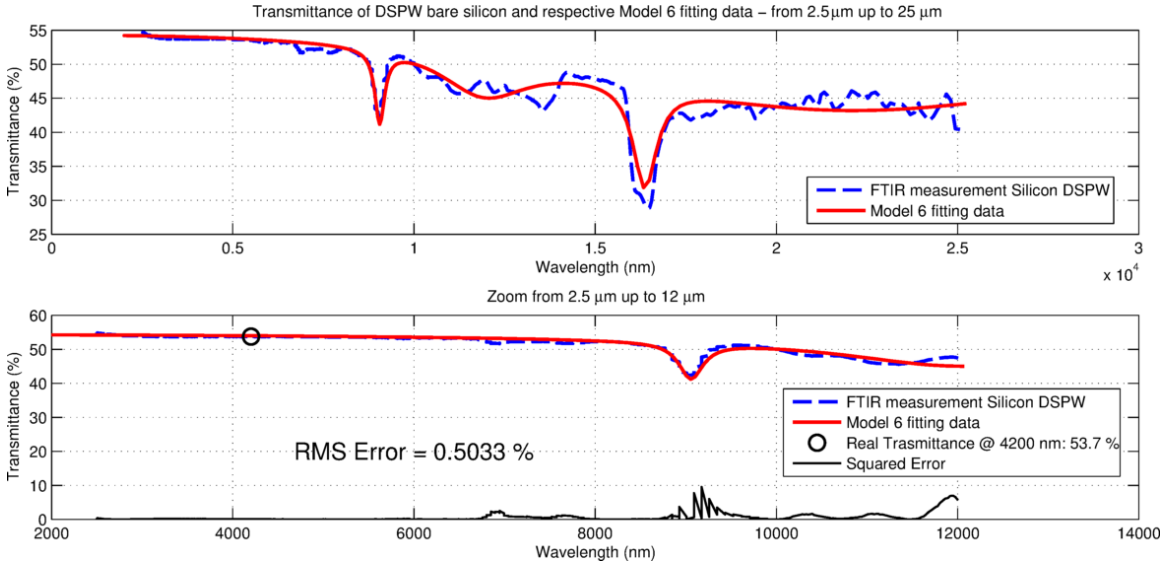

Figure 66: Bare silicon DSPW transmittance and the corresponding M6 fitting data, with RMS error of 0.5033 %.

6. Notice that our final goal is to find the optical constants of the SiN layer. To do this, we have used a special model like the one in Figure 63. This model creates the stack SiN-Si-SiN that refers to both M1 and M2. These models together have a total of 23 different parameters, which is a lot for the χ^2 minimization. The more parameters, the harder is to find an optimal solution. By fitting bare silicon data alone, we could first

find close enough M2 parameters, and afterwards try to fit the SiN model. In addition, remember that we have FTIR transmittance measurements for stacks with three different thicknesses for the SiN layer, that is, 1086 nm, 958 nm and 495 nm. We have then, created three special models (M3, M4 and M5), one for each stack and SiN layer thickness, and run the fitting, as in Figure 67.

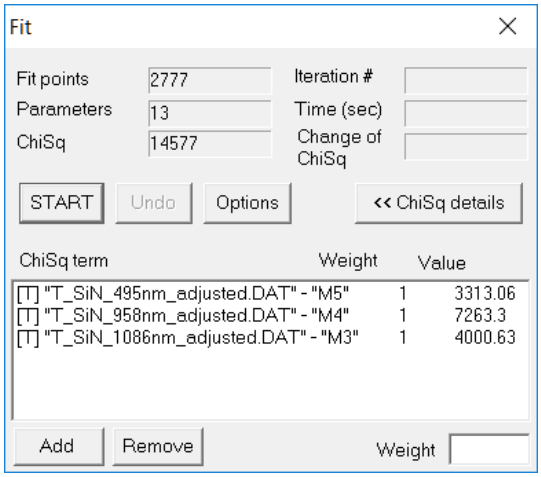

Figure 67: Multiple Chi-square fitting with equal weights between M5, M4 and M3 and their respective FTIR transmittance data.
Source: Adapted from [96].

7. The resulting SiN dielectric function was already presented in M1 of Figure 62. In contrast with the fitting performed in Figure 65, where M6 had only the substrate, for models M3, M4 and M5, we have allowed the SiN layer thicknesses to vary around the measured nominal value $\pm 10\ nm$, since there is always a thickness measurement error. Final model parameters after χ^2 minimization are shown in Figure 68. Note that resulting thicknesses for best fit were 1095.70 nm, 968.00 nm and 503.77 nm, for models M3, M4 and M5, respectively.

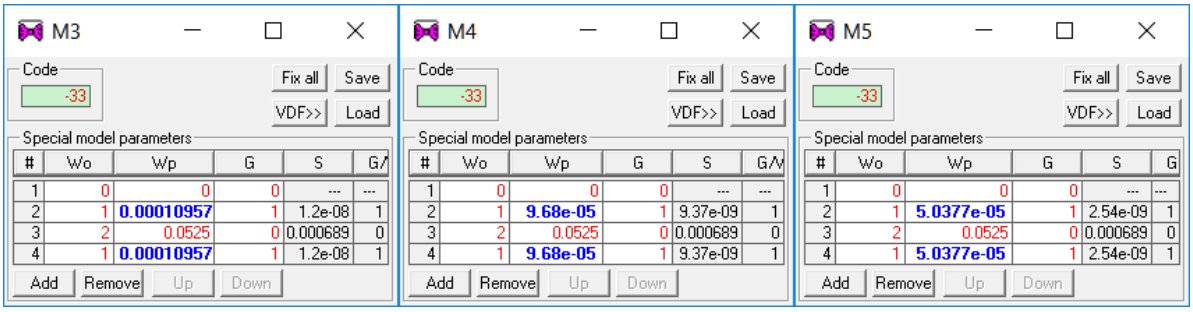

Figure 68: Special models M3, M4 and M5 and their respective final thicknesses.
Source: Adapted from [96].

8. Lastly, Figure 69, Figure 70, and Figure 71 show both measured and fitted data for each stack with different thickness. Clearly, fitted models presented much superior results than the (n, k) dataset from *Wicaksono et al., 2009*, used for previous simulations in TFCalc™ (Figure 56, Figure 57 and Figure 58).

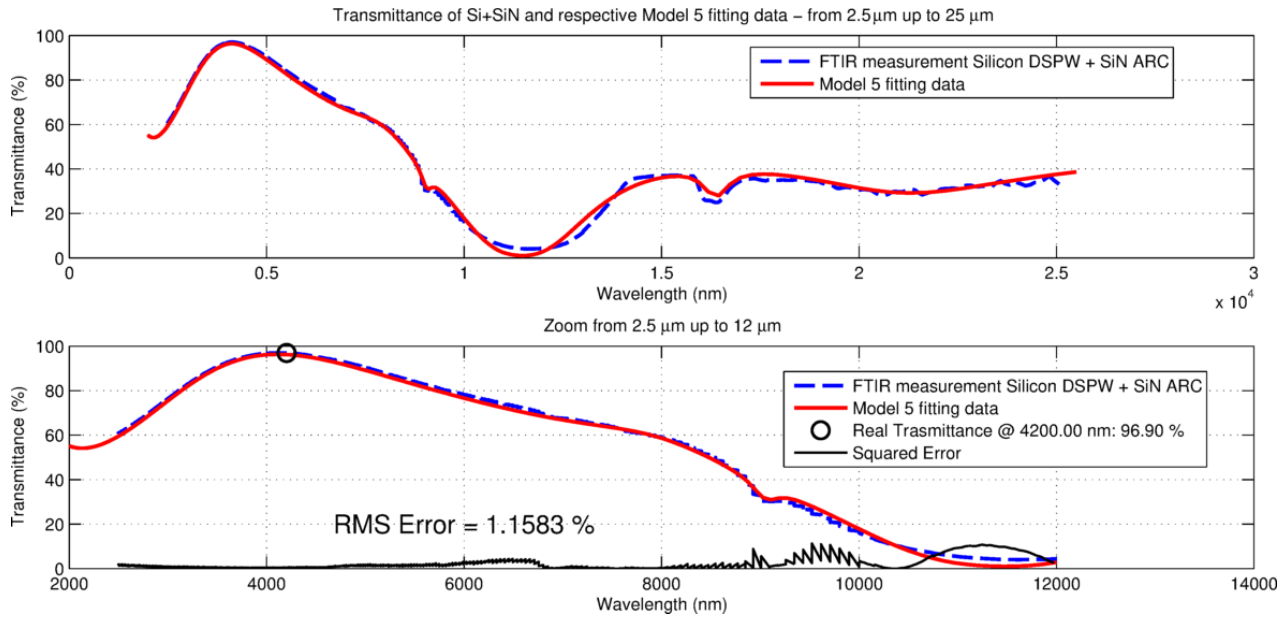

Figure 69: Silicon DSPW plus SiN ARC transmittance and the corresponding M5 fitting data, with RMS error of 1.1583 %.

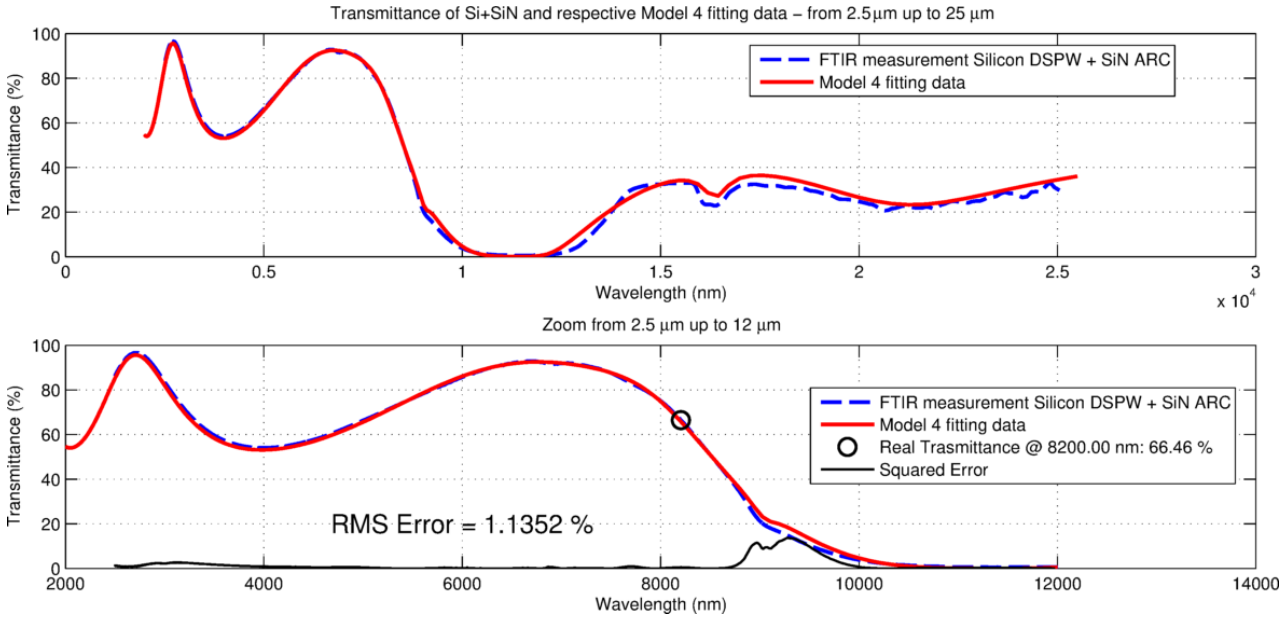

Figure 70: Silicon DSPW plus SiN ARC transmittance and the corresponding M4 fitting data, with RMS error of 1.1352 %.

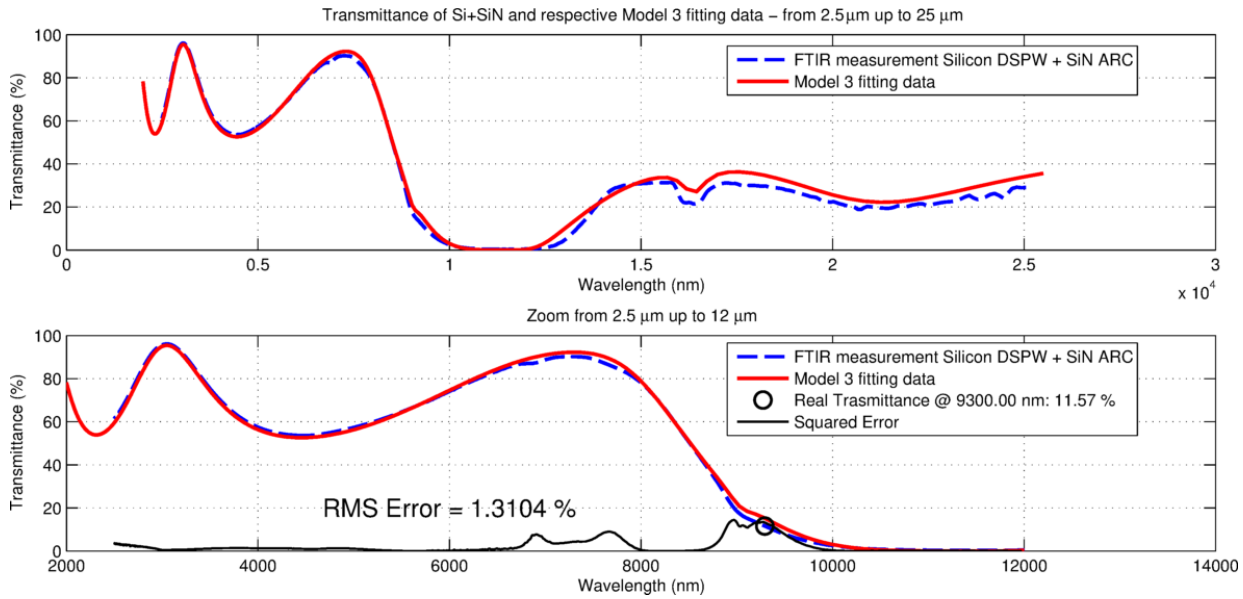

Figure 71: Silicon DSPW plus SiN ARC transmittance and the corresponding M3 fitting data, with RMS error of 1.3104 %.

9. Figure 72 shows the fitted optical constants (n, k) obtained from the SiN complex dielectric function model M1 of Figure 62. The (n, k) table is presented in the Appendix A.

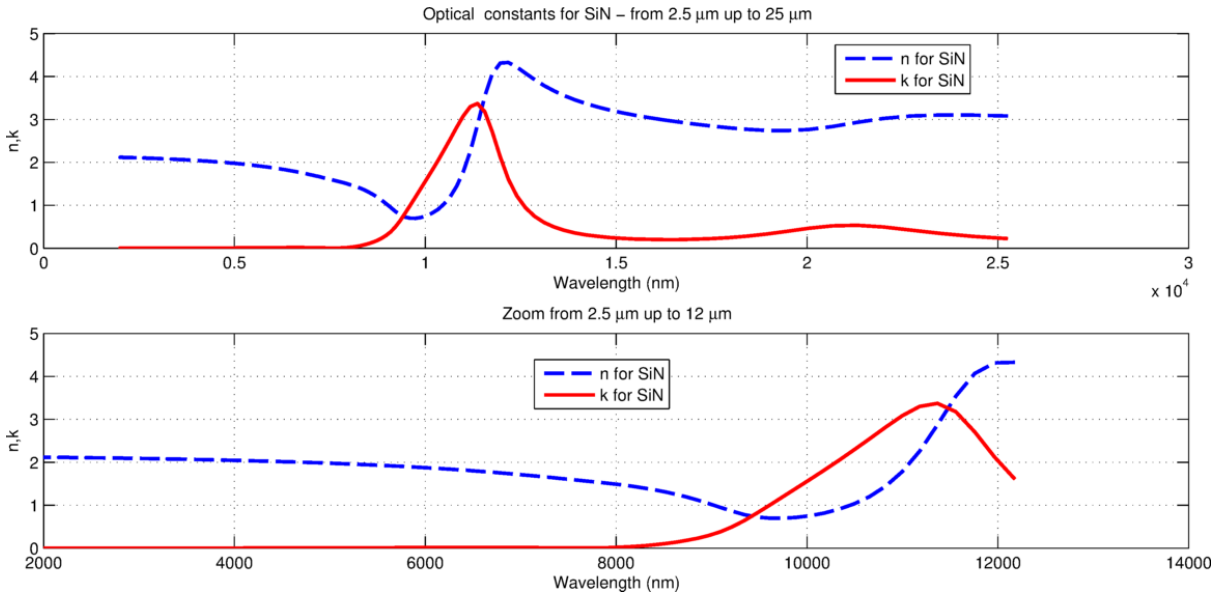

Figure 72: SiN resulting fitted optical constants (n,k).

10. Although RefFIT is a very powerful software, it does not offer the right tools for anti-reflective coatings design. Since our objective is to be able to design new coatings, we have imported the (n, k) dataset from both silicon (M1) and SiN (M2) generated from RefFIT into TFCalc™. Then, we have carried out simulations for the following thicknesses 1095.70 nm, 968.00 nm and 503.77 nm to compare with real and fitted data.

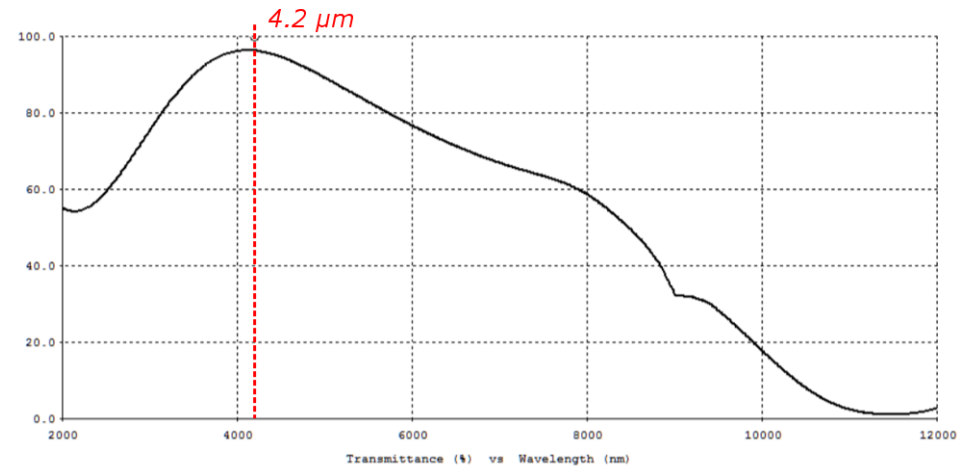

Figure 73: TFCalc™ simulation of 503.77 nm SiN with dataset from RefFIT.

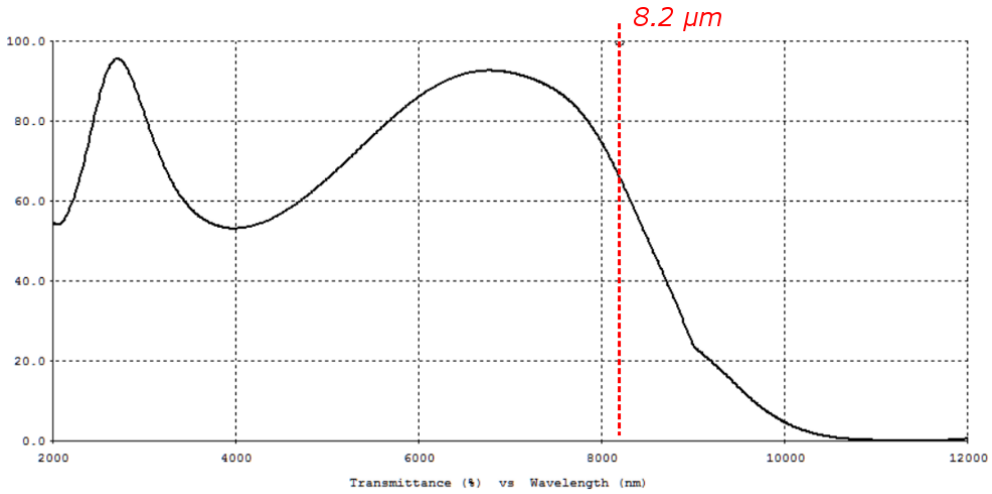

Figure 74: TFCalc™ simulation of 968.00 nm SiN with dataset from RefFIT.

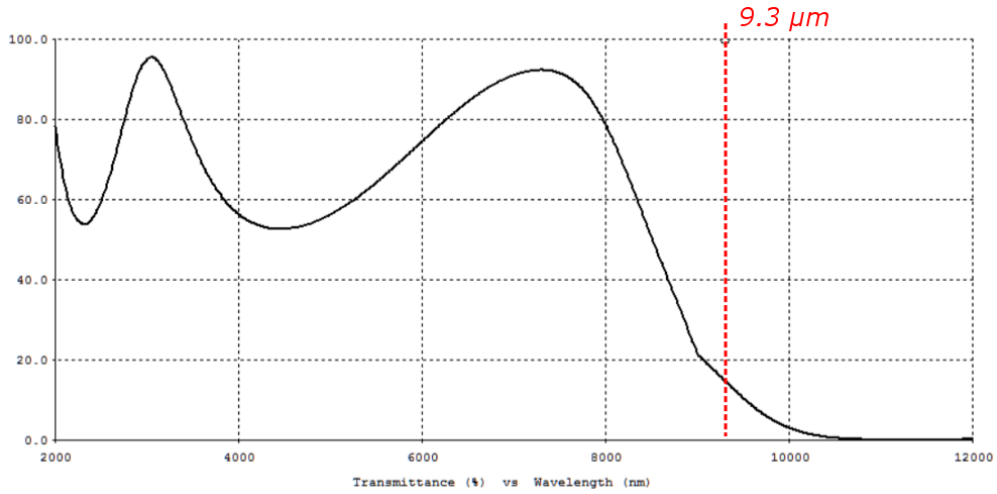

Figure 75: TFCalc™ simulation of 1095.70 nm SiN with dataset from RefFIT.

11. Table 11 shows the comparison between FTIR measured data for the DSPW samples, and simulations in TFCalc™, but now with new dataset from RefFIT. Agreement between

experiment and simulation is by far superior, compared to the previously used (n, k) dataset. Note that thicknesses used in the new simulations are different from the measured ones. This is because these values were the ones with greatest probability of approximating to real data, based on fitting information from RefFIT.

Table 11: Comparison between measured and simulated maximum transmittances for the three different thicknesses with new dataset from RefFIT.

Fitted thickness from RefFIT	Measured	Simulated in TFCalc™
503.77 nm	97.00 % @4.18 µm	96.32 % @4.13 µm
968.00 nm	92.91 % @6.70 µm	92.55 % @6.78 µm
1095.70 nm	90.22 % @7.25 µm	92.27 % @7.29 µm

So far, following all the 11 aforementioned steps, we started from raw FTIR transmittance measurements to the generation of a new dataset of optical constants, being able to reliably simulate in TFCalc™. These are really great results, but still not the end of the line, once our foremost objective is to design ARCs to be deposited over a wafer with the microlenses. The analyzed measurements up to now were for DSPW without any KOH etch. We have observed, however, that the roughness introduced by KOH may alter the optical constants of the stack SiN–Si–SiN to a little extent. Therefore, we need to investigate in RefFIT the situation where FTIR transmittance measurements were performed through a silicon wafer after KOH.

For that purpose, measuring data shown in Figure 61 were used. More specifically, measurements with the light beam at the flat surface of the wafer with the microlenses, for both bare silicon and silicon + SiN. Transmittance curves with the beam at the lenses were ignored, since their average lower profile are might not due to different optical constants, but rather because the lenses end up diverging the incoming light, as previously discussed. Therefore, we have followed the same 11 steps afore-listed, with the difference that we did not change the SiN model (M1), but instead used the same optical data obtained from previous measurements. In other words, we assumed that only the silicon substrate (model M2) would change its optical constants due to roughness. In fact, a new model for the silicon wafer after KOH etching was obtained through χ^2 minimization, as shown in Figure 76. The fitted optical constants (n, k) obtained for Silicon after KOH etching is presented in the Appendix B.

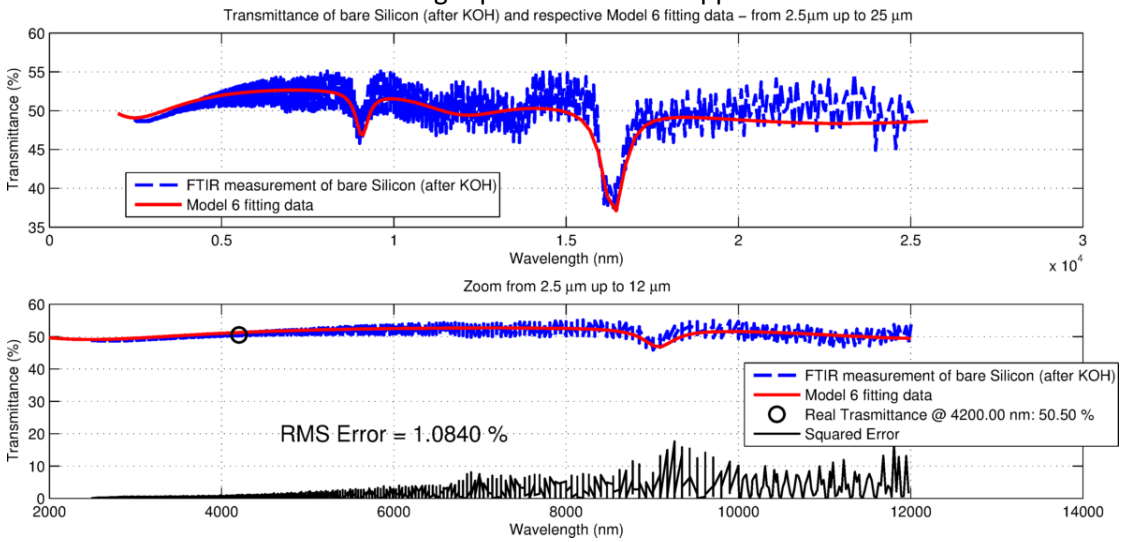

Figure 76: Bare silicon after KOH transmittance and the corresponding M6 fitting data, with RMS error of 1.0840 %.

Then, a special model (M5) for the stack of SiN-Si-SiN with measured thickness of 523 nm was created, and chi-square minimization was performed. We have once again allowed the thickness to vary around the nominal value $\pm 10\ nm$, and a final thickness of 533 nm was found, as shown in Figure 77. Note in M5 that we have used a thickness of 325 µm for the substrate, since we did not protect the back surface of the wafer, resulting in KOH thinning on both sides: $525 - 2 * 100 = 325\ \mu m$. Additionally, this time we have used the new model for silicon, shown in Figure 78.

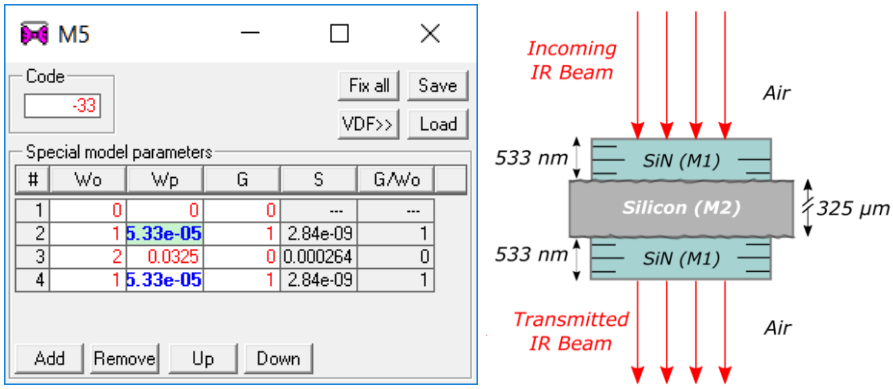

Figure 77: Special model M5 for stack SiN-Si-SiN and its respective final thickness. The cross-section of the sample is also shown.
Source: Adapted from [96].

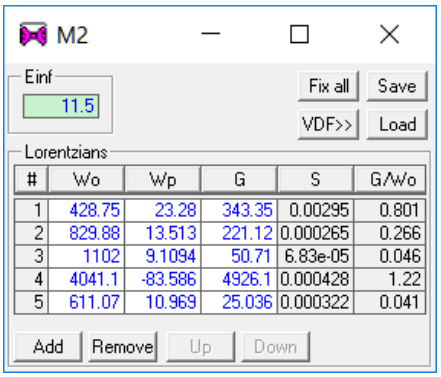

Figure 78: Lorentz Oscillator model for Silicon after KOH etching in RefFIT.
Source: Adapted from [96].

The results for fitting data of M5, compared to the FTIR measurement, are presented in Figure 79. Once again, we did not change the model M1 for the SiN material, but only the model M2 for the silicon material, that can be seen as a 'post-KOH-etching silicon'. Even without altering the SiN model, very good outcomes could be obtained.

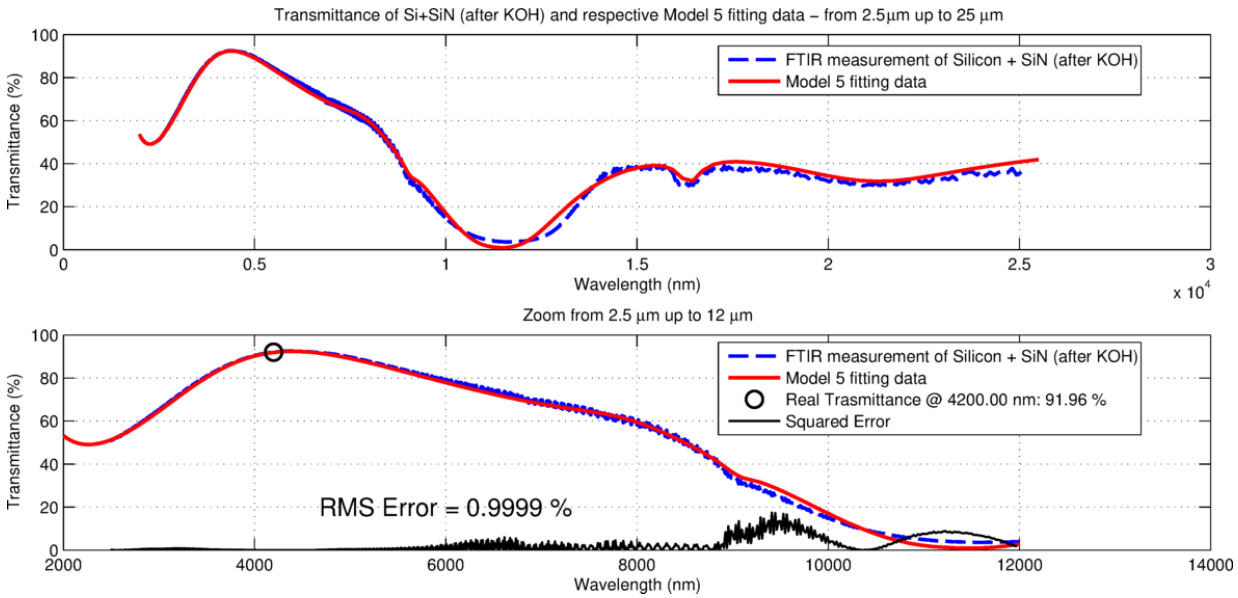

Figure 79: Silicon wafer after KOH etching plus SiN ARC transmittance, and corresponding M5 fitting data, with RMS error of 0.9999 %.

Lastly, the new (n, k) optical constant dataset for silicon (M1), generated in RefFIT, were imported into TFCalc™. Figure 80 shows how simulation from TFCalc™ now can be used to reliably predict the behavior of the SiN ARC, either over a DSPW, or a wafer with the microlenses, which undergone through KOH etching.

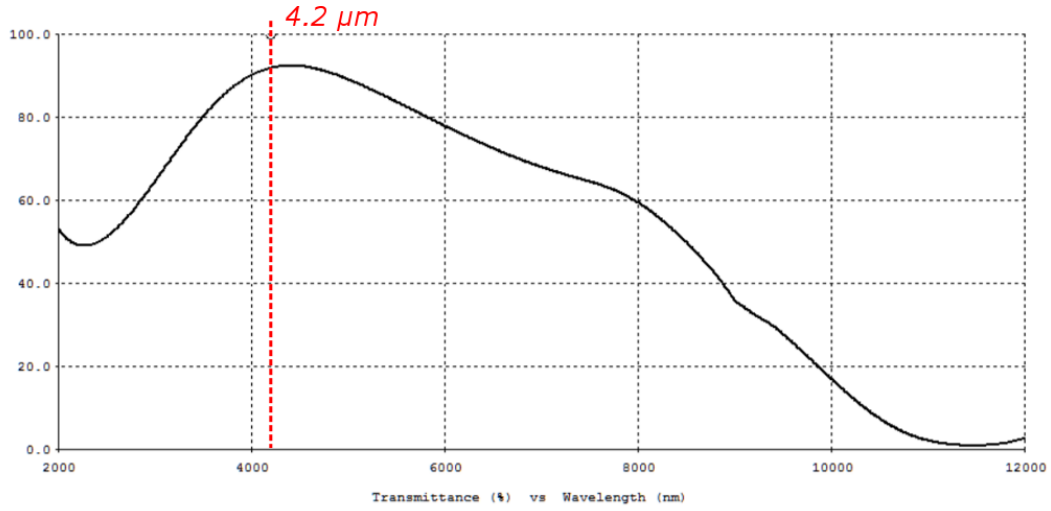

Figure 80: TFCalc™ simulation of 533 nm SiN over microlenses wafer, with dataset from RefFIT.

From Table 12, we can compare the maximum transmittance from FTIR measured data and simulated in TFCalc™. It is clear that the used model is very appropriate for future designs, due to its capability of predicting outcomes with good precision, with a RMS error between the fitted model (M5) and FTIR measurement of less than 1 %.

Table 12: Comparison between measured and simulated maximum transmittances for the SiN wafer with microlenses, with dataset from RefFIT.

Fitted thickness from RefFIT	Measured	Simulated in TFCalc™
533 nm	92.56 % @4.40 µm	92.41 % @4.42 µm

We know from Figure 72 that the extinction coefficient (k) starts to rise near 8 µm, where SiN will begin to absorb. To check the maximum wavelength for which SiN can be used as ARC, we have tried to optimize the thickness of the SiN layer aiming at maximum transmittance at 8 µm. Figure 81 shows the result of this optimization. We can see that it simply cannot reach the aimed wavelength with a peak in transmittance, but rather is held at a specific wavelength lower than 8 µm. The maximum transmittance at the right-hand side is 89.90 % and occurs at 7.63 µm.

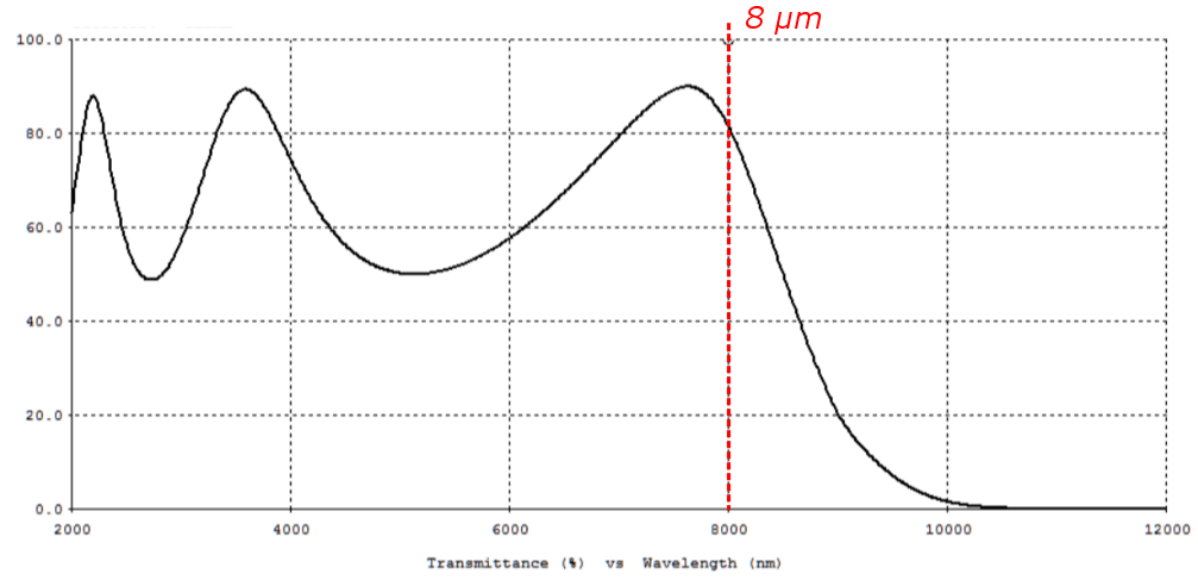

Figure 81: TFCalc™ simulation of 1300 nm SiN over microlenses wafer, with dataset from RefFIT.

It is known that insulators, like the SiN layer, can have its complex dielectric function subdivided in four regions, where the complex dielectric function is described by equation (60) – recalling its parameters as the plasma frequency (ω_p), the resonant frequency (ω_o), and the linewidth (γ), all given in wavenumber in $[cm^{-1}]$ units. As a rule of thumb, the boundaries of each region can be approximated as: $\omega < \omega_o - \gamma/2$ where transmission dominates; $\omega_o - \gamma/2 < \omega < \omega_o + \gamma/2$ for absorption; $\omega_o + \gamma/2 < \omega < \omega_p$ for reflection; and $\omega > \omega_p$ for transmittance again [97], [98]. Figure 82 shows the real (ε_1) and imaginary (ε_2) parts of the dielectric function, given by equation (60) and parameters of the SiN model (M1), presented in Figure 62. The real and imaginary parts presented in the graph were converted to wavelength, in $[\mu m]$ units – therefore, with reversed x-axis, in comparison with the $[cm^{-1}]$ unit.

We can see that absorption is the strongest at the dominant Lorentzian resonant frequency: $\omega_o = 860.9\, cm^{-1}$, or $\omega_o = 11.619\, \mu m$. But most importantly, note that the plasma frequency of this Lorentzian (corresponding to the last row of M1 in Figure 62), is $\omega_p = 1371.3\, cm^{-1}$, or $\omega_p = 7.292\, \mu m$. Any wavelength smaller than ω_p (in $[\mu m]$) is considered to be in the region where transmission dominates, and any wavelength larger than that will result in decreasing transmittance. Although the boundaries of these regions are an approximation, they are corroborated by the TFCalc™ simulation of Figure 81. Therefore, we can conclude that the range at which SiN can be used as an ARC, is limited by the minimum thickness that can be deposited with the technique, up to about a wavelength of 7.6 µm. Fortunately, this range is broad enough for many infrared applications, including the CO_2 gas detection system that will be presented at the end of this thesis, in section 4.4.

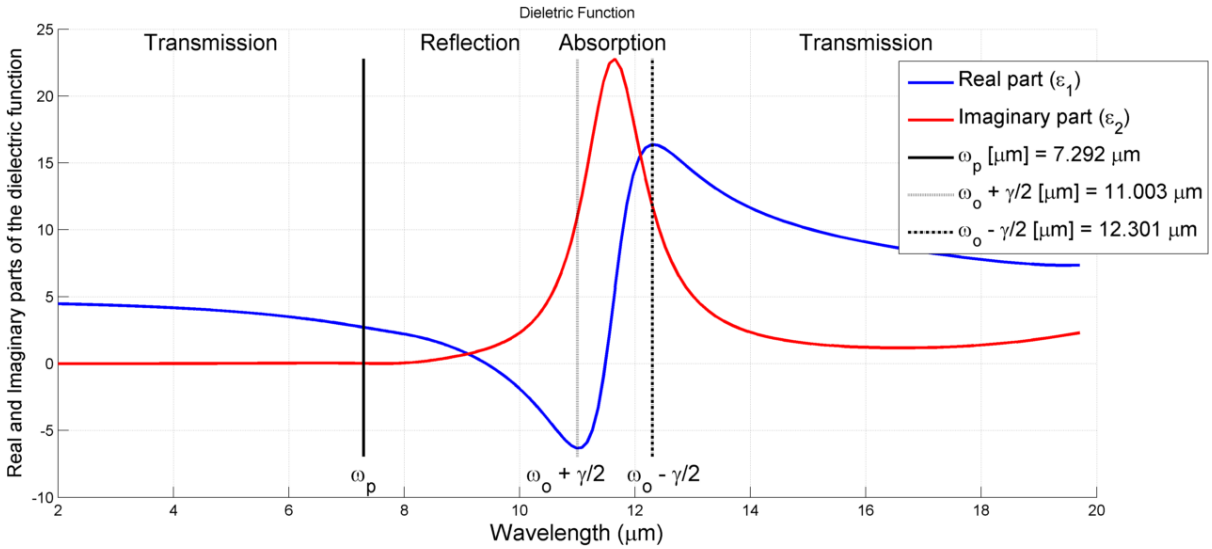

Figure 82: Real and Imaginary parts of the SiN dielectric function.

4.3 Circuit Analysis

4.3.1 Infrared photodetector modelling

So far we have demonstrated good results for the optical part of a gas detection system at a wavelength of 4.2 μm, namely the microlenses and the ARC. The other part of the system comprises the read-out pixel circuit, that is, the bouncing pixel presented in section 3.3. However, in order to detect IR light, one has to interface this circuit with an appropriate detector. For the purpose of this thesis, we have chosen a commercial detector, that was modelled to interface with the bouncing pixel to validate the idea. The chosen one was an *InAsSb* detector, model *P11120-201*, from Hamamatsu [100]. In addition, a light source compatible with a CO_2 detection system was chosen, namely, a Quantum Cascade Laser (QCL), model *L12004-2310H-C*, also from Hamamatsu [101]. The QCL has a peak emission wavelength of $\lambda_{QCL} = 4.33\ \mu m$.

A simplified diagram of the proposed example system – that was not actually implemented – is shown in Figure 83. The laser beam passes through a reference cell filled with a reference gas (N_2), where transmittance through the media does not vary considerably. The sample cell is constantly fed through with the sampling gas, which may or may not contain CO_2. When CO_2 is present, light is absorbed, rendering in lesser optical power falling over the detector. The difference between the detector photocurrent of the reference cell, and the one with the sampling cell is the actual signal to be measured. Also, the amount of light absorbed depends on the concentration of the CO_2 gas, and hence its partial pressure. This way, the detector plus readout circuitry must be able to distinguish between small differences in incoming IR radiation, since the measuring CO_2 gas pressure is unknown. Furthermore, note that the system can be physically composed of two cells – the reference and the sample one – or can have only one cell. In that case, the sample cell needs to be calibrated with the reference gas prior to the actual

measurements (when there is no CO_2), and will look exactly as the reference cell in that case. More information on calibration will be given along this and the next subsection.

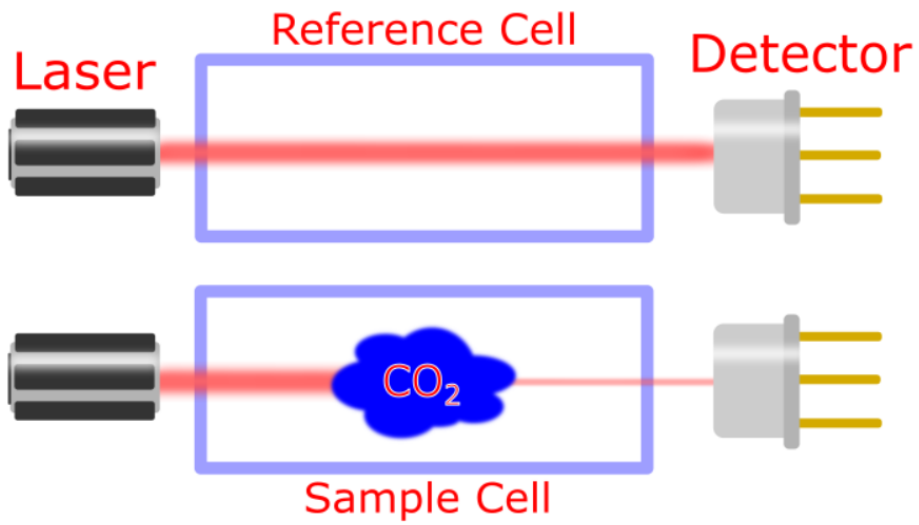

Figure 83: Detection system diagram.

The chosen IR detector comes with a two-stage thermoelectric (TE) cooling system, capable of operating at a minimum temperature of -30 °C, or 243 K, when room temperature is at 25°C. TE cooling is based on the Peltier Effect, that is, when a current flows through a semiconductor, one side of it is heated and the other is cooled. This effect can be used to cool down devices, such as infrared detectors [81]. The performance of the photodetector is greatly improved by cooling it down, for instance increasing its shunt resistance and reducing dark current. Figure 84 shows the dark current vs. reverse voltage across the photodetector for three different temperatures (from the datasheet [100]), namely -30 °C, -10 °C and +25 °C. For the same reverse voltage of $V_{DET} = 20\,mV$, dark current is almost 10 times smaller at -30 °C compared to the detector at +25 °C.

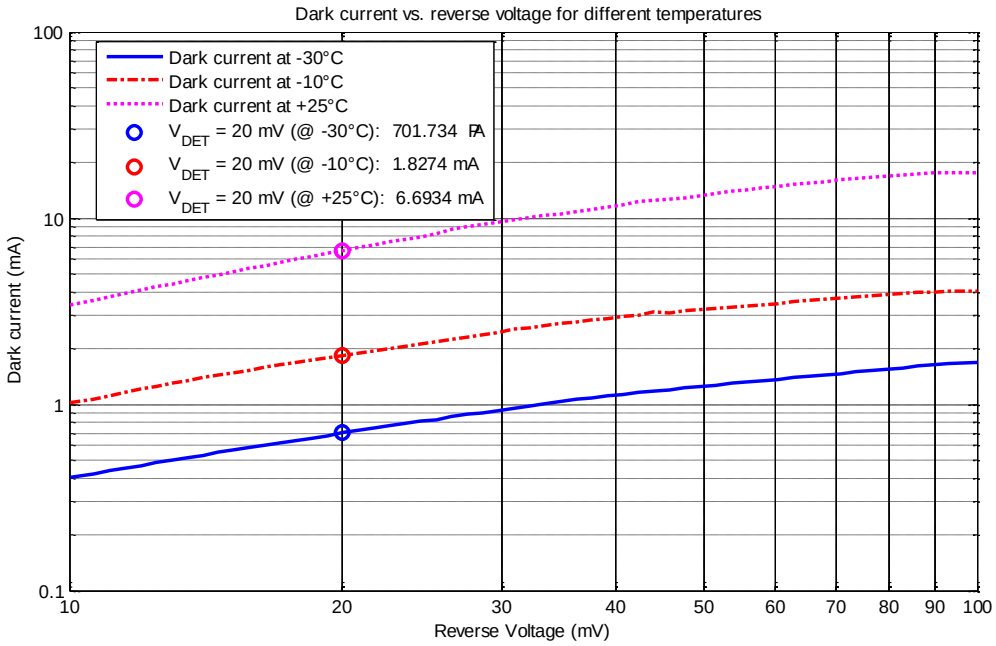

Figure 84: Dark current vs. reverse voltage of the IR detector model P11120-201.
Source: Adapted from [100].

Figure 85 shows how the shunt resistance of the photodetector decreases as temperature rises (from the datasheet [100]). From the graph, at -30 °C, the shunt resistance of the detector (R_0) is approximately 15 Ω, although the datasheet indicates a typical value of 13 Ω (and minimum of 10 Ω) for $V_{DET} = 10\ mV$ [100].

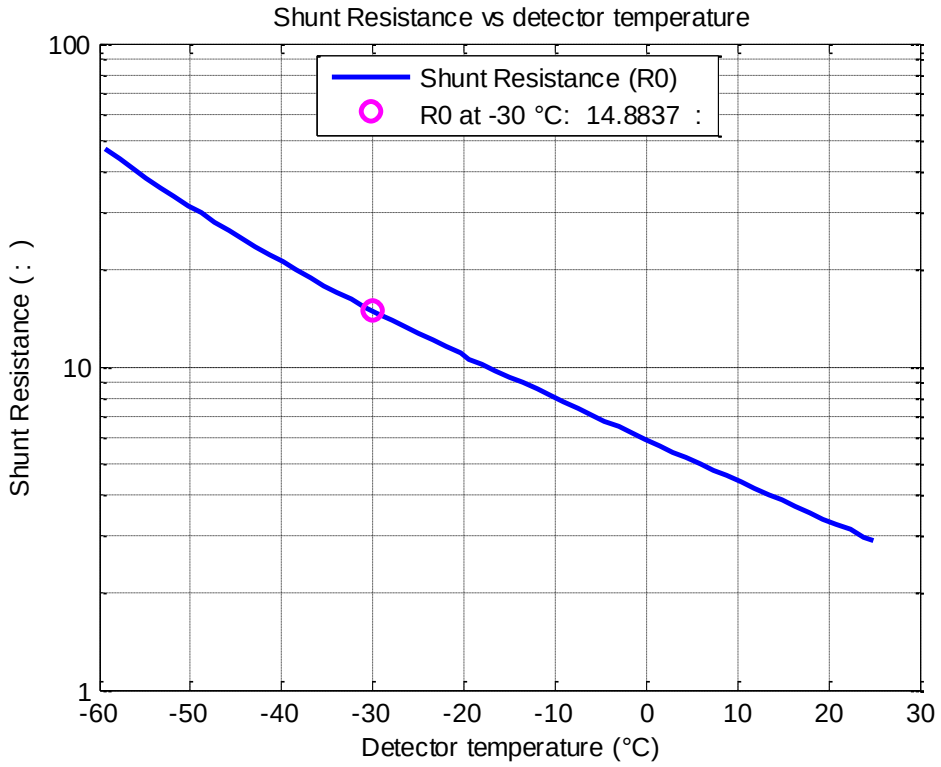

Figure 85: Shunt resistance of the IR detector model P11120-201.
Source: Adapted from [100].

For best performance, the detector should operate at the lowest temperature, and the smallest reverse bias voltage possible. The lowest temperature depends on the Peltier element, and is bounded to $-30\ °C$. Ideally, the lowest reverse voltage is $0\ V$, however, since we do not have information on dark current and shunt resistance for $V_{DET} < 10\ mV$, then we are going to consider the worst case, when $V_{DET} \cong 10\ mV$. In this situation, the shunt resistance is $R_0 \cong 13\ \Omega$ and dark current is $I_{dark} \cong 400\ \mu A$ [100]. From the datasheet, we can find that the IR detector can operate fairly linear for output photocurrents ranging from about 20 µA until 4 mA. In order to know what incident optical power corresponds to this current range, we have to know the responsivity of the photodetector, that is, the photocurrent [A] produced by the detector for a given input optical power [W]. The datasheet of the detector provides information on the spectral response, since the responsivity varies with the wavelength.

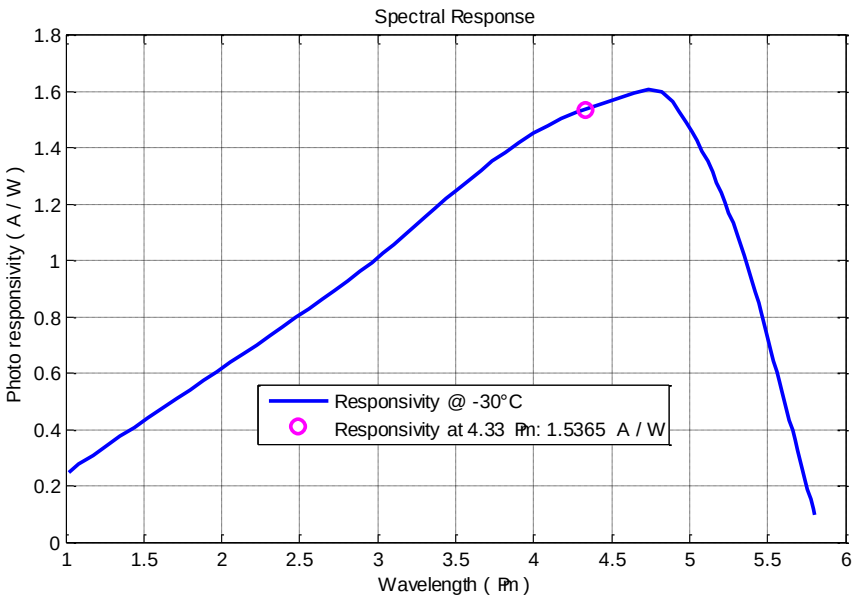

Figure 86: Responsivity of the IR detector model P11120-201.
Source: Adapted from [100].

The spectral response of the IR detector is shown in Figure 86 (from the datasheet [100]), where we can find the responsivity at $\lambda_{QCL} = 4.33\ \mu m$ as approximately 1.54 A/W. The optical power corresponding to the range from 20 µA until 4 mA, is approximately from 13 µW until 2.6 mW. This information is necessary so the laser can be tuned up to yield maximum power arriving on the detector at around 2.6 mW, in order to guarantee that the maximum photocurrent produced by the detector is about 4 mA. Therefore, the maximum current flowing through the detector will be given by $I_{sat} = I_{ph\_max} + I_{dark} = 4\ mA + 400\ \mu A = 4.4\ mA$, while the minimum current is given by the dark current itself: $I_{ph\_min} = I_{dark} = 400\ \mu A$.

From section 3.3.1, we know that for the bouncing pixel, the virtual operating voltage can be extended by a great extent, reaching a value of $V_{R\_max} = 2^{b_c}.\Delta V$, where b_c is the counter number of bits, and $\Delta V\ (= V_{max} - V_{min})$ is the stipulated voltage difference between the upper and lower reference voltages. Also, the maximum allowed sensitivity to be able to detect currents as high as I_{sat} is given by:

$$S_{max} = \frac{V_{R\_max}}{I_{sat}} = \frac{2^{b_c}.\Delta V}{I_{sat}} = \frac{2^8.1}{4.4x10^{-3}} \cong 58.2x10^3 \quad \left[\frac{V}{A}\right] \rightarrow S_{max} = 58.2 \quad \left[\frac{mV}{\mu A}\right] \qquad (61)$$

where we have used $b_c = 8\ bits$, and $\Delta V = 1\ V$. A value of $S_{max} = 58.2\ mV/\mu A$ means that a $1\ \mu A$ variation at the input of the circuit (through branch M_{N1} and M_{P1} of Figure 36) would result in an output voltage variation of $58.2\ mV$ at the integration capacitor. Note that a sensitivity smaller than S_{max} can be used, but will result in worse performance. For instance, for a conventional linear pixel circuit without voltage extension, the maximum sensitivity would be determined by the maximum operating voltage of the circuit (V_{DD}), that is usually very low. For a *180 nm CMOS technology* [25], typical V_{DD} is 1.8 V, and maximum sensitivity would be given by $S_{max} = \frac{V_{DD}}{I_{sat}} = \frac{1.8}{4.4x10^{-3}} \cong 0.41\ mV/\mu A$, meaning that the same $1\ \mu A$ current variation at the input would result in an output voltage too small, smaller than $1\ mV$. This has a direct impact

on the chosen ADC number of bits, since the smallest readable voltage necessary, then the larger the number of bits. This will be covered with more details in next section.

In order to implement such high sensitivity, one might look at equation (38) – of the sensitivity – to find appropriate values for the integration time and capacitance. From equation (49), we have for the worst case when the counter saturates, that the integration time is $\Delta T = 2^{b_c}.t_0$, where $f_{3dB} = 1/t_0$ represents the maximum frequency at which the circuit can operate. For $b_c = 8$ and an initial estimate of the 3dB frequency of the circuit as 1 MHz, then the minimum integration time is $\Delta T = 2^8/1MHz = 256\,\mu s$. Considering a current gain $\beta = 1$ and $\varepsilon = 1$, then:

$$S = \beta.\varepsilon.\frac{\Delta T}{C_{int}} \rightarrow \quad C_{int} = \frac{256x10^{-6}}{58.2x10^{3}} = 4.4x10^{-9} \quad [F] \qquad (62)$$

Data from the datasheet of the photodetector, along with calculated parameters of the bouncing pixel were incorporated in a CMOS integrated circuit (IC) simulation environment. Generally, IC simulation is done by using specific CMOS process technology. In the case of this thesis, a standard CMOS 180 nm technology from TSMC was employed, featuring 1 layer of polysilicon, 6 metal layers, both N-Well and P-Well, and MIM (Metal-Insulator-Metal) capacitors of 2 $fF/\mu m^2$. The *180 nm* value denotes for the smallest manufacturable size in this process, also designated as the *technology node*, as it dictates the minimum possible transistor gate length. As mentioned in section 2.5, transistors can be simulated with very reliable models provided by the foundry, while other components, like the photodetector, need to be additionally modelled. Simulations were performed using the Cadence® Virtuoso® software, with the AMS (Analog Mixed-Signal) Designer simulator.

Figure 87 shows the used equivalent circuit for the photodiode. The voltage across the detector is given by $V_{DET} = V_{FD} - V_{POL}$, where V_{POL} is generally connected to the ground reference (GND). Note that V_{DET} denotes the reverse voltage across the photodetector. The model contains one capacitor, two resistors, one diode, and three current sources. The capacitance of the detector (C_{det}) was estimated based on the detector time constant τ_{det} from the datasheet, for a load resistance of 50 Ω [100], that is: $C_{det} \cong \tau_{det}/50 = 0.4\,\mu s/50\,\Omega = 8\,nF$. Shunt Resistance is $R_0 = 13\,\Omega$, series resistance is $R_s = 10\,m\Omega$ [102], and the diode model has width and length equal to $1\,mm$, which is the size of the photosensitive area. In addition, at the left-hand side, there is the current source responsible for the DC current, i.e., the sum of dark and photogenerated current $\left(I_{DC} = I_{ph} + I_{dark}\right)$. Likewise, at the right-hand side, the two current sources account for the noise in the photodiode, namely the thermal and shot noise.

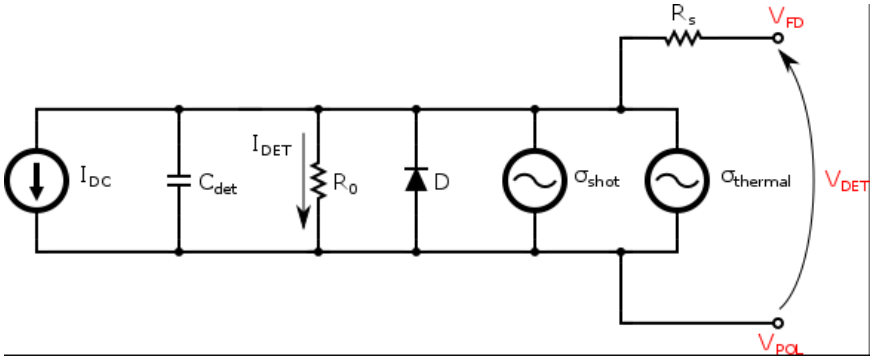

Figure 87: Equivalent circuit of the photodetector.

Noise can be accounted in the circuit in two ways: in a *frequency*- or *time-domain analysis*. Noise analysis in the frequency-domain is performed in a small-signal fashion, assuming that noise is too small to alter the DC operating point of the circuit. This small-signal simplification can be inaccurate when noise is large. Since this can be the case for an IR photodetection system, we have chosen the second option. Time-domain noise simulation is performed as any other usual transient analysis. The difference is that all noise sources inject random noise during each time step of the simulation. This random noise depends on the PSD of the noisy component, on the noise-equivalent bandwidth of the circuit, namely the maximum noise frequency, and on a *seed* value. In fact, simulation is pseudo-random, meaning that for the same seed, transient analysis will be exactly the same, which is convenient for simulation reproducibility. However, to produce different random simulations, different seeds must be used. In conclusion, a sufficiently large number (≥ 30) of recursive simulations – with different seeds – must be performed to extract statistical information of a given output.

Each of the noise current sources in Figure 87 contains information about the modelled white-noise PSD, given by $4kT/R_{eq}$ and $2qI_{DC} = 2q(I_{dark} + I_{ph})$, for thermal and shot noise, respectively, as discussed in section 2.5. These current sources have zero DC current, and standard deviation given by their respective RMS noise, described by equations (7) and (9). The term B in these equations is the noise-equivalent bandwidth, given by $B = f_{3dB}.\pi/2$. Consequently, knowing the bandwidth of the circuit is essential for the transient noise analysis.

4.3.2 Bouncing pixel simulation

4.3.2.1 The cascoded version of the bouncing pixel

In section 3.3.2, the basic pixel circuit architecture was presented. Regarding Figure 36 – of the bouncing pixel unit cell –, it is desirable that current I_2 at the branch M_{P2}-M_{N2} tracks current I_1 at M_{P1}-M_{N1}, that is, they need to be nearly identical. However, even if transistors are equally sized, have the same V_{GS} (Gate-to-Source voltage) and are operating in saturation, I_2 will not track I_1 if the V_{DS} (Drain-to-Source voltage) of the transistors are different. This is due to the channel length modulation (CLM) effect, in which transistors that are in saturation will slightly change their drain currents when V_{DS} varies (Early effect). Consider for instance the mirror formed by M_{P1} and M_{P2}. Transistor M_{P1} is in saturation, and its drain current is given by:

$$I_1 = \gamma.\left(\frac{W}{L}\right)_{MP1}.(V_{SGP1} - V_t)^2.[1 + \chi_{CLM}.V_{SDP1}] \quad [A] \qquad (63)$$

where γ is a constant, W and L are the width and length of the transistor, V_{SGP1} and V_{SDP1} are, respectively, the V_{SG} and V_{SD} of M_{P1}, V_t is the threshold voltage, and χ_{CLM} is the factor accounting for the CLM effect. If M_{P2} is also in saturation, then the ratio between their two currents is:

$$\frac{I_1}{I_2} = \frac{(1 + \chi_{CLM}.V_{SDP1})}{(1 + \chi_{CLM}.V_{SDP2})} \quad [1] \qquad (64)$$

From equation (64), I_2 will track I_1 only if $V_{SDP1} = V_{SDP2}$. The same condition is necessary for current mirror M_{N1}-M_{N2}, if voltage V_{DET} across the photodiode is small. Note that if currents are identical, V_{GS} of the transistors are very close, and transistors are equally sized, then M_{P2} and M_{N1} will always be in saturation, since their peers M_{P1} and M_{N2}, respectively, are diode-connected and operate permanently in saturation. Therefore, each transistor composing the current mirrors needs to be at the same DC operating point of its peer, that is, have the same drain current, V_{GS} and V_{DS}, if equally sized.

The bouncing pixel presented in Figure 36, however, features limitations regarding the tracking of currents for infrared photodetectors, mainly due to the low shunt resistance R_0 – shown in Figure 87 – and the relatively high values of current from the photodetector. The low R_0 implies in smaller injection efficiency (ε), causing a portion of the current I_{DET} to flow through the parasitic detector shunt resistance. For a given ε – as in equation (21) and Figure 18 – the larger the photocurrent, the larger I_{DET}, and consequently, the higher the absolute voltage across the detector. However, if V_{DET} is not close to zero, then transistors M_{N1} and M_{N2} will have different V_{GS}, rendering in different operating points. In other words, for such detector, $I_1 \neq I_2$ and V_{DET} will be away from zero.

The double-current mirror structure of Figure 36 helps trying to make V_{GS} to be approximately equal. If we can assure that the V_{DS} are also the same, then transistors will be at the same operating point, and I_2 will track I_1. A common way of making this happen is by employing a cascoded current mirror. Therefore, we have modified the original CMI block of the bouncing pixel from Figure 36 [31], adding a cascode stage to the double-current mirror, as in Figure 88. In this circuit schematic, the anode of the photodiode is connected to an external bias V_{POL}, where $V_{DET} = V_{FD} - V_{POL}$.

Since transistors M_{P1}-M_{P2} have always the same V_{SG}, it is easy to ensure that their V_{SD} are equal by employing the cascode stage. For M_{P1}, the source-to-drain voltage is $V_{SDP1} = V_{DD} - V_{P1}$, while for M_{P2} it is $V_{SDP2} = V_{DD} - V_X$. Therefore, $V_{SDP1} = V_{SDP2}$, only if $V_X = V_{P1}$, that is, if M_{P2} is also in saturation. At the node of V_{P2}, we have that $V_{P2} = V_X - V_{SGP7} = V_{P1} - V_{SGP6}$, where V_{SG6} and V_{SG7} account for the source-to-gate voltage of M_{P6} and M_{P7}, respectively. When $V_{SG6} = V_{SG7}$, then the wanted condition is achieved, and we ensure that both M_{P1} and M_{P2} have the same drain-to-source voltage. It is known that this condition is reached with proper sizing of the transistors, namely when: $\frac{(W/L)_{MP1}}{(W/L)_{MP2}} = \frac{(W/L)_{MP6}}{(W/L)_{MP7}}$. By employing all these transistors with the same W/L, it is possible to assure that current I_2 will track I_1 very closely.

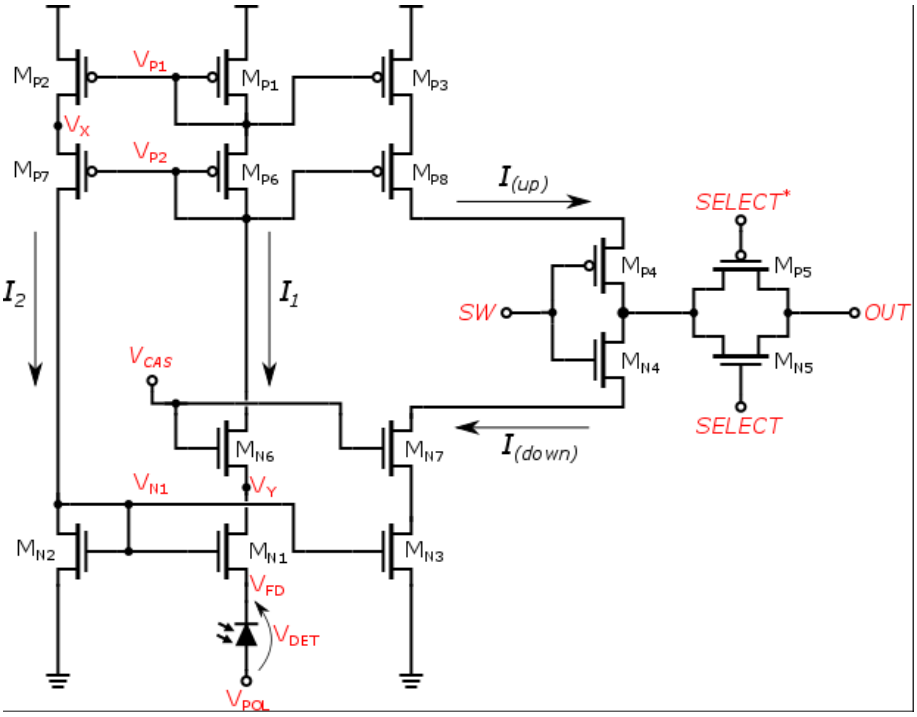

Figure 88: Bouncing pixel circuit with cascoded current mirrors.

In order to make the voltage across the detector as close to zero as possible, another kind of cascode stage was introduced, with transistor M_{N6} with gate bias equal to V_{CAS}. This bias is chosen so that the drain-to-source voltage of M_{N1} and M_{N2} are almost the same, i.e., $V_{DSN1} \cong V_{DSN2}$. In this case, the NMOS transistors are at the same DC operating point, that is, have the same drain current, V_{DS} and V_{GS}, which renders in very small V_{DET} (considering $V_{POL} = 0\ V$). This condition however, is only achievable for a given value of drain current. We could remove the need to bias the circuit with V_{CAS} if we have employed a cascode stage like the one with M_{P7} and M_{P6}. However, preliminary simulations did not yield satisfactory results for this case, and hence we have decided to employ the stage with M_{N6} and V_{CAS}, since it resulted in better performance.

We have found an optimal bias of $V_{CAS} = 1.152\ V$ (through using a parametric sweep), so that V_{DS} of M_{N1} and M_{N2} are almost the same, for the highest current flowing through the circuit, namely, for $I_1 = I_2 = 4.4\ mA$. When drain current is reduced until it reaches the minimum value of $I_{dark} = 400\ \mu A$, then the transconductance of M_{N1} will also be reduced, since it is given by $g_{mn1} = 2.I_1/(V_{GSN1} - V_t)$. As a consequence, the input resistance of the circuit will increase, making more current to flow through the parasitic shunt resistance of the detector, which in turn increases V_{DET}. Since the relation between the gate-to-source voltage of M_{N1} and M_{N2} is given by $V_{GSN1} = V_{GSN2} - V_{FD}$, or equivalently, $V_{GSN1} = V_{N1} - V_{DET}$ (for $V_{POL} = 0\ V$), and $I_1 \cong I_2$, then the higher the voltage across the photodiode, the smaller will be V_{GSN1}. On the other hand, for a given current I_1, if V_{GSN1} decreases, then V_{DSN1} has to increase. This will keep transistor M_{N1} always in saturation, since for the maximum drain current, $V_{DSN1} = V_{GSN1}$, while for the minimum value, $V_{DSN1} > V_{GSN1}$, and the condition for saturation is $V_{DSN1} \geq V_{GSN1} - V_t$.

Transistors M_{P8} and M_{N7} were employed in order to increase the output resistance of the mirroring stage. To illustrate the effect of the cascode stages on the bouncing pixel, two sets of simulations were performed: one with the conventional bouncing pixel of Figure 36, and the other with the cascoded current mirrors of Figure 88. A parametric sweep from $I_{dark} = 400\ \mu A$

until $I_{sat} = 4.4\,mA$ was performed for both circuit topologies, measuring the voltage V_{DET} across the photodetector (with $V_{POL} = 0\,V$), and the difference between the two branch currents, namely $\Delta I = I_2 - I_1$. Transient simulations had a reset time of 256 µs (calculated in section 4.3.1), and the average values of V_{DET} and ΔI along this time interval for each input photocurrent was measured. These simulations were carried out without taking noise into account, since the purpose is to compare the DC condition between the circuits.

Table 13: Comparison of ΔI and V_{DET} between the bouncing pixel with and without the cascoded current mirror stage, for different input DC currents $I_{DC}\ (= I_{ph} + I_{dark})$.

$\Delta I = I_2 - I_1 \quad [\mu A]$		$I_{DC} \quad [\mu A]$	$V_{DET} \quad [mV]$	
Cascoded	***Conventional***		***Cascoded***	***Conventional***
1.93	133.10	*420*	2.70	13.31
1.96	133.30	*435*	2.68	13.27
2.00	133.60	*462*	2.63	13.19
2.07	134.20	*510*	2.56	13.06
2.18	135.10	*594*	2.43	12.82
2.37	136.30	*742*	2.25	12.40
2.64	137.10	*1003*	1.94	11.69
2.97	135.50	*1464*	1.51	10.48
3.18	124.20	*2278*	0.94	8.43
2.66	83.96	*3713*	0.19	4.87
1.71	40.00	*4400*	-0.27	2.11
2.33 ± 0.47	***120.58 ± 30.78***	***Average***	***1.78 ± 1.06***	***10.51 ± 3.82***

Table 13 shows that the cascoded version of the circuit presents much better results than its conventional counterpart. The difference ΔI between the current at the branches of the CMI block is minimal for the cascode version, being 2.33 µA on average throughout the expected photocurrent range. For the conventional one the mean is around 120.58 µA, a much larger difference, and not negligible if compared to the minimum of 400 µA flowing through the circuit. These large values of ΔI result in a circuit in which transistors of the same current mirror are not at the same DC operating point. As a consequence, the voltage V_{DET} across the photodetector will be larger, and so will be the parasitic current I_{DET}. From Table 13, although the conventional circuit yields an average V_{DET} that is not so high, the cascoded version achieves the goal of keeping the reverse voltage as low as possible, and in addition, lower than $10\,mV$, necessary to guarantee that both the shunt resistance and dark current are kept close to the expected values, as discussed in previous section. Beyond that, the improved version shows a lower spreading of V_{DET} around the mean. Of course, the performance enhancement comes at the cost of a little bit more of complexity added to the circuit, with the addition of the cascoded transistors.

In fact, keeping V_{DET} as small as possible is a key prerequisite for infrared detectors with low R_0. Assuming that $V_{DET} = 10\,mV$, and $R_0 = 13\,\Omega$, then the current at the shunt resistance will be as high as $I_{DET} = V_{DET}/R_0 = 769.23\,\mu A$. Depending on the V_{DET} sign, I_{DET} can be either in the direction of being subtracted from the photodetector current $I_{DC}\ (= I_{ph} + I_{dark})$, or added to it. In other words, it can make injection efficiency either lower or higher than unity. Note that the best condition is when $\varepsilon \to 1$, and any other situation will result in nonlinearity to the circuit, since the integrated current in the bouncing pixel will not have a linear relationship with the input photodetector current.

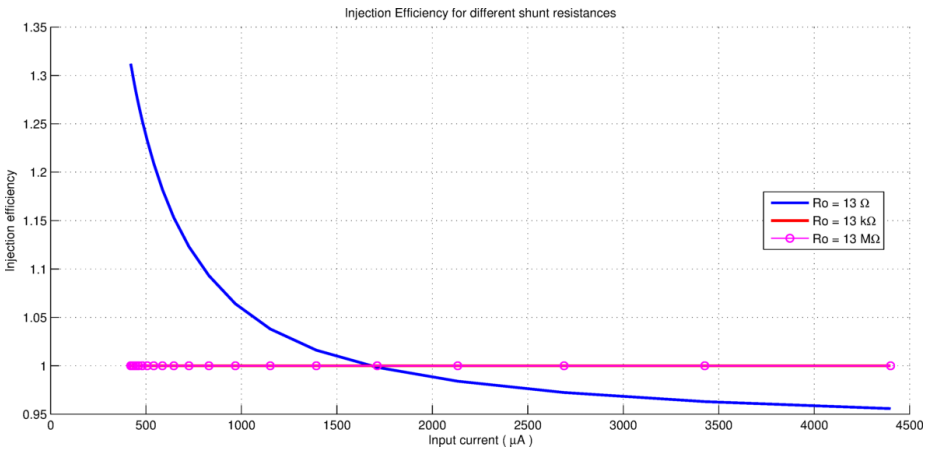

Figure 89: Shunt Resistance effect on injection efficiency.

Figure 89 illustrates the effect of the shunt resistance on injection efficiency, where the bouncing pixel fulfills its role of keeping ε close to the unity for higher shunt resistances. However, for the lowest value of R_0, ε does not keep steady at one, resulting in nonlinearity. Note that for a given photodetector material, the product $R_0 A$ is constant. Since this is a large detector, then its R_0 will consequently be smaller. If much smaller photodetectors were employed, for instance for matrices or linear-array applications, then the injection efficiency of the bouncing pixel would probably be near 100 %. Table 14 compares the cascoded bouncing pixel for two different values of R_0, namely for 13 Ω and 13 MΩ. In fact, for the highest value of R_0, both ΔI and V_{DET} are maintained closer to zero, if compared to the lowest R_0. That means that the cascoded bouncing pixel behaves properly for detectors with higher shunt resistances, but still works very well even for lower shunt resistances, since a decreasing in R_0 of about $10^6\ \Omega$, results in a detector bias only 3 times larger.

Table 14: Comparison of ΔI and V_{DET} between the cascoded bouncing pixel with high and low detector shunt resistances.

$\Delta I = I_2 - I_1$ $[\mu A]$		I_{DC} $[\mu A]$	V_{DET} $[mV]$	
$R_0 = 13\ \Omega$	$R_0 = 13\ M\Omega$		$R_0 = 13\ \Omega$	$R_0 = 13\ M\Omega$
1.93	1.02	*420*	2.70	2.07
1.96	1.04	*435*	2.68	2.02
2.00	1.06	*462*	2.63	1.95
2.07	1.11	*510*	2.56	1.82
2.18	1.17	*594*	2.43	1.63
2.37	1.24	*742*	2.25	1.33
2.64	1.28	*1003*	1.94	0.91
2.97	1.13	*1464*	1.51	0.33
3.18	0.44	*2278*	0.94	-0.47
2.66	-1.70	*3713*	0.19	-1.77
1.71	-3.86	*4400*	-0.27	-2.71
2.33 ± 0.47	***0.36 ± 1.64***	***Average***	***1.78 ± 1.06***	***0.65 ± 1.64***

One way of improving the linearity of the injection efficiency for the low R_0, while keeping ΔI and V_{DET} approximately at the same level, is by tuning the circuit with the bias V_{POL} at the anode of the photodiode. From Table 14, we can note that V_{DET} for $R_0 = 13\ \Omega$ is the highest

when input current is the lowest. This results in the highest parasitic current I_{DET}, and therefore makes the injection efficiency the furthest from the unity, since $\varepsilon = 1 + \frac{I_{DET}}{I_{DC}}$. The largest I_{DET} (hence the highest V_{DET}), should occur for the largest I_{DC}, so that the deviation from the unity in ε is the lowest.

Values of V_{DET} in Table 14 were taken for $V_{POL} = 0\ V$, therefore they are exactly equal to V_{FD}, since $V_{DET} = V_{FD} - V_{POL}$. By increasing V_{POL}, the voltage V_{DET} will decrease, since the bouncing pixel circuit tries to maintain V_{FD} constant for a given DC condition. As a consequence, by making V_{POL} near 2.7 mV – the highest value of V_{DET} in Table 14 for 13 Ω – will probably result in a lesser variation along the injection efficiency curve of Figure 89. Some simulations were performed by varying V_{POL}, and the injection efficiency was calculated. Figure 90 shows the effect of V_{POL} on the injection efficiency.

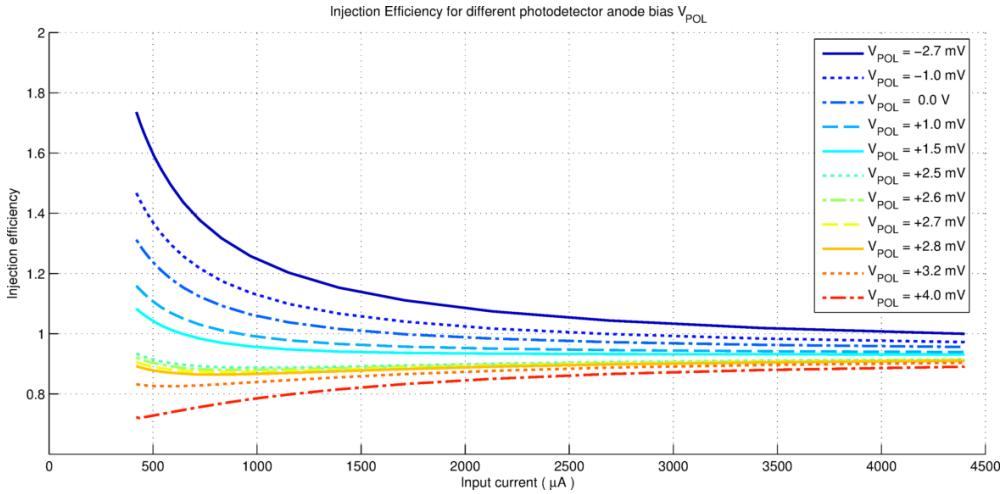

Figure 90: Effect of V_{POL} on the injection efficiency.

To maintain ε as constant as possible is more important than to increase its value, since this may interfere in the circuit linearity. In addition, if ε is lower than one, but is stable, the "current loss" to the parasitic resistance R_0 can be compensated with the bouncing pixel internal gain β. Therefore, among the curves of Figure 90, the one that yields the best performance is the curve that looks more like a constant line. To find which of them is the best one, each curve was approximated as a flat line, as presented in Figure 91.

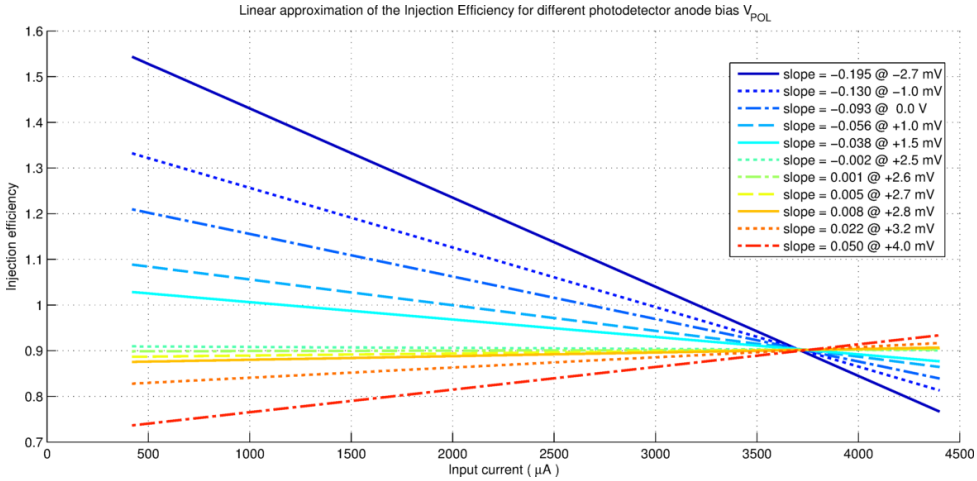

Figure 91: Linear approximation of injection efficiency.

From Figure 91, the curve with the minimum absolute slope is the one that resembles more a constant slopeless line. This happens for $V_{POL} = 2.6\,mV$, whose ε curve has a mean of $\varepsilon_{mean} \cong 0.91 \pm 0.08$, i.e., roughly 90 % of the input current from the IR photodetector will be integrated in the bouncing pixel, while the other 10 % will be lost at the parasitic shunt resistance.

Table 15: Comparison of *ΔI* and V_{DET} for V_{POL} equal 2.6 mV and for V_{POL} equal 0 V.

$\Delta I = I_2 - I_1 \quad [\mu A]$		$I_{DC} \quad [\mu A]$	$V_{DET} \quad [mV]$	
$V_{POL} = 0\,V$	$V_{POL} = 2.6\,mV$		$V_{POL} = 0\,V$	$V_{POL} = 2.6\,mV$
1.93	0.98	*420*	2.70	-0.43
1.96	1.00	*435*	2.68	-0.47
2.00	1.02	*462*	2.63	-0.53
2.07	1.06	*510*	2.56	-0.64
2.18	1.12	*594*	2.43	-0.82
2.37	1.20	*742*	2.25	-1.10
2.64	1.27	*1003*	1.94	-1.50
2.97	1.22	*1464*	1.51	-2.07
3.18	0.70	*2278*	0.94	-2.85
2.66	-1.09	*3713*	0.19	-4.06
1.71	-2.96	*4400*	-0.27	-4.95
2.33 ± 0.47	***0.50 ± 1.33***	***Average***	***1.78 ± 1.06***	***-1.77 ± 1.56***

Table 15 summarizes the results of ΔI and V_{DET}, comparing the cascoded bouncing pixel with V_{POL} at 0 V, and when set to 2.6 mV. In fact, by applying 2.6 mV, the maximum absolute voltage across the detector increases reaching almost 5 mV, but still smaller than 10 mV. The difference between currents I_2 and I_1 inside the bouncing pixel actually diminishes, which is beneficial to a better performance. This is because the average voltage V_{FD} $(= -1.77 + 2.60 = 0.83\,mV)$ at the source of M_{N1} in Figure 88 is closer to the ground reference than it was with zero bias.

However, to bias the circuit with such small voltage can be quite difficult. A precision voltage source can be used if available, either in the same chip as the bouncing pixel, or an external one. Note that the pixel circuit will keep working without this bias, but with more non-linearity. Also, for other photodetectors with larger shunt resistances, this precision bias would not be necessary.

4.3.2.2 AC Analysis

In the last subsections, the optimal parameters for DC operation were discussed, while this subsection will present the AC analysis of the bouncing pixel. Since noise will be introduced in the time-domain simulation, it is important to know the noise-equivalent bandwidth B of the circuit, obtained through the AC analysis.

There are two major factors limiting the maximum operational frequency of the circuit, namely, the photodetector capacitance and the integration capacitance. The first one actually restricts the minimum integration time, i.e. the lowest frequency at which the circuit may

operate. If current is too small, the integrated voltage will take a time longer than 5τ to integrate from V_{min} until V_{max}, where τ is the time constant of the photodetector, defined by its internal capacitance. Therefore, the voltage profile will reach a steady level before being detected by the comparator, and no bounce will ever occur. The latter limits the maximum frequency of operation, i.e. the bandwidth or, equivalently, the f_{3dB} frequency. For this case, we can assume that the circuit has just one pole limiting its bandwidth, that is, the noise-equivalent bandwidth will be given by $B = f_{3dB}.\pi/2$, where f_{3dB} is the bandwidth cut-off frequency [27].

In order to find the frequency response of the circuit, the simulation testbench was configured to perform an AC analysis. Switching signals like RESET and SELECT were set to be always at V_{DD} – so that the capacitor is free to integrate – allowing just a small-signal AC current source to operate. First, we have placed the current source in parallel with the photodetector of Figure 88. Then, we tried to measure the frequency response at the voltage node of the capacitor that is placed just outside the pixel, as in Figure 37. However, we have noticed that it does not yield the expected results, that is, it does not present a regular AC response with poles and zeros, but rather an almost flat line. This is probably due to the up-down switch that is in the middle of the path of the current, or due to the high capacitances of the MOSFETs. However, by setting the signal SW, for instance, always to logic high, we could successfully measure the current $I_{(up)}$ at the drain of transistor M_{P4} in the up-down switch.

Figure 92 shows the AC response for the CMI block for two input DC currents. The circuit considered for this calculation is also presented in the graph. The curves can be described as the transfer function $|H_{CMI}(\omega)|$ between the input small-signal current source $I_{AC}(\omega)$ and the AC measured current $I_{(up)}(\omega)$, that is, $|H_{CMI}(\omega)| = |I_{(up)}(\omega)/I_{AC}(\omega)|$. Both curves are normalized by their DC value, i.e., they are divided by $|H_{CMI}(0)|$. The f_{3dB} frequencies of both curves are 4.36 MHz and 24.12 MHz for the DC currents of 400 µA and 4.4 mA, respectively. This measurement, however, just takes into account the transfer function of the bouncing pixel, without the transfer gate and the external capacitor. Since the output node capacitance is quite large, then it might limit the circuit bandwidth, as mentioned before.

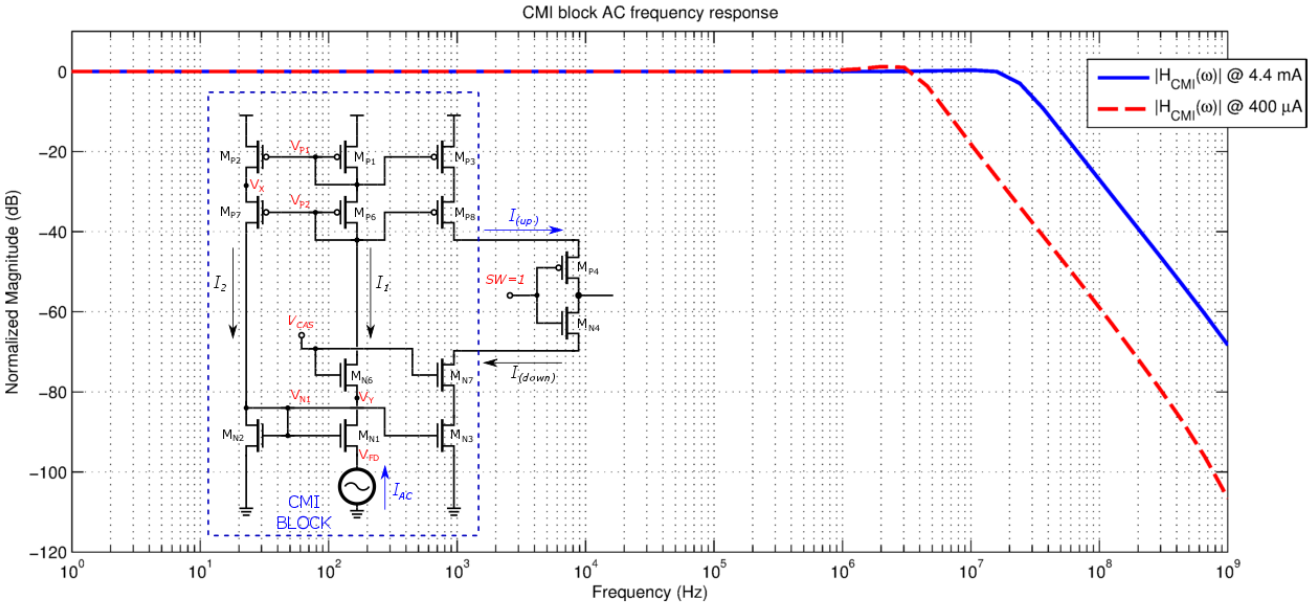

Figure 92: Normalized frequency response of the CMI block and circuit schematics.

To account for the frequency response of the output stage, i.e., for both transfer gate and external capacitor, we have moved the small-signal current source to be positioned after the up-down switch of the bouncing pixel. This way, we can evaluate how this stage limits the maximum frequency. Figure 93 shows the AC response of the output stage, where $\left|H_{output}(\omega)\right|$ represents the transfer function – normalized by $\left|H_{output}(0)\right|$ – between the input current source $I_{AC}(\omega)$ and the AC measured voltage $V_{CAP}(\omega)$, or equivalently, $\left|H_{output}(\omega)\right| = |V_{CAP}(\omega)/I_{AC}(\omega)|$. The f_{3dB} frequencies for 400 µA and 4.4 mA are, respectively, $344\ kHz$ and $1.58\ MHz$. As expected, the cut-off frequencies for this stage are smaller than for the CMI block. For the transient simulations, whatever the input current, we should consider the worst case, that is, the situation where the noise-equivalent bandwidth is the largest. Although the cut-off frequency is larger for the CMI block, the output stage will limit the whole circuit bandwidth. Then, the worst case happens for $f_{3dB} = 1.58\ MHz$, which results in $B = f_{3dB}.\pi/2 = 2.48\ MHz$. This value will be used in next sections for the noise-equivalent bandwidth in time-domain noise simulations.

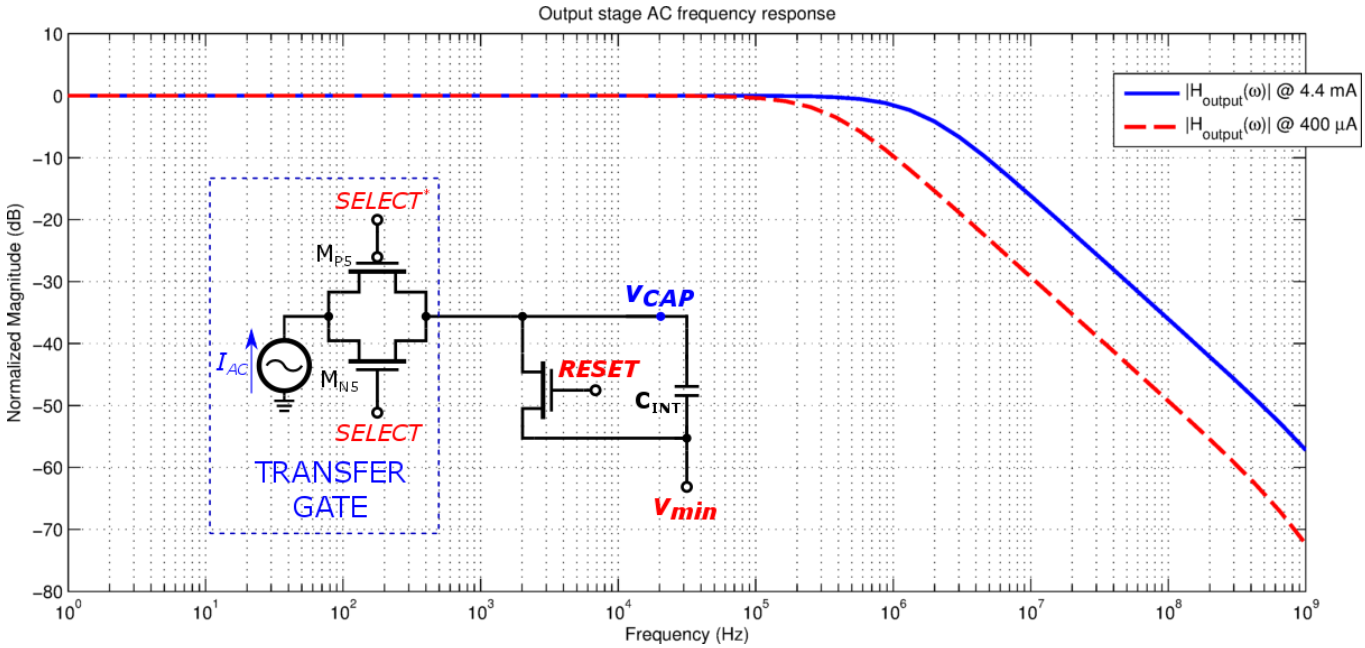

Figure 93: Normalized frequency response of the output stage and circuit schematics.

4.3.2.3 Time-domain transient noise analysis

The aim of this subsection is to simulate the circuit for the whole intended dynamic-range (DR), namely from 400 µA until 4.4 mA, extracting the achieved sensitivity, DR and SNR. For these simulations, each DC input current was integrated during a time interval of ΔT, producing a corresponding output voltage. However, since a noise component is now included, a given DC current will produce a set of output voltages, where each value of this set is produced by a different seed. The range of voltages is centered around an average value, and spread around it depending on the RMS noise in the circuit plus detector. The total number of different seeds used for simulations impacts directly on the accuracy of the measured RMS voltage variation.

For the following transient simulations, since the calculated f_{3dB} frequency from previous section is larger than the value of $1\ MHz$ estimated in section 4.3.1, then the circuit is capable of operating with $\Delta T = 256\ \mu s$. This way, we do not need to recalculate the capacitance, and we may employ $C_{int} = 4.4\ nF$. Table 16 compiles the most relevant parameters used for the following set of simulations. Some parameters, like the ADC number of bits, were included after the initial simulations, as will become clear later.

Table 16: Most important parameters used for transient simulations.

Parameter Name	*Abbreviation*	*Value*
CMOS Technology maximum voltage	V_{DD}	$1.8\ V$
Integration time	ΔT	$256\ \mu s$
Integration capacitance	C_{int}	$4.4\ nF$
NMOS cascode stage bias	V_{CAS}	$1.152\ V$
Bouncing pixel current gain	β	1
Average injection efficiency	ε_{mean}	~ 0.9
Detector anode bias	V_{POL}	$2.6\ mV$
Temperature	T	$-30\ °C$
Noise-equivalent bandwidth	B	$2.48\ MHz$
Upper reference voltage	V_{max}	$1.3\ V$
Lower reference voltage	V_{min}	$0.3\ V$
Voltage-difference boundaries	ΔV	$1\ V$
Detector Shunt-resistance	R_0	$13\ \Omega$
Dark current	I_{dark}	$400\ \mu A$
Number of seeds	$\#_{seeds}$	30
Counter number of bits	b_c	*8 bits*
ADC number of bits	b_a	*9 bits*
ADC reference voltage	V_{REF}	$1.8\ V$

In order to correctly perform all simulations, the width of the transistors needed to be extended to almost their maximum sizes – noticing that in the context of integrated circuits, the transistors feature a maximum recommended width that can be properly manufactured, for a given technology. Transistors with larger widths were necessary to handle larger currents ($\sim mA$), and to increase the injection efficiency ε of the circuit – since it depends on the input resistance, which in turn is directly affected by the transconductance of M_{N1}. Remember that all current that is not being injected into the circuit is flowing through the detector parasitic shunt resistance R_0. Since the detector under discussion has a very low R_0, then this current would result in an undesirably high voltage across the photodetector ($V_{DET} > 100\ mV$). And, as discussed in section 2.6, higher values of V_{DET} are very detrimental for IR optoelectronic sensors. Therefore, the very large transistors were used to keep V_{DET} low.

However, even with transistors with very large widths, the injection efficiency was not close to one ($\varepsilon \cong 0.9$), but could be held fairly constant, when employing the right V_{POL}. For this case, we could have used a current gain β that overcomes the low ε, for instance $\beta \cong 1/\varepsilon \cong 1.11$. In addition, we have mentioned in section 2.8.1 that an even higher current gain would render a noise factor close to one, meaning a circuit with less noise. Although larger values of β are very beneficial, note from Table 16 that we have employed $\beta = 1$, which does not help in any of the aforementioned aspects. The reason why we have used such low β, is because the current gain

is given by the ratio between the transistor dimensions, for instance: $\frac{I_{(down)}}{I_1} = \frac{(W/L)_{MN3}}{(W/L)_{MN1}}$. Since we have always used the minimum length for transistors of this CMOS technology, that is, $L_{min} = 180\ nm$, then $W_{MN3} = \beta . W_{MN1}$. If M_{N1} is already close to the maximum width, then M_{N3} cannot be made much larger – and neither β. Unquestionably, we could have disposed transistors in parallel, so that they are able to handle larger currents, resulting in an increased "equivalent width". Nonetheless, a *chip* has a limited area, and the width and number of the transistors cannot increase unrestrictedly.

Notice that all these issues are due to the low shunt resistance R_0 and larger currents derived from the employed detector. For a different photodetector, however, for instance in array applications, where the sensor size is much smaller, the transistors could be made much smaller as well. This is because the shunt resistances would be smaller (due to the constant $R_0 A$ product, for the same sensor material), and due to the also smaller currents coming from the detector. Therefore, in this case, smaller transistors would be necessary to feature $\varepsilon \to 1$ and $V_{DET} \cong 0\ V$, also allowing the use of higher values of β.Prior to presenting the pixel circuit results – varying the input DC current from 400 µA to 4.4 mA – some preliminary simulations are going to be shown, in order to illustrate one of the main advantages of the bouncing pixel: its high sensitivity, even for high current levels. Generally, when a current is too high, in order to prevent signal saturation, the integration time is shortened. Recalling equation (38) – $S = \beta . \varepsilon . \Delta T / C_{int}$ – if ΔT decreases, then S diminishes likewise. Since the bouncing pixel prevents the voltage signal from saturation, then the sensitivity may be preserved.

First, a transient simulation with a linear pixel circuit without the bouncing feature is presented. Then, the bouncing pixel is shown. Both simulations use the same input currents. However, the regular linear pixel needs to reduce the integration time so that signal does not saturate, as will become clear soon. For these preliminary results, a smaller number of seeds ($\#_{seeds} = 5$) were employed, compared to the 30 seeds from Table 16. This is because simulations are very computer-intensive, hence time-consuming. Unlike the example of Figure 35, the voltage signal now starts integrating from V_{max} instead of V_{min}, like presented in Figure 94. These simulation examples were taken for five different DC currents, namely: 400.00 µA, 401.25 µA, 402.50 µA, 403.75 µA, and 405.00 µA. The used seeds were 1, 2, 3, 4 and 5.

Figure 94 shows how a regular pixel circuit – with no bouncing feature – would deal with a high current level, emulating a situation where signal would saturate and ΔT would need to be shortened. Note that the integration time for this transient simulation was $\Delta T = 7.6\ \mu s$ (from $13.5\ \mu s$ up to $21.1\ \mu s$), rendering in a sensitivity of $S_{(7.6\ \mu s)} = \beta . \varepsilon . \Delta T / C_{int} = 1 * 0.9 * 7.6x10^{-6} / 4.4x10^{-9} = 1.56\ mV/\mu A$. This value is way smaller than the sensitivity for a larger integration time of $\Delta T = 256\ \mu s$, that is: $S_{(256\ \mu s)} = 52.4\ mV/\mu A$.

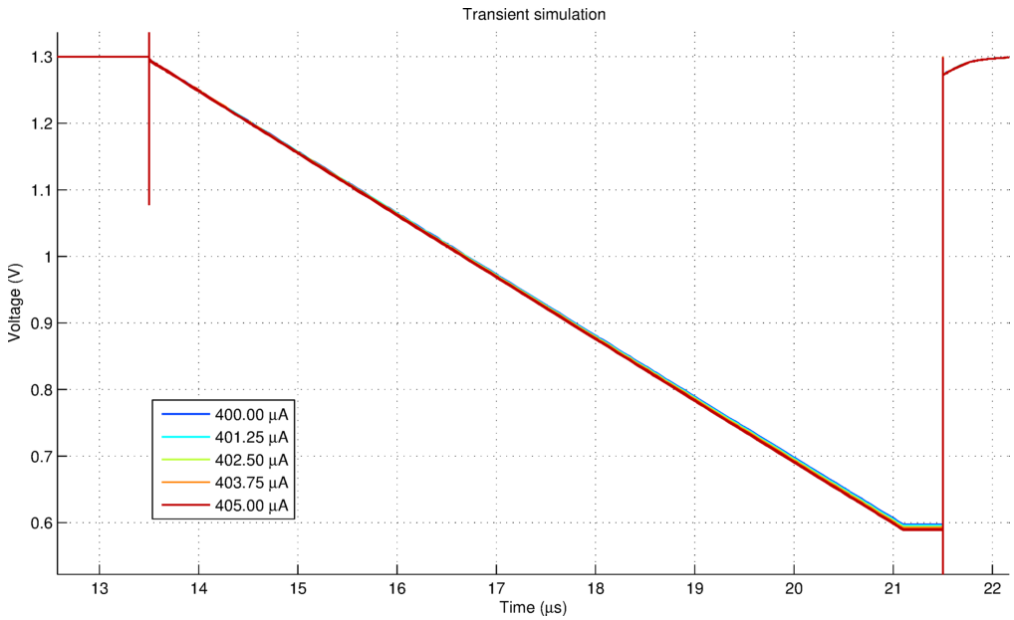

Figure 94: Transient noise simulation example for a regular linear pixel circuit with $\Delta T = 7.6\ \mu s$ and 5 seeds.

In order to better observe the differences between the output signals, Figure 95 shows a zoom in the sample & hold time interval (from $21.1\ \mu s$ up to $21.5\ \mu s$). It is possible to notice that even with just 5 seeds, the output signals for a given photocurrent are so spread that they are almost superimposed on other photocurrent responses, with a maximum dispersion on the final output voltages of about $\sim 10\ mV$.

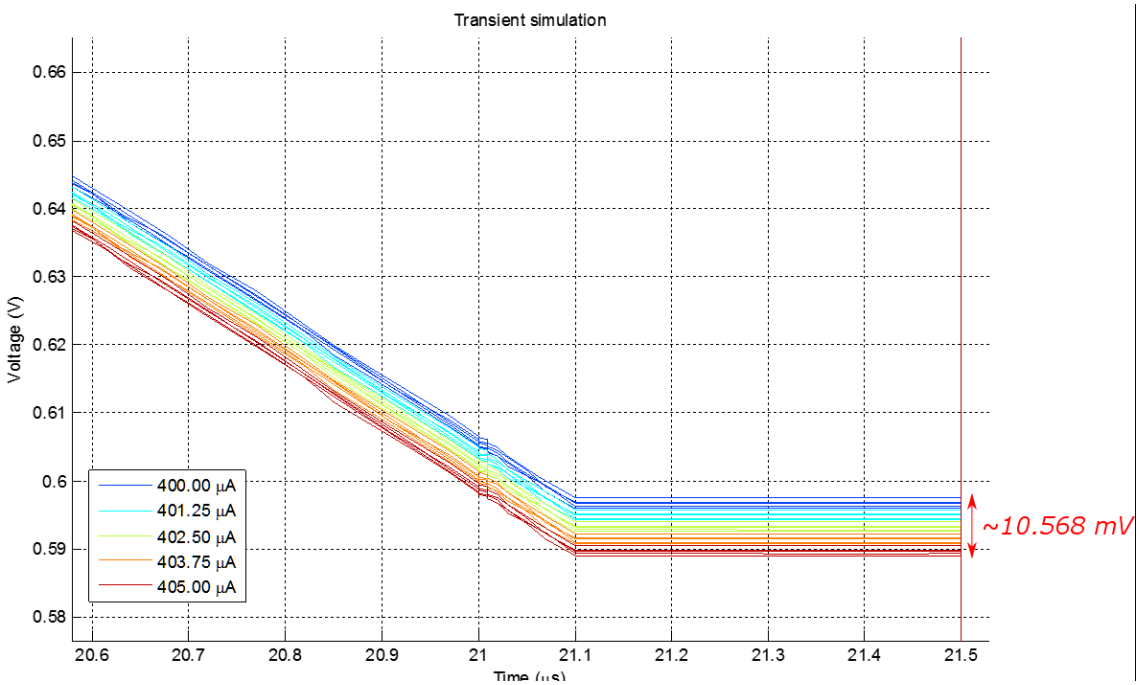

Figure 95: Zoom in for the transient noise simulation for a regular linear pixel circuit with $\Delta T = 7.6\ \mu s$ and 5 seeds.

Next, the simulations for the bouncing pixel are going to be presented, in order to demonstrate the increase in sensitivity. The simulations were performed with the same

configurations as in Figure 94 and Figure 95, except by the integration time, now set to $\Delta T = 256\ \mu s$. Figure 96 shows these transient simulations. Although the current values are very close to each other, their output voltages – read at the end of the integration time – are now very easy to discriminate from each other, even in the presence of noise. We can note that at the beginning of the integration time, all five sets of curves are indistinguishable, and after each re-bounce their relative differences start to be evident. This effect is due to the higher sensitivity, achievable with the bouncing pixel.

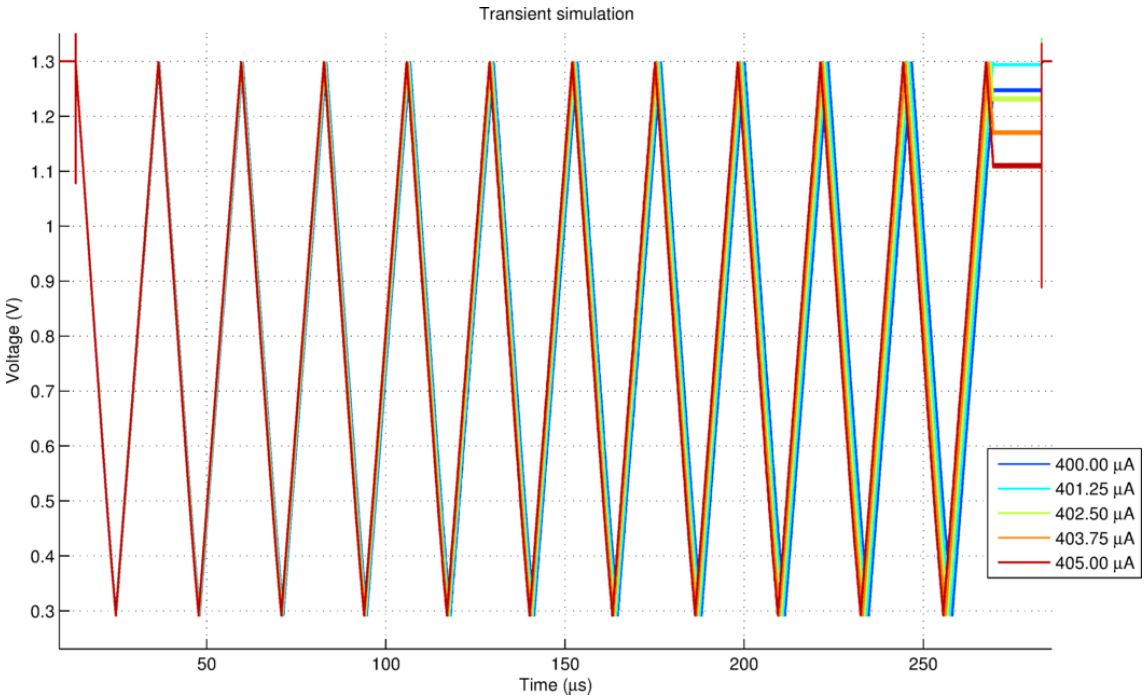

Figure 96: Transient noise simulation example for the bouncing pixel circuit with $\Delta T = 256\ \mu s$ and 5 seeds.

Figure 97 zooms in the sample & hold time interval (from $269.5\ \mu s$ up to $283.0\ \mu s$). It is clear that for the same set of different input DC currents, signal-to-noise ratio is much higher, since now the different set of output voltages do not intertwine due to noise, with a maximum dispersion of output voltages around $\sim 190\ mV$, a much higher value compared to the $\sim 10\ mV$ for the regular linear pixel.

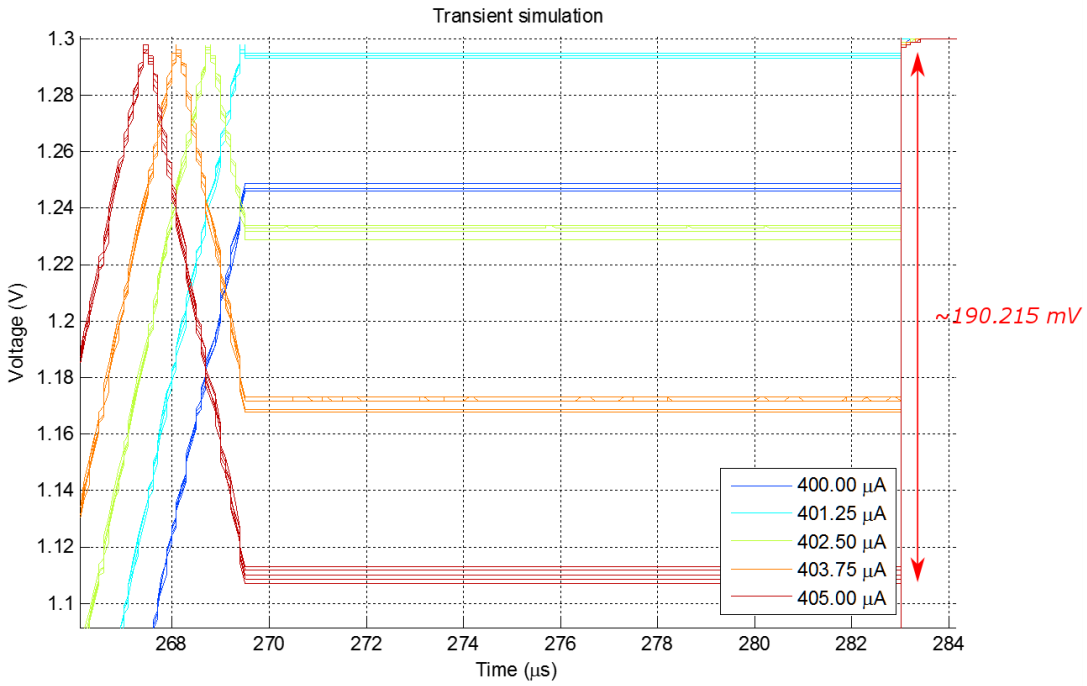

Figure 97: Zoom in for the transient noise simulation for the bouncing pixel circuit with $\Delta T = 256\ \mu s$ and 5 seeds.

During the sample & hold time interval, the voltage on the integration capacitor (V_{CAP}) is maintained constant, so that the ADC can read the desired signal. Consider that we would require to identify differences in the current signal of $\delta I = 1\ \mu A$, that is, a current resolution of δI. The corresponding required voltage resolution at the output would be given by $\delta V = S.\delta I$. We know that the voltage resolution of an ADC is determined by its number of bits (b_a), since $\delta V = \frac{V_{REF}}{2^{b_a}-1}$, where V_{REF} is the ADC conversion reference voltage. Therefore, the minimum number of bits would be given by: $b_a = \left\lceil \log_2\left(1 + \frac{V_{REF}}{\delta V}\right)\right\rceil = \left\lceil \log_2\left(1 + \frac{V_{REF}}{S.\delta I}\right)\right\rceil$, where $Y = [X]$, means that Y equals X, but rounded up to the nearest integer.

Choosing $V_{REF} = V_{DD} = 1.8\ V$, for $\Delta T = 256\ \mu s$, the required number of bits of the ADC would be $b'_a = \left\lceil \log_2\left(1 + \frac{1.8}{52.4\ x\ 10^{-3}}\right)\right\rceil = [5.1] = 6\ bits$. For $\Delta T = 7.6\ \mu s$, however, the number of bits is much larger: $b''_a = 11\ bits$. Table 17 compares the ADC number of bits required for the two analyzed different sensitivities, for different current resolutions.

Table 17: ADC required number of bits (b_a) for two different sensitivities, and a counter with $b_c = 8\ bits$.

Required number of bits for the ADC (b_a)		
Current resolution (δI)	$\Delta T = 7.6\ \mu s$ $S_{(7.6\ \mu s)} = 1.56\ mV/\mu A$	$\Delta T = 256\ \mu s$ $S_{(256\ \mu s)} = 52.4\ mV/\mu A$
100 µA	4	1
10 µA	7	3
1 µA	11	6
100 nA	14	9
10 nA	17	12
1 nA	21	16

In fact, the minimum readable current is a function of the number of bits of both the counter (b_c) and the ADC (b_a), and of the maximum current at the circuit just before the counter saturates (I_{sat}). The maximum current resolution is given by $\delta I_{min} = \frac{\delta V}{S_{max}}$, where $\delta V = \frac{V_{REF}}{2^{b_a}-1}$. Therefore, $\delta I_{min} = \frac{1}{S_{max}} . \frac{V_{REF}}{2^{b_a}-1}$. From equation (61), $S_{max} = \frac{2^{b_c}.\Delta V}{I_{sat}}$, resulting in:

$$\delta I_{min} = \frac{V_{REF}}{\Delta V} . \frac{I_{sat}}{2^{b_c}.(2^{b_a}-1)} \cong \frac{V_{REF}}{\Delta V} . \frac{I_{sat}}{2^{b_c}.2^{b_a}} \quad [A] \qquad (65)$$

where $\left(2^{b_a}-1\right) \cong 2^{b_a}$, since generally $2^{b_a} \gg 1$. Equation (65) can be rewritten as:

$$b_{sum} = b_c + b_a = \left\lceil \log_2 \left(\frac{I_{sat}}{\delta I_{min}} . \frac{V_{REF}}{\Delta V} \right) \right\rceil \quad [bits] \qquad (66)$$

in which b_{sum} accounts for the total number for bits, that is, the sum $(b_c + b_a)$. Table 18 summarizes the requirements for the counter and ADC, for different maximum current resolutions, employing $I_{sat} = 4.4\ mA$, $V_{REF} = 1.8\ V$, and $\Delta V = 1\ V$. We can see that the smaller the minimal readable current by the bouncing pixel, the higher is the required total number of bits. Also, as previously shown in equation (55), the maximum DR is extended when a higher b_{sum} is employed. Moreover, the required number of bits for the ADC is shown, for different counters. We can see that – without considering noise – the resolution can be increased as much as we want for a given maximum current I_{sat}, by wisely choosing the parameters: b_a, b_c, ΔV and V_{REF} in equation (65). Note that this current resolution is measurable even when it is embedded in a very strong background signal, around mA, for this case. This is one of the main advantages of the bouncing pixel, arisen from its high sensitivity, and yet, large charge storage capacity, with a fairly linear response.

Table 18: Required b_{sum} for each current resolution, for $I_{sat} = 4.4\ mA$, $V_{REF} = 1.8\ V$, and $\Delta V = 1\ V$. It also presents the maximum possible DR, and the required number of bits of the ADC (b_a), for different number of bits of the counter (b_c).

Current resolution (δI_{min})	$b_c + b_a$	*Maximum DR:* $20.log(2^{b_{sum}})$ *[dB]*	b_a for $b_c =$ 8 *bits*	b_a for $b_c =$ 10 *bits*	b_a for $b_c =$ 12 *bits*	b_a for $b_c =$ 16 *bits*
100 µA	7	42	0	0	0	0
10 µA	10	60	2	0	0	0
1 µA	13	78	5	3	1	0
100 nA	17	102	9	7	5	1
10 nA	20	120	12	10	8	4
1 nA	23	138	15	13	11	7
100 pA	27	163	19	17	15	11
10 pA	30	181	22	20	18	14
1 pA	33	199	25	23	21	17

The actual required current resolution we should employ will be based on the output voltage RMS noise. This value, however, will be estimated based on the following simulations that

employ a larger number of seeds, and hence, feature a more reliable calculation of the RMS noise.

With parameters presented in Table 16 we have performed a parametric sweep varying the photogenerated current I_{ph} from 1 µA until 4 mA. Since the dark current is $I_{dark} = 400\ \mu A$, then the total DC current flowing through the circuit $(I_{DC} = I_{dark} + I_{ph})$ varies from $401\ \mu A$ until $4.4\ mA$. These simulations scan the whole intended range of input currents, with 30 different seeds for each I_{DC} value. The used seeds were 1, 2, 3, … 30.

Since the voltage signal V_{CAP} now starts integrating from V_{max} instead of V_{min} (in comparison with section 3.3.1), then equation (47) for the output reconstructed voltage V_R had to be modified. We have simply inverted the condition for N odd and even, that is:

$$\begin{cases} V_R = N.\Delta V + (V_{CAP} - V_{min}) \ \ [V]: \quad for\ N\ odd \\ V_R = N.\Delta V + (V_{max} - V_{CAP}) \ \ [V]: \quad for\ N\ even \end{cases} \qquad (67)$$

By counting the number of bounces N and measuring the average V_{CAP} during the sample & hold time interval, we can reconstruct signal V_R. All calculations and data analysis after simulations were performed using MATLAB®. The number of bounces N was extracted from the serial signal, that is, the output of the shift register. Figure 98 shows an example of a serial signal waveform, for an input DC current of 400 µA, where the first two bits are the start signal (10_2), the third one is the bit MAX (0_2) – corresponding to a possible overflow of the counter – and the following eight (00010101_2) represent the decimal number $N = 21$. The serial waveform for each seed and each DC current is analyzed, and the corresponding decimal value for the number of bounces is stored.

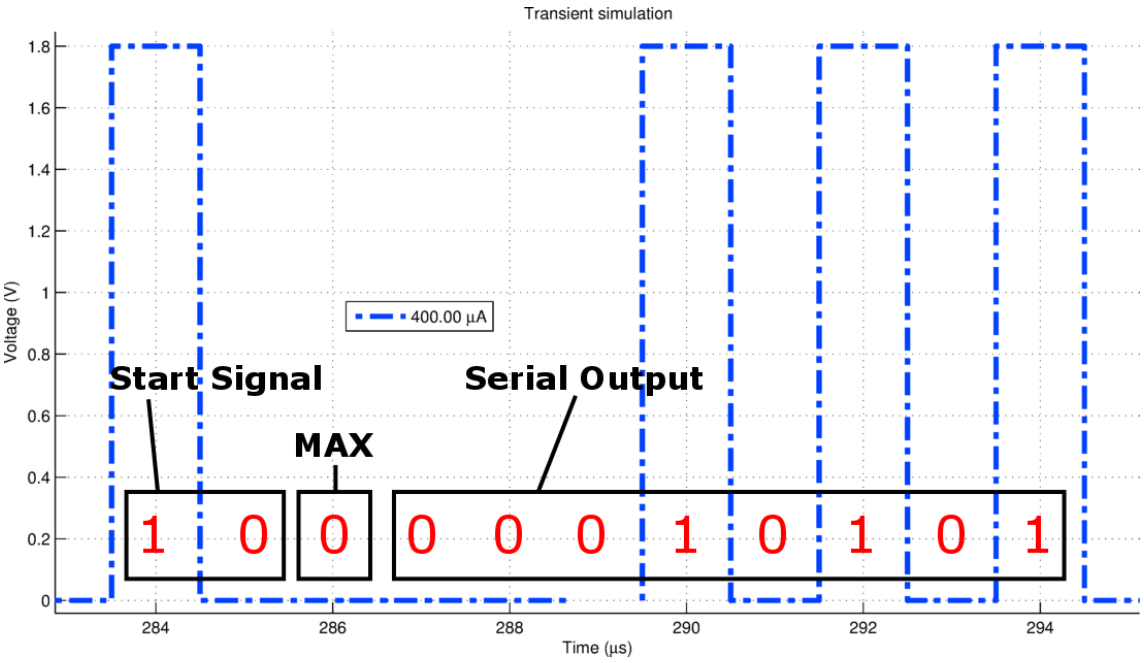

Figure 98: Serial signal and its corresponding bits, for $\Delta T = 256\ \mu s$ and $I_{DC} = 400\ \mu A$.

The average voltage on the capacitor V_{CAP}, during the whole sample & hold time is acquired, and its value is stored for each seed and input current. For the first simulations, no ADC was considered, meaning a quantization of the analog voltage without any rounding errors. An ADC will be considered afterwards. Then, knowing both V_{CAP} and N for every seed and input current, their corresponding reconstructed voltage V_R is calculated in MATLAB®, using equation (67). Next, for each DC current, the mean and standard deviation values for V_R were estimated. Figure 96 presents the average reconstructed voltage V_R for each analyzed input DC current.

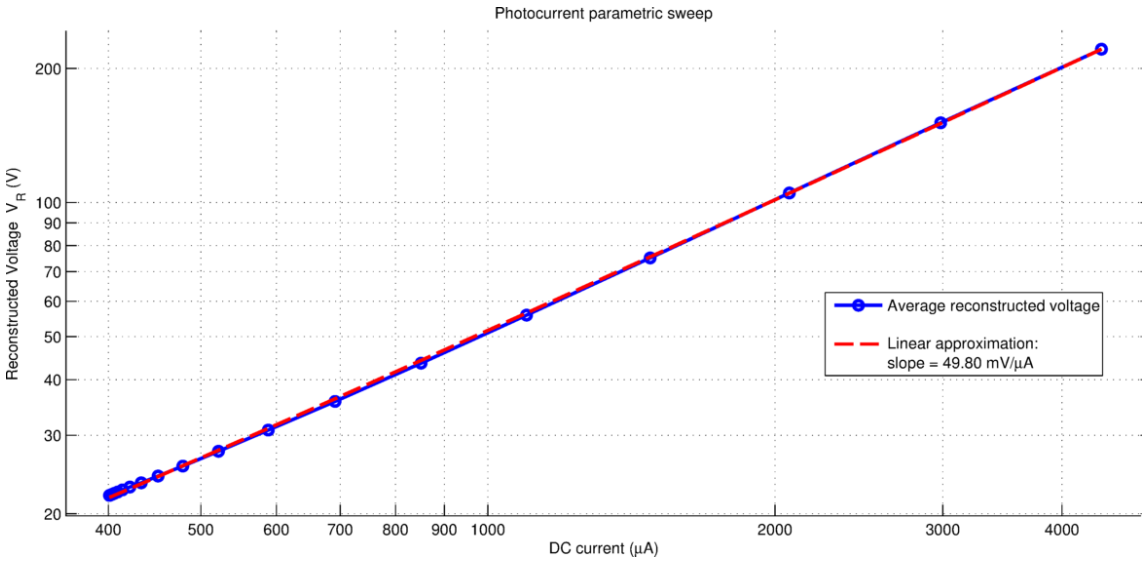

Figure 99: Average reconstructed voltage from 400 μA up to 4.4 mA, with $\Delta T = 256\ \mu s$.

Graph of Figure 99 also shows a linear approximation of the reconstructed signal V_R. The slope of this line corresponds to the simulated sensitivity, resulting in $S_{(graph)} = 49.8\ mV/\mu A$, compared to the calculated $S_{(256\ \mu s)} = 52.4\ mV/\mu A$. The difference between these values are not too large, but they may have differed due to parasitic capacitances not considered in the calculation of $S_{(256\ \mu s)}$, or due to the variation in injection efficiency ($\varepsilon = 0.91 \pm 0.08$).

Figure 100 shows the voltage signal V_{CAP} used for the reconstructed signal, for all input DC currents, and different seeds, totaling 600 different curves. The serial output signals at the end of the integration time were omitted, for clarity.

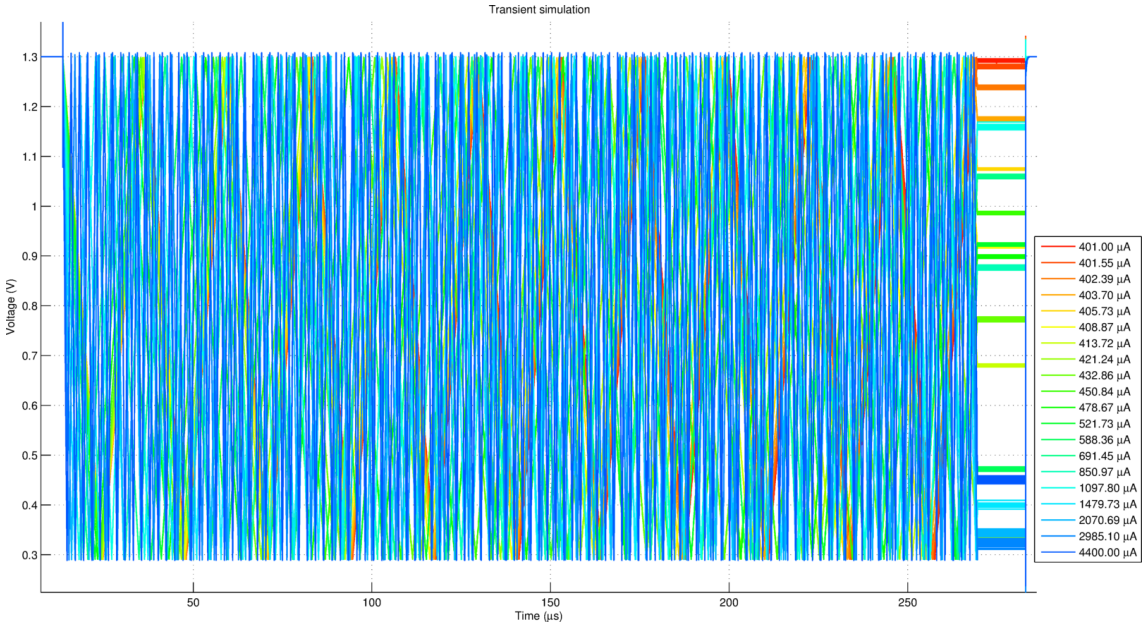

Figure 100: Transient noise simulation for the whole DR with 30 seeds.

Figure 101 zooms in the sample & hold time interval. We can notice that the sampled voltage appears to spread more around their respective means for larger DC currents. That makes sense, since from a SNR perspective, signal increases faster than noise. Even for 30 different seeds, the output noise voltage does not appear to be too large. But, even if it were larger, we could have boosted the maximum sensitivity S_{max} of the circuit (for the same saturation current I_{sat}) by using a larger number of bits (b_c) for the counter. Note that, from equation (66), the same current resolution would be maintained if the sum b_{sum} $(= b_c + b_a)$ was kept constant.

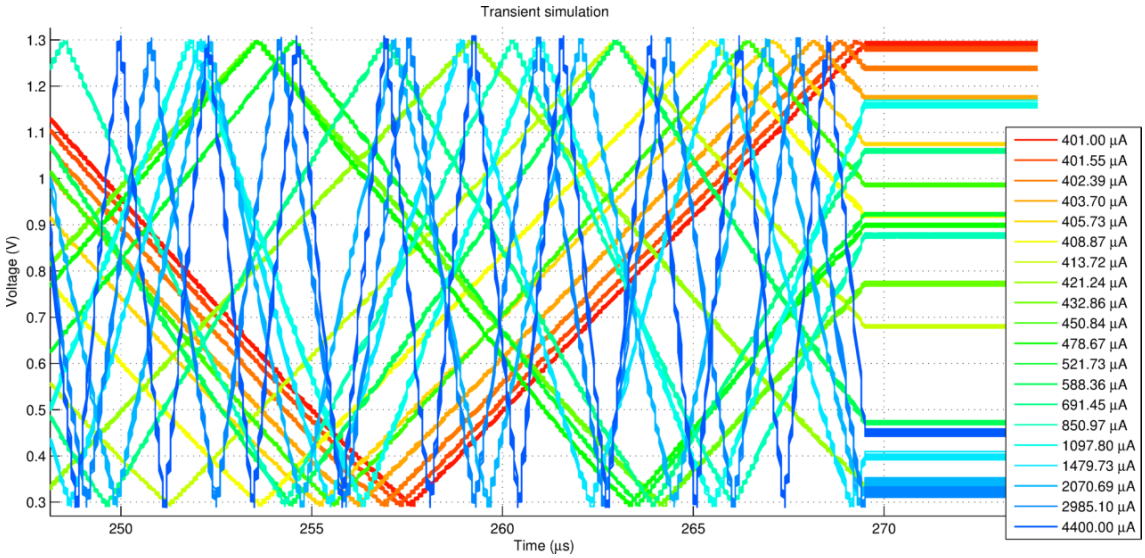

Figure 101: Zoom in transient noise simulation for the whole DR with 30 seeds.

The ratio between the average output voltage signal (V_{signal}) and the RMS voltage around it $\sigma_{V-noise}$, yields the signal-to-noise ratio at the circuit output, as in equations (27). Or, in a more convenient form, as discussed in sections 2.8.1 and 2.9: $SNR_{out-dB} = 20 * \log(V_{signal}/\sigma_{V-noise})$:

$$SNR_{out-dB} = 20 * \log\left(\frac{V_{signal}}{\sigma_{V-noise}}\right) = 20 * \log\left(\frac{I_{ph}.S}{\sqrt{\sigma_{V-pixel}^2 + S^2.\sigma_{noise}^2}}\right) \quad [dB] \qquad (68)$$

where S is the circuit sensitivity, σ_{noise} is the RMS noise from the photodetector – as in equation (12) – $\sigma_{V-pixel}$ is the RMS noise introduced by the pixel circuit, and I_{ph} is the useful input photocurrent signal. The voltage V_{signal} is correlated with the photocurrent I_{ph} $(= I_{DC} - I_{dark})$, therefore, the curve of interest for analysis is given by SNR_{out-dB} vs. I_{ph}, where the photocurrent is swept form $1\ \mu A$ until $4\ mA$. Figure 102 presents both the input and output signal-to-noise ratios, or SNR_{in-dB} and SNR_{out-dB} , respectively. The SNR_{in-dB} was theoretically calculated based on equation (13) and parameters from Table 16. On the other hand, the SNR_{out-dB} – represented by equation (68) – was actually extracted from simulated data. The useful signal voltage was estimated using $V_{signal} = I_{ph}.S$, where the sensitivity was measured for each input DC current I_{DC} and average reconstructed voltage V_R as: $S = V_R/I_{DC}$. The standard deviation of the output reconstructed voltage, i.e. the RMS noise voltage $\sigma_{V-noise}$, was measured for each DC input current, considering all 30 seeds.

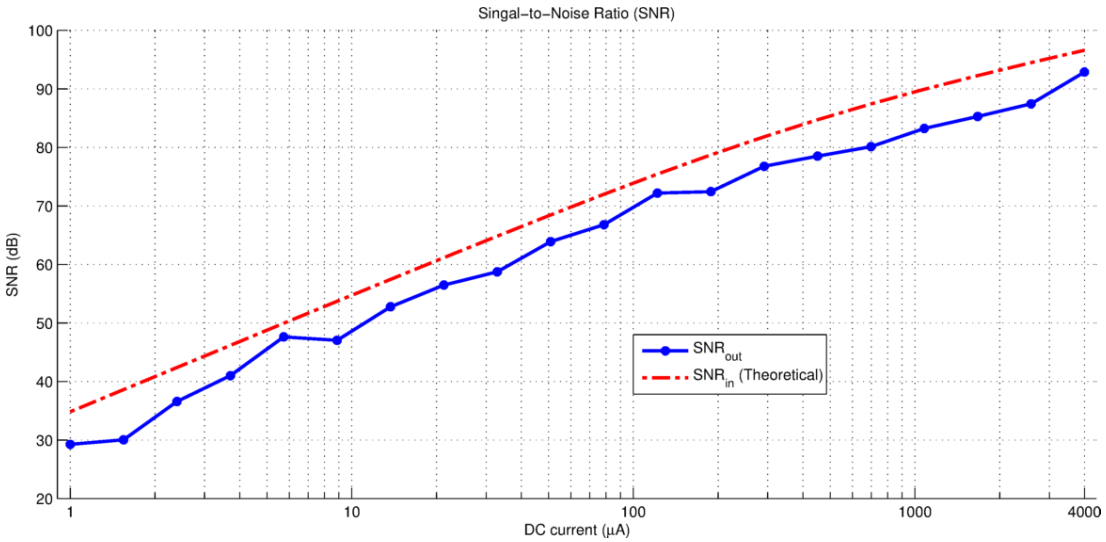

Figure 102: Simulated SNR without ADC and input SNR in *dB*.

Figure 102 shows that the circuit introduces a considerable amount of noise comparable to the one from the detector, since SNR_{out-dB} is much smaller than SNR_{in-dB}. The average noise factor, given by equation (28) – (with SNR not in dB units) – is: $F = SNR_{in}/SNR_{out} = 3.83 \pm 1.29$. This is an indeed large noise factor, which has a direct impact on the desirable input current resolution, since noise limits the minimum measurable signal. From equations (32) and (68), we have that the RMS noise introduced by the pixel circuit can be estimated by:

$$\sigma_{V-pixel} = \sqrt{F-1}\,.(S.\sigma_{noise}) \quad [V] \qquad (69)$$

where $\sigma_{V-pixel}$ was calculated for the values of *F*, σ_{noise}, and S $(= V_R/I_{DC})$ for each current. Table 19 summarizes the RMS noise components for each photocurrent of the parametric sweep, namely the measured $\sigma_{V-noise}$ from simulated data, the integrated noise from the

photodetector $\sigma_{V-nose-PD} = S.\sigma_{noise}$, and the estimated noise $\sigma_{V-pixel}$ from the readout circuit.

Table 19: Output RMS noise voltage components.

Photocurrent (μA)	$\sigma_{V-noise}\ (mV)$	$\sigma_{V-noise-PD}\ (mV)$	$\sigma_{V-pixel}\ (mV)$
1.0	1.890	0.995 (27.715 %)	1.607 (72.295 %)
1.5	2.674	0.996 (13.874 %)	2.482 (86.155 %)
2.4	1.945	0.996 (26.223 %)	1.670 (73.721 %)
3.7	1.807	0.998 (30.503 %)	1.507 (69.552 %)
5.7	1.306	1.000 (58.629 %)	0.841 (41.467 %)
8.9	2.163	1.003 (21.502 %)	1.917 (78.547 %)
13.7	1.730	1.007 (33.882 %)	1.407 (66.145 %)
21.2	1.740	1.013 (33.894 %)	1.414 (66.039 %)
32.9	2.063	1.021 (24.494 %)	1.792 (75.453 %)
50.8	1.752	1.036 (34.966 %)	1.413 (65.045 %)
78.7	1.927	1.057 (30.088 %)	1.611 (69.892 %)
121.7	1.585	1.092 (47.466 %)	1.149 (52.551 %)
188.4	2.361	1.146 (23.560 %)	2.064 (76.424 %)
291.5	2.185	1.224 (31.381 %)	1.810 (68.621 %)
451.0	2.742	1.342 (23.954 %)	2.392 (76.100 %)
697.8	3.507	1.512 (18.588 %)	3.164 (81.396 %)
1079.7	3.777	1.748 (21.418 %)	3.348 (78.574 %)
1670.7	4.612	2.064 (20.028 %)	4.125 (79.996 %)
2585.1	5.552	2.470 (19.792 %)	4.972 (80.198 %)
4000.0	4.567	2.972 (42.348 %)	3.467 (57.630 %)

Table 19 shows that the estimated noise from the pixel circuit is larger than the corresponding integrated noise from the photodetector, for almost all photocurrents. Indeed, the best performance would be achieved if $\sigma_{V-noise} \to \sigma_{V-noise-PD}$, where $F \to 1$, and therefore $\sigma_{V-pixel}$ becomes negligible according to equation (69). By increasing the internal current gain β of the circuit, the sensitivity is also increased by the same rate, and from equation (32) – $F = 1 + \sigma_{V-pixel}^2/(S.\sigma_{noise})^2$ – if we have employed a current gain of $\beta = 10$, then the noise factor would be decreased down to $F = 1.03 \pm 0.01$.

If we recall the fact that we want to distinguish between small optical power differences on the photodetector, depending on the desired impinging power resolution, decreasing noise can be crucial. Suppose an input optical power P_W focusing over the photodetector. This will be transduced as a DC current given by: $I_{DC} = P_W.R_{(4.33\ \mu m)}$, where $R_{(4.33\ \mu m)} = 1.54\ A/W$ is the responsivity of the photodetector @4.33 µm. However, since noise is added to the signal from both the circuit and the detector, then the optical power P_W would result in a measured current $(I_{measured})$ in the following interval, with 99.7 % probability:

$$I_{DC} - 3.\sigma_{input-noise} \leq I_{measured} \leq I_{DC} + 3.\sigma_{input-noise} \quad [A] \qquad (70)$$

where $\sigma_{input-noise}$ is the input-referred RMS noise, given by $\sigma_{input-noise} = \sigma_{V-noise}/S$, as in equation (33).

Therefore, the $\pm 3.\sigma_{input-noise}$ interval, i.e. $\delta I_{min} = 6.\sigma_{input-noise}$, determines the minimum current resolution the photodetection system should have in order to distinguish between two closest values of input optical power. In other words, $\delta I \leq 6.\sigma_{input-noise} = 6.\sigma_{V-noise}/S$. This minimum current was calculated based on the $\sigma_{V-noise}$ from Table 19 and on the sensitivity for each input current $S\ (= V_R/I_{DC})$, leading to $\delta I_{min} = 143.06\ nA$. This current resolution corresponds to an optical power minimum difference of $\delta P_{min} = \delta I_{min}/R_{(4.33\ \mu m)} = 92.90\ nW$.

For such value of current resolution, the minimum required ADC number of bits (b_a) would be given by equation (66), that is:

$$b_a = b_{sum} - b_c = \left\lceil \log_2\left(\frac{I_{sat}}{\delta I_{min}}.\frac{V_{REF}}{\Delta V}\right)\right\rceil - b_c = \lceil 15.76 \rceil - 8 = 8 \quad [bits] \qquad (71)$$

in which $I_{sat} = 4.4\ mA$, $V_{REF} = 1.8\ V$, $\Delta V = 1\ V$, $b_c = 8\text{-}bit$, and $\delta I_{min} = 143.06\ nA$.

If we had employed a current gain of $\beta = 10$, the sensitivity would also have increased tenfold. This would result in a larger $\sigma_{V-noise-PD}$, so that $\sigma_{V-noise} \cong \sigma_{V-noise-PD}$, namely $\sigma_{V-noise} \cong 13.54\ mV$ and $\sigma_{V-noise-PD} \cong 13.35\ mV$ – if $\sigma_{V-pixel}$ is considered not to vary as much. Despite the increase in output noise, as previously mentioned, the noise factor F decreases down to $F = 1.03 \pm 0.01$. In this case, we can consider $\sigma_{input-noise} \cong \sigma_{noise}$, where σ_{noise} was calculated using equation (12), resulting in a minimum value of $\sigma_{noise} = 18.14\ nA$. Therefore, the minimum current resolution should be given by $\delta I_{min} = 6.\sigma_{noise} = 108.86\ nA$.

In order to guarantee this current resolution, the ADC number of bits should be $b_a = b_{sum} - b_c = \lceil 16.15 \rceil - 8 = 9\ bits$. Certainly, it would be necessary to re-simulate the circuit in order to know if noise introduced by the pixel would not increase much, and accurately calculate the noise factor F and the necessary current gain to maintain it close to the unity. However, since transient noise simulations are very computer-intensive, we have considered this approximation as good enough for the purposes of this thesis. Simulations with high current gains would certainly improve the results, especially the SNR, and will be left for future work. In addition, the maximum number of bits (b_{sum}) necessary for the circuit happens when $SNR_{in} = SNR_{out}$, and was calculated as $b_{sum} = 17\ bits$, resulting in a necessary ADC of $b_a = 9\ bits$.

Regarding the analog-to-digital conversion, we have reanalyzed the simulation data, with no current gain $(\beta = 1)$, by emulating an ADC quantization in MATLAB®. This was done by employing the following quantizer operator:

$$V_{CAP-quantized} = \delta V.\left\lfloor \frac{V_{CAP}}{\delta V} + \frac{1}{2}\right\rfloor \quad [V] \qquad (72)$$

where V_{CAP} is the analog voltage on the capacitor read during the sample & hold time interval, δV is the voltage resolution of the ADC, and $V_{CAP-quantized}$ is the quantized value of the V_{CAP}. Note that $Y = \lfloor X \rfloor$ means that Y equals X, but rounded down to the nearest integer, and that $\delta V = \frac{V_{REF}}{2^{ba}-1}$, where V_{REF} is the ADC voltage reference, given in Table 16. In order to

reconstruct the output voltage signal, the difference by employing the ADC is that we have now used $V_{CAP-quantized}$ instead of V_{CAP} in equation (67).

Figure 103 presents the reconstructed voltage V_R vs. input DC current with and without ADC. We can note that the resulting curves are virtually the same, with an RMS error between them of only $\sim 220\ \mu V$. Even the slope of the linear approximation are the same, for the given number of decimal digits.

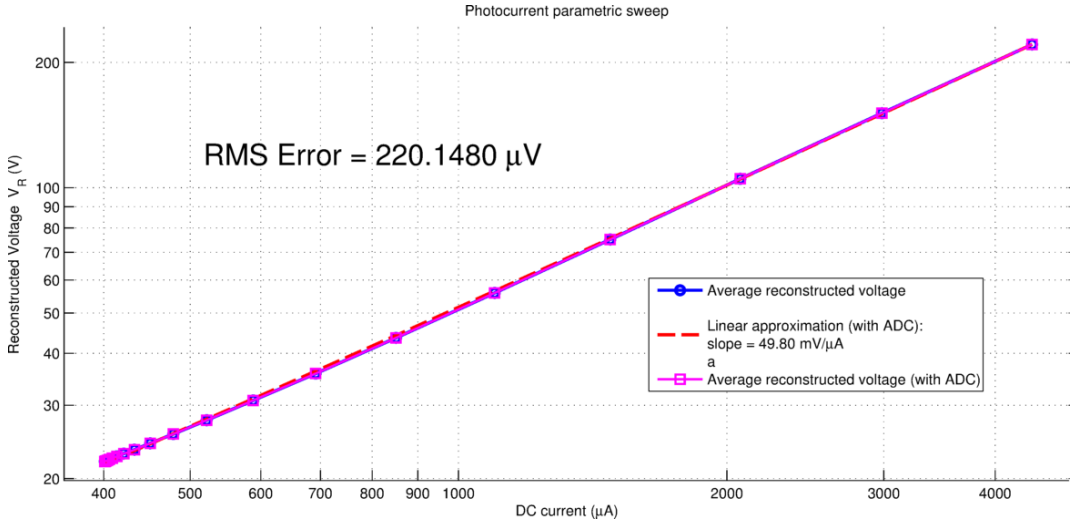

Figure 103: Average reconstructed voltage from 400 µA up to 4.4 mA, with $\Delta T = 256\ \mu s$, and with and without ADC.

However, from Figure 104 we can see that the employment of the ADC impacts the SNR curve by a larger extent, with a RMS error between the curves of approximately $1\ dB$. It resulted in a worse noise factor of $F_{ADC} = 4.65 \pm 1.58$ (compared to $F = 3.83 \pm 1.29$). This corroborates the demand to study, in future work, the behavior of the circuit with higher current gains, since a large β could improve the SNR of the circuit, even with the ADC. Certainly, the use of a higher resolution ADC could also result in better outcomes, at least approaching the curve to the one without ADC.

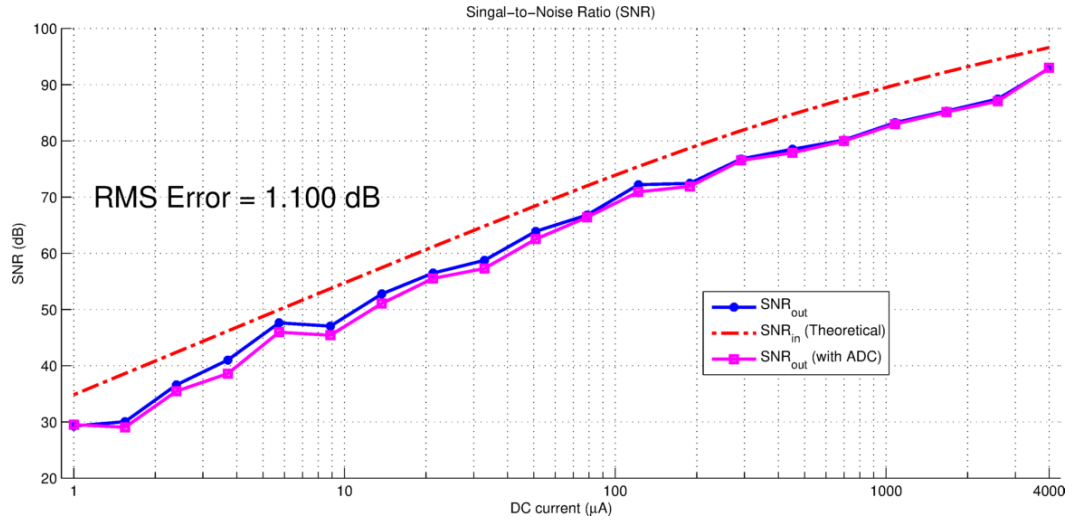

Figure 104: Output and input SNR in dB, and output SNR considering the ADC. The RMS error between the curve with and without ADC is also presented.

A curious detail from Figure 104 is the tendency that the curves with and without ADC tend to converge for higher input currents. This is because the higher the input current, the less important the analog voltage V_{CAP} – read by the ADC – becomes. From equation (67), we know that the reconstructed voltage depends on both the analog V_{CAP} and on the digital N (converted to decimal from the serial signal stream of bits). For larger DC currents, the term $(N.\Delta V)$ becomes larger than the terms $(V_{CAP} - V_{min})$ or $(V_{max} - V_{CAP})$, since these last two terms are always smaller than ΔV $(= V_{max} - V_{min})$. Therefore, for higher DC currents, the rounding of V_{CAP}, executed by the ADC, starts to become less noticeable, and does not affect the SNR by a great degree.

Therefore, that means that the circuit is able to register differences in a large extent of input currents, either by the ADC or by the counter. The ADC is responsible mainly for the low photocurrents, while the counter plays its role for higher input DC currents. Yet, for high current signals, the ADC will still help to sense small differences between photocurrents. In fact, any current difference slightly larger than the RMS noise floor can be sensed, even when the two currents are very high. Consider, for instance, two current values, namely, I_A and I_B, where $I_B = I_A + \delta I_{min}$. The corresponding output reconstructed voltages are, respectively, $V_A = I_A.S$, and $V_B = I_B.S = (I_A + \delta I_{min}).S = V_A + \delta V_{min}$, where S is the sensitivity, and δV_{min} is designed to be the voltage resolution of the ADC. These voltages will be converted to the digital to domain, resulting in two adjacent values of the ADC, since their voltage difference is equal to δV_{min}, that is still true even for larger current values for I_A and I_B. The main reason why the current resolution is high even for larger currents, is that the sensitivity is also kept very high for these current values, and the minimum readable current in inversely proportional to the sensitivity $(\delta_{min} \propto 1/S)$. For a regular linear pixel circuit, however, the sensitivity has to be decreased when sensing large currents (decreasing the integration time), which therefore, affects the resolution in this range.

The measure of the extent of input photocurrent signals the circuit is able to sense is called the dynamic range (DR), defined in section 2.8.2. The DR of the bouncing pixel is given by equation (37), rewritten here in the simplest form for convenience:

$$DR = 20 * \log\left(\frac{I_{ph\_max}}{I_{ph\_min}}\right) \quad [dB] \qquad (73)$$

where I_{ph_max} is the maximum signal, and I_{ph_min} is the RMS noise, that is, the input-referred noise, but in dark conditions. The maximum signal was designed to be $I_{ph\_max} = 4\ mA$, while $I_{ph\_min} = \sigma_{input-noise}$, was already estimated based on the output(voltage)-referred noise: $\sigma_{input-noise} = 23.84\ nA$. From equation (73), the DR of the circuit will be:

$$DR = 20 * \log\left(\frac{4x10^{-3}}{23.84x10^{-9}}\right) \cong 105 \quad [dB] \qquad (74)$$

Such large value of $DR = 105\ dB$ was achieved with a fairly linear curve response, as demonstrated by Figure 99, and with a high sensitivity. This allows that differences of $\delta I_{min} = 143.06\ nA$ in the input current, or equivalently in optical power units, $\delta P_{min} = 92.90\ nW$, can be perceived for both low and high photocurrents.

4.4 Application example: The CO_2 gas detection system

The purpose of this section is to bring all designed parts of the system together, aiming at the specific application of a CO_2 gas detection system. Note that this is only an application example, that has not actually been implemented. Many others could have been explored, like Lab-on-Chip (LoC) for fluorescence detection, or a soldering monitoring tool in industry, or a headlike-detection sensor at night, or solar observation, and so on. The CO_2 gas in particular finds many applications, for instance as an indicator of combustion efficiency, since it is the primary product emission from combustion [103]. Depending on the application, it would be necessary to quantify the exact amount of CO_2 in a mixture, or at least be able to identify variations on its concentration. We are going to present here a model for CO_2 detection that exemplifies how the electronic and optical parts of the system could be used in order to identify very small concentration changes.

In order to extract any information from the sampled gas, a laser is pointed at the detector passing through the gas mixture, like illustrated on Figure 83. The chosen laser for the modelled detection system was a commercial Quantum Cascade Laser (QCL), model L12004-2310H-C [101]. At nominal temperature and forward bias, it features a peak emission at $\nu_{QCL} = 2310\ cm^{-1}$, where ν_{QCL} is the wavenumber given in $[cm^{-1}]$. The equivalent wavelength is given by: $\lambda_{QCL} = 1x10^7/\nu_{QCL} = 4329.00\ nm$. Controlling the device operating temperature through a Peltier element, and by its forward bias current, it is possible to tune the peak emission in the range of $\pm 1.0\ cm^{-1}$, that is, $\nu_{QCL} = 2310\ \pm 1.0\ cm^{-1}$, or equivalently: $4327.13\ nm \leq \lambda_{QCl} \leq 4330.88\ nm$ ($\lambda_{QCL} \cong 4329.004 \pm 1.874\ \mu m$). This is an important feature, since the peak absorption of gases is very sensitive to the wavelength.

Initially, the envisioned system should be calibrated with a reference cell, which contains N_2 at controlled temperature and pressure, for instance $T_{ref} = 296\ K$ and total pressure of $p_{total} = 1\ atm$. Since nitrogen is an inert gas, the laser power arrives at the detector without significant absorbance. The system can be tuned, so that the maximum power impinging over the detector would be $\sim 2.6\ mW$, as designed in section 4.3.1. This can be done by adding filters, or by adding lenses with proper control of the beam aperture. We assume here that the laser beam is collimated, and it does not suffer from the inverse square law for the conducted optical path, that is, the laser intensity will not fall off with distance. This is because the laser beam does not diverge, keeping the same diameter for the whole optical path, and the intensity (E) will be held constant, since $E = (Laser\ Power)/(Laser\ beam\ cross\ section\ area)$. Note that we can either use the reference cell only during calibration, or in real time during measurements, comparing the measurement from the sample and reference. The last option can be more expensive, but may lead to better results due to possible variations in the measuring environment. However, we consider here that both temperature and pressure of N_2, in the gas mixture to be measured (N_2:CO_2), are well controlled, therefore the situation where just the sample cell is measured will be considered.

The nominal laser output power is $P_{LASER} = 20\ mW$, passing through a circular optical window with radius $r_{win} = 2.2\ mm$, ideally leading to an intensity of $E_{source} = \frac{P_{LASER}}{\pi . r_{win}^2} = 1.32x10^3\ W/m^2$. The chosen *InAsSb* detector (model P11120-201) [100] has a squared area with lateral length $L_{DET} = 1\ mm$. This results in a total optical power impinging over the

detector of $P_{DET} = E_{source}.L_{DET}^2 = 1.32\ mW$. By placing a microlens in front of the detector, the impinging power over it would increase by a factor of OG, where OG is the optical gain. With a proper design of the optical gain, the necessary power of $2.6\ mW$ focusing over the photodetector could be provided, recalling from section 2.9.2 that $OG = A_{microlens}/A_{DET}$, where $A_{microlens}$ is the microlens area, and $A_{DET} = L_{DET}^2$ is the detector area. The optical gain should be $OG = 2.6\ mW/1.32\ mW = 1.97$, and the microlens should have an area of $A_{microlens} = OG.A_{DET}$. On the other hand, the microlens area is given by $A_{microlens} = \pi\left(\frac{D_{Lens}}{2}\right)^2$, where D_{Lens} is the lens diameter. Therefore, $D_{LENS} = 2.\sqrt{OG.A_{DET}/\pi} = 1.58\ mm$.

The designed and fabricated microlens array, shown in section 4.1, contains microlenses with diameters of $1.41\ mm$, and therefore would not provide the necessary optical gain. Nonetheless, the previously shown microfabricated lens array had the purpose of demonstrating the fabrication process, and although successful, it had not then been designed for this specific setup. Therefore, another lens mask pattern, just like the one of Figure 34, could be designed, using the computer program developed by *de Lima Monteiro*, in order to fit this specific application. In addition, since the lens is made of silicon, it needs to be covered with an ARC, as discussed in section 2.9.3. This ARC layer has a transmittance that depends on wavelength, and will affect the OG calculation. However, in order to determine the exact transmittance through the lens, we need to know the exact wavelength at which the laser will operate, within its peak emission range. This operating point depends on the peak absorbance of the CO_2 to be considered, that will be discussed next. Therefore, the lens specifications will be discussed with more details later on.

Since the objective of this system is to infer how much of the target gas is present on a volume, considering the amount of absorbed light though the optical path, then we must know the relation between the transmittance through the gas and its concentration. The gas absorption theory is based mainly on the Beer-Lambert law [103]–[110]:

$$T(\nu_0) = \left(\frac{P_T}{P_0}\right)_{\nu_0} = e^{-\alpha(\nu_0).D_{PL}} \quad [1] \qquad (75)$$

where P_0 is the incident power, P_T is the power transmitted through the path containing the gas mixture, $T(\nu_0)$ and $\alpha(\nu_0)$ are, respectively, the transmittance and the absorption coefficient for a given wavenumber ν_0, and D_{PL} the optical path length. Note that equation (75) is very similar to equation (3). The only difference is that we have used the wavenumber ν_0 instead of the wavelength λ, and power P instead of the photon flux Φ.

For gases, however, the absorption coefficient $\alpha(\nu_0)$ for a monochromatic light – for instance, a laser – may be written as:

$$\alpha(\nu_0) = S_L.g(\nu - \nu_0).N_M.p_a \quad [cm^{-1}] \qquad (76)$$

in which S_L is the molecular line intensity, $g(\nu - \nu_0)$ is the lineshape function, N_M is the total number of molecules of the absorbing gas per cm^3 per atm, and p_a is the partial pressure of the gas in the mixture given in atm.

A given gas-absorption spectra may contain many different absorption lines at different wavenumbers. These lines correspond to specific rotational or vibrational modes of the gas that may enter in resonance with an incoming light at the same frequency. The spectral lines are the absorption signature of a specific gas. The shape of each line can be fitted by normalized *Lorentzian*, *Gaussian* or *Voigt* curves $g(\nu - \nu_0)$ [103]–[107], centered around ν_0, while its intensity is given by S_L. The term S_L is given in $cm/molecule$ and depends on temperature. At $T = 296\,K$, the parameter N_M equals the *Loschmidts'* number (N_L), that is: $N_L = 2.479x10^{19}\ molecules.cm^{-3}.atm^{-1}$ [107]–[110]. The lineshapes were considered to be given by Lorentzian curves, like:

$$g(\nu - \nu_0) = \frac{1}{\pi}.\frac{\gamma_p}{\left[(\nu - \nu_0)^2 + \gamma_p^2\right]} \quad [cm] \qquad (77)$$

where γ_p is called the pressure-broadened halfwidth at half-maximum $(HWHM)$, given in $[cm^{-1}]$. This term controls the width of the curve, and depends on the total pressure and temperature, as: $\gamma_p = g_{air}.\left(\frac{296}{T_{ref}}\right)^n.p_{total}$, in which n is a fitting constant, p_{total} the total pressure, and g_{air} is called the air-broadened halfwidth.

Note that gas absorption line spectra have been already extensively investigated, and all the necessary parameters for the wavenumbers of interest can be found in the HITRAN database [111]. Therefore, we have implemented the Beer-Lambert model in order to find the transmittance through CO_2 gas, for specific partial pressures, on a N_2:CO_2 gas mixture. Equations (75), (76), (77), and the HITRAN database were used in order to find the transmittance $T(\nu_0)$ for a wavenumber of ν_0:

$$T(\nu_0) = \exp\left(-N_L.p_a.D_{PL}.\Sigma_j\left[S_{L_j}.g\left(\nu_j - \nu_0\right)\right]\right) \quad [1] \qquad (78)$$

where S_{L_j} and $g\left(\nu_j - \nu_0\right)$ correspond to the j^{th} line intensity, and j^{th} lineshape, respectively. The sum Σ_j in equation (78) is necessary because more than one different line may affect the absorbance at the same wavenumber ν_0, depending on the lineshape broadening and centering.

In addition, different gases may also affect the absorbance, and therefore the transmittance, at the same wavenumber. Therefore, for the N_2:CO_2 gas mixture, both nitrogen and carbon dioxide were considered in the calculation of the transmittance through the optical path (although N_2 does not affect the absorbance very much):

$$T(\nu_0) = \exp\left(-D_{PL}\left[\alpha_{CO_2}(\nu_0) + \alpha_{N_2}(\nu_0)\right]\right) \quad [1] \qquad (79)$$

The considered line intensities from HITRAN database varied from $346\,cm^{-1}$ up to $12784\,cm^{-1}$ for CO_2, and from $11.8\,cm^{-1}$ up to $9354\,cm^{-1}$ for N_2, covering the target wavenumbers. Figure 105 shows the spectral molecular line intensities (S_L) around the laser peak emission $(2310\,cm^{-1}\ or\ 4329\,\mu m)$ for CO_2.

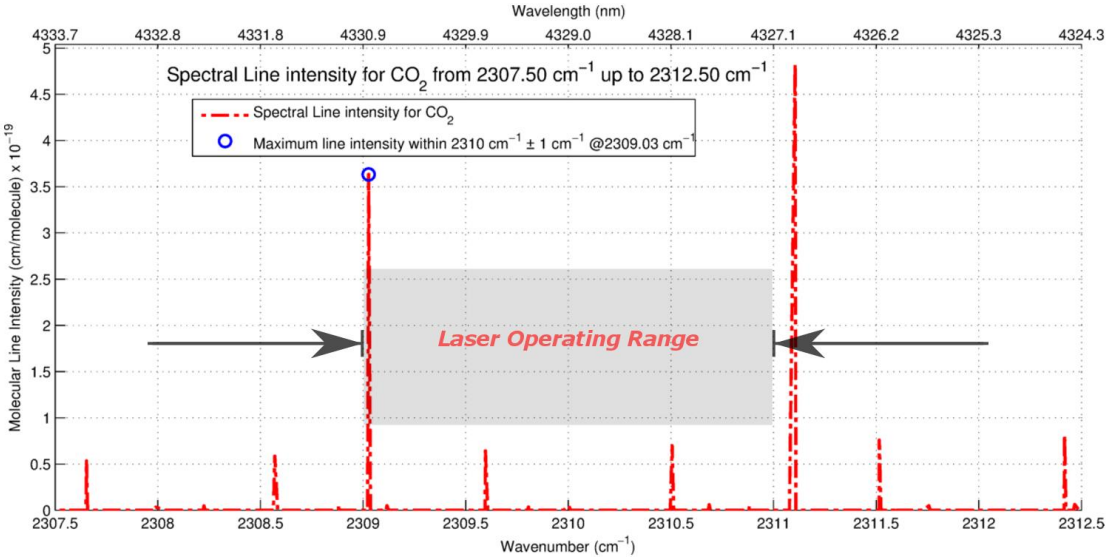

Figure 105: Carbon dioxide CO_2 spectral lines around the laser peak emission wavenumber. The wavelengths are also shown in the top horizontal axis.

From Figure 105, we can see that the highest line intensity within the range at which the laser can operate $(\nu_{QCL} = 2310 \pm 1.0\ cm^{-1})$ is $\nu_{LASER} = 2309.03\ cm^{-1}$, and therefore, we should consider that the laser was tuned for this wavenumber, since this is the condition for which the gas will absorb more light. This wavenumber corresponds to a wavelength of $\lambda_{LASER} = 4330.82\ nm$.

We have mentioned before, that the desired laser output power at the detector could be achieved by employing the microlenses, that would provide the calculated optical gain. In reality, this optical gain should be adjusted due to the transmittance of the microlenses with the ARC, that reduces the actual power arriving over the photosensitive area. Figure 79 presented the transmittance measurements for the SiN layer deposited over the surface of the microlenses, with thickness of $533\ nm$. That graph highlighted the transmittance at $4200\ nm$. However, since the light source for the discussed application is centered at $\lambda_{LASER} = 4330.82\ nm$, then the transmittance profile is presented again in Figure 106, but highlighting this wavelength. We can see that the transmittance through the silicon wafer with the SiN ARC, at the laser peak emission wavelength, is $T(\lambda_{LASER}) = 92.49\ \%$.

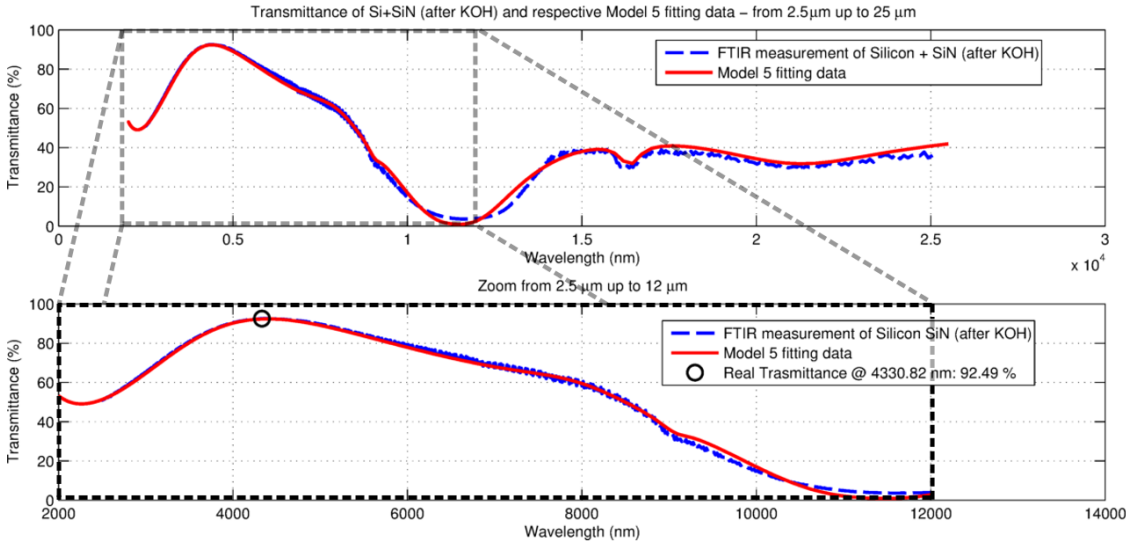

Figure 106: Transmittance of 533 nm thickness SiN layer deposited over KOH micromachined wafer (FTIR measurement and fitted model).

Now that we have a well-adjusted model for the optical constants of SiN, we could have optimized a new design with optimal thickness targeting λ_{LASER}. However, this is not necessary, since the value for transmittance @λ_{LASER} is already sufficiently close to the maximum of 92.56 % (from Table 12). Consequently, in order to achieve an optical power of $2.6\ mW$ at the detector, the value of optical gain should be modified to: $OG' = OG/T(\lambda_{LASER})$, resulting in $OG' = 2.13$. As a consequence, the needed diameter for the microlens should also be adjusted to $D'_{LENS} = 2.\sqrt{OG'.A_{DET}/\pi} = 1.65\ mm$.

Beyond the diameter, we must specify the new sagitta s' of the lens and its respective radius of curvature r'_c. The latter depends on the desired focal length, given by the Lens-Maker's Formula – in equation (58). In this context, the detector datasheet states that the distance between its entrance window of the packaging and the photosensitive area is $f' = 3.1\ mm$. If we determine that the lens would be placed on the entrance window through a pin hole to control the aperture to have the same size of the lens, then the new lens design should have a focal length around f'. In fact, since the goal is to maximize the power over the detector, it does not necessarily need to be in focus. Figure 107 *A)* shows the case when the focal length of the lens (f_{LENS}) is exactly equal to f'. However, as demonstrated in Figure 107 *B)* and *C)*, when it is out of focus, either with larger or smaller focal length, still the amount of power arriving over the detector photosensitive area is maximized. The range in which the focal length can vary is: $f' - \Delta f_2 \leq f_{LENS} \leq f' + \Delta f_1$, that is, $f'/(1 + L_{DET}/D'_{LENS}) \leq f_{LENS} \leq f'/(1 - L_{DET}/D'_{LENS})$. From equation (58), we can find that the radius of curvature should be $r'_c = 7.52\ mm$, considering $f_{LENS} = f' = 3.1\ mm$. The sagitta depends on both D'_{LENS} and r'_c, and is given by equation (57) rewritten here as: $s' = r'_c - \sqrt{r'_c - (D'_{LENS}/2)^2}$. This results in a sagitta of $s' = 45.38\ \mu m$.

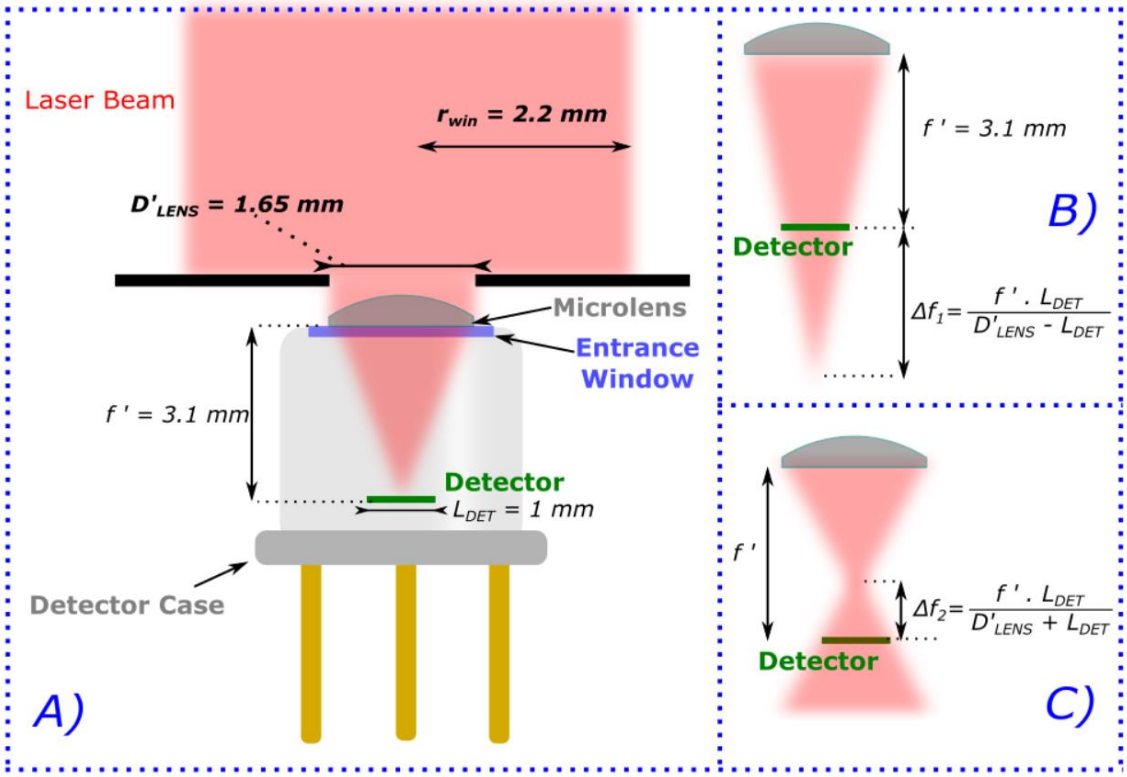

Figure 107: *A)* Schematics of the optical part of the system with focal length of the microlens equal to f'; *B)* with f_{LENS} larger than f', and C) smaller than f'.

By knowing D'_{LENS} and s', we can use *de Lima Monteiro*'s program to generate the custom lithography pattern mask for this lens, and then follow the process described in section 4.1, in order to micromachine the desired lens, on a double-side polished silicon wafer. Note that the sagitta value is compatible with the described process in this thesis, being much smaller than the 4-inch wafer thickness, i.e. 525 µm. The microfabricated wafer can then have the SiN layer deposited over its front and back surfaces, that will act as an ARC. Any other optical component, that may need to be added to the optical path, could be designed using the same procedure.

Up to now, we have ensured that the maximum optical power focusing over the detector would be around $P_{max} = 2.6\ mW$, when no CO_2 is present in the optical path, that is, for the cell filled up just N_2. That condition is necessary so that the bouncing pixel can read the input signals without saturating the digital counter – designed to have 8 bits. For the maximum optical power P_{max}, the input DC current in the circuit should be $I_{sat} = 4.4\ mA$. When CO_2 is present, it will absorb an amount of light depending on the *gas mole fraction* x_{CO_2} (related to the concentration), leaving a residual power of approximately $P(\nu_{LASER})$ to be detected, where $P(\nu_{LASER}) = P_{max}.T(\nu_{LASER})$, and $T(\nu_{LASER})$ is the transmittance through the optical path, given by equation (79) – centered at the laser peak emission wavenumber $\nu_{LASER} = 2309.03\ cm^{-1}$. The residual power $P(\nu_{LASER})$ will then be transduced by the detector, and can be written in units of electric current as $I_{ph} = P(\nu_{LASER}).R_{(\nu_{LASER})}$, where $R_{(\nu_{LASER})} = 1.54\ A/W$ is the detector responsivity, and I_{ph} the photogenerated current signal. The DC current flowing through the circuit will be given by $I_{DC} = I_{ph} + I_{dark}$. This current will then be integrated and converted to the reconstructed voltage V_R of the bouncing pixel, as: $V_R = I_{DC}.S$, where S is the sensitivity. In conclusion, the output voltage V_R can be related to the absorbing gas mole fraction x_{CO_2}.

From equations (77) and (78), we know that the transmittance depends on the temperature T_{gas} of the gas mixture, the total pressure on the sample cell p_{total}, on the path length D_{PL} and on the gases partial pressures. We will assume here that both the temperature and total pressure will be controlled, and maintained very close to the calibration reference, respectively $T_{gas} = T_{ref} = 296\,K$, and $p_{total} = 1\,atm$. The path length, i.e., the distance between the laser and the detector, is considered to be $D_{PL} = 1\,cm$. The only variables left are the partial pressures of CO_2 and N_2 on the gas mixture, referred as p_{CO_2} and p_{N_2}, respectively. However, since N_2 practically does not absorb, then the only actual variable is the term p_{CO_2}. Summing both partial pressure components results in the total pressure, i.e., $p_{total} = p_{CO_2} + p_{N_2}$, and for an ideal gas mixture, which we assume here, since the gases do not chemically interact with each other, we can state that the $x_{CO_2} = p_{CO_2}/p_{total}$. Therefore, by knowing the gas partial pressure, we can infer its mole fraction, given in mol/mol, or ppm (*parts per million*).

Figure 108 presents the discussed gas mixture calculated transmittance for various different partial pressures of CO_2, varying the wavenumber in the range from $2306.50\,cm^{-1}$ up to $2311.50\,cm^{-1}$. The partial pressures of CO_2 vary in the range of: $1\,torr \leq p_{CO_2} \leq 10\,torr$ – where $1\,atm = 760\,torr$. Note that a small variation in p_{CO_2} – and therefore in x_{CO_2} – of less than $10\,torr$, or equivalently, less than $13x10^{-3}\,atm$, results in a large variation in transmittance, primarily at the absorption lines of Figure 105.

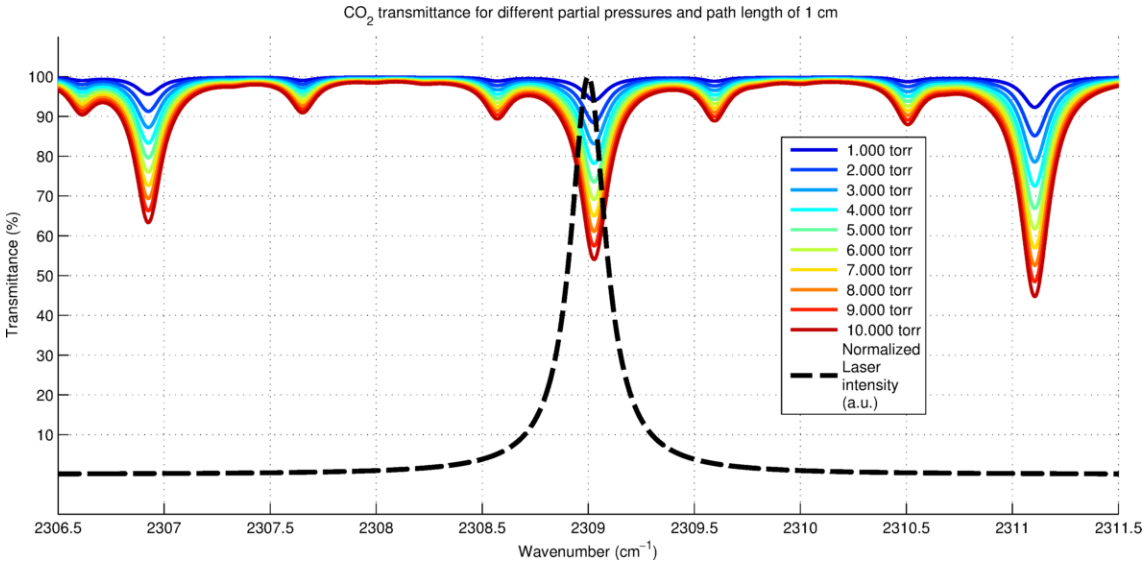

Figure 108: Carbon dioxide CO_2 transmittance for partial pressures in the range of $1\,torr \leq p_{CO_2} \leq 10\,torr$ and path length of 1 cm. The normalized laser intensity is shown as well.

In addition, we have also plotted in Figure 108 the laser spectral intensity $\left(I_{norm[\nu]}\right)$ normalized by its maximum intensity, presented in arbitrary units (a.u.). Its datasheet provides information about the fullwidth at half-maximum $(FWHM)$, being it equal to $0.2\,cm^{-1}$. The $FWHM$ is related to the halfwidth at half-maximum $(HWHM)$ as $HWHM = FWHM/2$. We have then used a Lorentzian, as in equation (77), to model the laser emission, where $\gamma_p = HWHM = 0.1\,cm^{-1}$.

By doing $P(\nu_{LASER}) = P_{max}.T(\nu_{LASER})$, we are considering that the laser emits solely at ν_{LASER}. This simplification may render in inaccurate results when the absorption line is as narrow as the laser intensity profile. Therefore, in order to predict with more fidelity the amount

of power being transmitted through the gas mixture, we need to know the laser spectral intensity (in $[W/cm^{-1}]$) being transmitted $\left(I_{T[\nu]}\right)$, that is, the product of the incoming laser spectral intensity $\left(I_{max[\nu]}\right)$ and the transmittance spectra of the gas $\left(T(\nu)\right)$. The incoming power (in $[W]$) is given by the integral of its spectral intensity, for the whole wavenumber range: $P_{max} = \int I_{max[\nu]}.d\nu$. The term $I_{max[\nu]}$ can be written as $I_{\max[\nu]} = \rho.I_{norm[\nu]}$, where ρ is a constant, and $I_{norm[\nu]}$ is the normalized spectrum of the laser shown in Figure 108. The constant ρ is given by: $\rho = \int I_{max[\nu]}/\int I_{norm[\nu]} = 2.6\,mW/306.16\,mW = 8.49x10^{-3}$. Therefore, the incoming spectral intensity is given by the Lorentzian function, but with peak intensity of $8.49\,mW/cm^{-1}$.

The transmitted power will also be given by the integral of its corresponding spectral intensity, or $P_T = \int I_{T[\nu]}.d\nu$, but we know that $I_{T[\nu]} = I_{max[\nu]}.T(\nu)$. Figure 109 shows the incoming maximum laser intensity (corresponding to $P_{max} = 2.6\,mW$), and other intensities for different partial pressures of CO_2, along with their respective transmitted optical powers. We can see that the area of the transmitted laser spectral intensity diminishes with p_{CO_2}, and therefore, resulting in lesser power to be detected.

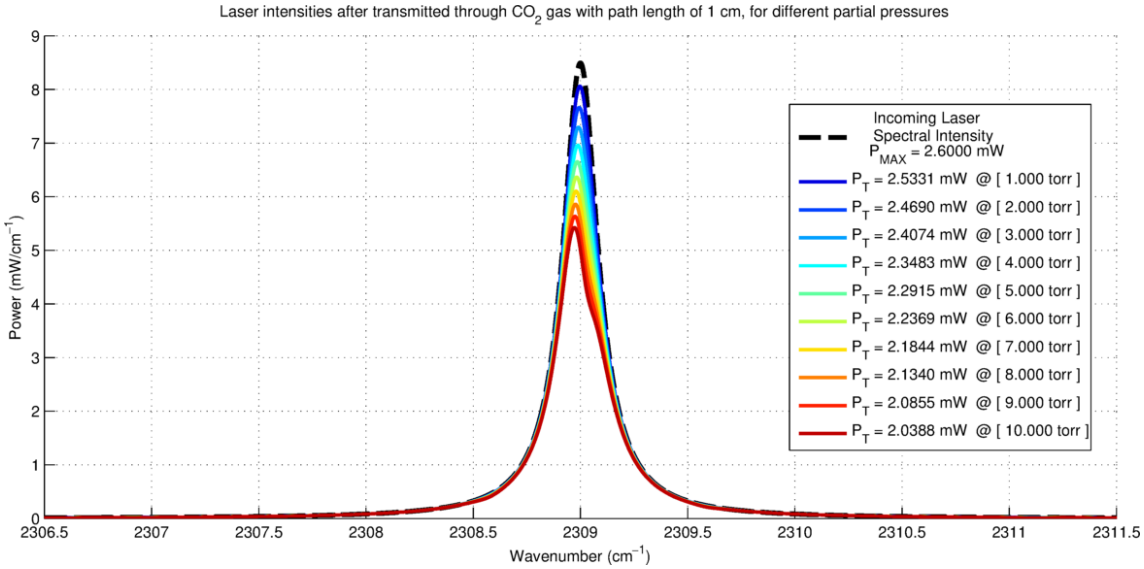

Figure 109: Transmitted laser intensity through the gas mixture and respective transmitted power for partial pressures in the range of $1\ torr \leq p_{CO_2} \leq 10\ torr$ and path length of 1 cm.

Figure 110 shows the mole fraction and their corresponding transmitted power in a natural logarithm scale, where the mole fraction is given in ppm, as: $x_{CO_2} = p_{CO_2}/p_{total} * 1x10^6\ \ [ppm]$. Obviously, the more CO_2 is present in the sample cell, the more light is absorbed, and the smaller is the transmitted power. We can see from equation (78) that applying the natural logarithm to the transmittance, and hence to the transmitted power $(\ln(T) = \ln(P_T) - \ln(P_{MAX}))$, then $\ln(P_T)$ should present a linear scale in relation to the mole fraction x_{CO_2}. In a practical system, the measured signals related to the detected power should be converted to the ln scale to present linearity.

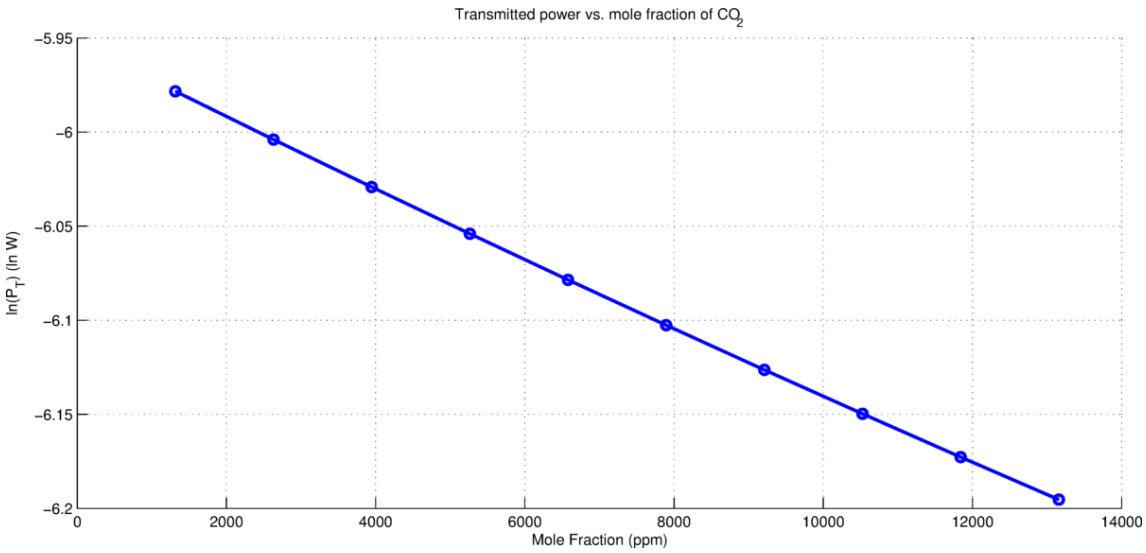

Figure 110: Transmitted power vs. mole fraction in ppm for partial pressures in the range of $1\ torr \leq p_{CO_2} \leq 10\ torr$ and path length of 1 cm.

In order to contextualize with the bouncing pixel simulations, the partial pressure was swept in a range compatible with the one presented in section 4.3.2.3, i.e., with photocurrent signals varying between $1\ \mu A$ and $4\ mA$. This range corresponds to optical powers of $0.649\ \mu W$ up to $2.6\ mW$. Figure 111 shows spectral transmittance of the gas mixture, which varies very widely, due to the broad range of different partial pressures.

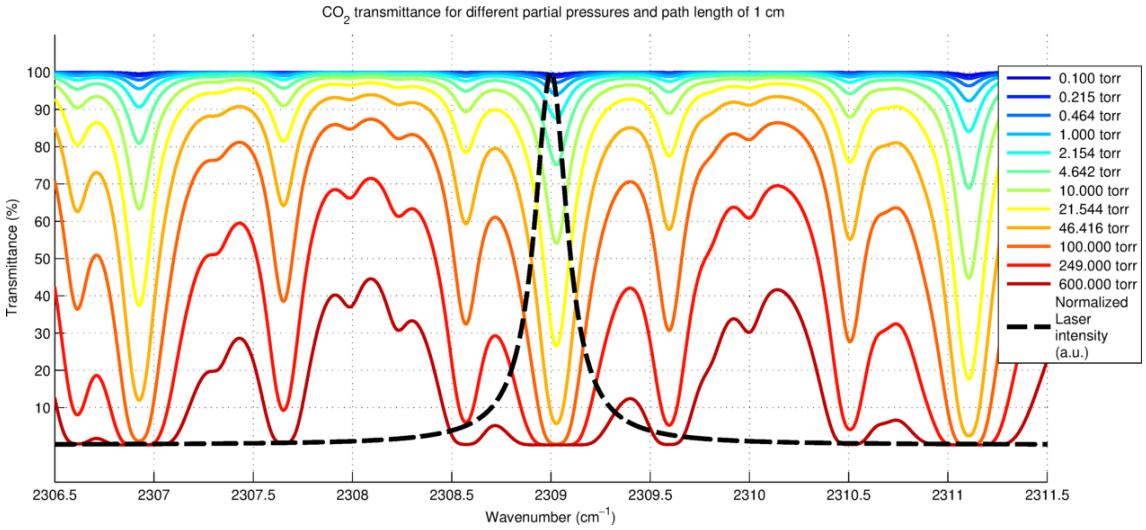

Figure 111: Carbon dioxide CO_2 transmittance for partial pressures in the range of $0.1\ torr \leq p_{CO_2} \leq 600\ torr$, and path length of 1 cm. The normalized laser intensity is shown as well.

Figure 112 presents the transmitted laser power for the same p_{CO_2} range, varying from almost P_{MAX}, for $p_{CO_2} = 0.1\ torr$, until a nearly opaque gas mixture with $P_T = 0.7\ \mu W$, for $p_{CO_2} = 600\ torr$.

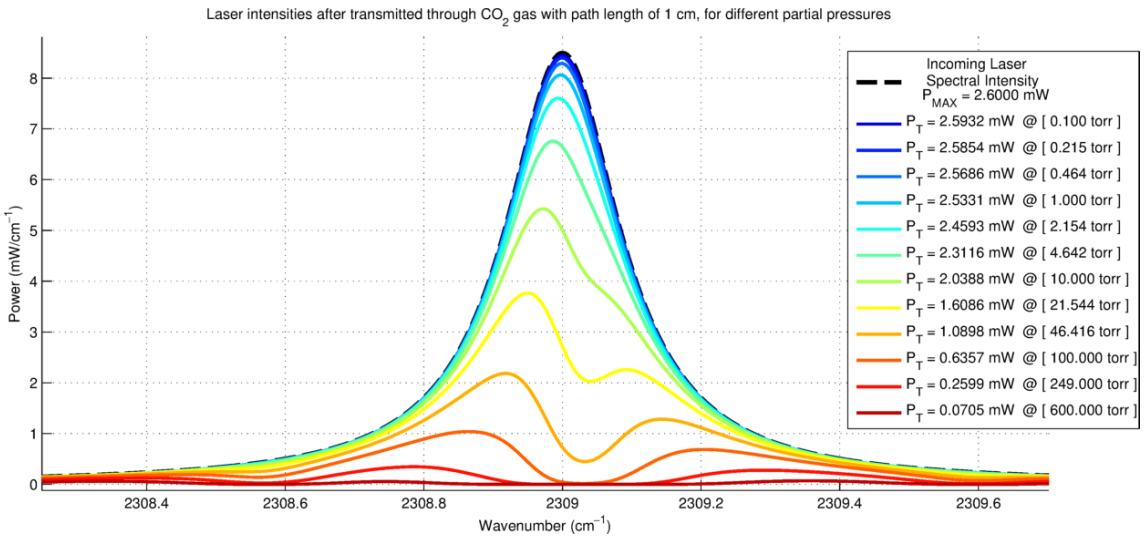

Figure 112: Transmitted laser intensity through the gas mixture and respective transmitted power for partial pressures in the range of $0.1\ torr \leq p_{CO_2} \leq 600\ torr$ and path length of 1 cm.

Figure 113 shows the variation of the transmitted power for the different mole fractions (in ppm) of CO_2. In addition, the path length was varied to show how it impacts the transmitted optical power. Certainly, shorter optical paths render in more linear responses, like for $D_{PL} = 0.1\ cm$. However, such small length can be unfeasible in a practical system. On the other hand, for a longer optical path length, the stronger is the absorption for the same mole fraction of the gas, as can be seen from the increasing absolute slopes of the lines in Figure 113. A path length of $D_{PL} = 1\ cm$ presents a good tradeoff, due to its better linearity in comparison with $D_{PL} > 1\ cm$, and relying on the fact that the bouncing pixel has a high sensitivity, being able to sense smaller variations of transmitted power.

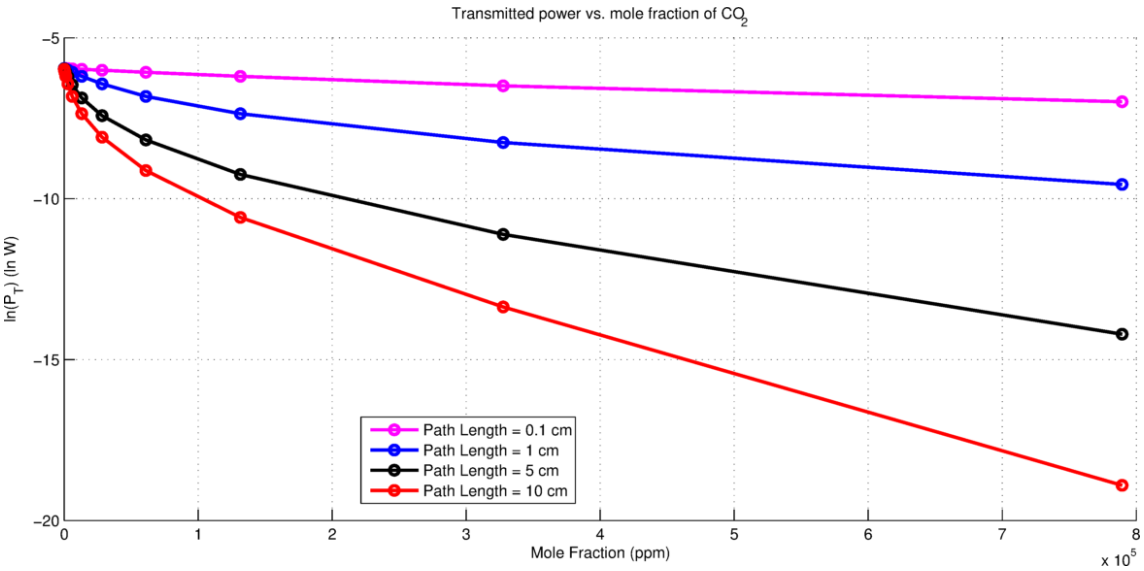

Figure 113: Transmitted power P_T vs. mole fraction in ppm for partial pressures in the range of $0.1\ torr \leq p_{CO_2} \leq 600\ torr$ and path length of 0.1, 1, 5, and 10 cm.

The transmitted power shown in Figure 112 and Figure 113, for $D_{PL} = 1\ cm$, would arrive at the photodetector and be converted to a photocurrent, which in turn would be integrated (along with dark current) resulting in the reconstructed voltage V_R. Figure 110 presents an estimate of

the voltage V_R that the bouncing pixel would have produced (in ln scale) for each mole fraction of CO_2 (in ppm) in the sample cell. This estimate was performed by considering $V_R = S.\left(I_{dark} + P_T.R_{(\nu_{LASER})}\right)$, where S is the sensitivity, I_{dark} the dark current, P_T the transmitted power, and $R_{(\nu_{LASER})}$ the responsivity. In addition, it presents the corresponding voltage signal V_{signal}, i.e., the portion of V_R equivalent only to the photocurrent I_{ph}. The voltage signal can be found by subtracting the component V_{dark} $(= S.I_{dark})$ from V_R, that is, $V_{signal} = V_R - V_{dark}$. The "dark voltage" component can be measured during calibration of the photodetector, by covering its photosensitive area, protecting from radiation.

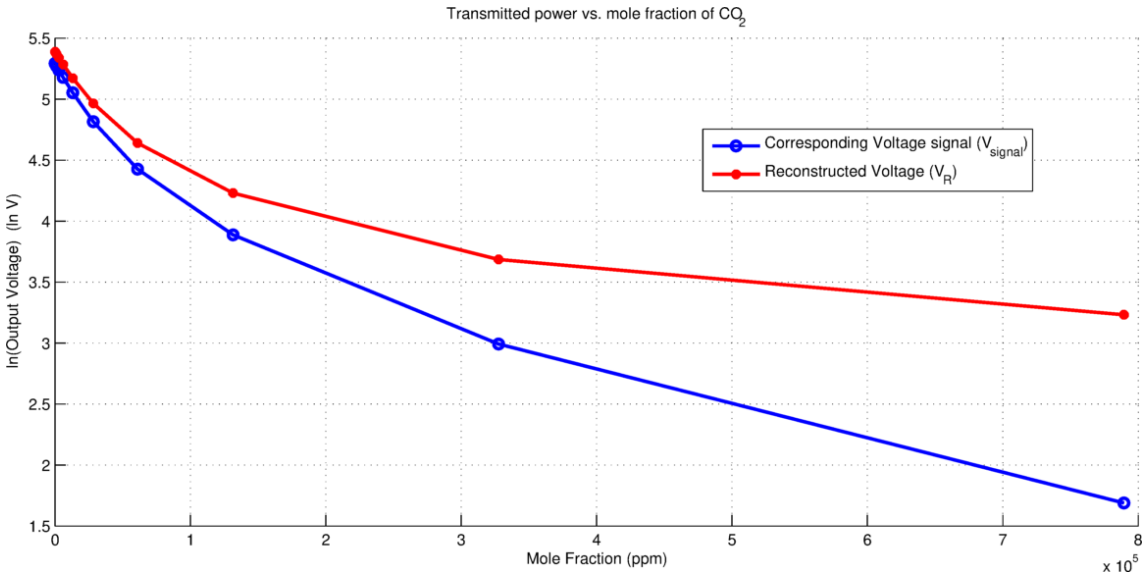

Figure 114: Output voltages V_R and V_{signal} vs. mole fraction in ppm for partial pressures in the range of $0.1\ torr \leq p_{CO_2} \leq 600\ torr$ and path length of $1\ cm$.

Note that there is a direct relation between $\ln\left(V_{signal}\right)$ and the mole fraction x_{CO_2}. When the laser intensity does not fall to zero – as happens for the last three curves in Figure 108 – this relation is linear, and given by:

$$\ln\left(V_{signal}\right) = Q_1 - Q_0.x_{CO_2} \quad [\ln V] \qquad (80)$$

where Q_0 and Q_1 are constants. Excluding the last three points of Figure 114, then the relation between $\ln\left(V_{signal}\right)$ and x_{CO_2} will be linear, indeed.

Equation (80) can be derived based on equations (75) and (76), and by recalling that $V_{signal} = S.P_T.R_{(\nu_{LASER})}$. If the application requires knowing the exact values of x_{CO_2}, then the constants Q_0 and Q_1 need to be estimated. This can be done during calibration, by feeding the sample cell with a known CO_2 mole fraction, and measuring V_{signal} for different values of x_{CO_2}. Next, a linear regression can be performed in order to find both Q_0 and Q_1. This linear model is suitable for $x_{CO_2} < 1x10^5\ ppm$, for the considered optical path length. A quadratic model could be used for a wider range of x_{CO_2}. However, if the purpose of the application is to detect leakage, for instance, then we need only to monitor V_{signal} for arbitrary abrupt changes in concentration of CO_2 inside the sample cell. Table 20 summarizes the set of values presented in Figure 112, Figure 113 $(D_{PL} = 1\ cm)$, and Figure 114.

Table 20: Modelled data for partial pressures in the range of $0.1\ torr \leq p_{CO_2} \leq 600\ torr$ and path length of $1\ cm$.

Mole Fraction (ppm)	*Partial Pressure* $(torr)$	$P_T\ (mW)$	$V_R\ (V)$	$ln(V_R)\ (ln\ V)$	$V_{signal}\ (V)$	$ln(V_{signal})\ (ln\ V)$
131.5789	0.100	2.5932	218.7969	5.3881	198.8769	5.2927
283.4782	0.215	2.5854	218.1966	5.3854	198.2766	5.2897
610.7354	0.464	2.5686	216.9130	5.3795	196.9930	5.2832
1315.7895	1.000	2.5331	214.1921	5.3669	194.2721	5.2693
2834.7825	2.154	2.4593	208.5300	5.3401	188.6100	5.2397
6107.3537	4.642	2.3116	197.1988	5.2842	177.2788	5.1777
13157.8947	10.000	2.0388	176.2802	5.1721	156.3602	5.0522
28347.8249	21.544	1.6086	143.2904	4.9649	123.3704	4.8152
61073.5373	46.416	1.0898	103.5004	4.6396	83.5804	4.4258
131578.9474	100.000	0.6357	68.6728	4.2294	48.7528	3.8868
327631.5789	249.000	0.2599	39.8541	3.6852	19.9341	2.9924
789473.6842	600.000	0.0705	25.3292	3.2320	5.4092	1.6881

Finally, consider the case in which we want to detect very small changes in the mole fraction x_{CO_2}. Consider as well that the present concentration of CO_2 is relatively small, i.e., that the transmitted power through the gas is high, producing an equally large current in the read-out circuit. Recalling the previous section, we have estimated the minimum readable current by the circuit as $\delta I_{min} \cong 143\ nA$, that is, the minimum step between two adjacent input DC currents. This value renders in a detectable power difference of $\delta P_{min} \cong 92.9\ nW$, for the discussed detector.

In order to find the minimum variation in the mole fraction that we could be read with the proposed system, i.e., the differential δx_{CO_2}, we have considered an arbitrary range of different partial pressures – but very close to each other – that results in fairly strong transmitted powers. This range is shown in Figure 115, namely: $12.015\ torr \leq p_{CO_2} \leq 12.065\ torr$, that is, a variation of only $50\ mtorr$. Their corresponding transmittance profiles are so close that they cannot be distinguished without a zoom – presented in the figure.

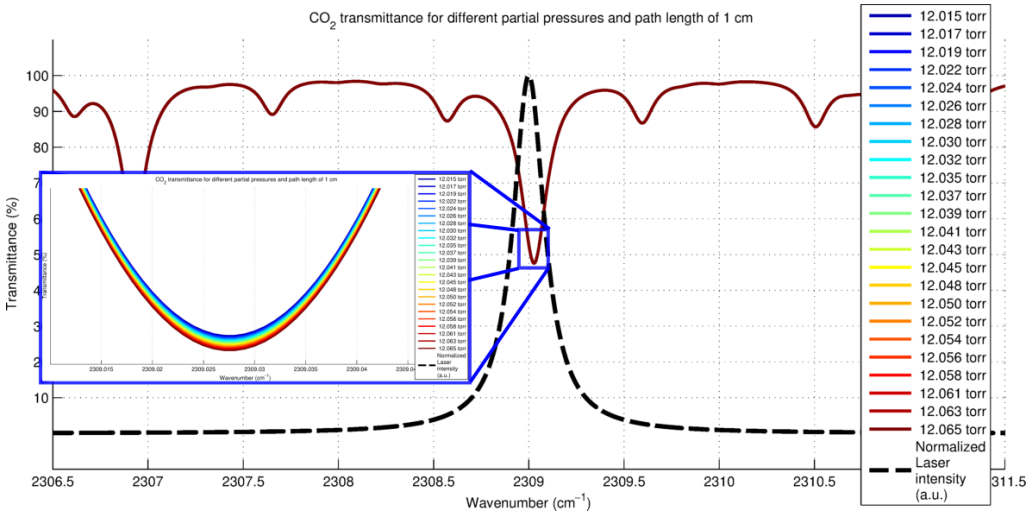

Figure 115: CO_2 transmittance for very close partial pressures in the range of $12.015\ torr \leq p_{CO_2} \leq 12.065\ torr$, and path length of 1 cm. The normalized laser intensity is shown as well.

From the model that we have implemented, the given pressure range corresponds to the following transmitted optical powers: $1.94788\,mW \leq P_T \leq 1.95000\,mW$, with a total variation of $2.12\,\mu W$. The transmitted intensity profiles of the laser are shown in Figure 116. Once again, the different curves cannot be distinguished without the help of a zoom.

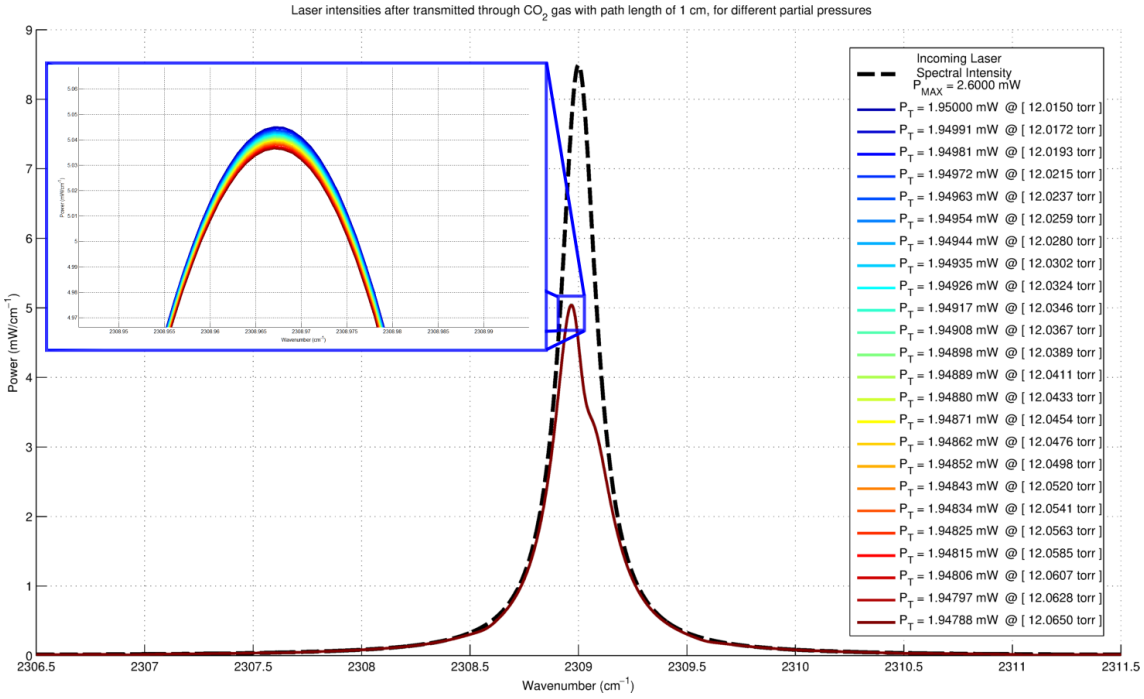

Figure 116: Transmitted laser intensity through the gas mixture and respective transmitted power for partial pressures in the range of $12.015\ torr \leq p_{CO_2} \leq 12.065\ torr$ and path length of 1 cm.

From data of Figure 116, we can plot the natural log of the transmitted laser power $\ln(P_T)$ by the mole fraction x_{CO_2}, as shown in Figure 117.

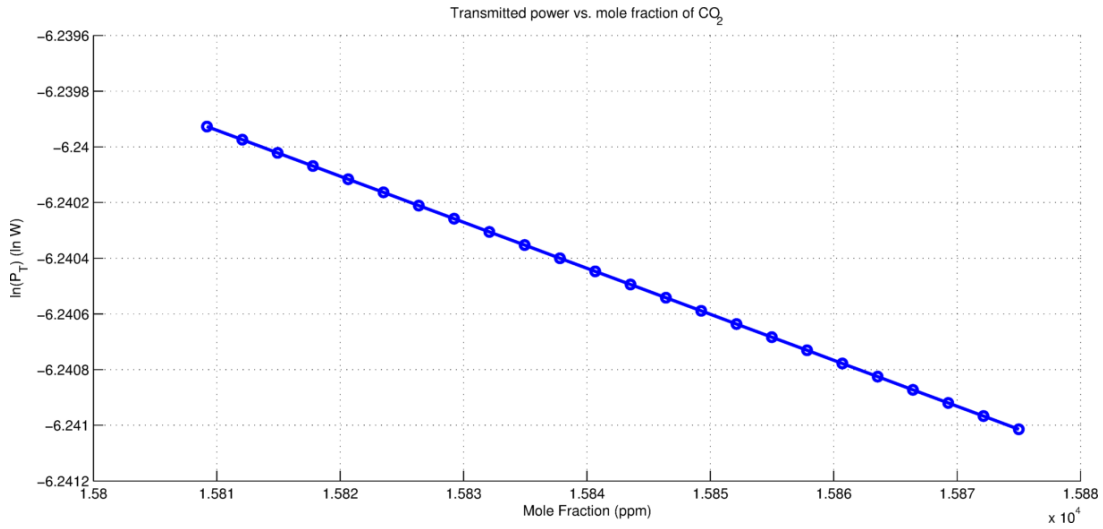

Figure 117: Output voltages V_R and V_{signal} vs. mole fraction in ppm for partial pressures in the range of $12.015\ torr \leq p_{CO_2} \leq 12.065\ torr$ and path length of $1\ cm$.

The implemented model describes how the transmitted power decays with the mole fraction, and can also be used to predict how V_{signal} will behave, as discussed earlier, since $V_{signal} = S.I_h$, and $I_{ph} = P_T.R_{(\nu_{LASER})}$. The analyzed range of optical powers can be converted

to photocurrents, as: $2.9997\ mA \leq I_{ph} \leq 3.0030\ mA$; or equivalently: $3\ mA\ -\ 0.3\ \mu A \leq I_{ph} \leq 3\ mA + 3\ \mu A$. However, in order to predict the behavior of V_{signal} with x_{CO_2}, we should re-simulate the bouncing pixel for the desired photocurrent range, for two reasons: First, because we need to guarantee that the bouncing pixel is actually capable of sensing differences as small as δI_{min} ($\cong\ 143\ nA$), even at high currents, considering that the added noise can overcome the difference between the signals; and second, because we need to evaluate the average sensitivity (S in $[mV/\mu A]$) at this range, that dictates the linearity of V_{signal} in relation to the photocurrent I_{ph}. We know from equation (38) – $S = \beta.\varepsilon.\Delta T/C_{int}$ – that the sensitivity is a function of the injection efficiency ε, i.e., the ratio between the amount of current that is fed into the bouncing pixel and the current produced by the photodetector. Although the average injection efficiency is $\varepsilon \cong 0.90$, we have seen in section 4.3.2.1, that ε may vary depending on the photocurrent range, mainly due to the low shunt resistance of the detector. As a result, the value for the sensitivity may be slightly different, if compared to the one calculated in the previous section ($S = 49.80\ mV/\mu A$).

For the following simulation, we have employed the same parameters described in Table 16, except by the number of seeds, using the following ones: 1,2,3 ... 10 ($\#_{seed} = 10$). In this transient simulation, the photocurrent varied from $3\ mA$ with steps of $\delta I_{min} = 143\ nA$, and a total of 11 steps, that is: $3\ mA \leq I_{ph} \leq 3\ mA + 1.43\ \mu A$. Figure 118 shows a zoom in at the end of the integration time. We can see that noise can be really detrimental for this situation, where small current differences need to be detected. Therefore, in order to accurately sense the correct signal, it would be mandatory for the practical system to perform redundant measurements.

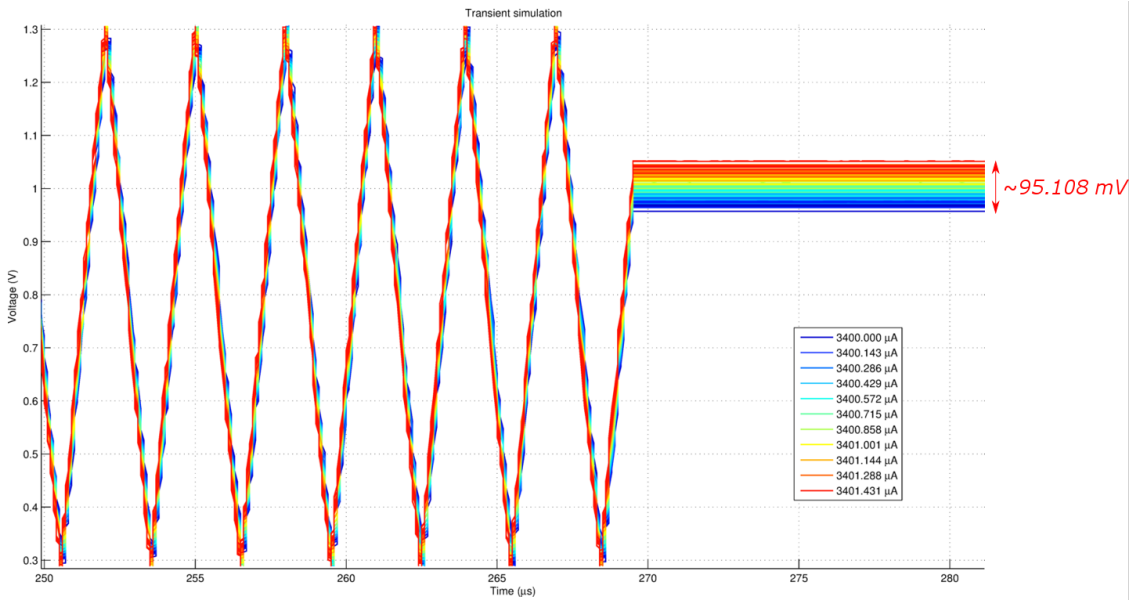

Figure 118: Zoom in the transient noise simulation for a narrow DR with 10 seeds.

Figure 119 shows the average reconstructed voltage V_R extracted from the analog voltage V_{CAP} of Figure 118, and the serial signal (omitted here). We can see that the signal V_R presents good linearity, even with the apparently large noise of Figure 118, and with the ADC of 9 bits.

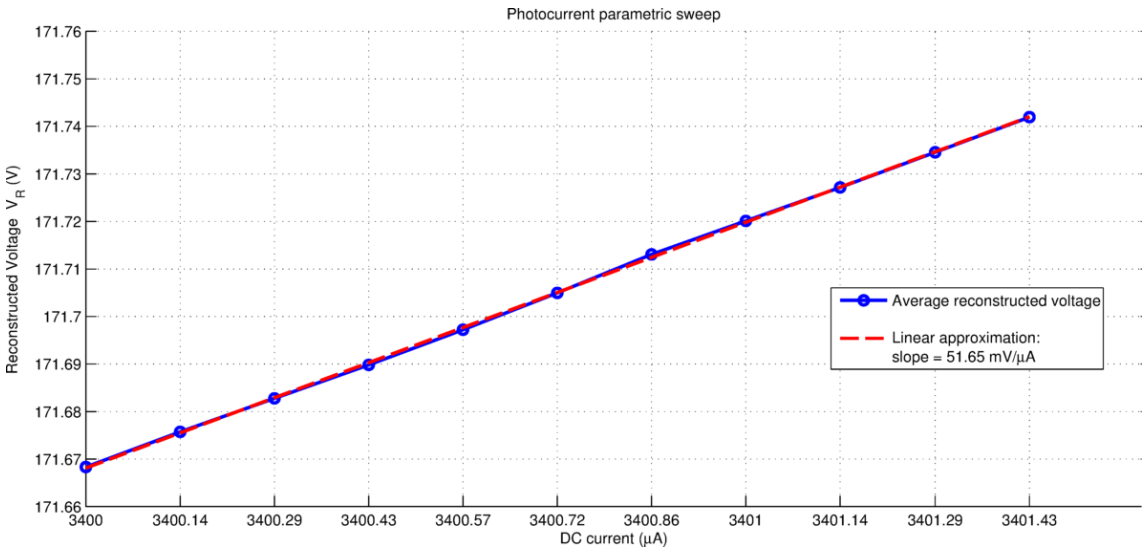

Figure 119: Average reconstructed voltage from $3.4\ mA$ up to $3.4\ mA + 1.43\ \mu A$, with $\Delta T = 256\ \mu s$, and ADC of 9 bits.

The average and standard deviation of the reconstructed voltages of Figure 119, for each DC current, are presented in Table 21. Note that the minimum difference between two values of V_R is $7.0450\ mV$, that is larger than the voltage resolution of the ADC with 9 bits $(3.5225\ mV)$. If we had employed the ADC with 8 bits (resolution of $7.0588\ mV$), then it would have generated more quantization errors, since its resolution is not sufficient. This demonstrates that we have properly designed the number of bits for the circuit, suitable for this application. In addition, the maximum standard deviation $(6.8566\ mV)$ is smaller than the minimum difference between adjacent signals. Therefore, we can conclude that the bouncing pixel, assuming a number of redundant samples, is able to detect input photocurrent differences as small as $\delta I_{min} = 143\ nA$, even when the DC input current is relatively high.

Table 21: Output voltage averages and standard deviations.

Photocurrent (μA)	$V_R\ (mV)$	$\sigma_{V-noise}\ (mV)$
3400.0000	171.6683	6.5269
3400.1430	171.6757	6.4312
3400.2860	171.6828	6.4312
3400.4290	171.6898	6.4312
3400.5720	171.6972	6.3122
3400.7150	171.7050	6.4419
3400.8580	171.7131	6.2573
3401.0010	171.7201	6.2573
3401.1440	171.7272	6.2573
3401.2880	171.7346	6.8566
3401.4310	171.7420	6.8061

That said, we can extract the sensitivity estimation from Figure 119, given by the slope of the linearized curve $(S' = 51.65\ mV/\mu A)$, in order to apply it to our model of CO_2 gas absorption, for the discussed photocurrent range. Note that the photocurrent range considered for the model is almost two times larger than the simulated, since simulations are very computer-intensive. However, we assume here that the value of S'does not change much, since

this photocurrent range is too small to result in a significant variation of the injection efficiency (ε) – recalling that S scales linearly with ε. Therefore, applying $V_{signal} = S'.I_{ph}$, we can plot the curve of $\ln(V_{signal})$ vs. x_{CO_2}, shown in Figure 120.

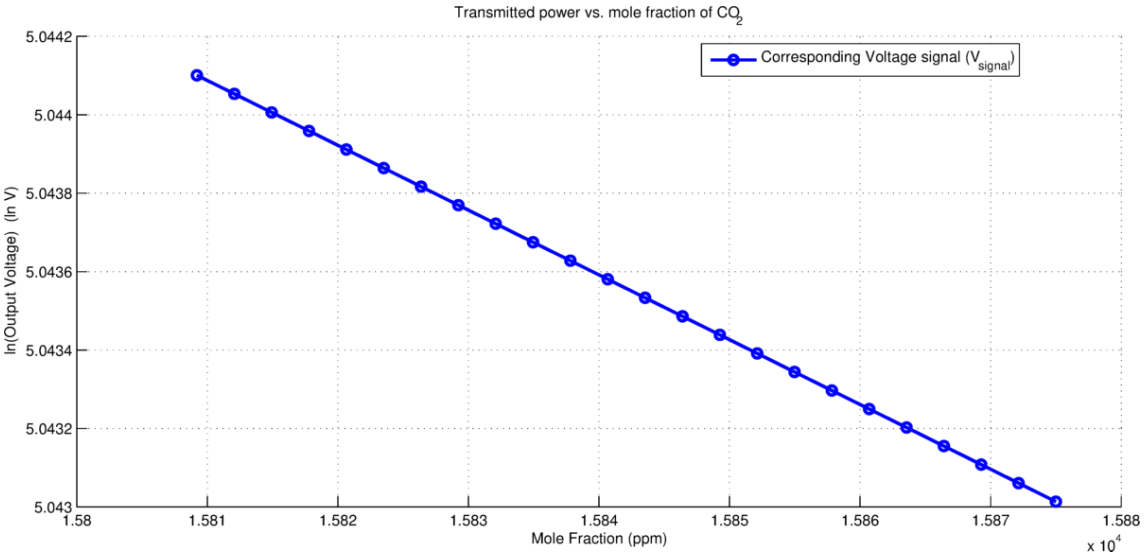

Figure 120: Output voltages V_R and V_{signal} vs. mole fraction in ppm for partial pressures in the range of $12.015\ torr \leq p_{CO_2} \leq 12.065\ torr$ and path length of $1\ cm$.

This curve can be modelled by equation (80), and could be used for calibration of the detection system, in order to help quantifying the exact mole fraction of CO_2, for a given condition. Furthermore, the curve of Figure 120 varies the mole fraction with a step of approximately $2.86\ ppm$, while the average step of the voltage signal corresponds to $\delta V = 7.3\ mV$. As a conclusion, if properly calibrated, the system could be designed to detect differences of mole fractions as small as $\delta x_{CO_2} = 2.86\ ppm$. This measuring resolution is at least 10 times better than that of a typical high accuracy commercial detector ($\delta x_{CO_2} \cong 30\ ppm$) [112].

5. CONCLUSION

The main purpose of this thesis was to address and propose better solutions to two main parts of a photodetection system at the infrared spectra: the optics and the readout electronics. For that matter, the demands for both parts were carefully analyzed in the section on theoretical background, and their behavior examined through tests and simulations. Ultimately, they were joined together in an idealized example of application at the infrared, that is, the CO_2 gas detection system, with absorption around $\lambda = 4.2\ \mu m$.

In order to read the signal from the photodetector, a readout circuit has to be used. However, generally IR photodetectors feature low shunt resistance, large dark current, and increased noise, posing a great difficulty for readout. We have seen that a pixel circuit capable of circumventing these limitations would be one that feature a high-dynamic range (HDR) capability, since it is able to accommodate very large currents, and because SNR increases with the input current. However, general pixel circuits for the IR that have a large charge storage capacity present low sensitivities, especially for large currents. And, since the signal is embedded on a relatively large current, a high sensitivity is desired. The bouncing pixel has the advantage of presenting both HDR capability and high sensitivity, being able to detect very small differences in the input signal, even for large currents. This is done by preventing that the integrated photocurrent saturates the output voltage signal at the maximum operating voltage (V_{DD}). In fact, the bouncing pixel virtually enlarges the saturation voltage to a much larger value, which depends on the number of bits of the deployed counter.

Beyond that, for infrared detectors, the pixel circuit should offer an injection efficiency close to the unity, that is, all current from the photodetector should be injected into the circuit. Any current not being injected flows through the shunt resistance of the detector, which in turn, raises the voltage across its terminals (V_{DET}). Ideally, the detector bias should be very stable, and as close to zero as possible. The bouncing pixel circuit – and especially the cascoded version – is capable of interfacing with IR photodetectors with very small shunt resistances, and yet feature very stable and small bias across the detector.

From an optics point of view, the microlenses offer a good solution to improve the *SNR* of photodetection systems, mainly when the detectors or pixels present a low fill-factor (ratio between the photosensitive area and the total area). At the infrared spectra, silicon proves to be a suitable material, mainly due to its low cost and high refractive index. In addition, we have used a microfabrication technique proposed by *de Lima Monteiro et al., 2003* that utilizes only a single binary mask and two steps of KOH etching, that is less expensive than the conventional microlenses fabrication processes. The fabricated lenses presented low roughness for the considered wavelengths. Also, the estimated focal lengths were very short, being suitable for compact systems. Afterwards, the actual focal points of the fabricated microlenses array were measured, proving to be sharp focal points. This last measurement was done at NIR (Near-Infrared), and should be performed at other spectral bands within the IR, like at MWIR ($3 - 8\ \mu m$). Therefore, to methodically analyze the microlenses at other spectral ranges are a good topic for future work.

The last component of the optical part is the anti-reflective coating (ARC), necessary to reduce the reflectance of silicon at the target wavelengths. It is also important to note that silicon double-side polished wafers (DSPW) are essential for optical applications, and therefore all tested ARC layers were deposited over DSPW. Simulations and depositions of SiN thin-film layers were performed, providing good results for SWIR and MWIR spectral ranges. Lastly, the appropriate (n, k) dataset, for future ARC designs was extracted from the measured samples (and presented in the Appendix A).

At the end, to exemplify a real application, a model was implemented describing the wavelength-dependent CO_2 gas absorption. The detection system was conceived by considering all parts of the system, including the commercial QCL laser and IR detector, the silicon optics plus ARC, and the readout bouncing pixel. For this case, the light source – the laser – is pointed straight at the detector passing through the gas, and any amount of CO_2 in the path would absorb a portion of the laser power. Without the presence of CO_2, the current produced by the detector is the maximum one. To detect small amounts of CO_2 in the cell, the pixel circuit should be able to sense very small differences in the large current produced by the detector, that corresponds to the small variation in the transmitted laser power. With the considered modelled system, very small changes in the mole fraction of CO_2 inside a gas sample cell, of $\sim 2.86\ ppm$, could be detected. Such high detection resolution is achievable only due to the bouncing pixel, which has a very high sensitivity even for large currents, allowing it to detect small signals embedded on larger background levels. To experimentally implement such system is a topic for future work.

Summarizing, both the silicon microlens and the designed ARC layer showed to be a good solution to enhance the SNR of IR detection systems. And mostly, the proposed bouncing pixel circuit presented all desired features that makes it a perfect solution suitable for infrared applications, that is, HDR capability, high sensitivity, high injection efficiency, and bias stability. Five different chips have already been fabricated with the bouncing pixel structure on it. Appendix C summarizes all of these chips, and to experimentally test them is a topic for future work.

6. REFERENCES

[1] O. Yadid-Pecht, R. Etienne-Cummings, and J. M. Walker, *CMOS Imagers - From Phototransduction to Image Processing*. Springer Science & Business Media, 2007.

[2] "Dark frame subtraction," 1998.

[3] "Efficient dark current subtraction in an image sensor," 2001.

[4] A. Rogalski, "Infrared detectors: status and trends," 2003.

[5] A. Rogalski, *Infrared Detectors*. 2011.

[6] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, "Wide-Dynamic-Range CMOS Image Sensors, a Comparative Performance Analysis," *Electron Devices, IEEE Trans.*, vol. 56, no. 11, pp. 2446–2461, 2009.

[7] B. O. Hultgren and D. W. Hertel, "Megapixel mythology and photospace: estimating photospace for camera phones from large image sets," *Proc. SPIE*, vol. 6808, pp. 680810–680818, 2008.

[8] A. Karim and J. Y. Andersson, "Infrared detectors: Advances, challenges and new technologies," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 51, p. 12001, 2013.

[9] M. J. Riedl, *Optical design fundamentals for infrared systems*, vol. 48. SPIE press, 2001.

[10] M. Diakides, J. D. Bronzino, and D. R. Peterson, *Medical infrared imaging: principles and practices*. CRC press, 2012.

[11] M. Kumar, T. B. Pandian, E. Sreekiran, and S. Narayanan, "Benefit of dual layer silicon nitride anti-reflection coating," in *Conference Record of the Thirty-first IEEE Photovoltaic Specialists Conference, 2005.*, pp. 1205–1208.

[12] INCT-DISSE, "INCT-DISSE." [Online]. Available: http://inct.cnpq.br/web/inct-disse.

[13] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. John wiley & sons, 2006.

[14] S. M. Sze, *Semiconductor Devices: Physics and Technology*. 2006.

[15] S. L. Pyshkin and J. M. Ballato, *ADVANCED MATERIALS AND DEVICES Edited by Sergei L . Pyshkin*. InTech, 2013.

[16] R. G. Newton, "Scattering Theory of Waves and Particles."

[17] J. Chou, *Hazardous gas monitors: a practical guide to selection, operation and applications*. McGraw-Hill Professional Publishing, 2000.

[18] S. Kong, "Infrared micro-spectrometer based on a multi-slit grating," 2002.

[19] F. Jutzi, D. H. B. Wicaksono, G. Pandraud, N. de Rooij, and P. J. French, "Far-infrared sensor with LPCVD-deposited low-stress Si-rich nitride absorber membrane—Part 1. Optical absorptivity," *Sensors Actuators A Phys.*, vol. 152, no. 2, pp. 119–125, Jun. 2009.

[20] J. W. Beletic *et al.*, "Teledyne Imaging Sensors : Infrared imaging technologies for Astronomy & Civil Space," *Proc. SPIE*, vol. 7021, no. 1, pp. 1–14, 2008.

[21] Á. D. B. Maia, “Crescimento, fabricação e teste de fotodetectores de radiação infravermelha baseados em pontos quânticos,” Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo, São Paulo, 2012.

[22] T. Gebhard, D. Alvarenga, M. P. Pires, K. Unterrainer, P. S. S. Guimaraes, and P. L. De Souza, “Superposition of Positive and Negative Contributions to the Photocurrent Spectrum of InAs/InAlGaAs/InP Quantum Dot Infrared Photodetectors,” in *ECS Transactions*, 2009, vol. 23, no. 1, pp. 521–526.

[23] D. R. Alvarenga *et al.*, “Dual Sign Photocurrent in Quantum Dot Structures for Infrared Photodetection,” in *ECS Transactions*, 2010, vol. 31, no. 1, pp. 207–211.

[24] D. R. Alvarenga *et al.*, “Exceptionally Narrow-Band Quantum Dot Infrared Photodetector,” *IEEE J. Quantum Electron.*, vol. 48, no. 10, pp. 1360–1366, Oct. 2012.

[25] R. Baker, *CMOS: circuit design, layout, and simulation*. 2008.

[26] S. Alexander, *Optical communication receiver design*. 1997.

[27] T. J. Sobering, “Equivalent Noise Bandwidth,” *SDE Consult. Technote*, vol. 1, 2008.

[28] M. Vollmer and K.-P. Möllmann, *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons, 2010.

[29] E. R. F. and B. Pain, “Infrared Readout Electronics for Space Science Sensors: State of the Art and Future Directions,” pp. 1–24.

[30] European Machine Vision Association, “EMVA Standard 1288: Standard for characterization of Image Sensors and Cameras,” p. 82, 2012.

[31] P. N. A. Belmonte, D. W. de Lima Monteiro, F. S. Torres, and P. J. French, “Linear high-dynamic-range bouncing pixel with single sample,” in *Proceedings of 2013 International Image Sensor Workshop, ISSW*, 2013, pp. 12–16.

[32] Z. Y. Chong and W. Sansen, *Low-noise wide-band amplifiers in bipolar and CMOS technologies*, vol. 117. Springer Science & Business Media, 2013.

[33] J. Ohta, *Smart CMOS image sensors and applications*. CRC press, 2007.

[34] M. Johnson, *Photodetection and Measurement: Making Effective Optical Measurements for an Acceptable Cost*. McGraw Hill Professional, 2003.

[35] G. Davis and R. Jones, *Yamaha-The Sound Reinforcement Handbook*. 1990.

[36] L. Yao, “CMOS readout circuit design for infrared image sensors,” in *International Symposium on Photoelectronic Detection and Imaging 2009*, 2009, vol. 7384, p. 73841B–73841B.

[37] A. Rogalski, “Analysis of the R0A product in n+-p Hg1−xCdxTe photodiodes,” *Infrared Phys.*, vol. 28, no. 3, pp. 139–153, 1988.

[38] J. Xu and M. Tacke, “Temperature dependence of R0A product for PbSe photodiodes fabricated by MBE and diffusion,” *Infrared Phys.*, vol. 33, no. 3, pp. 151–157, 1992.

[39] T. Yuan, S.-J. Chua, and Y. Jin, “Calculation of the R0A product in n+–n–p and p+–p–n GaInAsSb infrared detectors,” *Infrared Phys. Technol.*, vol. 45, no. 3, pp. 181–189, 2004.

[40] E. R. Fossum, S. Mendis, and S. E. Kemeny, “Active pixel sensor with intra-pixel charge

transfer," 28-Nov-1995.

[41] T.-P. Sun, Y.-C. Lu, L.-L. Kang, and H.-L. Shieh, "A buffer direct injection and direct injection readout circuit with mode selection design for infrared focal plane arrays," *Infrared Phys. Technol.*, vol. 63, pp. 140–146, 2014.

[42] M. Tepegoz and T. Akin, "A readout circuit for QWIP infrared detector arrays using current mirroring integration," *… , 2003. ESSCIRC'03. Proc. …*, no. Cmi, pp. 133–136, 2003.

[43] N. Yoon, B. Kim, H. Lee, and C. Kim, "High injection efficiency readout circuit for low resistance infra-red detector," *Electron. Lett.*, vol. 35, no. 18, pp. 1507–1508, 1999.

[44] H. Kulah and T. Akin, "A CMOS current mirroring integration readout structure for infrared focal plane arrays," in *State Circuits Conference, 1998. ESSCIRC'98 …*, 1998, vol. 855, pp. 8–11.

[45] H. Kulah and T. Akin, "A current mirroring integration based readout circuit for high performance infrared FPA applications," *IEEE Trans. Circuits Syst. II Analog Digit. Signal Process.*, vol. 50, no. 4, pp. 181–186, 2003.

[46] N. Yoon, B. Kim, H. Lee, and H. Shin, "New unit cell of current mirroring direct injection circuit for focal plane arrays," *AeroSense'*, 1997.

[47] J. Longo, D. Cheung, and A. Andrews, "Infrared focal planes in intrinsic semiconductors," *on Electron Devices*, no. 9, pp. 6–10, 1978.

[48] A. Sedra and K. Smith, *Microelectronic circuits*. 1998.

[49] D. Pozar, *Microwave Engineering Fourth Edition*. 2005.

[50] A. El Gamal, "EE392B Classnotes: Introduction to Image Sensors and Digital Cameras," *Isscc 2002*. 2001.

[51] S. Decker, D. McGrath, K. Brehmer, and C. G. Sodini, "A 256×256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output," *IEEE J. Solid-State Circuits*, vol. 33, no. 12, pp. 2081–2091, 1998.

[52] N. Akahane, S. Sugawa, S. Adachi, K. Mori, T. Ishiuchi, and K. Mizobuchi, "A Sensitivity and Linearity Improvement of a 100-dB Dynamic Range CMOS Image Sensor Using a Lateral Overflow Integration Capacitor," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp. 851–858, Apr. 2006.

[53] Y. Hagihara, "Logarithmic-converting CMOS area image sensors with an FPN cancel circuit," *KYOKAI JOHO IMEJI ZASSHI*, vol. 55, no. 7, pp. 1039–1044, 2001.

[54] N. Akahane, R. Ryuzaki, S. Adachi, K. Mizobuchi, and S. Sugawa, "16.7 A 200dB Dynamic Range Iris-less CMOS Image Sensor with Lateral Overflow Integration Capacitor using Hybrid Voltage and Current Readout Operation," 2006.

[55] E. C. Fox, J. Hynecek, and D. R. Dykaar, "Wide-dynamic-range pixel with combined linear and logarithmic response and increased signal swing," in *Electronic Imaging*, 2000, pp. 4–10.

[56] K. p. Frohmader, "A novel MOS compatible light intensity-to-frequency converter suited for monolithic integration," *IEEE J. Solid-State Circuits*, vol. 17, no. 3, pp. 588–591, Jun. 1982.

[57] X. Wang, W. Wong, and R. Hornsey, “A High Dynamic Range CMOS Image Sensor With Inpixel Light-to-Frequency Conversion,” *IEEE Trans. Electron Devices*, vol. 53, no. 12, 2006.

[58] X. Guo, X. Qi, and J. G. Harris, “A Time-to-First-Spike CMOS Image Sensor,” *IEEE Sens. J.*, vol. 7, no. 8, pp. 1165–1175, Aug. 2007.

[59] V. Brajovic and T. Kanade, “New massively parallel technique for global operations in embedded imagers,” in *1995 IEEE Workshop on CCDs and Advanced Image Sensors*, 1995.

[60] O. Yadid-Pecht and E. Fossum, “Image sensor with ultra-high linear-dynamic range utilizing dual output CMOS active pixel sensors,” *IEEE Trans. Electron Devices*, vol. 44, no. 10, pp. 1721–1723, 1997.

[61] M. Mase, S. Kawahito, M. Sasaki, Y. Wakamori, and M. Furuta, “A wide dynamic range CMOS image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic A/D converters,” *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2787–2795, Dec. 2005.

[62] T. Hamamoto and K. Aizawa, “A computational image sensor with adaptive pixel-based integration time,” *IEEE J. Solid-State Circuits*, vol. 36, no. 4, pp. 580–585, Apr. 2001.

[63] O. Yadid-Pecht and A. Belenky, “In-pixel autoexposure cmos aps,” *IEEE J. Solid-State Circuits*, vol. 38, no. 8, pp. 1425–1428, Aug. 2003.

[64] J. Piotrowski and A. Rogalski, “High-operating-temperature infrared photodetectors,” 2007.

[65] W. J. Smith, *Modern Optical Engineering*. 2000.

[66] J. E. Greivenkamp, *Field guide to geometrical optics*. 2004.

[67] P. N. A. Belmonte, D. W. de Lima Monteiro, R. F. de Oliveira Costa, P. J. French, and G. Pandraud, “Microfabrication and characterization of single-mask silicon microlens arrays for the IR spectra,” in *SPIE Photonics Europe*, 2014, p. 91300D–91300D.

[68] D. W. de Lima Monteiro, O. Akhzar-Mehr, P. M. Sarro, and G. Vdovin, “Single-mask microfabrication of aspherical optics using KOH anisotropic etching of Si,” *Opt. Express*, vol. 11, no. 18, p. 2244, Sep. 2003.

[69] Y. Jianhua, O. Wen, and O. Yi, “Fabrication of a 100% fill-factor silicon microlens array,” *J. Semicond.*, 2012.

[70] Z. D. Popovic, R. A. Sprague, and G. A. Neville Connell, “Technique for monolithic fabrication of microlens arrays,” *Appl. Opt.*, vol. 27, no. 7, p. 1281, Apr. 1988.

[71] “Gray-tone lithography using optical diffusers,” 2008.

[72] J. Yao *et al.*, “Coding gray-tone mask for refractive microlens fabrication,” *Microelectron. Eng.*, vol. 53, no. 1, pp. 531–534, 2000.

[73] S. Franssila, *Introduction to Microfabrication*. 2010.

[74] D. Kendall, G. De Guel, and S. Guel Sandoval, Chemically etched micromirrors in silicon, *Appl. Phys.*, 1988.

[75] D. Kendall and W. Eaton, “Micromirror arrays using KOH: H2O micromachining of silicon

for lens templates, geodesic lenses, and other applications," *Optical*, 1994.

[76] S. Schiller, G. Beister, W. Sieber, G. Schirmer, and E. Hacker, "Influence of deposition parameters on the optical and structural properties of TiO2 films produced by reactive dc plasmatron sputtering," *Thin Solid Films*, 1981.

[77] J. Kischkat, S. Peters, B. Gruska, and M. Semtsiv, "Mid-infrared optical properties of thin films of aluminum oxide, titanium dioxide, silicon dioxide, aluminum nitride, and silicon nitride," *Applied*, 2012.

[78] D. Poenar, "Thin film colour sensors," 1996.

[79] A. Emadi, "Linear-variable optical filters for microspectrometer application," 2010.

[80] Software Spectra Web Site, "TFCalc™." [Online]. Available: http://www.sspectra.com/.

[81] H. Photonics, "Characteristics and use of Infrared Detectors," 2004.

[82] P. French, P. Sarro, R. Mallée, and E. Fakkeldij, "Optimization of a low-stress silicon nitride process for surface-micromachining applications," *Sensors Actuators A*, 1997.

[83] H. Murray, "Overview and comparison of portable spectroscopy techniques: FTIR, NIR and Raman FTIR Spectroscopy Instrument design."

[84] C. A. de Moraes Cruz, P. N. A. Belmonte, and D. W. de Lima, "Linear Mode High Dynamic Range Bouncing Pixel with Single Transistor," *Int. Image Sens. Work. ISSW 2015*, pp. 0–3, 2015.

[85] S. Wolf, "Silicon processing for the VLSI era," in *LATTICE*, 1995, pp. 559–581.

[86] D. Didascalou, M. Dottling, N. Geng, and W. Wiesbeck, "An approach to include stochastic rough surface scattering into deterministic ray-optical wave propagation modeling," *IEEE Trans. Antennas Propag.*, vol. 51, no. 7, pp. 1508–1515, Jul. 2003.

[87] N. Pinel, C. Bourlier, and J. Saillard, "DEGREE OF ROUGHNESS OF ROUGH LAYERS: EXTENSIONS OF THE RAYLEIGH ROUGHNESS CRITERION AND SOME APPLICATIONS," *Prog. Electromagn. Res. B*, vol. 19, pp. 41–63, 2010.

[88] NCSU Nanofabrication Facility, "Report on KOH Process Module Etch Characteristics and Design Guide," 2005.

[89] P. Enoksson, B. K. E, S. H. M. K. K. K. and P. A. Offereins H L, K. K. N. R. and S. H. Csepregi V L, S. G. and S. E. Enoksson P, and S. G. and S. E. Enoksson P, "New structure for corner compensation in anisotropic KOH etching," *J. Micromechanics Microengineering*, vol. 7, no. 3, pp. 141–144, Sep. 1997.

[90] A. N. Belbachir, Ed., *Smart Cameras*. Boston, MA: Springer US, 2010.

[91] *Foculus FO442SB/SC Firewire Camera*. .

[92] "Filters - Ir longpass." [Online]. Available: http://www.visionlighttech.com/products.php?hoofdgroep=filters&subgroep=ir-longpass.

[93] M. A. Green and M. J. Keevers, "Optical properties of intrinsic silicon at 300 K," *Prog. Photovoltaics Res. Appl.*, vol. 3, no. 3, pp. 189–192, 1995.

[94] E. H. Korte and A. Röseler, "Infrared reststrahlen revisited: commonly disregarded

optical details related to n<1," *Anal. Bioanal. Chem.*, vol. 382, no. 8, pp. 1987–1992, Aug. 2005.

[95] A. B. Kuzmenko, "Kramers–Kronig constrained variational analysis of optical spectra," *Rev. Sci. Instrum.*, vol. 76, no. 8, p. 83108, 2005.

[96] U. of G. Alexey Kuzmenko, "RefFIT." [Online]. Available: http://optics.unige.ch/alexey/reffit.html.

[97] M. J. Mageto, C. M. Maghanga, and M. Mwamburi, "The Lorentz Oscillator Model Simulation Illustrating a Broad Maximum in the Bulk Reflectance for Frequencies Just Above the Resonance Frequency," *African Rev. Phys.*, vol. 7, pp. 95–105, 2012.

[98] I. F. Almog, M. S. Bradley, and V. Bulovi, "The Lorentz Oscillator and its Applications," 2011.

[99] C. Kittel, "Introduction to Solid State Physics," *Am. J. Phys.*, vol. 35, no. 6, p. 547, 1967.

[100] HAMAMATSU, "InAsSb photovoltaic detector (model P11120-201)."

[101] J. Faist, F. Capasso, and C. Sirtori, "Quantum Cascade Laser," 1994.

[102] M. Razeghi, "Longwavelength InAsSb infrared photodetectors," DTIC Document, 1995.

[103] K. Wu, F. Li, X. Cheng, Y. Yang, X. Lin, and Y. Xia, "Sensitive detection of CO2 concentration and temperature for hot gases using quantum-cascade laser absorption spectroscopy near 4.2 μm," *Appl. Phys. B*, vol. 117, no. 2, pp. 659–666, Nov. 2014.

[104] Q. Huang, F. Wang, H. Zhang, J. Yan, M. Ni, and K. Cen, "In-situ CO measurement of gas and oil combustion flame using near infrared tunable diode laser with direct and modulated absorption signals," *Opt. Commun.*, vol. 306, pp. 99–105, 2013.

[105] C. W. Schneider, Z. Kucerovsky, and E. Brannen, "Carbon dioxide absorption of He–Ne laser radiation at 42 μm: characteristics of self and nitrogen broadened cases," *Appl. Opt.*, vol. 28, no. 5, p. 959, Mar. 1989.

[106] B. H. Winters, S. Silverman, and W. S. Benedict, "Line shape in the wing beyond the band head of the 4·3 μ band of CO2," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 4, no. 4, pp. 527–537, 1964.

[107] R. M. Goody and Y. L. Yung, *Atmospheric radiation: theoretical basis*. Oxford University Press, 1995.

[108] A. G. Maki and J. S. Wells^, "New wavenumber calibration tables from heterodyne frequency measurements," *J. Res. Natl. Inst. Stand. Technol. Res. Natl. Inst. Stand. Technol*, vol. 97, no. 97.

[109] B. J. Soukup, "A fiber optic array for the detection of sub-surface carbon dioxide at carbon sequestration sites," Montana State University - Bozeman, College of Engineering, 2014.

[110] K. Repasky and Kevin, "Development of a 1 x N Fiber Optic Sensor Array for Carbon Sequestration Site Monitoring," Pittsburgh, PA, and Morgantown, WV (United States), Feb. 2014.

[111] L. S. Rothman *et al.*, "The HITRAN 2004 molecular spectroscopic database," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 96, no. 2, pp. 139–204, 2005.

[112] “CO2 Meter - Portable Battery Operated Carbon Dioxide Meter - CO2 Meter by AZ Instruments.” [Online]. Available: http://www.co2meter.com/collections/portable/products/psense-portable-co2-meter. [Accessed: 08-Mar-2017].

APPENDIX A

Table A: Optical constants (n, k) for the SiN (LPCVD-deposited low-stress Si-rich nitride technique) with *wavenumbers* varying in the range from 396 to 5000 cm^{-1}

Wave-number (cm^{-1})	n	k (x 10^{-3})	Wave-number (cm^{-1})	n	k (x 10^{-3})	Wave-number (cm^{-1})	n	k (x 10^{-3})	Wave-number (cm^{-1})	n	k (x 10^{-3})
395.89	3.08	220.90	**1656.89**	1.87	11.71	**2917.89**	2.07	1.79	**4178.89**	2.11	0.57
410.56	3.10	274.79	**1671.55**	1.88	11.44	**2932.55**	2.07	1.76	**4193.55**	2.11	0.56
425.22	3.10	342.82	**1686.22**	1.88	11.16	**2947.21**	2.07	1.73	**4208.21**	2.11	0.55
439.88	3.08	421.06	**1700.88**	1.89	10.89	**2961.88**	2.07	1.70	**4222.87**	2.11	0.55
454.55	3.02	493.58	**1715.54**	1.90	10.62	**2976.54**	2.08	1.68	**4237.54**	2.11	0.54
469.21	2.93	532.87	**1730.21**	1.90	10.35	**2991.20**	2.08	1.65	**4252.20**	2.11	0.54
483.87	2.84	519.73	**1744.87**	1.91	10.09	**3005.87**	2.08	1.62	**4266.86**	2.11	0.53
498.53	2.77	464.19	**1759.53**	1.91	9.83	**3020.53**	2.08	1.60	**4281.52**	2.11	0.52
513.20	2.74	394.02	**1774.19**	1.92	9.58	**3035.19**	2.08	1.57	**4296.19**	2.11	0.52
527.86	2.74	330.31	**1788.86**	1.92	9.33	**3049.85**	2.08	1.55	**4310.85**	2.11	0.51
542.52	2.77	280.83	**1803.52**	1.93	9.09	**3064.52**	2.08	1.53	**4325.51**	2.11	0.51
557.19	2.81	245.59	**1818.18**	1.93	8.85	**3079.18**	2.08	1.50	**4340.18**	2.11	0.50
571.85	2.85	222.14	**1832.84**	1.94	8.62	**3093.84**	2.08	1.48	**4354.84**	2.11	0.50
586.51	2.89	207.93	**1847.51**	1.94	8.40	**3108.50**	2.08	1.46	**4369.50**	2.11	0.49
601.17	2.94	201.04	**1862.17**	1.94	8.18	**3123.17**	2.08	1.43	**4384.16**	2.11	0.49
615.84	2.99	200.19	**1876.83**	1.95	7.97	**3137.83**	2.08	1.41	**4398.83**	2.11	0.48
630.50	3.04	204.64	**1891.50**	1.95	7.76	**3152.49**	2.08	1.39	**4413.49**	2.11	0.48
645.16	3.10	214.13	**1906.16**	1.96	7.56	**3167.16**	2.08	1.37	**4428.15**	2.11	0.47
659.82	3.16	228.78	**1920.82**	1.96	7.37	**3181.82**	2.08	1.35	**4442.82**	2.11	0.47
674.49	3.22	249.09	**1935.48**	1.96	7.18	**3196.48**	2.08	1.33	**4457.48**	2.11	0.46
689.15	3.29	276.03	**1950.15**	1.97	7.00	**3211.14**	2.08	1.31	**4472.14**	2.11	0.46
703.81	3.37	311.10	**1964.81**	1.97	6.82	**3225.81**	2.09	1.29	**4486.80**	2.11	0.45
718.48	3.46	356.57	**1979.47**	1.97	6.65	**3240.47**	2.09	1.27	**4501.47**	2.11	0.45
733.14	3.55	415.81	**1994.13**	1.98	6.49	**3255.13**	2.09	1.25	**4516.13**	2.11	0.44
747.80	3.66	493.85	**2008.80**	1.98	6.32	**3269.79**	2.09	1.24	**4530.79**	2.11	0.44
762.46	3.79	598.29	**2023.46**	1.98	6.17	**3284.46**	2.09	1.22	**4545.45**	2.11	0.43
777.13	3.93	740.69	**2038.12**	1.99	6.02	**3299.12**	2.09	1.20	**4560.12**	2.11	0.43
791.79	4.08	938.56	**2052.79**	1.99	5.87	**3313.78**	2.09	1.18	**4574.78**	2.11	0.43
806.45	4.22	1217.22	**2067.45**	1.99	5.73	**3328.45**	2.09	1.17	**4589.44**	2.11	0.42
821.11	4.33	1607.23	**2082.11**	1.99	5.59	**3343.11**	2.09	1.15	**4604.11**	2.11	0.42
835.78	4.31	2123.56	**2096.77**	2.00	5.46	**3357.77**	2.09	1.13	**4618.77**	2.11	0.41
850.44	4.07	2706.55	**2111.44**	2.00	5.33	**3372.43**	2.09	1.12	**4633.43**	2.11	0.41
865.10	3.54	3177.56	**2126.10**	2.00	5.20	**3387.10**	2.09	1.10	**4648.09**	2.11	0.41
879.77	2.87	3370.10	**2140.76**	2.00	5.08	**3401.76**	2.09	1.09	**4662.76**	2.11	0.40
894.43	2.26	3299.68	**2155.43**	2.01	4.96	**3416.42**	2.09	1.07	**4677.42**	2.11	0.40
909.09	1.79	3084.84	**2170.09**	2.01	4.85	**3431.09**	2.09	1.06	**4692.08**	2.11	0.39
923.75	1.44	2818.73	**2184.75**	2.01	4.74	**3445.75**	2.09	1.04	**4706.74**	2.11	0.39
938.42	1.20	2546.99	**2199.41**	2.01	4.63	**3460.41**	2.09	1.03	**4721.41**	2.11	0.39
953.08	1.03	2286.78	**2214.08**	2.01	4.53	**3475.07**	2.09	1.02	**4736.07**	2.11	0.38
967.74	0.91	2042.47	**2228.74**	2.02	4.43	**3489.74**	2.09	1.00	**4750.73**	2.11	0.38
982.41	0.82	1813.36	**2243.40**	2.02	4.33	**3504.40**	2.09	0.99	**4765.40**	2.11	0.38
997.07	0.75	1596.94	**2258.06**	2.02	4.23	**3519.06**	2.09	0.98	**4780.06**	2.12	0.37
1011.73	0.71	1390.34	**2272.73**	2.02	4.14	**3533.72**	2.09	0.96	**4794.72**	2.12	0.37
1026.39	0.70	1191.03	**2287.39**	2.02	4.05	**3548.39**	2.10	0.95	**4809.38**	2.12	0.36
1041.06	0.70	997.74	**2302.05**	2.02	3.96	**3563.05**	2.10	0.94	**4824.05**	2.12	0.36

1055.72	0.73	812.21	**2316.72**	2.03	3.88	**3577.71**	2.10	0.93	**4838.71**	2.12	0.36
1070.38	0.78	641.51	**2331.38**	2.03	3.80	**3592.38**	2.10	0.91	**4853.37**	2.12	0.35
1085.04	0.86	496.38	**2346.04**	2.03	3.71	**3607.04**	2.10	0.90	**4868.04**	2.12	0.35
1099.71	0.95	382.59	**2360.70**	2.03	3.64	**3621.70**	2.10	0.89	**4882.70**	2.12	0.35
1114.37	1.04	296.80	**2375.37**	2.03	3.56	**3636.36**	2.10	0.88	**4897.36**	2.12	0.34
1129.03	1.12	231.88	**2390.03**	2.03	3.49	**3651.03**	2.10	0.87	**4912.02**	2.12	0.34
1143.70	1.19	181.61	**2404.69**	2.04	3.42	**3665.69**	2.10	0.86	**4926.69**	2.12	0.34
1158.36	1.25	141.70	**2419.35**	2.04	3.35	**3680.35**	2.10	0.85	**4941.35**	2.12	0.34
1173.02	1.30	109.38	**2434.02**	2.04	3.28	**3695.01**	2.10	0.84	**4956.01**	2.12	0.33
1187.68	1.35	82.95	**2448.68**	2.04	3.21	**3709.68**	2.10	0.82	**4970.67**	2.12	0.33
1202.35	1.39	61.34	**2463.34**	2.04	3.15	**3724.34**	2.10	0.81	**4985.34**	2.12	0.33
1217.01	1.43	43.91	**2478.01**	2.04	3.09	**3739.00**	2.10	0.80	**5000.00**	2.12	0.32
1231.67	1.46	30.23	**2492.67**	2.04	3.03	**3753.67**	2.10	0.79			
1246.33	1.48	19.95	**2507.33**	2.05	2.97	**3768.33**	2.10	0.78			
1261.00	1.51	12.69	**2521.99**	2.05	2.91	**3782.99**	2.10	0.77			
1275.66	1.53	8.00	**2536.66**	2.05	2.85	**3797.65**	2.10	0.77			
1290.32	1.55	5.38	**2551.32**	2.05	2.80	**3812.32**	2.10	0.76			
1304.99	1.57	4.34	**2565.98**	2.05	2.74	**3826.98**	2.10	0.75			
1319.65	1.59	4.38	**2580.65**	2.05	2.69	**3841.64**	2.10	0.74			
1334.31	1.61	5.11	**2595.31**	2.05	2.64	**3856.30**	2.10	0.73			
1348.97	1.62	6.20	**2609.97**	2.05	2.59	**3870.97**	2.10	0.72			
1363.64	1.64	7.44	**2624.63**	2.05	2.54	**3885.63**	2.10	0.71			
1378.30	1.66	8.67	**2639.30**	2.06	2.50	**3900.29**	2.10	0.70			
1392.96	1.67	9.79	**2653.96**	2.06	2.45	**3914.96**	2.10	0.70			
1407.62	1.69	10.78	**2668.62**	2.06	2.41	**3929.62**	2.10	0.69			
1422.29	1.70	11.59	**2683.28**	2.06	2.36	**3944.28**	2.10	0.68			
1436.95	1.72	12.25	**2697.95**	2.06	2.32	**3958.94**	2.10	0.67			
1451.61	1.73	12.75	**2712.61**	2.06	2.28	**3973.61**	2.10	0.66			
1466.28	1.75	13.11	**2727.27**	2.06	2.24	**3988.27**	2.10	0.66			
1480.94	1.76	13.36	**2741.94**	2.06	2.20	**4002.93**	2.10	0.65			
1495.60	1.77	13.49	**2756.60**	2.06	2.16	**4017.60**	2.11	0.64			
1510.26	1.78	13.55	**2771.26**	2.06	2.12	**4032.26**	2.11	0.63			
1524.93	1.79	13.52	**2785.92**	2.07	2.09	**4046.92**	2.11	0.63			
1539.59	1.80	13.44	**2800.59**	2.07	2.05	**4061.58**	2.11	0.62			
1554.25	1.81	13.31	**2815.25**	2.07	2.02	**4076.25**	2.11	0.61			
1568.91	1.82	13.15	**2829.91**	2.07	1.98	**4090.91**	2.11	0.61			
1583.58	1.83	12.95	**2844.57**	2.07	1.95	**4105.57**	2.11	0.60			
1598.24	1.84	12.73	**2859.24**	2.07	1.91	**4120.23**	2.11	0.59			
1612.90	1.85	12.49	**2873.90**	2.07	1.88	**4134.90**	2.11	0.59			
1627.57	1.86	12.24	**2888.56**	2.07	1.85	**4149.56**	2.11	0.58			
1642.23	1.86	11.98	**2903.23**	2.07	1.82	**4164.22**	2.11	0.57			

APPENDIX B

Table B: Optical constants (n, k) for Silicon after KOH etching with *wavenumbers* varying in the range from 385 to 5000 cm^{-1}

Wave-number (cm^{-1})	n	k ($x\ 10^{-3}$)	Wave-number (cm^{-1})	n	k ($x\ 10^{-3}$)	Wave-number (cm^{-1})	n	k ($x\ 10^{-3}$)	Wave-number (cm^{-1})	n	k ($x\ 10^{-3}$)
384.62	3.39	0.58	1561.09	3.39	0.04	2737.56	3.39	0.06	3914.03	3.39	0.05
407.24	3.39	0.58	1583.71	3.39	0.04	2760.18	3.39	0.06	3936.65	3.39	0.05
429.86	3.39	0.57	1606.33	3.39	0.04	2782.81	3.39	0.06	3959.28	3.39	0.05
452.49	3.39	0.53	1628.96	3.39	0.04	2805.43	3.39	0.06	3981.90	3.39	0.05
475.11	3.39	0.49	1651.58	3.39	0.04	2828.05	3.39	0.06	4004.52	3.39	0.05
497.74	3.39	0.45	1674.21	3.39	0.04	2850.68	3.39	0.06	4027.15	3.39	0.05
520.36	3.39	0.41	1696.83	3.39	0.04	2873.30	3.39	0.06	4049.77	3.39	0.05
542.99	3.39	0.38	1719.46	3.39	0.04	2895.93	3.39	0.06	4072.40	3.39	0.05
565.61	3.39	0.39	1742.08	3.39	0.04	2918.55	3.39	0.06	4095.02	3.39	0.05
588.24	3.39	0.55	1764.71	3.39	0.04	2941.18	3.39	0.06	4117.65	3.39	0.05
610.86	3.39	1.41	1787.33	3.39	0.04	2963.80	3.39	0.06	4140.27	3.39	0.05
633.48	3.39	0.51	1809.95	3.39	0.04	2986.43	3.39	0.06	4162.90	3.39	0.05
656.11	3.39	0.30	1832.58	3.39	0.04	3009.05	3.39	0.06	4185.52	3.39	0.05
678.73	3.39	0.25	1855.20	3.39	0.04	3031.67	3.39	0.06	4208.14	3.39	0.05
701.36	3.39	0.23	1877.83	3.39	0.04	3054.30	3.39	0.06	4230.77	3.39	0.05
723.98	3.39	0.22	1900.45	3.39	0.04	3076.92	3.39	0.06	4253.39	3.39	0.05
746.61	3.39	0.23	1923.08	3.39	0.04	3099.55	3.39	0.06	4276.02	3.39	0.05
769.23	3.39	0.23	1945.70	3.39	0.05	3122.17	3.39	0.06	4298.64	3.39	0.05
791.86	3.39	0.24	1968.33	3.39	0.05	3144.80	3.39	0.06	4321.27	3.39	0.05
814.48	3.39	0.24	1990.95	3.39	0.05	3167.42	3.39	0.06	4343.89	3.39	0.05
837.10	3.39	0.23	2013.57	3.39	0.05	3190.05	3.39	0.06	4366.52	3.39	0.05
859.73	3.39	0.21	2036.20	3.39	0.05	3212.67	3.39	0.06	4389.14	3.39	0.05
882.35	3.39	0.19	2058.82	3.39	0.05	3235.29	3.39	0.06	4411.76	3.39	0.05
904.98	3.39	0.17	2081.45	3.39	0.05	3257.92	3.39	0.06	4434.39	3.39	0.05
927.60	3.39	0.15	2104.07	3.39	0.05	3280.54	3.39	0.06	4457.01	3.39	0.05
950.23	3.39	0.13	2126.70	3.39	0.05	3303.17	3.39	0.06	4479.64	3.39	0.05
972.85	3.39	0.12	2149.32	3.39	0.05	3325.79	3.39	0.06	4502.26	3.39	0.05
995.48	3.39	0.11	2171.95	3.39	0.05	3348.42	3.39	0.06	4524.89	3.39	0.05
1018.10	3.39	0.11	2194.57	3.39	0.05	3371.04	3.39	0.06	4547.51	3.39	0.04
1040.72	3.39	0.11	2217.19	3.39	0.05	3393.67	3.39	0.06	4570.14	3.39	0.04
1063.35	3.39	0.14	2239.82	3.39	0.05	3416.29	3.39	0.06	4592.76	3.39	0.04
1085.97	3.39	0.23	2262.44	3.39	0.05	3438.91	3.39	0.06	4615.38	3.39	0.04
1108.60	3.39	0.27	2285.07	3.39	0.05	3461.54	3.39	0.06	4638.01	3.39	0.04
1131.22	3.39	0.16	2307.69	3.39	0.05	3484.16	3.39	0.06	4660.63	3.39	0.04
1153.85	3.39	0.10	2330.32	3.39	0.05	3506.79	3.39	0.06	4683.26	3.39	0.04
1176.47	3.39	0.08	2352.94	3.39	0.05	3529.41	3.39	0.06	4705.88	3.39	0.04
1199.10	3.39	0.07	2375.57	3.39	0.05	3552.04	3.39	0.06	4728.51	3.39	0.04
1221.72	3.39	0.06	2398.19	3.39	0.05	3574.66	3.39	0.06	4751.13	3.39	0.04
1244.34	3.39	0.06	2420.81	3.39	0.05	3597.29	3.39	0.06	4773.76	3.39	0.04
1266.97	3.39	0.05	2443.44	3.39	0.05	3619.91	3.39	0.06	4796.38	3.39	0.04
1289.59	3.39	0.05	2466.06	3.39	0.05	3642.53	3.39	0.06	4819.00	3.39	0.04
1312.22	3.39	0.05	2488.69	3.39	0.05	3665.16	3.39	0.06	4841.63	3.39	0.04
1334.84	3.39	0.05	2511.31	3.39	0.05	3687.78	3.39	0.06	4864.25	3.39	0.04
1357.47	3.39	0.05	2533.94	3.39	0.05	3710.41	3.39	0.06	4886.88	3.39	0.04
1380.09	3.39	0.05	2556.56	3.39	0.05	3733.03	3.39	0.06	4909.50	3.39	0.04

1402.71	3.39	0.05	2579.19	3.39	0.05	3755.66	3.39	0.06	4932.13	3.39	0.04
1425.34	3.39	0.05	2601.81	3.39	0.05	3778.28	3.39	0.06	4954.75	3.39	0.04
1447.96	3.39	0.04	2624.43	3.39	0.05	3800.90	3.39	0.06	4977.38	3.39	0.04
1470.59	3.39	0.04	2647.06	3.39	0.05	3823.53	3.39	0.05	5000.00	3.39	0.04
1493.21	3.39	0.04	2669.68	3.39	0.05	3846.15	3.39	0.05			
1515.84	3.39	0.04	2692.31	3.39	0.05	3868.78	3.39	0.05			
1538.46	3.39	0.04	2714.93	3.39	0.06	3891.40	3.39	0.05			

APPENDIX C

Table C: Fabricated Chips with the bouncing pixel structure, and respective chip description.

Name	General description of the chip	Bouncing pixel structures on each chip
Chip "IR2" (2013)	Analog and Mixed Signal Integrated Circuits for IR readout. AMS CMOS 350 nm Technology	Layout of 2 pixels with internal photodiodes (PD). Capable of operating with external photodetectors.
Chip "VCHIP1" (2014)	Mixed Signal Integrated Circuits for HDR Sensors. IBM 180 nm Technology	Small matrix of 15 pixels with internal PDs. There are 3 types of different PDs in the matrix. Additional pixel without PD, capable of operating with external detector.
Chip "TSMC1" (2016)	Mixed Signal 32x32 HDR Matrix Image Sensor. TSMC 180 nm Technology | EUROPRACTICE mini@sic 2016	Low-resolution camera with 1024 pixels. Camera with HDR capability.
Chip "TSMC2" (2016)	Analog Integrated Circuit for IR readout. TSMC 180 nm Technology	Same technology as the last one, but with single pixels capable of operating with external photodetectors. Additional PDs for testing.