Delft University of Technology

Shining light on human breath analysis with quantum cascade laser spectroscopy

Reyes Reyes, Adonis

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

SHINING LIGHT ON HUMAN BREATH ANALYSIS WITH QUANTUM CASCADE LASER SPECTROSCOPY

Adonis Reyes Reyes

SHINING LIGHT ON HUMAN BREATH ANALYSIS WITH QUANTUM CASCADE LASER SPECTROSCOPY

Adonis Reyes Reyes

SHINING LIGHT ON HUMAN BREATH ANALYSIS

WITH QUANTUM CASCADE LASER SPECTROSCOPY

Shining light on human breath analysis

with Quantum Cascade Laser Spectroscopy

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
vrijdag 20 januari 2017 om 12:30 uur

door

Adonis Reyes Reyes

Master of Science in Applied Physics,
Delft University of Technology, The Netherlands,
geboren te Puebla, Mexico.

This dissertation has been approved by the

promotor: Prof. dr. H. P. Urbach
copromotor: Dr. N. Bhattacharya

Composition of the doctoral committee:

Rector Magnificus,	chairman
Prof. dr. H. P. Urbach	Delft University of Technology
Dr. N. Bhattacharya	Delft University of Technology

Independent members:

Prof. dr. J. Dankelman	Delft University of Technology
Prof. dr. ing. P. H. C. Eilers	Erasmus University Medical Center
Prof. dr. J. C. de Jongste	Erasmus University Medical Center
Dr. O. Vaittinen	University of Helsinki, Finland
Dr. F. J. M. Harren	Radboud University
Prof. dr. ir. L.J. van Vliet	Delft University of Technology, reserve member

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Printed in The Netherlands.

To Ireri, David and Braulio

We don't stop playing because we grow old;
we grow old because we stop playing.

George Bernard Shaw

SUMMARY

In the search for new non-invasive diagnostic methods, healthcare researchers have turned their attention to exhaled human breath. Breath consists of thousands of molecular compounds in very low concentrations, in the order of parts per million by volume (ppm_v), parts per billion by volume (ppb_v) and parts per trillion by volume (ppt_v). They are the result of the different biological process taking place inside the body. When a disease is present the production of specific molecules is altered. In this work in particular, we investigate two cases. One being the concentration of acetone on minors with type 1 diabetes (T1D). In the second case we compare the breath of three groups to establish significant differences and to identify relevant molecules. The groups under study are healthy children, children with asthma and children with cystic fibrosis (CF).

The main challenges in human breath research are the detection of concentration changes in small quantities and the establishment of a direct relation between specific molecules and particular diseases. We use quantum cascade lasers (QCLs), a multipass cell and Mercury Cadmium Telluride (MCT) detectors to study the absorption of the molecular components of breath. We improve the identification of molecules by applying a multiline fitting algorithm.

The different molecules present in breath have a strong absorption signature in the mid-infrared. For this reason we use QCLs emitting in the region between 832 and 1262.55 cm^{-1}. In this region each molecular species has a unique absorption fingerprint that allows its identification. The absorption is magnified by increasing the interaction distance between the light of the QCLs and the gas sample. We use a multipass cell with two astigmatic mirrors. The multiple reflections in the mirrors provide an effective interaction distance of 54.36 meters inside a volume of only 0.6 liters. For the detection we use MCT detectors directly because the QCLs emit a very specific wavenumber at a time. The absorption spectra are built by scanning the wavenumber of the QCLs twice: first with the multipass cell empty, to build a reference, and then with the breath sample to measure the absorption.

The scan of the QCLs eliminates the need of extra elements to separate the wavenumbers to build the absorption spectra. However, scanning over a broad wavenumber region introduces a new challenge: to guarantee its repeatability. This includes the assurance that the QCLs emit the same wavenumbers with the same intensities in every single scan. Only by minimizing the variability between independent scans we can create reliable absorption spectra and improve the sensitivity of the setup. We use two MCT detectors to monitor the intensity fluctuations. One detector is dedicated to monitor the intensity fluctuations of the QCLs while the other detector measures the intensity of the QCLs after the light has crossed the multipass cell. The variation of the wavenumber emission produces that independent scans are warped and uncorrelated with respect to each other. We implement two methods to correlate the measurements taken with the empty multipass cell and the measurements with the breath samples: a scan correlation

using selected wavenumbers and a scan correlation using semiparametric time warping. Both methods are successful in obtaining a meaningful absorption spectrum. The selection of wavenumbers is more adequate to study molecules with a smooth profile and the semiparametric time warping method is more suitable for molecules with sharp absorption features. The result of the wavenumber and intensity corrections give the system a noise equivalent absorption sensitivity (NEAS) of $2.99 \times 10^{-7} \text{cm}^{-1}\text{Hz}^{-1/2}$. With this NEAS we can detect ppb_v concentrations of acetone in presence of 2% of water in the same wavenumber region.

The complexity of the gas mixture in breath makes the identification of specific molecular components difficult. We implement a multiline fitting algorithm to analyze specific molecules and determine their concentrations. We use this method to study the concentration of acetone and methane in the exhaled breath of healthy children. For acetone we use its absorption signature in the 1150 - 1250 cm^{-1} region. Our results show that the production of acetone in healthy children is below the standard range established for healthy adults, between 0.39 and 1.09 ppm_v. But the information and studies in this regard are limited and therefore more studies should be performed. In the case of methane we use its absorption fingerprint between 1258 and 1262.5 cm^{-1}. The methane concentration in the breath of the participants is below 1 ppm_v, which classifies them as non-producers. Given the small number of participants, eleven, this result is in accordance with previous reports establishing that only 10% to 20% of the children are methane producers.

We perform a specific study to investigate the acetone concentration in the exhaled breath of T1D patients. We analyze the breath of two minors and one adult T1D patient, and the breath of one healthy volunteer. Simultaneously, we measure the blood glucose and ketone concentrations in blood to inspect their relation with acetone in exhaled breath. For each volunteer, we performed a series of measurements over a period of time, including overnight fasting of 11 ± 1 hours and during ketosis-hyperglycemia events for the minors. The results highlight the importance of performing personalized studies because the response of the minors to the presence of ketosis was consistent but unique for each individual. As in the case of healthy children mentioned above, we also find that the acetone concentration in the breath of T1D minors in stable conditions is lower than the standard range for healthy adults. This emphasizes the need to perform more studies with children and specifically with T1D minors. We strongly believe that a better understanding of the production of acetone in exhaled breath can help to develop new diagnostic methods. For example, it can be used to detect chronic ketosis, which is a condition that many children present in the early stages of T1D. In many cases children live with chronic ketosis for years before being diagnosed with T1D. By detecting abnormal concentrations of acetone we can help to diagnose T1D earlier.

In a separate study we explore the clinical applicability of our spectroscopic setup by comparing the exhaled breath of 35 healthy children, 39 children with stable asthma and 15 with stable CF. Their age range is 6 – 18 years. We collect two to four exhaled breath samples in Tedlar bags and obtain their absorption spectrum in the region between 832 and 1262.55 cm^{-1}. The results show a poor repeatability (Spearman's ρ = 0.36 to 0.46) and agreement of the complete profiles. However, we identify wavenumber regions where the profiles are significantly different. Using these regions and the information from

two molecular databases we present a list of molecules that can be used to discriminate between healthy children and children with asthma or CF. Our suggestion is to perform more studies and use the identified molecules as basis to understand the underlying inflammatory processes of asthma and CF. This study shows that the identification of the molecular components of exhaled breath is important and may be useful to develop new personalized treatments.

Because scientists like to dream about the future, we also explore the future possibilities in exhaled breath research. We strongly believe the next generation of exhaled breath systems will be a hybrid between optical detection systems, electrochemical methods and nanotechnology. This idea is firmly supported by the latest developments in small hollow waveguides for lasers and the most advanced pre-concentration and filtering methods for gas samples. Furthermore, the growing interest in new, non-invasive medical systems is making exhaled breath research a very important player in the global economy. We cannot foresee all the benefits exhaled breath research can offer to society but without doubt its value is immense.

SAMENVATTING

In de zoektocht naar nieuwe niet-invasieve diagnostische methoden hebben onderzoekers in de gezondheidszorg hun aandacht gericht op uitgeademde menselijke adem. Deze bevat duizenden molecuulgroepen in zeer lage concentraties, in de orde van delen per miljoen op volumebasis (ppm$_v$), delen per miljard op volumebasis (ppb$_v$), en delen per biljoen op volumebasis (ppt$_v$). Deze zijn het resultaat van de verschillende biologische processen die in het lichaam plaats vinden. Als er sprake is van ziekte vinden er veranderingen plaats in de productie van bepaalde moleculen. In dit onderzoek hebben wij twee gevallen onderzocht. Een betreft de concentratie van aceton bij minderjarigen met type 1 diabetes (T1D). In het tweede vergelijken wij de adem van drie groepen teneinde significante verschillen vast te stellen en om relevante moleculen te identificeren. De bestudeerde groepen zijn: gezonde kinderen, kinderen met astma en kinderen met taaislijmziekte (CF).

De belangrijkste uitdagingen in onderzoek naar de menselijke adem zijn de detectie van concentratieveranderingen in kleine volumes en de vaststelling van een directe relatie tussen specifieke moleculen en bepaalde ziektes. We maken gebruik van quantum cascade lasers (QCLs), een meetcel die meerdere malen wordt doorlopen en Kwik-Cadmium-Telluride (MCT) detectoren om de absorptie van de moleculaire componenten in adem te bestuderen. We verbeteren de identificatie van moleculen door toepassing van een meer-lijns curve-fitting algoritme.

De moleculen in adem vertonen een sterke absorptie in het midden-infrarood. Om deze reden maken wij gebruik van QCLs die straling produceren tussen 832 en 1262,55 cm^{-1}. In dit golflengtegebied heeft iedere molecuul een unieke absorptie-vingerafdruk waarmee identificatie mogelijk wordt. De absorptie wordt verveelvoudigd door de afstand waarover interactie plaats heeft tussen het licht van de QCLs en het gasmonster te vergroten. We maken gebruik van een meetcel die meerdere malen wordt doorlopen en die is voorzien van twee astigmatische spiegels. De meervoudige reflecties aan deze spiegels verschaffen een effectieve interactielengte van 54,36 meter binnen een volume van slechts 0,6 liter. Voor de detectie maken we direct gebruik van MCT detectoren omdat QCLs op ieder moment straling produceren met een specifiek golfgetal. De absorptiespectra worden opgebouwd door het golfgetal van de QCLs twee maal te scannen: eenmaal met een lege multi-pass meetcel waarmee een referentie wordt verkregen, en daarna met de meetcel gevuld met het ademmonster om daarmee de absorptie te meten.

Doordat de QCLs worden gescand vervalt de noodzaak om extra elementen in te zetten om de golfgetallen te scheiden teneinde de absorptiespectra te kunnen opbouwen. Scanning over een breed bereik van golfgetallen introduceert echter een nieuwe uitdaging: het garanderen van de reproduceerbaarheid. Dit is inclusief de zekerheid dat de QCLs dezelfde golfgetallen uitzenden met dezelfde intensiteit voor elke afzonderlijke scan. Slechts door de variatie tussen de onafhankelijke scans zo klein mogelijk te hou-

den kunnen we betrouwbare absorptiespectra creëren en de gevoeligheid van de mee-
topstelling verbeteren. We maken gebruik van twee MCT detectoren om de fluctuaties in
de intensiteit te bemonsteren. Een van de detectoren is bestemd voor bemonstering van
fluctuaties in de intensiteit van de QCLs terwijl de andere detector de intensiteit van de
QCLs meet nadat hun licht de multi-pass meetcel heeft doorlopen. Variatie van de golf-
getalemissie leidt ertoe dat onafhankelijke scans vervormd zijn en niet gecorreleerd ten
opzichte van elkaar. Wij implementeren twee methoden om de metingen die zijn uitge-
voerd met de lege multi-pass meetcel en de metingen metingen aan de ademmonsters te
correleren: een scancorrelatie voor geselecteerde golfgetallen en een scancorrelatie op
basis van semi-parametrische vervorming van de tijd. Beide methoden zijn succesvol
voor het verkrijgen van een betekenisvol absorptiespectrum. De selectie van golfgetal-
len is meer geschikt voor bestudering van moleculen waarvan het spectrum een glad
verloop heeft terwijl de methode gebaseerd op semi-parametrische vervorming van de
tijd beter geschikt is voor moleculen met scherpe pieken in het absorptiespectrum. Cor-
recties op het golfgetal en op de intensiteit resulteren in een ruis-equivalente absorp-
tiegevoeligheid (NEAS) voor het systeem van $2,99 \times 10^{-7} \text{cm}^{-1}\text{Hz}^{-1/2}$. Met deze NEAS
kunnen we ppb_v concentraties van aceton detecteren in aanwezigheid van 2% water bij
hetzelfde golfgetalgebied.

De complexiteit van het gasmengsel in menselijke adem maakt het moeilijk om spe-
cifieke moleculaire componenten te detecteren. Wij implementeren een meer-lijns curve-
fitting algoritme om specifieke moleculen te analyseren en hun concentraties vast te stel-
len. Wij gebruiken deze methode om de concentratie van aceton en methaan in de adem
van gezonde kinderen te bestuderen. Voor aceton maken we gebruik van het kenmer-
kende absorptiespectrum in het 1150-1250 cm^{-1} gebied. Onze resultaten laten zien dat
de productie van aceton van gezonde kinderen beneden de standaard ligt die is vast-
gesteld voor gezonde volwassenen, tussen 0,39 en 1,09 ppm$_\text{v}$. De verkregen informatie
en de studies op dit gebied blijven echter beperkt en daarom zouden meerdere studies
moeten worden uitgevoerd. Voor methaan gebruiken we het kenmerkende absorptie-
spectrum in het gebied tussen 1258 en 1262,5 cm^{-1}. De methaanconcentratie in de adem
van de deelnemers ligt onder de 1 ppm$_\text{v}$, hetgeen deze groep als niet-producenten clas-
sificeert. Gezien het lage aantal deelnemers, elf, is dit resultaat in overeenstemming met
eerdere rapporten waarin is vastgesteld dat slechts 10% tot 20% van de kinderen me-
thaan produceert.

We hebben een specifieke studie uitgevoerd ten einde de acetonconcentratie in de
adem van T1D patiënten te onderzoeken. We analyseren daarbij de adem van twee min-
derjarige en een volwassen T1D patiënt. Tegelijkertijd hebben we de bloedsuikerspiegel
en de ketonconcentraties in het bloed gemeten om het verband daarvan met aceton in
de adem te kunnen onderzoeken. Voor iedere vrijwilliger hebben we een serie metin-
gen verricht over een zekere periode, inclusief een nachtelijke nuchtere periode van 11
± 1 uur, en gedurende ketose-hyperglycemie gebeurtenissen voor de minderjarigen. De
resultaten laten duidelijk zien dat het belangrijk is op de persoon toegesneden studies
uit te voeren omdat de respons van de minderjarigen op de aanwezigheid van ketose
consistent maar ook uniek was voor ieder individu. Evenals in het geval van bovenge-
noemde gezonde kinderen vinden wij bovendien dat de concentratie aceton in de adem
van T1D minderjarigen in stabiele toestand lager is dan de standaard range voor gezonde

volwassenen. Dit benadrukt de noodzaak om meer studies uit te voeren met kinderen en speciaal met T1D minderjarigen. Wij zijn er stellig van overtuigd dat een beter inzicht in de productie van aceton van nut kan zijn bij de ontwikkeling van nieuwe diagnostische methoden. Dat kan bijvoorbeeld worden gebruikt om chronische ketose te detecteren, een conditie die vele kinderen hebben in vroege stadia van T1D. In vele gevallen leven kinderen jaren met chronische ketose voordat voor hen de diagnose T1D wordt gesteld. Door abnormale concentraties van aceton vast te stellen kunnen wij bijdragen aan een vroege diagnose van T1D.

In een afzonderlijke studie onderzoeken wij de klinische toepasbaarheid van onze spectroscopische opstelling door een vergelijking te maken tussen de adem van 35 gezonde kinderen, die van 39 kinderen met stabiele astma, en die van 15 kinderen met stabiele CF. Hun leeftijd ligt tussen de 6 en 18 jaar. Wij verzamelen twee tot vier ademmonsters in Tedlar zakken en meten hun absorptiespectra in het gebied van 832 tot 1262,55 cm^{-1}. De reproduceerbaarheid van de resultaten is matig (Spearman's $\rho = 0{,}36$ tot $0{,}46$) en de complete profielen stemmen niet goed overeen. We hebben echter golfgetalgebieden kunnen identificeren waarvoor de profielen significante verschillen vertonen. Voor deze gebieden, en gebruik makend van informatie uit twee databestanden voor molecuulgegevens, presenteren wij een lijst van moleculen die kunnen worden gebruikt om onderscheid te maken tussen gezonde kinderen en kinderen met astma of CF. Wij bevelen aan meer studies uit te voeren en de geïdentificeerde moleculen te gebruiken als basis voor een beter begrip van de onderliggende ontstekingsprocessen van astma en CF. Deze studie laat zien dat de identificatie van de moleculaire componenten in menselijke adem belangrijk is, en mogelijk bruikbaar kan zijn voor de ontwikkeling van op maat gesneden behandelingen.

Omdat wetenschappers van toekomstdromen houden, verkennen wij ook de toekomstige mogelijkheden van onderzoek naar de menselijke adem. Wij zijn er stellig van overtuigd dat de volgende generatie van sytemen voor ademanalyse een hybride zal zijn van optische detectiesystemen, electrochemische methoden, en nano-technologie. Deze gedachte wordt stevig ondersteund door de meest recente ontwikkelingen op het gebied van kleine holle golfgeleiders voor lasers en de meest geavanceerde methoden voor concentratie vooraf en filtering van gasmonsters. Bovendien zorgt de groeiende interesse in nieuwe niet-invasieve medische systemen ervoor dat onderzoek op het gebied van ademanalyse een belangrijke speler wordt in de wereldeconomie. Wij kunnen niet alle voordelen die ademanalyse kan bieden voorzien, maar zonder twijfel is de waarde ervan enorm.

CONTENTS

LIST OF ACRONYMS

BAC   Blood alcohol concentration
CAGR   Compound annual growth rate
CRDS   Cavity ring down spectroscopy
CF   Cystic fibrosis
CW   Continuous wave
DALY   Disability adjusted life year
DFB   Distributed feedback
DFB-QCL   Distributed feedback - quantum cascade laser
DIAL   Differential absorption lidar
DOAS   Differential optical absorption spectroscopy
EC   External cavity
EC-QCL   External cavity - quantum cascade laser
FC-CEAS   Frequency comb - cavity enhanced absorption spectroscopy
FeNO   Fractional exhaled nitric oxide
FTIR   Fourier transform infrared spectroscopy
FWHM   Full width at half maximum
GC   gas chromatography
GC-FID   Gas chromatography - flame ionization detection
GC-MS   Gas chromatography - mass spectrometry
HITRAN   High Resolution Transmission
HWG   Hollow waveguide
ICOS   Integrated cavity output spectroscopy
iHWG   Substrate-integrated hollow waveguide
IUPAC   International union of pure and applied Chemistry
IMS   Ion mobility spectroscopy
LC-MS   Liquid chromatography – mass spectrometry
LIDAR   Light detection and ranging
LOD   Limit of detection
MDA   Minimum detectable absorbance
MDA_{ps} Minimum detectable absorption per scan
MDA_{pp} Minimum detectable absorption per point
MS   Mass spectrometry
MCT   Mercury cadmium telluride
NaN   Not-a-Number
NEAS   Noise equivalent absorption sensitivity

NICE-OHMS	Noise-immune cavity-enhanced optical heterodyne spectroscopy
NMR	Nuclear magnetic resonant
PAS	Photoacoustic spectroscopy
PC	Principal component
PCA	Principal component analysis
pd-VI	Peak detection virtual instrument
PNNL	Pacific northwest national laboratory
ppb_v	Parts per billion by volume
ppm_v	Parts per million by volume
ppt_v	Parts per trillion by volume
PTR-MS	Proton-transfer reaction - mass spectrometry
PTW	Parametric time warping
QCL	Quantum cascade laser
QEPAS	Quartz-enhanced photoacoustic spectroscopy
SIFT-MS	Selected ion flow tube - mass spectrometry
STW	Semiparametric time warping
TDLAS	Tunable diode laser absorption spectroscopy
T1D	Type 1 diabetes
VOC	Volatile organic compound
VSL	Dutch metrology institute

1

INTRODUCTION

You should dream more, (...)
Reality in our century is not something to be faced.

Graham Greene, Our man in Havana.

Our existence would not be possible without the Earth's atmosphere. We live so immersed in it, we never stop to think what constitutes this atmosphere and how we modify it with our lifestyles. In recent years the effects of global warming and pollution have drawn our attention to how certain gases, like CO_2, modify our environment [1, 2]. The change in the concentration of these gases can greatly affect the ecosystems and the weather, to the point that it poses a threat to our way of life. It is specially interesting because these gases are considered trace gases, which means that they only represent a extremely small portion of a gas mixture, in this case the atmosphere [3]. In the case of CO_2 the global concentration was 400 parts per billion by volume (ppb_v) in November 2015 and is projected to increase in the coming years affecting the temperature of the oceans and the weather [2]. Fortunately, we are becoming conscious that our current economical and social activities play a crucial role in how the composition of the atmosphere is being modified, as shown in the 2014 report of the Intergovernmental panel on climate change [4].

We still need to become aware of our interaction with the atmosphere at a much individual scale, through our breath. Each of us is constantly inhaling the molecules in the air around us, not only oxygen but all the molecules, including the toxic ones. Through our metabolic processes we change some of these molecules and exhale a modified gas mixture. From this common process many questions arise: what are we breathing?, what can we learn about our metabolism from the products of our breath?, how are our activities affecting the quality of the air we inhale?

The issues mentioned previously along with the potential applications of trace gas analysis have increased the interest in this research area in recent years. One of the goals of this work is to acquaint the reader with the fascinating world of trace gas analysis and its applications, in particular in human breath research.

1.1. APPLICATIONS OF TRACE GAS ANALYSIS

Since its appearance, trace gas analysis has been very important and influential in many different fields. In the eighteenth century it was crucial to understand the composition of air [5]. In the nineteenth century the mining industry used it to analyze the air inside the mines and the iron industry employed it to improve their combustion techniques [6].

In the last century we observed an explosion in the number of applications of trace gas analysis that has continued until our days. The petroleum industry has always had a great interest in the separation of gases [7] and nowadays is very keen on improving the accuracy in the measurement of the oil components. The food, chemical and pharmaceutical industries have incorporated trace gas detection in their production systems. A particularly interesting example is the monitoring of ethylene in the fruit industry to control ripening [8]. In the product packaging area the analysis of the head-space is indispensable for quality control and to define the shelf-life of the products [9]. In the lithography and semiconductor industry the detection of contaminants is crucial for the fabrication processes [10–12]. In the security and defense fields the detection of explosive materials is quickly improving thanks to the rise in the sensitivity of trace gas analysis techniques [13–15]. One of the most relevant applications of trace gas analysis is air quality control, which is directly used to monitor the pollution in the atmosphere. This includes open spaces [16, 17], hazardous environments and stack emissions [18].

The advances in trace gas analysis are helping to improve the production of primary standard gas mixtures. These are essential to define the future standards, to calibrate metrology and industrial devices, to improve the spectroscopic databases and to produce higher quality ultrapure materials. In the healthcare field trace gas analysis has proven to have a lot of potential. In this area the most successful application is the detection of Nitric Oxide to diagnose and monitor asthma [19]. Following this example, the research community is working to identify other biomarkers that can be used in medical applications, from diagnosis, monitoring to treatment of different diseases.

1.2. TECHNIQUES

The ideal spectroscopic setup for trace gas analysis would have a large spectral bandwidth, high spectral resolution, high sensitivity, low detection limit[1], fast acquisition time, low cost, small size and should be easy to use. Over the years many improvements have been reached in the different spectroscopic systems, but the research community has not found yet the recipe to build the ideal spectroscopic device. The systems developed so far are able to tackle specific requirements, but they have their own limitations.

The early volumetric and reactive methods were perfect to understand the compositions of air and specific atmospheres, for example inside mines [6], but they are compound specific and their sensitivities were not good enough to detect concentrations in the ppb_v level. Their subsequent improvements allowed to obtain highly accurate gas mixture preparations, specially with the development of gravimetric blending.

Also interesting is the use of thermoanalytical methods for trace gas detection [22]. They have shown a good capability to identify and quantify the components of a gas sample. But these methods can only be used in solid samples, which requires to transform the gas samples to solids using suitable chemical reactions, which is not possible for all gaseous components.

The use of the Fourier transform formalism to analyze the spectrum of a light source allowed the development of a new commercial method for gas analysis in the 1950's: the Fourier Transform Infrared spectroscopy (FTIR). It allows the study of complex gas mixtures using the absorption signature of the molecular components [23]. Although the resolution of the FTIR provides a very high discrimination level of the molecular components of a gas mixture, the sensitivity of FTIR is limited to hundreds of ppb_v, limiting its application in trace gas analysis.

In the last decades the electrochemical analysis methods have experienced an accelerated development. Among them the best known method is the electronic nose (eNose) [24] and the most popular application is the breathalyzer to measure blood alcohol using the exhaled breath [25]. In general, these systems relay on the change of the electronic properties of their detectors in the presence of chemical components with a particular molecular structure or reactivity. This interaction can include a chemical reaction as in the case of the breathalyzer. Some advantages of the eNose are its compactness and easy to use. Among its disadvantages are its low sensitivity and non-specificity [25].

---

[1]There is a common misuse of the terms detection limit and sensitivity. The definition given by the International Union of Pure and Applied Chemistry (IUPAC) establishes that sensitivity is related only to the strength of a signal, while the detection limit is the smallest amount of material detectable (3σ-criterion), which depends on the signal strength (or sensitivity) and the signal stability [20, 21].

1.2.1. GC AND MS BASED METHODS

Without doubt the most sensitive gas analysis methods with the lowest detection limits are based on gas chromatography (GC) and mass spectrometry (MS). Some of them have the capability to detect concentration in the order of few parts per trillion by volume (ppt_v). GC and MS based methods are a fine example where the combination of different injection, separation, ionization and detection techniques have contributed to the improvement of the sensitivity. A quick description of GC and MS is given below together with a summary of the most common combined methods used in gas analysis, specially in human breath analysis.

In GC the sample is mixed with an inert carrier gas. It travels through a column whose wall is made of an inert polymer or a liquid. The walls of the column retain the molecules of the sample and release each type of molecule at different times [26].

In MS the molecules are ionized, sent through electric and magnetic fields and captured with a detector array. Each ion follows a specific trajectory due to its mass-charge ratio, reaching a specific place in the detector. The position and the number of ions indicate the type of molecule and the concentration [27].

Gas chromatography - mass spectrometry (GC-MS) uses GC to make a first selection of the molecules and then MS to identify each molecular species. In gas chromatography - flame ionization detection (GC- FID) the molecules are first separated using GC and then the ions formed during the combustion of organic compounds in a hydrogen flame are detected. Here the number of ions is defined by the organic compounds concentration.

In the selected ion flow tube - mass spectrometry (SIFT-MS) the gas sample reacts with known precursor ions which are selected using a mass filter. The amount of reaction products and remaining precursor ions are measured using MS. As these quantities are related through a reaction rate coefficient, the concentration of the original compounds in the gas sample can be determined. Proton-transfer reaction - mass spectrometry (PTR-MS) is a technique similar to SIFT-MS but the sample is not diluted. It only uses one type of precursor ions, eliminating the need of a mass filter. These two factors increase the sensitivity of the system compared to SIFT-MS.

Another interesting technique which has been coupled to GC and MS with very interesting results is ion mobility spectroscopy (IMS). It ionizes the sample molecules and measures how fast the ions travel through a neutral atmosphere in the presence of an uniform electric field. Each ion has a specific velocity allowing to identify it. Once the ions are recognized the original molecules are determined.

Each of the GC and MS methods fulfill specific needs with high accuracy. Some of them are commercially available and in a stage where there is no need for highly skilled personnel to operate the device. However, there are applications for which GC and MS methods are still not adequate. Some of the main drawbacks are the requirement for compactness, on-line operation and remote sensing.

1.2.2. LASER BASED METHODS

The invention of the laser opened up a new world for trace gas analysis thanks to the wavenumber selectivity that lasers provide [28]. The interaction between the molecules and the laser light strongly depends on the wavenumber of the laser, $\widetilde{\nu}$, as we explain in

detail in Chapter 2. The interaction is governed by the Beer-Lambert law [29]. It establishes that the strength of the interaction is defined by the absorbance of the molecular species, $A(\widetilde{v})$, which depends on the molecular concentration, C, the interaction distance, l, and the molar absorption coefficient, $\epsilon(\widetilde{v})$, through the equation

$$A(\widetilde{v}) = \epsilon(\widetilde{v})Cl. \tag{1.1}$$

All laser spectroscopy techniques use this relation to select the most adequate laser source to interact, detect and identify the molecules of interest. By choosing the right wavenumber a strong absorption signal can be obtained.

In atmospheric monitoring a precise knowledge of the absorption fingerprint of the relevant molecules allows us to determine the concentrations. One of the main challenges is to establish a reliable reference because the baseline of the spectra fluctuate continuously. Two techniques have been used for several years to do remote sensing: differential optical absorption spectroscopy (DOAS) [30] and differential absorption lidar[2] (DIAL) [31]. They locally define the baseline around each absorption feature.

In laboratory conditions there are more options to increase the absorption signal. Tunable diode laser absorption spectroscopy (TDLAS) uses wavelength and frequency modulation methods [32, 33] to reduce the noise level [34]. Photoacoustic spectroscopy (PAS) can detect concentrations in the ppb_v level because it uses the photoacoustic effect [35]. In this case, the absorption of light changes the pressure inside the sample and produces acoustic waves that can be detected using sensitive microphones or high-Q crystal elements. The latter method is known as quartz-enhanced photoacoustic spectroscopy (QEPAS) [36].

Another way to increase the absorption signal and hence improve the detection limit and sensitivity is increasing the interaction distance between the laser and the sample. There is a group of laser spectroscopic techniques that enhances the interaction using high quality cavities. Among these methods we find cavity ring down spectroscopy (CRDS) [37], integrated cavity output spectroscopy (ICOS) [38], noise-immune cavity-enhanced optical heterodyne spectroscopy (NICE-OHMS) [39] and frequency comb - cavity enhanced absorption spectroscopy (FC-CEAS) [40]. All of these methods have reported concentration sensitivities in the ppb_v level, with high repeatability, but often they still require laboratory conditions to operate.

1.3. Mid infrared laser sources

The strongest molecular absorptions are in the mid-infrared region, as we explain in Section 2.1. For this reason there are ongoing efforts concerted to produce reliable laser sources emitting in the mid-infrared.

Promising techniques include the use of nonlinear frequency conversion methods to expand the emission spectra of near-infrared lasers. Among these methods are difference frequency generation [41], optical parametric oscillation [42] and optical parametric amplification [43]. They have successfully produced high power laser sources covering broad ranges of the mid-infrared spectrum, which makes them strong candidates for

---

[2]LIDAR stands for light detection and ranging

future generations of laser spectrometers. However, at this moment their applicability is constrained to laboratory use because of their stringent operation conditions.

More robust mid-infrared lasers generate the emission directly from the gain medium. In general, these laser sources are compact and easier to use because their operation conditions are more flexible. They provide very stable emission and adequate power for trace gas analysis. Their main limitation when compared to nonlinear laser systems is that their spectral range is much narrower. The best known examples of this type of lasers are CO and CO_2 gas lasers [44, 45]. Other lasers use solid materials doped with transition metals and rare earths as gain medium, like Cr^{2+} II-VI [46]. But the most versatile mid-infrared lasers are based on semiconductors. In semiconductor lasers the gain medium can be engineered to emit in any specific mid-infrared region. The best studied mid-infrared semiconductor lasers are the lead-salt laser [47] and the quantum cascade laser (QCL) [48]. In Figure 1.1 we present a schematic of the available mid-infrared sources and the wavenumber range in which laser emissions are possible today.
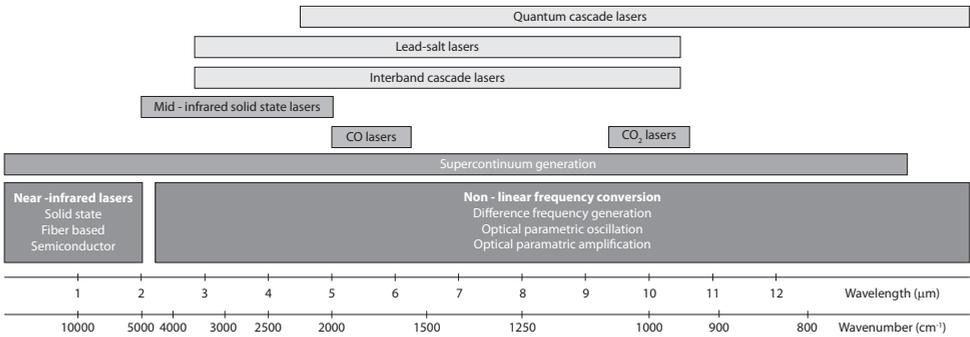
Figure 1.1: Mid-infrared laser sources and their emission regions. The information for this diagram comes from references 34, 49, 50 and 51.

The design of the gain media in QCLs is very flexible as we show in Section 3.2. This has opened an amazing world of possibilities for trace gas applications [52]. In short, the wavenumber emitted by a QCL is defined by the layer structure of the semiconductor instead of its bandgap [48]. Thanks to this flexibility, QCLs are currently providing laser sources from the 2 μm (5000 cm^{-1}) region all the way to the terahertz domain (~ 3000 μm, ~ 3.3 cm^{-1}) [53, 54].

The selectivity of the wavenumber emitted is mainly done via distributed feedback (DFB) or external cavities (EC). In DFB-QCLs a grating is embedded in the laser cavity to select and narrow the wavenumber emission. As a result the QCLs can have a linewidth as narrow as 1-3 MHz in continuous wave (CW) operation and down to 250 MHz in pulsed mode [55]. The main limitation of DFB-QCLs is their limited wavenumber tuning range, typically the maximum they can be tuned is 10-20 cm^{-1} using thermal tuning [54]. In the case of EC-QCLs a diffraction grating is used to select the wavenumber of the laser emission [53], as we explain in Section 3.2.2. In principle this configuration allows to tune over the complete emission range of the QCLs, which can cover hundreds of cm^{-1}. However, EC-QCLs suffer from mode hopping due to the competition of modes in the laser cavity. This limits their performance and make their power emission less stable

in comparison to DFB-QCLs. A lot of work is being done in this respect and today QCLs with up to 60 cm^{-1} of mode-hop-free tuning are available [54].

Another important feature offered by QCLs is their relatively high emission power. For trace gas detection sufficient high power is necessary to increase the sensitivity and lower the detection limit. Specially for techniques that enhance the interaction using high quality cavities such as the previously mentioned CRDS, ICOS, NICE-OHMS and FC-CEAS. Power emissions for CW operation are now in the hundreds of milliwatts [34, 54, 56], while pulsed operation provide peak powers of few tens of milliwatts with duty cycles of less than 1% [55]. There are even reported power emissions at the watt level [57, 58].

The operating temperature of QCLs makes them a very strong candidate for field applications. The latest thermoelectric cooling systems and the advances in the QCL designs allow QCLs to operate at room temperature. This is routinely done in pulsed QCLs while in CW operation it is still a challenge, although it has already been achieved in very specific QCLs [48]. The importance of producing CW operation at room temperature lies in the achievable narrower linewidth, which will improve the sensitivity of trace gas analysis.

The next generation of mid infrared laser sources consists of frequency comb lasers based on QCLs [59, 60]. Frequency combs laser are coherent sources whose spectrum consist of discrete, equally-spaced lines. The linewidth of each spectral line is below the MHz range and they have a very high intensity stability. Also the comb structure make them a perfect ruler to determine the wavenumber emission of each line. Until recently the direct generation of frequency combs was limited to the near infrared, below 3 μm (3333.3 cm^{-1}). The expansion to the mid infrared range between $3-12$ μm ($833.3-3333.3$ cm^{-1}) was possible using difference frequency generation, optical parametric oscillation, optical parametric amplification and microresonators. However, all these methods require extra laser sources and laboratory conditions that restrict their use. It was again the versatility in the design of the gain media of the QCLs which allowed to induce the nonlinear processes needed to generate the comb emission directly in the semiconductor structure. Now we find frequency comb lasers with spectral bandwidth of 60 cm^{-1} centered at 7 μm (1428.5 cm^{-1})[59]. More recently the generation of a high power frequency comb laser centered at 9 μm (1111.1 cm^{-1}), with a bandwidth of 65 cm^{-1} and an output power of 180 mW for ~ 176 comb lines was reported [60]. The compactness of these laser sources will allow to develop highly accurate spectroscopic systems for field applications in the near future.

In this thesis we explore the capabilities of pulsed EC-QCLs and direct absorption spectroscopy for trace gas analysis. We focus in the mid infrared region between 850 and 1262.5 cm^{-1}. In particular, we push the boundaries of direct EC-QCL spectroscopy by improving the repeatability of the laser scans. For this we correlate independent scans using two approaches. The first one selects specific emission wavenumbers and the second approach applies semiparametric time warping techniques. Our findings corroborate that direct EC-QCL spectroscopy is appropriate to study human exhaled breath when the molecules under scrutiny are known and their absorption signature is strong. We also demonstrate that using the adequate criteria and statistical analysis we can use broad band QCL spectroscopy to identify meaningful differences between different

study groups, however the sensitivity and the detection limits obtainable with direct EC-QCL spectroscopy is not sufficient.

1.4. THESIS OUTLINE

The main goal of this work is to study exhaled human breath using direct EC-QCL spectroscopy. To give the reader a clear idea of the concepts and methods employed the thesis is divided in seven chapters. This introduction has described in a broad sense the areas where gas analysis is applied today and the principal techniques that are used. It specially emphasizes QCL techniques to put in context the EC-QCL method we use.

In Chapter 2 we present the theory behind trace gas detection. It presents the origin of the molecular absorption spectrum and its different components, specifically its vibration, rotation, rovibration and overtone signatures. It also explains the Beer-Lambert law, which is at the basis of all the measurements performed in this thesis.

In Chapter 3 we discuss the details of the spectroscopic setup. We dedicate specific sections to explain the theory and the characteristics of the principal components of the setup: the QCL, the multipass cell and the infrared Mercury Cadmium Telluride detectors. We continue with the explanation of how we obtain an absorption measurement and the application of the Beer-Lambert law. At this point we also determine the noise, the detection limits and the sensitivity of the system. We end the chapter with the demonstration of the measured absorption spectra of acetone and ethanol in the part per million by volume (ppm_v) level. We also show the absorption spectra of CO_2 and water in lab air and exhaled breath samples.

Chapter 4 presents the data acquisition and the data processing used in this work. We explain the steps implemented to configure the acquisition system, to detect the peaks and valleys in the signal in order to determine the intensity of each pulse emitted by the EC-QCL; and to check and correct the information collected. In Sections 4.2 and 4.3 we explain two methods to correlate the spectral information obtained from independent QCL scans. The first method uses selected wavenumber points while the second method uses semiparametric time warping. In Section 4.4 we show how to calibrate the wavenumber of the measured spectra using absorption lines of known molecules like CO_2. The last part of the chapter explains the multiline fitting method used to determine the molecular concentrations when multiple molecular species are present in a gas sample.

Our study of exhaled acetone in type I diabetes is presented in Chapter 5. It starts with the state of the art of breath analysis in diabetes studies and points out the lack of studies of the exhaled acetone concentrations in healthy children and children with type I diabetes. In particular, we analyze the potential of breath analysis to detect hyperglycemia and ketosis. We continue by describing the methodology used to study the exhaled breath of the volunteers and the calibration of the system for acetone detection. In the results and discussion section we investigate the reference acetone levels in breath after an overnight fasting period and during the occurrence of hyperglycemia - ketosis events. Our conclusions show that more studies are necessary to establish reliable and accurate reference levels of the acetone in the exhaled breath of children because they are lower than the reference levels reported for adults.

Chapter 6 presents the results of our collaboration with the Department of Pediatric Pulmonology of the Erasmus Medical Center - Sophia Children's Hospital and the Bioin-

formatics Department of the Erasmus Medical Center of Rotterdam. The chapter is divided in two parts. In the first half we explore the feasibility of our QCL spectroscopic system in the differentiation between healthy children and minors with asthma and cystic fibrosis. In the second half we study the reference level of exhaled acetone and methane in healthy children. In the healthy, asthma, cystic fibrosis comparison we describe the characteristics of each study group and the methodology used in the research. The data analysis evaluates the repeatability of the measurements and establishes statistical differences in the absorption spectra of the different study groups. With these differences we were able to identify and suggest molecules that should be investigated further. Also in this analysis we show why a careful selection of the discrimination criteria is necessary in order to avoid the exclusion of relevant information. The study of the exhaled acetone and methane in healthy children showed how the data processing optimizes the absorption analysis depending on the specific needs of the study. In this case, for acetone, it was better to use the scan correlation by point selection method described in Section 4.2, while for methane, the use of the semiparametric time warping of Section 4.3 is more adequate. The result of the exhaled acetone corroborates that for healthy children the reference levels are lower than for adults. For exhaled methane we corroborate the general tendency reported in previous studies, but we recognize the need of performing extra measurements with an independent system.

Chapter 7 gives an outlook of the future applications of QCL spectroscopy. The valorisation presented in this chapter is not restricted to the use of QCL sources in spectroscopy. It explores the potential of future hybrid systems for trace gas analysis and the impact they can have in science and industry. The last chapter gives the general conclusions of this thesis. At the end the limit to the possibilities is in our own dreams.

2

THEORY OF GAS DETECTION

If, in some cataclysm, all of scientific knowledge were to be destroyed,
and only one sentence passed on to the next generations of creatures,
what statement would contain the most information in the fewest words?
(...) all things are made of atoms.

Richard Feynman

And the LORD said: 'Behold, they are one people,
and they have all one language; and this is what they begin to do;
and now nothing will be withholden from them,
which they purpose to do.

Genesis 11:6

The complexity of the molecular structure is reflected in the molecular absorption spectrum. Over the years a complete and comprehensive theoretical description of the spectrum has been developed. In this work we will only give a general overview based on the detailed explanations presented in the Physical Chemistry book written by Atkins, et. al. [61] and the Molecular Physics books written by Demtroder [62] and Bransden et. al. [63]. We encourage the interested reader to consult the books published by Herzberg [64–66] for an excellent and very complete study of the properties of polyatomic molecules.

We start with a revision of the basic principles behind the rotation and vibration spectra of the molecules, including the selection rules applicable under the harmonic approximation. We continue with a brief study of the rovibrational spectrum caused by the coupling of the rotation and vibration motions of the molecules and we end with a discussion of the Beer-Lambert law and its practical implications.

Before fully submerging in the study of the molecular structure we consider relevant to make a remark on the units used to characterize the radiation interacting with the absorbing medium. The radiation is well characterized by its energy, but for different historical and practical reasons different fields opted to report the radiation using different quantities such as the frequency, ν, the wavelength, λ, or the wavenumber $\widetilde{\nu}$. This is justified because there is a direct relation between these quantities and the energy through the equality

$$E = h\nu = hc\frac{1}{\lambda} = hc\widetilde{\nu} \qquad (2.1)$$

here h is the Planck constant and c is the speed of light in vacuum. Our remark is important because sometimes the different quantities are used indistinctly, which may lead to confusion.

In this work we will mainly use the wavenumber to characterize the radiation following the common practice of molecular spectroscopy and Physical Chemistry. The wavenumber is defined as the number of wavelengths per unit distance, with the centimeter as the preferred unit length, and therefore the wavenumber is expressed in inverse centimeters, cm^{-1}. The choice of the wavenumber in infrared spectroscopy is based on the fact that in this region the wavenumber takes values in the order of $10^2 - 10^3$ which are easier to manipulate compare to the orders of $10^{-6} - 10^{-9}$ of the corresponding wavelengths in meters or the $10^{13} - 10^{14}$ order of the frequencies in Hertz.

We would like to clarify that in our treatment of the theory of the molecular spectrum we prefer to refer to the energy levels rather than the wavenumber levels to avoid confusion with well established descriptions, the reader only needs to keep in mind the relations described by equation 2.1 and apply the right proportionality constants to switch between quantities.[1]

---

[1]Using h = 6.626070040 $\times 10^{-34}$ J s and c = 299792458 m s^{-1} [67] we get the proportionality constant hc = 1.986445824 $\times 10^{-23}$ J cm.

2.1. MOLECULAR SPECTRUM

The study of the molecular spectrum is greatly simplified by the Born-Oppenheimer approximation [68, 69]. This approximation relies on the fact that nuclei are much heavier than electrons, allowing us to split the wavefunction of the molecules in an electronic and a nuclear component when using the center of mass framework. Furthermore, the motion of the nuclei can be separated in its vibration and rotation movements with high accuracy. With this approximation, the total energy E_{mol} is the sum of the electronic E_e, vibrational E_v and rotational E_r energies.

$$E_{mol} = E_e + E_v + E_r \tag{2.2}$$

The different energies add structure to the molecular spectrum because their values are in different ranges. The electronic energies are around 10000 cm^{-1}, in the visible and near infrared (NIR) ranges, the vibrational energies are in the region between 500 and 4161 cm^{-1},[2] while the rotational energies are in the region expanding from 1 to 100 cm^{-1}, which fall in the far infrared (FIR) and microwave regime.

As in the case of atoms, the electronic levels of the molecules are well defined by the distribution of their electron clouds. As an example, in Figure 2.1 we schematically show the potential and the energy levels of a diatomic molecule. We can observe how each electronic state may contain different vibration states, which in turn may contain different rotational states.
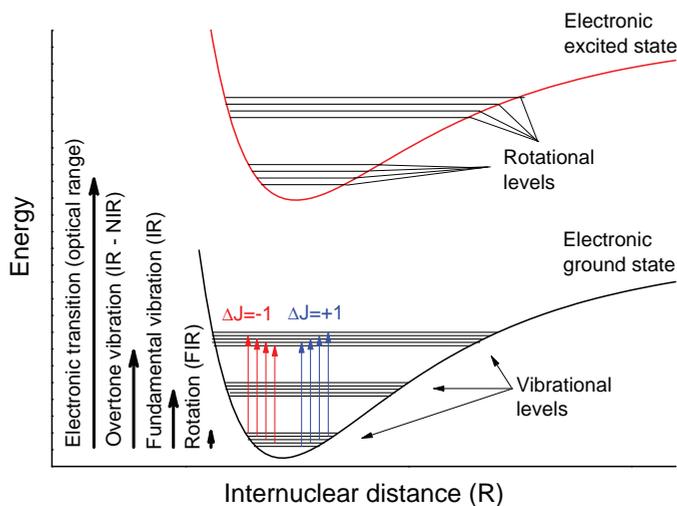
Figure 2.1: Schematic representation of the energy structure of a diatomic molecule. The level of the electric, vibration and rotational energies are in different orders of magnitude. The arrows illustrated transitions between different vibration levels. Reproduced with permission from reference 70

---

[2]This value corresponds to the vibration of the molecular hydrogen, which is the lightest diatomic molecule and therefore presents the highest vibrational frequency.

From the electromagnetic theory we know that in order to have an interaction between a molecule and an electromagnetic field with the creation or absorption of a photon with frequency ν, the molecule needs to have a dipole oscillating at the same frequency, at least transiently. This defines if the energy transition is allowed and is the basis behind the selection rules.

2.1.1. ROTATIONAL SPECTRUM

The pure rotational energy levels of a molecule can be calculated assuming the molecule is a rigid rotor with its internuclear distances fixed. With this approximation we can use the expression from classical mechanics for the kinetic energy, E_r, of the corresponding rigid body in the principal axis system [71]

$$E_r = \frac{\mathfrak{J}_a^2}{2I_a} + \frac{\mathfrak{J}_b^2}{2I_b} + \frac{\mathfrak{J}_c^2}{2I_c} \tag{2.3}$$

with the inertial momenta I_a, I_b, I_c about the axes a, b and c, and the components of the angular momentum $\mathfrak{J}_a, \mathfrak{J}_b, \mathfrak{J}_c$ about the same axes. Then we use the inertial momentum and the symmetry of the specific molecule under study to further simplify the expression of its rotational energy. Finally we import the quantum mechanical properties of the angular momentum into the equations to obtain the rotational energy values.

As an example, let us study the case of a molecule whose corresponding inertial momentum are the same as a spherical rotor, such as methane (CH_4) and Sulfur hexafluoride (SF_6). In this cases all the inertial momenta are equal, $I_a = I_b = I_c = I$. Using the relation $\mathfrak{J}^2 = \mathfrak{J}_a^2 + \mathfrak{J}_b^2 + \mathfrak{J}_c^2$ equation 2.3 becomes

$$E_r = \frac{1}{2I}\mathfrak{J}^2. \tag{2.4}$$

Finally we can apply the quantum mechanical operator by replacing $\mathfrak{J}^2$ by $J(J+1)\hbar^2$, where J is the rotational quantum number,

$$E_J = \frac{\hbar^2}{2I}J(J+1). \tag{2.5}$$

As discussed at the beginning of this chapter, in spectroscopy it is common to denote the energy levels in terms of wavenumbers. Using equation 2.1 we obtain the corresponding wavenumber

$$\widetilde{F}(J) = \frac{E_J}{hc} = BJ(J+1); \quad B = \frac{\hbar}{4\pi cI} \tag{2.6}$$

In a similar fashion we can calculate the rotational energies of molecules with more complex geometries, we just need to write their rotational momentum components in the right form to translate them into the correct quantum operator.

Pure rotational spectra are only observed in polar molecules because they are the only molecules that comply with the selection rules. When the polar molecules are rotating, their permanent dipole acts as an oscillating dipole that allows the interaction with electromagnetic radiation. As previously mentioned, the range of energies where pure rotational transitions occur lies in the Far Infrared and the microwave region between 1 to $100\,\text{cm}^{-1}$.

In the previous description of the molecular rotation based on a rigid rotor there is no possibility to re-arrange the electric charges of the molecule only with rotations. In Section 2.1.3 we study how the vibrations of the molecule change the distance between nuclei, allowing the presence of an electric dipole, introducing corrections to the rotational levels and inducing a rovibrational coupling.

2.1.2. VIBRATION SPECTRUM

The vibrations of polyatomic molecules is accurately described by the vibrational modes of the molecules. For simplicity we will first examine the vibration of a diatomic molecule and then we will extend the analysis to more complex molecules.

DIATOMIC MOLECULES

In this case there is only one vibrational mode along the axis connecting the conforming atoms. A typical potential curve for diatomic molecules is depicted in Figure 2.2.
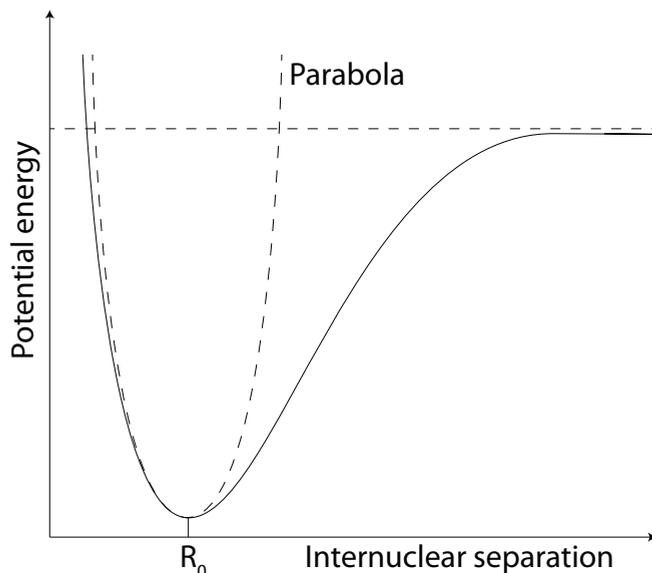
Figure 2.2: Schematic representation of the potential of a diatomic molecule.

The equilibrium position is marked as R_0 and the potential energy of the molecule around this minimum can be approximated using the Taylor series

$$V(R) = V(R_0) + \left(\frac{dV}{dR}\right)_{R_0} (R - R_0) + \frac{1}{2}\left(\frac{d^2V}{dR^2}\right)_{R_0} (R - R_0)^2 + \dots \tag{2.7}$$

The first term on the right, $V(R_0)$, is just an offset corresponding to the depth of the potential. The first derivative of the potential is 0 at the minimum. Therefore, the first non-zero element is proportional to the square of the displacement. For small displacements we can ignore the higher order terms and obtain an expression similar to the potential of a harmonic oscillator

$$V(R) = \frac{1}{2}\left(\frac{d^2V}{dR^2}\right)_{R_0} (R - R_0)^2 = \frac{1}{2}k(R - R_0)^2 \tag{2.8}$$

where k is the force constant. Using this potential to solve the Schrodinger equation we find the energies corresponding to the vibrational levels of the diatomic molecule.

$$E_p = \left(p + \frac{1}{2}\right)h\nu; \quad \nu = \frac{1}{2\pi}\left(\frac{k}{m_{eff}}\right)^{1/2}; \quad p = 0, 1, 2, \dots \tag{2.9}$$

where p is the vibrational quantum number and ν is the oscillator frequency. The frequency is determined by the force constant, k, and the effective mass of the diatomic molecule, m_{eff}, defined by the masses of the individual atoms m_1 and m_2 by the relation

$$m_{eff} = \frac{m_1 m_2}{m_1 + m_2}. \tag{2.10}$$

As discussed in the introduction of this chapter, it is common to express the energies of the vibrational states in wavenumbers, $\widetilde{G}(p)$, known as the vibrational terms of the molecule. From equation 2.1 we get the relation between $\widetilde{G}(p)$ and the energy $E_\nu = hc\widetilde{G}(p)$, then using equation 2.9 we get an expression for the vibrational terms

$$\widetilde{G}(p) = \left(p + \frac{1}{2}\right)\widetilde{\nu}; \quad \widetilde{\nu} = \frac{1}{2\pi c}\left(\frac{k}{m_{eff}}\right)^{1/2} \tag{2.11}$$

ANHARMONIC OSCILLATIONS

The treatment of the previous section is only valid in the harmonic approximation. For real systems we need to take into account higher order terms of equation 2.7. P. M. Morse proposed an empirical potential that represents better the potential of a diatomic molecule.

$$V_M(R) = D_e \left[e^{-2\alpha(R-R_0)} - e^{-\alpha(R-R_0)} \right] \tag{2.12}$$

where D_e and α are constants for a given molecule and $(R - R_0)$ is a displacement measured from the equilibrium position R_0. By expanding equation 2.12 in terms of $R - R_0$ we get

$$V_M(R) = D_e \left[-1 + \alpha^2 (R-R_0)^2 + ... \right] \tag{2.13}$$

For small displacements this potential should match the harmonic approximation. Comparing equations 2.13 and 2.7 and using equation 2.13 we find a relation with the force constant k given by

$$D_e \alpha^2 = \frac{1}{2} k. \tag{2.14}$$

The lower energy levels are know given by the expression

$$E_p = \left[\left(p + \frac{1}{2} \right) - \beta \left(p + \frac{1}{2} \right)^2 \right] h\nu; \tag{2.15}$$

with $\beta << 1$ and ν corresponding to the frequency of the energy levels of the harmonic case. The term $\beta\nu$ is known as the anharmonicity constant and is given by

$$\beta\nu = \frac{h\nu^2}{4D_e} \tag{2.16}$$

By using equation 2.1 we can write the energy levels in wavenumbers

$$\widetilde{G}(p) = \left[\left(p + \frac{1}{2} \right) - \beta \left(p + \frac{1}{2} \right)^2 \right] \widetilde{\nu}; \quad \widetilde{\nu} = \frac{1}{2\pi c} \left(\frac{k}{m_{eff}} \right)^{1/2} \tag{2.17}$$

POLYATOMIC MOLECULES

For polyatomic molecules the number of vibrational modes is defined by their geometry. A molecule formed by N atoms has 3N - 5 independent vibrational modes if it is a linear molecule and 3N - 6 independent vibrational modes if it is a non-linear molecule [71]. Each mode, q, acts as an independent harmonic oscillator under the assumption of small displacements. Therefore, each mode has a series of vibrational terms or energies (in wavenumber terms)

$$\widetilde{G}_q(p) = \left(p + \frac{1}{2}\right)\widetilde{v}_q; \quad \widetilde{v}_q = \frac{1}{2\pi c}\left(\frac{k_q}{m_q}\right)^{1/2}, \tag{2.18}$$

where $\widetilde{v}_q$ is the wavenumber of mode q, which is defined by the force constant k_q and the effective mass m_q of the mode. The effective mass of the mode measures the mass that is displaced by the vibration. In general, it is a complicated function of the masses of the atoms.

A good example to study the vibrational modes is carbon dioxide, CO_2. It is also an important molecule for our study because its concentration in human breath is considerable. CO_2 has four vibrational modes depicted in Figure 2.3. The symmetric and antisymmetric stretching modes have the corresponding wavenumbers $\widetilde{v}_1 = 1388$ cm^{-1} and $\widetilde{v}_3 = 2349$ cm^{-1}. Due to the symmetry of the molecule the two perpendicular bending modes share the same wavenumber in their lower level, $\widetilde{v}_2 = 667$ cm^{-1}. A good notation to identify the vibrational levels of CO_2, including the superposition of modes is $(p_1 p_2 p_3)$, where p_q is the vibrational quantum number associated with the wavenumber $\widetilde{v}_q$.

CO$_2$ vibrational modes
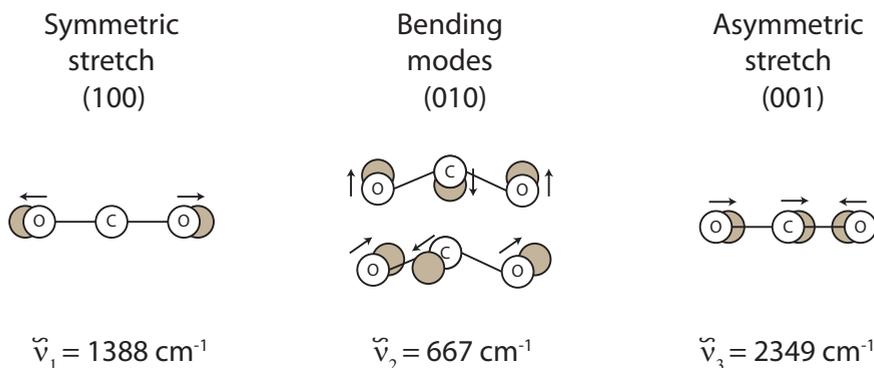
Figure 2.3: Vibrational modes of carbon dioxide.

Just as in the case of the rotational spectrum, a change in the vibrational states of a molecule with the absorption or emission of a photon should be accompanied by a

change of the dipole moment. For this reason the interaction of a molecule with an external electromagnetic field is subject to the selection rule $\Delta p = \pm 1$ in the harmonic approximation. In the case of CO_2 the symmetric stretch is not associated with a change in the dipole moment, which explains the absence of absorption lines in the $\tilde{\nu}_1 = 1388$ cm^{-1} region as can be seen in the absorption spectrum depicted in Figure 2.4. As can be seen in Figure 2.4 the spectrum is not composed of single transitions, the extra structure is due to the coupling of the rotation and vibration of the molecule as we will explain in Section 2.1.3. This spectrum is obtained using the information provided by the High Resolution Transmission (HITRAN) database [72].
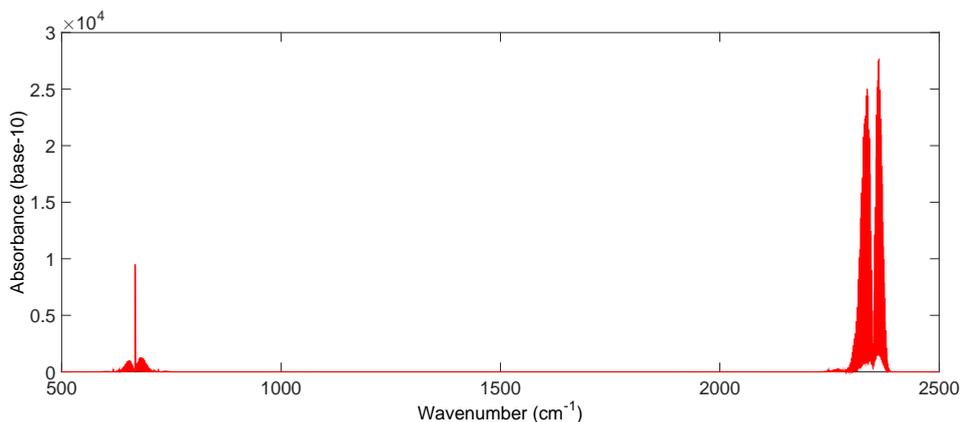
Figure 2.4: Absorption spectrum of CO_2. Information from HITRAN database.

In general the fundamental vibrational modes have wavenumbers in the range between 500 and 4161 cm^{-1}.

HOT BANDS

In the molecules vibrational transitions also occur between excited states. These transitions are known as hot bands[73, 74]. For such a transition to happen at least one excited states most be populated, the lower level in the case of absorption and the upper level in the case of emission. The most common form of excitation is thermal, which is the origin of the name: hot band. The population of an excited states N with respect to the ground state N_0 is given by the Boltzmann distribution

$$\frac{N}{N_0} = e^{-\frac{\Delta E}{k_B T}} \tag{2.19}$$

ΔE is the energy difference between the states, k_B is the Boltzmann constant and T is the temperature. By increasing the temperature the population of the excited state gets larger and the probaility of the transition increases, giving as result an increment in the intensity of a hot band.

In Figure 2.5 we show a diagram of the vibrational energy levels of CO_2 [44, 45, 75–77]. We show the hot bands that result from the transitions between the energy level (001) and the levels (100) and (020).These are the transitions used in CO_2 lasers. The wavenumber corresponding to the transition (001) → (100) has a central value of 961.0 cm^{-1} (10.4 μm) and the central wavenumber for the transition (001) → (020) is 1063.8 cm^{-1} (9.4 μm). These hot band transitions are important for our research because all our studies were performed in the region between 850 and 1250 cm^{-1}, specially for our research on exhaled human breath.
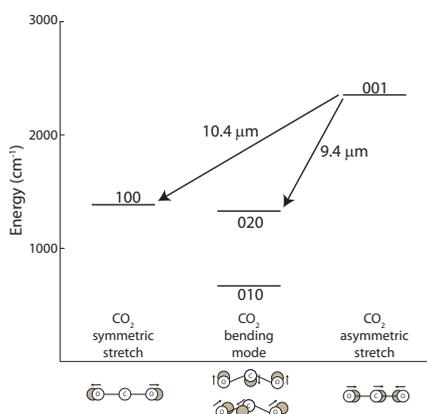
Figure 2.5: Energy level diagram of the vibration modes of CO_2.

In Figure 2.6 we show the main transitions and the hot bands of CO_2. In this figure we clearly observe the difference in strength of the hot bands. It is clear that the intensity of the main transitions is higher than the hot bands by 5 orders of magnitude at room temperature. Also in Figure 2.6 the spectrum has a complex structure which is the result of the coupling of the rotation and vibration of the molecule as we will explain in Section 2.1.3.
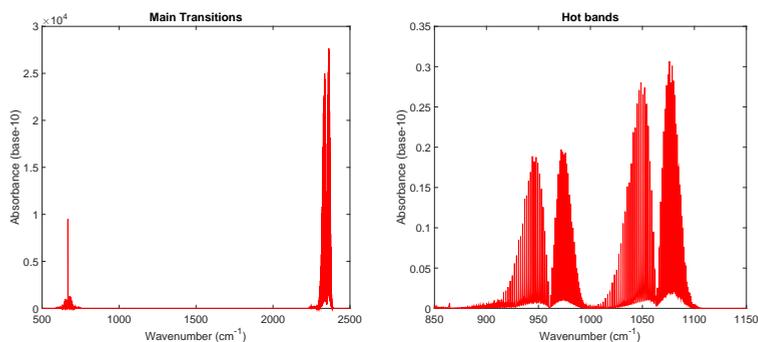
Figure 2.6: Absorption spectrum of the main transitions and hot bands of CO_2. Information from HITRAN database.

OVERTONE VIBRATIONS

In practice, the absorption spectrum of a molecule presents features at lower wavenumbers than its harmonics. They correspond to small perturbations of the main vibrational modes $\widetilde{G}_q(p)$, similar to the anharmonic case of diatomic molecules discuss in Section 2.1.2.

The energy values obtained so far are only approximate because they are based on a parabolic approximation of the potential energy for each vibrational mode, but when higher vibrational excitations are present higher terms of the Taylor expansion of equation 2.7 should be taken into account. Due to this higher order terms the vibration becomes anharmonic. This extra components explain the appearance of the additional weak absorption lines which correspond to transitions with $\Delta p = \pm 2, \pm 3, \dots$ known as overtones. In principle the overtones are forbidden by the selection rule $\Delta p = \pm 1$, but this rule is derived from the properties of the harmonic oscillator wavefunction, which are only valid for the harmonic case. In this sense, the selection rules are only an approximation and all values of Δp are allowed but the intensity of the overtone transitions are weaker than the intensity of the transitions with $\Delta p = \pm 1$.

2.1.3. ROVIBRATIONAL SPECTRUM

For real vibrating molecules the internuclear distance changes, which modifies the inertial momentum of the molecules. Therefore, the vibration and rotational motions are coupled. For simplicity we will present again the case of a diatomic molecule. The basic idea can be extended for polyatomic molecules and can be found in multiple reference such as the books of Demtroder [62] and Bransden [63].

In a more accurate treatment of the molecular spectrum, the potential used to solve the Schrodinger equation should take into account the anharmonicity of the vibration and the centrifugal distortion of the molecule. The anharmonic vibration is well described by the Morse potential of equation 2.12. The centrifugal distortion is found using the inertial momentum of a diatomic molecule $I = m_{eff} R^2$, where m_{eff} is the effective mass of the molecule and R is the separation between the molecules. Then, using equations 2.5 and 2.12 we get the effective potential

$$V_{eff}(R) = V(R) + \frac{h^2}{8\pi^2} \frac{J(J+1)}{m_{eff} R^2} \tag{2.20}$$

To find the lower energy levels we can expand equation 2.21 about the minimum of the potential V_0 at $R = R_1$. Keeping terms up to the fourth order $(R - R_1)^4$ we get

$$V_{eff}(R) = V_0 + \frac{1}{2}\widehat{k}(R - R_1)^2 + c_1(R - R_1)^3 + c_2(R - R_1)^4 \tag{2.21}$$

where the new force constant $\widehat{k}$ and the coefficients c_1 and c_2 can be written in terms of the quantum number, J, and the constants D_e, α and R_0. Finally, by treating the coefficients c_1 and c_2 as perturbations we get the energy levels of the molecule.

$$E_{p,J} = -D_e + \left[\left(p + \frac{1}{2}\right) - \beta\left(p + \frac{1}{2}\right)^2\right]h\nu + \frac{h^2}{8\pi^2 m_{eff}R_0^2}J(J+1)$$
$$-a\left(p + \frac{1}{2}\right)J(J+1) - bJ^2(J+1)^2 \tag{2.22}$$

where

$$a = \frac{3h^3\nu}{16\pi^2 m_{eff}\alpha R_0^3 D_e}\left(1 - \frac{1}{\alpha R_0}\right)$$
$$b = \frac{h^4}{64\pi^4 m_{eff}^2 \alpha^2 R_0^6 D_e} \tag{2.23}$$

The first three terms of equation 2.22 correspond to the potential depth, D_e, the harmonic and anharmonic vibrations of equation 2.15 and the rigid rotator energy of equation 2.5. The last two terms correspond to the coupling of the rotation and vibration of the molecule and a correction to the rotation motion due to the change in the distance between the nuclei.

The effect of the rovibrational coupling in the absorption spectrum of the molecule can be explained by studying closely the term $\left(p + \frac{1}{2}\right)J(J+1)$. When a molecule absorbs a photon it takes one unit of angular momentum, therefore each vibrational transition is accompanied by a change in the rotational state, which represents a change in the rotational quantum numbers J by $\Delta J = \pm 1$. For each allowed ΔJ the absorption lines form a group called branch. The P branch corresponds to all the transitions with $\Delta J = -1$ and the R branch is a collection of all the transitions with $\Delta J = +1$. In specific cases a change $\Delta J = 0$ is allowed, in which case the Q branch is observed. In Figure 2.7 we present the rovibrational spectrum of CO_2 centered at 667 cm^{-1} with its P, Q and R branches.
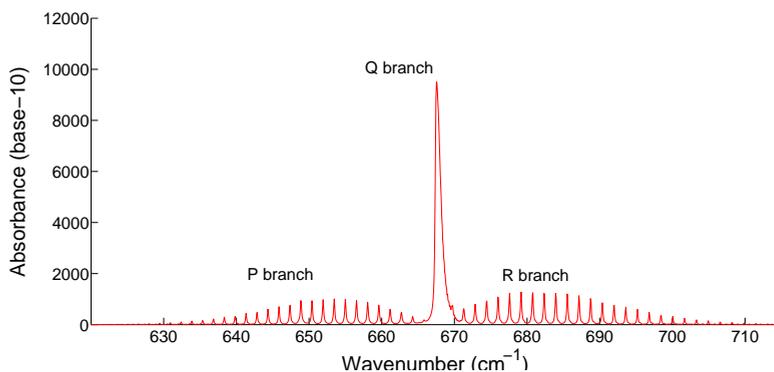
Figure 2.7: P, Q and R branches of the rovibrational spectrum of CO_2 centered at 667 cm^{-1}. Information from HITRAN database.

The rovibrational coupling is not restricted to the main molecular transitions and similar structures can be observed in the hot bands of some molecules. In Figure 2.8 we present the hot bands of CO_2 in the spectral region between 850 and 1250 cm^{-1}, where our studies were performed.
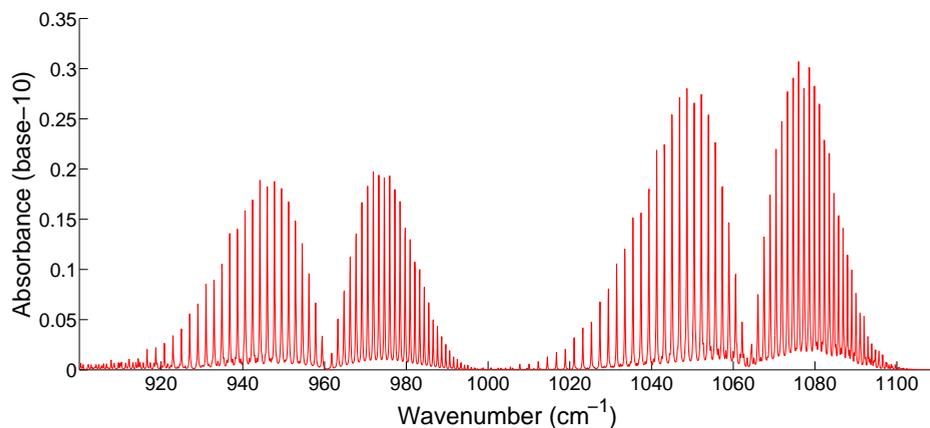
Figure 2.8: Hot bands of CO_2. Information from HITRAN database.

In general the rovibrational coupling of polyatomic molecules can be more complex and produce an absorption spectrum where the transitions are so tightly packed that they cannot be resolved by any spectroscopic device and therefore it is better to treat them as continuous.

2.1.4. Molecular identification

From the discussion of the previous sections we have established the origin of the complex structure of the absorption spectrum of molecules. A remarkable feature in the vibrational spectrum of a molecule is the appearance of very well defined absorption lines which are produced by small groups of atoms and specific bonds within a molecule, such as the C-H, O-H and C=C bonds. The vibration of these atoms dominate the normal modes of the whole molecules and give rise to very specific absorption features as shown in Figure 2.9. This figure depicts the regions where the vibration of specific bonds originates specific absorption profiles.

By looking at the region and comparing the strength of the absorption it is possible to identify the specific bond present in a molecule. However, sometimes different molecules share similar bonds, in which case it is necessary to study the specific fingerprint of the molecule in the region between 500 and 1500 cm^{-1} to uniquely identify the molecule. In this thesis we work in part of the fingerprint region between 850 and 1250 cm^{-1}, which allows us to make a full identification of the molecules.
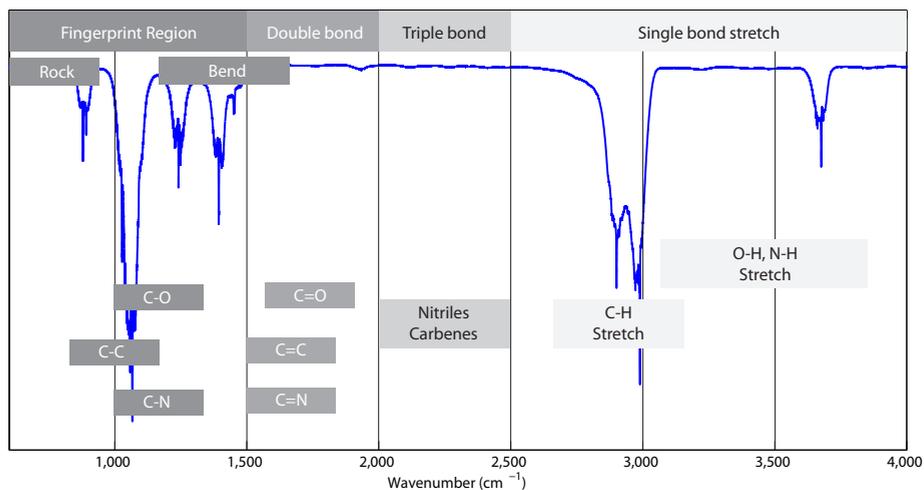
Figure 2.9: Regions where specific bonds induced absorption. Rock and bend refer to two of the different vibration modes of the atoms within a molecule.

2.2. BEER-LAMBERT LAW

In practice the molecular spectrum is obtained from the analysis of the interaction between a molecular sample and a light source with initial intensity I_0 at a specific wavenumber $\tilde{v}$. In our specific case the source is a laser. When the laser light passes through a layer of thickness dz, its intensity decreases by an amount dI. This reduction is determined by the incident intensity I, the absorption cross-section σ of the molecule and the molecular density n of the sample.

$$dI = I\sigma n \, dz. \tag{2.24}$$

The absorption cross-section is in fact a function of the wavenumber, $\sigma(\tilde{v})$, and therefore contains all the information related to the different absorption lines. It can be expressed as the product of the absorption intensity $S_i(\tilde{v})$ and the normalized line shape $g(\tilde{v})$.

$$\sigma(\tilde{v}) = S_i(\tilde{v})g(\tilde{v}) \tag{2.25}$$

This approach has the advantage that the total absorption of individual transitions is included in the parameter $S_i(\tilde{v})$, while the line shape $g(\tilde{v})$ takes into account the thermal distribution of the molecules. For gases, the line shape is properly described by the convolution of the Lorentzian line profile related to the lifetime of the transitions and the Gaussian profile produced by the Doppler broadening [62, 78]. The result of this convolution is known as the Voigt profile and has the form shown in Figure 2.10.
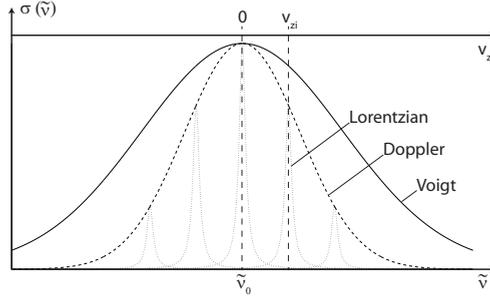
Figure 2.10: Voigt Profile. The inhomogeneous broadening is given by the distribution of the variable v_z, which can be, for example, the speed of the molecules along a given direction.

By integrating (2.24) over the total interaction distance, l, we obtain the Beer Lambert law [29],

$$-\ln\left(\frac{I(\widetilde{v})}{I_0(\widetilde{v})}\right) = \sigma(\widetilde{v})nl. \tag{2.26}$$

Following the recommendations of the International Union of Pure and Applied Chemistry (IUPAC) we express the Beer-Lambert law in a decadic base [79], obtaining directly the absorbance, $A(\widetilde{v})$, which is widely use in the spectroscopic databases and in particular in the Pacific Northwest National Laboratory (PNNL) database [80], which we have referred to for most of the molecular spectra relevant for this work.

$$A(\widetilde{v}) = -\log_{10}\left(\frac{I(\widetilde{v})}{I_0(\widetilde{v})}\right). \tag{2.27}$$

The change of base uses the equation $\log_{10} x = \ln x / \ln 10$, therefore from equation 2.26 we get

$$A(\widetilde{v}) = \frac{\sigma(\widetilde{v})}{\ln 10} nl. \tag{2.28}$$

We can rewrite the absorbance in terms of the molecular concentration C and the molar absorption coefficient $\epsilon(\widetilde{v})$. The molecular concentration is equal to $C = n/N_A$, where N_A is the Avogadro number. The molar absorption coefficient is given by $\epsilon(\widetilde{v}) = \frac{N_A \sigma(\widetilde{v})}{\ln 10}$. Therefore the absorbance takes the form

$$A(\widetilde{v}) = \epsilon(\widetilde{v})Cl. \tag{2.29}$$

In this work we will determine the molecular concentration by measuring the absorbance of different gas samples. This is possible due to the molecular absorption coefficient information provided by the molecular databases, like PNNL and HITRAN, and because we can accurately measure the interaction distance, as it will be shown in Chapter 3.

3

EXPERIMENTAL SETUP

To invent, you need a good imagination and a pile of junk.

Thomas A. Edison

In this chapter we describe in detail the experimental setup used in our research. We start with a general description of the experimental setup. We then provide a summary of its main characteristics and specifications. In subsequent sections we explain the working principle and the details of each component of the experimental setup. Section 3.2 provides the background of the quantum cascade laser (QCL) emission and the wavenumber tuning mechanism. Section 3.3 describes the details of the multipass cell. The working principle and advantages of the detectors are discussed in Section 3.4. We conclude by showing gas measurements performed to demonstrate the capabilities of the system.

3.1. GENERAL DESCRIPTION

The experimental setup incorporates a quantum cascade laser (QCL) source, a multipass cell to enhance the molecular absorption and two detectors. One of which monitors the laser intensity and the other measures the absorption signal inside the multipass cell. A visible laser distance-meter (DISTO D2) from Leica is co-aligned with the laser. It is used as guiding beam and to measure the interaction distance inside the multipass cell. A schematic overview of the experimental setup is shown in Figure 3.1.
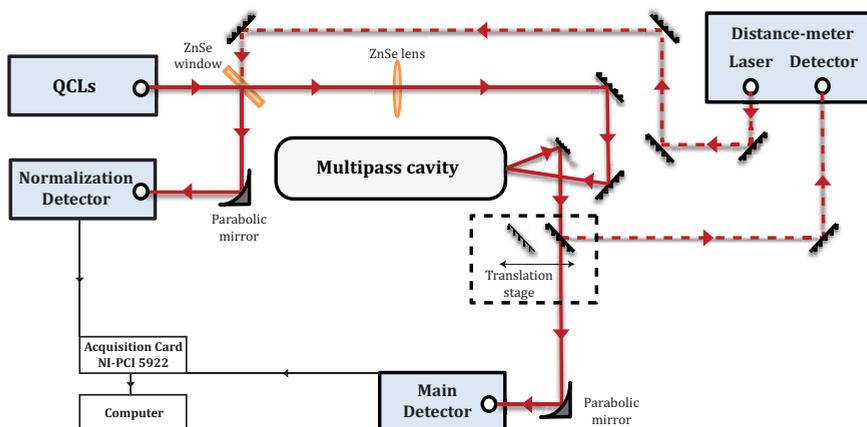
Figure 3.1: Schematic overview of the experimental setup. The translation stage allows to switch between the multipass cell length determination and the measurement configurations.

The QCL source is part of a LaserScope unit manufactured by Block engineering [82]. It consists of two co-aligned pulsed tunable QCLs, one QCL covers the region between 832.2 cm^{-1} and 1010 cm^{-1} and the second QCL emits in the region between 1010 cm^{-1} and 1263.4 cm^{-1}. Table 3.1 contains the specification of the QCL source provided by Block engineering.

The multipass cell is based on the modified Herriot configuration described by McManus et. al. [83]. The design and construction of this cell was performed in our research group. It uses two astigmatic mirrors to obtain an effective interaction distance of 54.36 meters between the gas samples and the QCL light in a volume of only 0.6 liters.

Table 3.1: Specification of the QCL source

	Values	Units
Tunable range	832.2 - 1263.4	cm^{-1}
	12.0 - 7.9	μm
Pulse repetition time	320 - 16000	ns
Pulse duration	32 - 496	ns
Duty cycle	5	% max
Line width	< 1	cm^{-1}
Average output power	0.5 - 12	mW

For the detection we use two thermo-electric cooled (TEC) Mercury Cadmium Tel-luride (MCT) detectors and a fast acquisition card (NI-PCI 5922). As depicted in Figure 3.1 the main detector collects the signal emerging from the multipass cell and a normal-ization detector is used to monitor the laser intensity directly outside the QCL source and normalize the intensity of each laser pulse. In Section 3.4 we explore the characteristics of the MCT detectors and show how we improved the performance of the normalization detector, which is part of the LaserScope unit from Block Engineering.

Additionally, we use a pressure controller (Bronkhorst, P-702CV-1K1A-RAD-22-V) and a diaphragm vacuum pump (Vacuubrand MD 1) to control the pressure inside the multipass cell. The pressure measurements are particularly important for molecules whose spectrum contain sharp absorption lines. Such lines can suffer broadening be-cause of different effects, for example collisional or Doppler broadening. To control the injection of samples inside the cell we use a set of mass flow controllers (Bronkhorst, F-201CV-1K0-RAD-22-V), which allows us to prepare samples with the desired concen-trations using single and double dilution techniques [84].

3.2. QUANTUM CASCADE LASER

A quantum cascade laser is a type of semiconductor laser where the emitted wavenum-ber is determined by the layer structure rather than the energy bandgap of the con-stituent materials [48]. This particular feature has allowed the fabrication of laser sources with central emission wavenumber over a very wide range, from the mid-infrared region up to the terahertz range, 5000 cm^{-1} - 40 cm^{-1}, (2 μm - 250 μm) [53, 54].

The lasing mechanism was first proposed in 1960 by Lax [85]. It was based on the transitions between the Landau levels in a strong magnetic field. Experimentally, the first demonstration occurred in 1994 in the Bell Laboratory [86] and since then many im-provements have been achieved. Among them we can mention continuous wave (CW) operation at room temperature [87] and power emissions in the watt level [57].

3.2.1. THEORY

In a conventional semiconductor laser the energy transitions occur between the con-duction and the valence band. The recombination of electrons and holes in the active region is the responsible for the photon production. In a QCL the energy states of the conduction band are quantized and the energy transitions occur between the resulting

sub-bands. The electrons are bound to move only in the plane of the semiconductor layers thus leading to energy quantization. Any movement of charge carriers between layers can only occur in discreet steps, this creates a series of quantum wells in which only specific energy levels are allowed. The depth and width of the wells is defined by the thickness of the layers. Therefore, by properly designing the superlattice we can choose the wavenumber of the laser emission.

Another difference between conventional semiconductor lasers and QCLs is the linewidth of the laser emission. In Figure 3.2 a) the reciprocal space of a conventional semiconductor laser is depicted. The interband transitions feature a broad linewidth because of the opposite curvatures of the conduction and valence band. In contrast, the subband of QCLs have the same curvature as shown in Figure 3.2 b). If we neglect the non-parabolicity the joint density of states is very high, as in the case of atomic transitions, and the linewidth is very narrow. Under these circumstances the linewidth only depends indirectly on temperature through collision processes and many body effects.
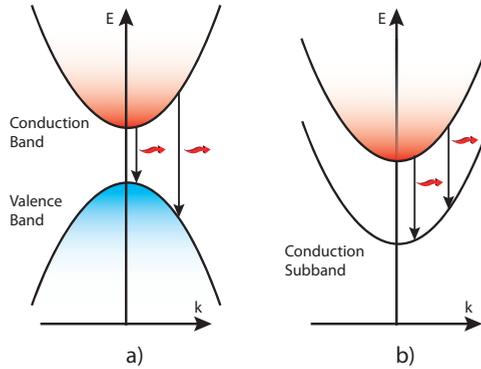
Figure 3.2: Reciprocal space of a) a conventional semiconductor laser and b) the intersubband transitions in a QCL.

QCLs consist of a series of alternating active and injector regions to obtain an efficient laser emission. The laser emission itself happens in the active regions. The injector regions align the ground state of one active region with the upper states of the next active region as shown in Figure 3.3. A voltage is applied to move the electrons from one active region to the next. This way each single electron emits multiple photons in a cascade mechanism. The injector and active regions are fabricated by properly designing the semiconductor's superlattice.

A critical condition for laser emission is population inversion. By approximating each active region to a three level system where light emission occurs between the upper two levels, 3 and 2, we can find the conditions to fulfill population inversion. In Figure 3.3 we present an energy diagram of two periods of the semiconductor structure.

The injector region is used to transport electrons from the ground energy level 1 of one active region to the excited level 3 of the next active region. The current injected in level 3, J_3, is determined by the charge of the electron, e, the electron density in level 3, n_3, and the lifetime of the electron in that level, τ_3, through the relation
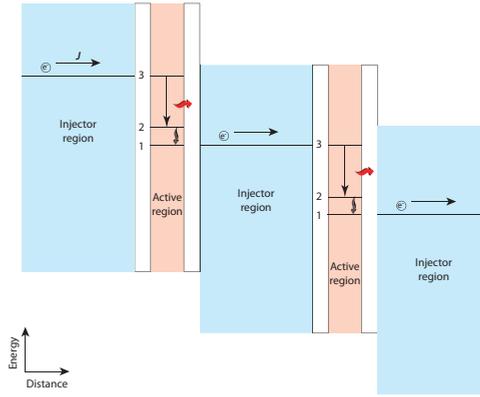
Figure 3.3: Cascade effect. The electrons move along the QCL supperlatice, emitting laser light in each active region. The ground level of an active region and the upper state of the next active region are aligned by the injector regions. The alignment of the energy levels is caused by an applied electric field.

$$J_3 = \frac{n_3 e}{\tau_3}. \tag{3.1}$$

Using this current we can define the injection efficiency, η_i, which is one of the most important parameters for QCL operation. It is defined as the ratio between the current injected in level 3 and the total current, J [48],

$$\eta_i = \frac{J_3}{J}. \tag{3.2}$$

In the ideal case the injection efficiency is equal to 1. But direct injection to level 1 or 2 and thermal excitation into the continuum and highly excited states can lower the efficiency. Using a narrow quantum well after the injection barrier, the injection into level 3 can be enhanced and the coupling with levels 1 and 2 is reduced. The thermal excitation can be avoided by optimizing the design of the active regions and using sufficient high energy barriers. Using equations (3.1) and (3.2) the electron density in level 3 can be written as

$$n_3 = \frac{\eta_i J \tau_3}{e}. \tag{3.3}$$

Assuming all the electron transitions occur from level 3 to 2, the electron density in level 2 is given by

$$n_2 = n_3 \frac{\tau_2}{\tau_{32}}, \tag{3.4}$$

where τ_{32} is the mean transition time from level 3 to 2 and τ_2 is the lifetime of the electron in level 2. In this case the population difference is given by

$$n_3 - n_2 = \frac{\eta_i J}{e} \tau_3 \left(1 - \frac{\tau_2}{\tau_{32}}\right). \tag{3.5}$$

Assuming that all the electrons flow directly from level 3 to 2, then $\tau_{32} = \tau_3$ and the population inversion gets the form

$$n_3 - n_2 = \frac{\eta_i J}{e} (\tau_3 - \tau_2).$$ (3.6)

When the condition $\tau_3 > \tau_2$ is satisfied, population inversion and optical gain are achieved. A common way to reduce the lifetime τ_2 is making the transition between levels 2 and 1 resonant with the phonon energy of the lattice structure. This way the electrons scatter quickly to level 1 after the emission. To prevent accumulation of electrons in level 1 the exit barrier of the active region is thin. This allows rapid tunnelling into an adjacent injector region. Because a single electron can emit multiple photons it is possible to get quantum efficiencies above 1.

In a similar way the energy levels of the valence band can be quantized, in which case only holes are involved in the light emission. This is another unique characteristic of QCLs: unipolarity. Only one type of carriers, electrons of holes, is involved in the energy transitions.

3.2.2. Wavenumber tuning

The QCL wavenumber can be tuned using different methods. For small ranges the tuning can be done by changing the conditions of the resonant cavity in a controlled way. A good example is the wavenumber tuning performed by changing the temperature of the semiconductor. In this case sensitivity of the resonant cavity to temperature changes is increased by building a grating on top of the laser structure or by cleaving both ends of the semiconductor structure to create a Fabry-Perot resonator.

The QCL source used in this work employs an external cavity to tune the wavenumber over a broader range [82]. It consists of a diffraction grating in a Littrow configuration [29, 78] with back extraction as shown in Figure 3.4.
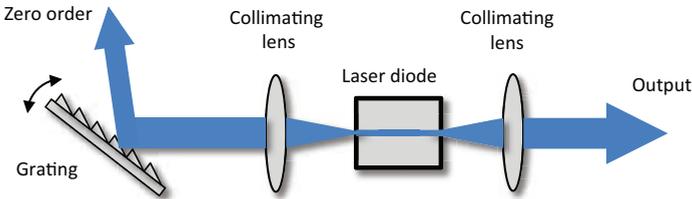
Figure 3.4: Littrow configuration with back extraction. Reproduced with permission from reference 88.

In this configuration the beam is sent to a diffraction grating from the back side of the laser. The first order of diffraction provides optical feedback to the laser diode. The angle of incidence on the grating determines the wavenumber injected back to the diode laser. By changing this angle the emitted wavenumber can be tuned over a wide range. In the specific case of our laser source, the angular position of the grating is controlled using a piezoelectric component. The voltage applied to the piezo is directly related

to the emitted wavenumber, but because of the non-linear response of the piezo the repeatability of the wavenumber tuning in the Block Engineering system is 0.1 cm^{-1}.

In a normal Littrow configuration the zero order of diffraction is used as the laser output. In this case the direction of the laser output depends on the angular position of the grating. The configuration with back extraction used in our laser source reduces this effect because the laser output comes directly from the diode and is not affected by any moving component. However, by studying the direction of the laser output we discovered a small displacement that we eventually corrected as explained in Section 3.2.3.

3.2.3. LASER BEAM DISPLACEMENT

In this section we report our findings on the wavenumber dependence of the direction of the output beam from the laser. Although in the back extraction Littrow configuration this dependence is small we preferred to have a full characterization of the system before sending the laser beam to the multipass cell. In our setup the laser beam is injected into a multipass cell. The performance of this type of cell depends on the direction of the laser beam at the entrance of the cell. For this reason, it is important to know how the laser output direction changes as a function of the wavenumber and minimize this variation as much as possible. We minimize the deviation by improving the co-alignment of the two QCLs embedded in the LaserScope unit.

Using an infrared camera (Micro-epsilon, TIM400) we observed the displacement of the QCL beams on a plane perpendicular to the optical axis. Figure 3.5 shows the displacement of the laser beam when the wavenumber is scanned over the entire spectral region.
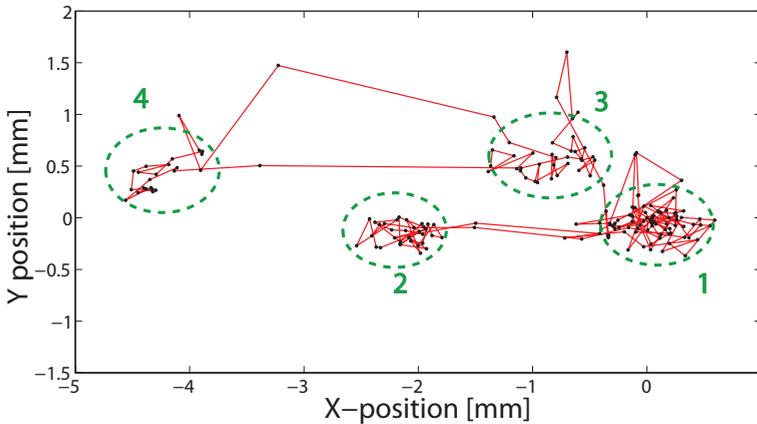
Figure 3.5: Central position of the laser beam over a full wavenumber scan. Reproduced with permission from reference 89.

There are four main zones where the center of the beam concentrates. Table 3.2 indicates the wavenumbers for each zone [89]. The QCL covering the region between 832.2 and 1010 cm^{-1} is denoted as Laser 1 and the QCL emmiting in the 1010 to 1263.4 cm^{-1}

region is named Laser 2. The beam of Laser 1 is emitted in the horizontal space between zones 1 and 2 while the emission of Laser 2 falls in the horizontal space between zones 3 and 4. The horizontal spread of the laser emission is directly determined by the emission of the diode. The vertical displacement between the emission of the two lasers is cause by a poor co-alignment.

Table 3.2: Laser beam direction distribution

Laser No.	Zone	Wavenumber (cm^{-1})
1	1	832.2 - 894±1
		967±2 - 1010
	2	894±1 - 967±2
2	3	1010.2 - 1108±1
		1198±2 - 1263.4
	4	1108±1 - 1198±2

With a careful alignment of the laser we were able to improve the vertical displacement and to reduce the horizontal space over which the laser beams move. The remaining spread in the beam position does not affect the performance of the multipass cell as we demonstrate in Section 3.3.2. The improvement of the QCLs' co-alignment also improved the signal used to monitor the QCLs intensity, as we show in Section 3.4.

3.2.4. QCL SPECTRUM

A direct measurement of the laser spectrum is shown in Figure 3.6. Clearly, the intensity of the laser is wavenumber dependent. In particular, the intensity jumps abruptly at 1010 cm^{-1}, here is where the LaserScope unit switches between the two QCLs.
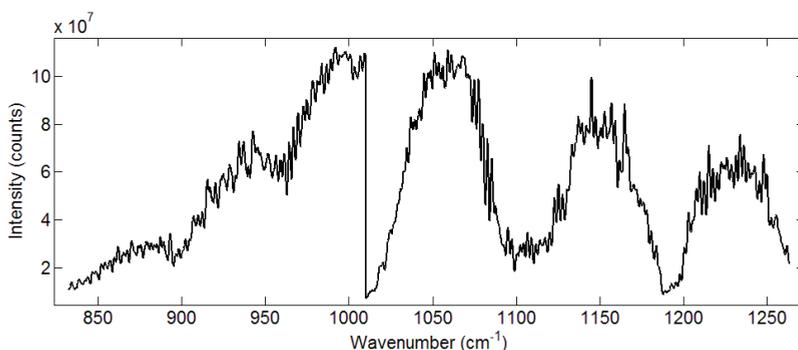
Figure 3.6: Emission spectrum of the QCL sources.

3.3. MULTIPASS CELL

One of the challenges in gas analysis using QCL spectroscopy is to get a large interaction distance between the gas and the laser light in small volumes, of less than a liter. Multipass cells offer an excellent solution because they are stable, robust and can provide

interaction distances of tens of meters. They consist of a pair of mirrors in which the light is reflected many times. The interaction length is defined by the design, the curvature of the mirrors and the distance between them [70, 90–93]. In Figure 3.7 we observe an open multipass cell built with two spherical mirrors.
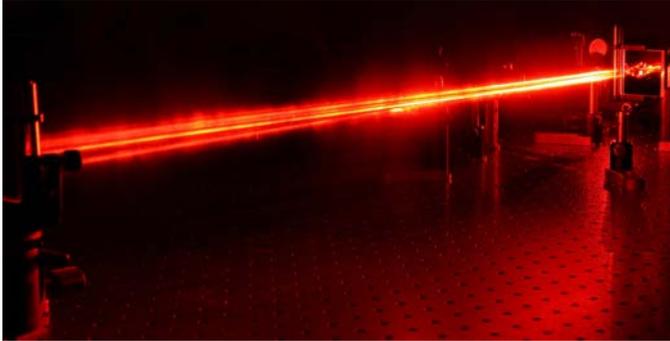
Figure 3.7: Multipass cell built with two spherical mirrors. Reproduced with permission from reference 70.

Among the available configurations we chose a modified Herriot cell designed by McManus, et. al. [83]. This cell is based on astigmatic mirrors and can provide interactions lengths up to 100 meters in a volume of only 0.6 liters.

3.3.1. THEORY

Astigmatic mirrors have two perpendicular radii of curvature, $R_{a,b}$. When two identical mirrors are placed in a multipass cell the laser reflections produces Lissajou patterns in the mirror's surface as shown in Figure 3.8.
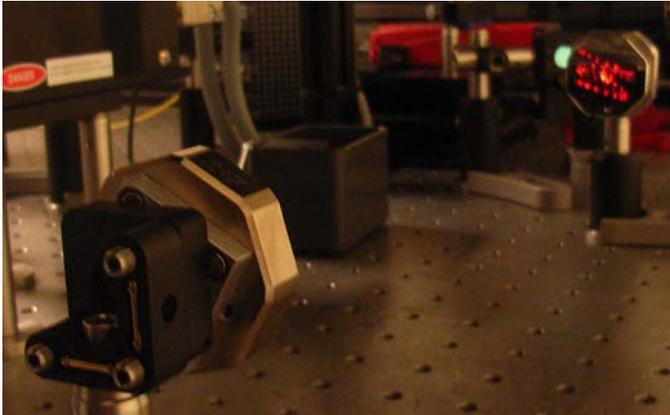
Figure 3.8: Astigmatic mirrors used for a multipass cell. The mirror in the upper right corner shows a Lissajous pattern formed by the multiple reflections of a He-Ne laser.

Due to the astigmatism of the mirrors the patterns are considerable more dense than cells designed with spherical or cylindrical mirrors. In the special situation when the

principal axis of curvature of the two mirrors are parallel, the position of nth reflection of the beam in one of the mirrors follows the sinusoidal relation [83]

$$x_n = X_0 \sin(n\Theta_x)$$
$$y_n = Y_0 \sin(n\Theta_y)$$

(3.7)

where X_0 and Y_0 are the size of the overall pattern and Θ_x and Θ_y are defined as

$$\Theta_x = \cos^{-1}\left(1 - \frac{D}{R_x}\right)$$
$$\Theta_y = \cos^{-1}\left(1 - \frac{D}{R_y}\right)$$

(3.8)

with D corresponding to the distance between the center of the mirrors, we will call this distance simply the mirror separation.

In a McManus cell [83] the principal axis of curvature of the mirrors are rotated around the optical axis as shown in Figure 3.9. In this case the principal axis of one of the mirrors are in the x and y directions, while the principal axis of the second mirror are in the x' and y' directions.
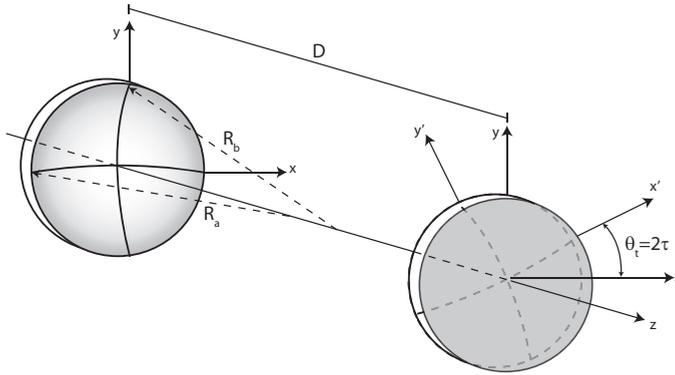
Figure 3.9: Multipass cell with two astigmatic mirrors with their principal axis rotated around the optical axis.

In this situation the components of the beam coordinates become more complex. The laser reflections still follow a Lissajou pattern, but with a different density of reflections. In this case the coordinates of the nth reflection of the beam spot in one of the mirror depends on the coordinates of the $(n-1)^{th}$ and $(n-2)^{th}$ reflections through the relations.

$$x_n = \left(A^2 - 2 - e^2\right)x_{n-1} - x_{n-2} + \epsilon y_{n-1}$$
$$y_n = \left(B^2 - 2 - e^2\right)y_{n-1} - y_{n-2} + \epsilon x_{n-1}$$

(3.9)

with

$$A = 2 - \left(\alpha_a D \cos^2(\tau) + \alpha_b D \sin^2(\tau)\right)$$
$$B = 2 - \left(\alpha_a D \sin^2(\tau) + \alpha_b D \cos^2(\tau)\right)$$
$$e = \left(\frac{1}{2}\right)(\alpha_a - \alpha_b) D \sin(2\tau) \tag{3.10}$$
$$\epsilon = -\left(\frac{1}{4}\right)(\alpha_a - \alpha_b)^2 D^2 \sin(4\tau)$$

where $\alpha_a = 2/R_a$, $\alpha_b = 2/R_b$ and $\tau = \theta_t/2$ is half of the angle θ_t at which the mirrors are rotated with respect to each other, as can be seen in Figure 3.9. By adjusting the angle between mirrors and their separation distance we can modify the density of the pattern inside the multipass cell and as a consequence the interaction distance.

3.3.2. CHARACTERIZATION

The multipass cell used in our experimental setup was built using two 2.5 in. (1 in. = 2.54 cm) astigmatic mirrors from Aerodyne Research, Inc. and standard vacuum components. Figure 3.10 shows a photograph of the multipass cell components before assembly. The details of the multipass cell elements are described in Appendix A.

Figure 3.10: Multipass cell components

The mirrors have a reflectivity $r > 0.983$ over the full spectral range of our QCL source [88]. The mirror separation inside the cell is 32.4 cm. The mirrors are placed on custom made mounts inside a stainless steel vacuum cell. The cell has an input ZnSe window (Thorlabs WG71050-F) with a transmission $t > 0.95$ in the working spectral range. The number of reflections N inside the multipass cell determines the interaction pathlength achievable but it also defines the amount of light coming out of the multipass cell through the intensity reduction coefficient k,

$$k = t^2 r^N. \tag{3.11}$$

Even though this type of cavities allow interaction lengths of up to 100 meters [83] we preferred a configuration with an interaction length of 54.36 meters, which is long

enough to detect molecules in concentrations levels of few parts per million by volume (ppm$_v$) in our measurement conditions and at the same time provides a strong output signal. The power coming out of the multipass cell is 5.5% of the laser power sent into the cell. The multiple reflections and the compact configuration allows the laser to travel through most of the volume of the cell, which allows a homogeneous interaction with the entire gas sample.

ROBUSTNESS

In most multipass cavities the allowed input angle of the incoming laser beam has a small tolerance. The McManus configuration increases the tolerance using a 500 mm ZnSe lens in front of the multipass cell. As explained in Section 3.2.2, the broadband tunability of our QCLs is obtained using a Littrow configuration with back extraction. The alignment improvement we explained in Section 3.2.3 allowed us to reduce the wavenumber dependance of the direction of the output laser beam. With this improvement the output beam has an angular displacement of less than 3 mrad. To demonstrate that the performance of our multipass cell is not affected by this small angular variation we produced pulses with a temporal width of 112 ns and measured the normalization signal and the signal from the multipass cell using only one detector. For this experiment the configuration of Figure 3.1 was modified to place both signals on the main detector as shown in Figure 3.11. We only used one detector to have the same detection response and avoid any electronic delay. The signal was analyzed using a RIGOL DS1102E oscilloscope [94].
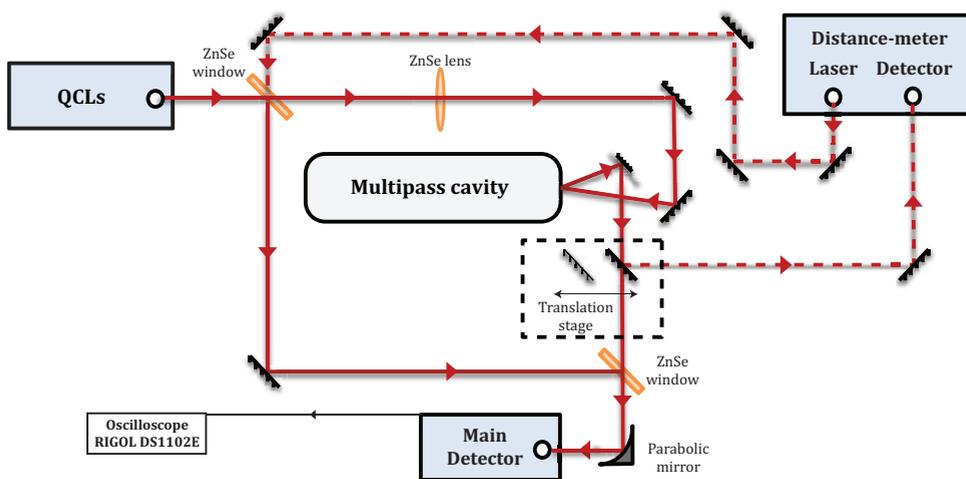
Figure 3.11: Modified experimental configuration to study the robustness of the multipass cell. A ZnSe window is used to send the normalization signal and the signal from the multipass cell into the same detector.

We repeated this measurement for three different wavenumbers, 950 cm^{-1}, 1000 cm^{-1} and 1150 cm^{-1}, as shown in Figure 3.12. Each wavenumber corresponds to a different angular output direction of the laser beam as determined in Section 3.2.3. In all

cases the delay between signals is 185 ns when the multipass cell is empty, confirming that pulses with different wavenumbers travel the same distance inside the multipass cell. From Section 3.2.3 we know that the angular displacement between the pulses with wavenumber $1000\,\mathrm{cm}^{-1}$ and $1150\,\mathrm{cm}^{-1}$ (Figures 3.12(b) and 3.12(c)) is maximum, which confirms the tolerance of the multipass cell.
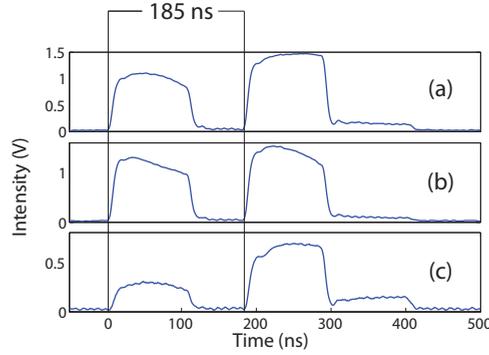
Figure 3.12: Delay time between the normalization and the main signals for three different wavenumbers. (a) $950\,\mathrm{cm}^{-1}$, (b) $1000\,\mathrm{cm}^{-1}$ and (c) $1150\,\mathrm{cm}^{-1}$.

INTERNAL PATH-LENGTH DETERMINATION

To measure the path-length inside the multipass cell we used two methods. First we used the laser distance meter, averaged 50 measurements and subtracted the path the laser travels outside the multipass cell. The result in this case is a path-length of $l_{dm} = 54.36$ meters. When we divide this path length by the mirror separation ($D_m = 32.4$ cm) we obtain the number of single passes inside the multipass cell. In principle this should be an even integer number because in our multipass cell the entrance and exit port is the same. However the result gives us 167.78 passes, which is clearly not an integer. To understand this result we should consider the geometry of the mirrors. If the mirrors were flat then the mirror separation would be the shortest pass inside the multipass cell, but the curvature of the mirrors allows even shorter passes. For this reason we can round the number of passes to $n_p = 168$. Using this number of passes we can get the average length of a single pass from the ratio, $\frac{l_{dm}}{n_p} = 32.36$ cm, whose 400 μm difference with the mirror separation is comparable to the uncertainty that we can associate with the assembly of the multipass cell.

The second method to measure the path-length inside the multipass cell uses the 185 ns pulse delay between the normalization and the main signal shown in Figure 3.12. Subtracting the distance traveled by the laser outside the multipass cell we got a measured distance of $l_p = 54.33$ meters. The 3 cm difference between this measurement and the one performed with the laser distance meter represents a delay of 100 ps which we cannot resolve using our oscilloscope. Furthermore, when we repeat the analysis to obtained the number of passes we get $\frac{l_p}{D_m} = 167.68$ passes, which we can round again to $n_p = 168$ passes. In this case the average length of a single pass is $\frac{l_p}{n_p} = 32.34$, which

gives us a 600 μm difference with the mirror separation. Again this is comparable to the uncertainty in the multipass cell construction.

The previous discussion shows the consistency of the independent methods used to measure the total path-length inside the multipass cell. The 3 cm difference between the two methods has different components. These are, one the uncertainty associated with the assembly of the multipass cell, another component is related to the overlap of the beam of the laser distance meter and the beams from the QCLs.

3.4. MERCURY CADMIUM TELLURIDE DETECTORS

The detectors used in our experiment use Mercury Cadmium Telluride (MCT) in their detection area. MCT is a semiconductor whose bandgap matches the radiation in the infrared region[95]. It can be optimized to operate at a wide range of the IR spectrum by choosing the proper composition of Cadmium and Mercury. Because emission of our QCL matches the peak wavenumber of thermal radiation at room temperature we need to cool down the MCT detector to minimize thermally induced currents. An elegant solution used in the latest MCT detectors is thermoelectric cooling using two and three stages.

The normalization detector of the setup is part of the LaserScope unit from Block Engineering [82, 96]. In the original configuration the QCL light was focused using a lens and then redirected to the MCT detector using a mirror. However, the multiple reflections in the lens produced interference fringes in the spectrum as shown by the blue spectrum in Figure 3.13. We replaced the lens and mirror by a parabolic mirror to eliminate the fringes. The new spectrum does not present the strong interference as shown in red in Figure 3.13.
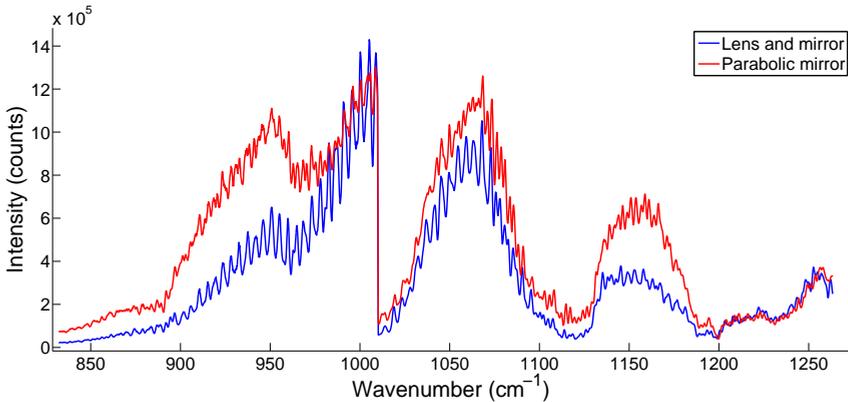
Figure 3.13: QCL spectrum obtained with the normalization detector. The fringes in the blue spectrum are caused by a lens and mirror, which are replaced by a parabolic mirror to obtained the red spectrum.

Another difference we can observe between the normalization spectra presented in Figure 3.13 is the overall strength of the spectrum. This is a result of the alignment improvement discussed in Section 3.2.3. The intensity increase is particularly clear in the regions around 950 and 1150 cm^{-1}.

The main detection of our experiment is done by a PVI-4TE-10.6-0.3x0.3-TO8-BaF2 detector from VIGO System. It has a 100 MHz cut-off frequency with reverse bias and a four stage thermoelectric cooling system. The normalization detector, from Block Engineering mentioned above, has a cut-off frequency range of 10 Hz to 20 MHz and a thermoelectric cooling system, but the number of cooling stages is not specified [96]. The spectral response of the detectors is different because of the difference in their specifications, but because their spectral response is constant we can perform a reliable normalization. In combination with the 10 mega samples per second (MS/s) sampling rate of the NI-PCI 5922 acquisition card, the MCT detectors allow us to record the peak power of each QCL pulse. This way the signal to noise ratio (SNR) of the measurement increases [55]. The SNR can be further improved if both signals are recorded using one single MCT detector [97], however this introduces new challenges for the recording system. In Chapter 4 we will explain in detail the data acquisition process, at this point we only want to emphasize the advantages provided by the use of MCT detectors.

In Figure 3.14 (a) and (b) the spectra detected in the normalization and main detectors are shown. The signal in the main detector was recorded when the multipass cell was empty. Figure 3.14 (c) shows the difference between the monitor and the main signal, which provides information of how the beam is modified by the multipass cell. The ratio plotted in Figure 3.14 (d) is the normalized signal we use to obtained the absorption spectrum of the samples. This last spectrum does not depend on the intensity fluctuations produced inside the QCL source, which is corroborated by the absence of the intensity jump at 1010 cm^{-1}. The result of the normalization was consistent when we repeated the measurements, which demonstrates that the difference in the spectral response of the detectors is constant and does not affect our measurements.
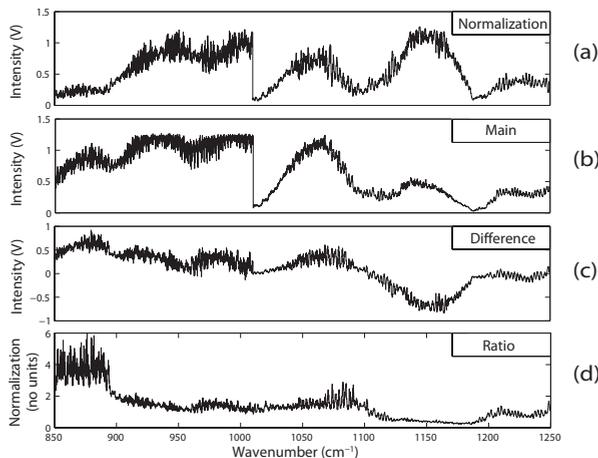
Figure 3.14: The QCL spectrum is measured with two different detectors. In (a) the normalization detector is used while in (b) and the main detector takes into account the effects introduced by the multipass cell. Subfigure (c) shows the difference between the two signals which make explicit the influence of the multipass cell. In subfigure (d) the ratio between the main signal over the normalization signal is presented, this signal should not depend on intensity variations in the QCL emission.

3.5. Measurements

For the gas absorption measurements we first build a reference spectra $I_0(\widetilde{\nu})$, where $\widetilde{\nu}$ is the wavenumber. For that we record the transmission of a laser scan through the empty multipass cell or filled with a neutral gas which has no absorption lines in the spectral region of our device. Then we average the result of ten laser scans. Immediately after, we introduce the gas sample inside the multipass cell to minimize the effects on our measurements induced by atmospheric variations outside the multipass cell. Then we scan the laser ten times and average the signal to construct a measurement spectrum $I(\widetilde{\nu})$. Using the Beer-Lambert law [79] we obtain the absorbance $A(\widetilde{\nu})$ of the sample,

$$A(\widetilde{\nu}) = -\log_{10}\left(\frac{I(\widetilde{\nu})}{I_0(\widetilde{\nu})}\right). \qquad (2.27 \text{ revisited})$$

As discussed in Section 2.2 the absorbance can also be written in terms of the molecular concentration C in parts per million by volume (ppm$_v$), the interaction path-length between the light and the molecules ($l = 54.36$ meters in the case of the present device) and the molecular absorption coefficient $\epsilon(\widetilde{\nu})$ in ppm$_v^{-1}$m^{-1}

$$A(\widetilde{\nu}) = \epsilon(\widetilde{\nu})Cl. \qquad (2.29 \text{ revisited})$$

Using this last relation we can use the strength of the absorbance to determine the molecular concentration.

3.5.1. Noise determination and sensitivity analysis

The noise level of our system is determined by comparing two reference spectra. As they are taken with an empty multipass cell the variations are only induced by the laser and detector fluctuations. The ratio between two reference spectra is a good comparison and as we are interested in the absorbance we can directly use equation 2.27. From this result the noise level is established by taking the standard deviation of the absorbance. For our system it has a value of 0.02.

To determine the performance of our system we used the formalism developed by Moyer et. al. [98]. It defines the minimum detectable absorbance, MDA, as

$$MDA = \left(\frac{\Delta P}{P}\right)_{n_s} \frac{1}{\sqrt{BW}}. \qquad (3.12)$$

Here $\Delta P/P$ corresponds to the noise level determined using n_s scans and BW is the relevant bandwidth of the system. The bandwidth allows us to study specific characteristics of the system and provide us a way to compare different spectroscopic systems. For example, if we want to compare two systems that determine the absorption based on multiple scans and whose scanning times are different, we can use the minimum detectable absorption per scan (MDA_{ps}). In this case BW is defined as the inverse of the total time needed to obtained the information. In other words the inverse of the product of the time a single scan takes, T_s and the number of integrated scans, n_s.

$$MDA_{ps} = \left(\frac{\Delta P}{P}\right)_{n_s} \sqrt{n_s}\sqrt{T_s}. \qquad (3.13)$$

In our particular case n_s is equal to 10 and T_s is equal to 5.3 s. From the noise level determined before, the MDA_{ps} of our system has a value of 0.14 $Hz^{-1/2}$.

To compare systems that scan over a narrower or broader spectral region we can define a minimum detectable absorption per point (MDA_{pp}). In this case the appropriate bandwidth of equation 3.12 must take into account the number of data points reported in the absorption spectrum. In our specific system there is a well-defined number of data points in the spectrum, $\#pts = 8000$. Therefore we can define the relevant bandwidth as the inverse of the data acquisition rate $\frac{\#pts}{n_s T_s}$ to obtain

$$MDA_{pp} = \left(\frac{\Delta P}{P} \right)_{n_s} \sqrt{n_s} \sqrt{T_s} \frac{1}{\sqrt{\#pts}}. \tag{3.14}$$

Using equation 3.13 we get a direct relation between the MDA_{pp} and the MDA_{ps}

$$MDA_{pp} = \frac{MDA_{ps}}{\sqrt{\#pts}} = 1.6 \times 10^{-3} Hz^{-1/2}. \tag{3.15}$$

The comparison with systems with different interaction path-lengths is better done using the noise equivalent absorption sensitivity (NEAS) [98, 99]. Since the interaction distance inside the multipass cell, l, was determined to be 54.36 meters the NEAS has a value of

$$NEAS = \frac{MDA_{pp}}{l} = 2.99 \times 10^{-7} cm^{-1} Hz^{-1/2} \tag{3.16}$$

3.6. Gas detection using QCL

In this section we present a series of measurements to show the performance of our setup. In these measurements the reference spectrum was taken with a pressure of less than 0.10 mbar inside the multipass cell. In all cases we use an atmosphere of N_2 as buffer for the sample gas mixture because N_2 does not have absorption lines in the spectral range of the present device. The first two measurements correspond to controlled mixtures of acetone and ethanol. The mixtures were obtained using a double dilution process [84]. This way we could cross-check the concentration of the samples. We choose these molecules because they have strong and smooth absorption profiles in our spectral region. Then we prepared a single dilution mixture of CO_2 to study the wavenumber resolution and the sensitivity of our system, because CO_2 has a lower absorption in our region. Finally we present the measurement of two complex mixtures to show the applications of our method.

3.6.1. Acetone, ethanol and CO_2 spectra

In Figure 3.15 (a) the measured spectrum of acetone with a concentration of 22 ppm_v at 900 mbar in an atmosphere of N_2 is shown in red. The black line corresponds to the acetone spectrum in the same wavenumber region from the Pacific Northwest National Laboratory (PNNL) database [80]. In this case there are broad features that can be easily identified, they are centered at 900 cm^{-1} and 1218 cm^{-1}.

In Figure 3.15 (b) we present the measured spectrum of ethanol with a concentration of 60 ppm_v at 900 mbar in an atmosphere of N_2, in red. The black line corresponds to

the ethanol spectrum in the same wavenumber region from the PNNL database. In this spectrum the experimental setup captures the finer spectral details, such as the peaks at 1028, 1038, 1058 and 1066 cm^{-1}. In this case there is a mismatch in the region between 980 cm^{-1} and 1020 cm^{-1} corresponding to the transition between the two QCLs. But this is not a limitation to detect the desired molecules since the main spectroscopic features are clearly identifiable.
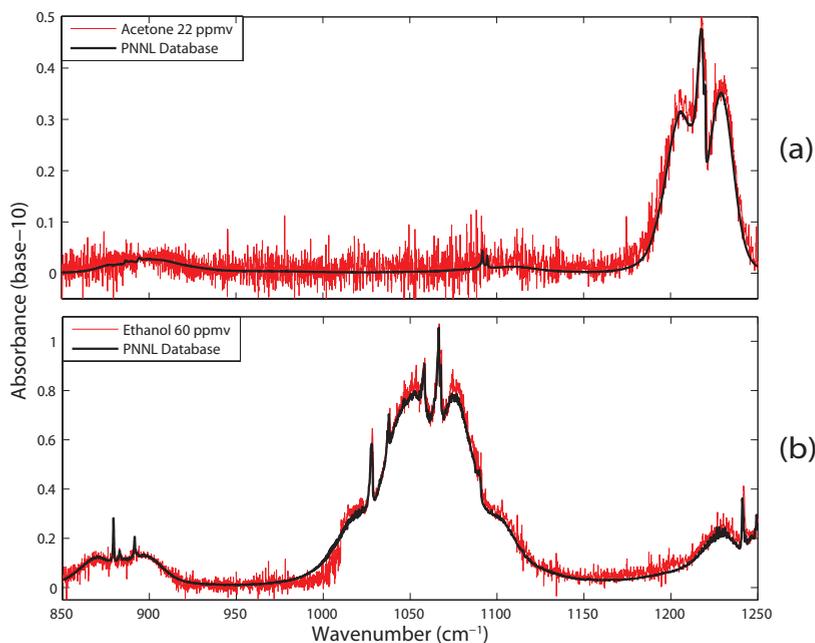
Figure 3.15: Absorption spectrum of 22 ppm_v of acetone and 60 ppm_v of ethanol. Both at 900 mbar in an atmosphere of N_2.

Furthermore, from a measurement of a 15% concentration of CO_2 at 500 mbar in an atmosphere of N_2 we could resolve the rovibrational spectrum as shown in Figure 3.16 (a). However, when we compare the spectrum with the theoretical information contained in the High Resolution Transmission (HITRAN) database [72] we notice that the strength of the absorbance does not match with the information contained in the database, Figure 3.16(b). The main reason for the mismatch is the wavenumber inaccuracy of our device, 0.1 cm^{-1}, arising from the nonlinearity of the piezo element used to scan the wavenumber and the long pulses of the QCL, which heat up the laser chip and chirped the wavelength [48]. In Chapter 4 we explore in depth the possibilities to improve the wavenumber accuracy of our system. Up to this point the wavenumber inaccuracy is too large to correctly retrieve the profile of each absorption line. Therefore it is not possible to directly measure the strength of each absorption line and the sample concentration. To determine the CO_2 concentration in the sample we first fit a Voigt profile to each individual absorption line, then we compare the maxima of the peaks with the HITRAN database using equation 2.29 to obtain a concentration value for each peak.

We thus have several concentration values calculated for different lines of the same molecule. The mean value of these concentrations give us the sample concentration. From this analysis we obtained a 15% concentration of CO_2 with a standard deviation of 5%.
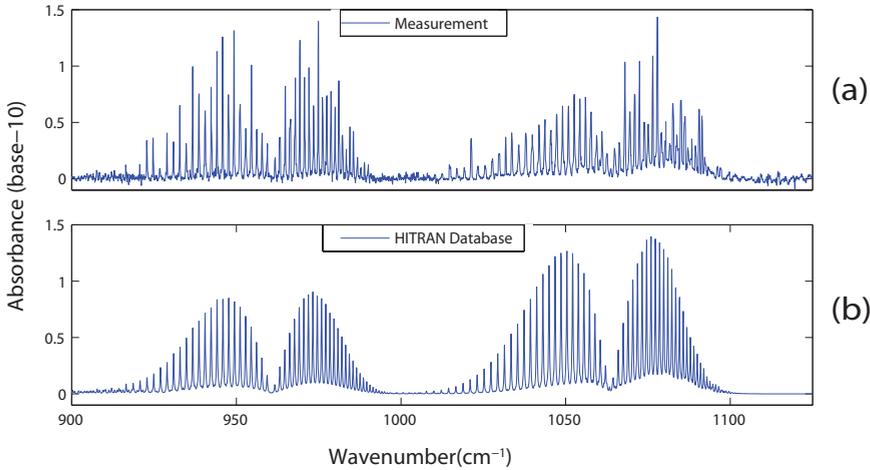
Figure 3.16: Spectrum of 15% of CO2 in an atmosphere of N_2 at 500 mbar. (a) Experimental measurement. (b) HITRAN database.

These measurements show some of the system's advantages and limitations. The broadband coverage of the system allows the molecular identification of molecules with a smooth spectral profile because all the available absorption features are used, this permits to determine the sample concentration in minutes. For molecules with sharp spectral features the identification is possible but the concentration determination strongly depends on the absorption strength.

3.6.2. LAB AIR SPECTRUM

An easy test for our setup is the measurement of the environment from our lab. For this purpose we filled the multipass cell with lab air and obtained the absorbance depicted in Figure 3.17. In this spectrum we can clearly observe five absorption lines of water at 1136 cm^{-1}, 1175 cm^{-1}, 1187 cm^{-1}, 1212 cm^{-1} and 1225 cm^{-1}. As in the case of CO_2 there is a mismatch between the strength from some measured spectral lines and the HITRAN database. Again the main reasons for this mismatch is the fact that we are observing absorption lines three orders of magnitude weaker than the strongest water absorption lines and the wavenumber inaccuracy mentioned previously. However, using the method described above we determined a concentration of 1% with a standard deviation of 0.7%, which reflects how the magnitude of the absorption strength of the molecule improves the concentration determination.

To obtained an independent measurement of the water concentration in the lab air we used a Vaisala transmitter (PTU300). It allowed us to measure the atmospheric pressure $P = 1001.0$ mbar, the relative humidity $RH = 31.0\%$ and the temperature $T = 23.0°C$ in the lab. With this information we calculated the saturated pressure of water vapour

p_w^* using the relation [100],

$$p_w^* = \left[1.0007 + (3.46 \times 10^{-6} P)\right] \times (6.1121) \exp\left[\frac{17.502\,T}{240.97 + T}\right] \qquad (3.17)$$

and then we calculated the concentration in percentage,

$$C_w = \frac{p_w^* RH}{P}. \qquad (3.18)$$

From this independent measurement we obtained a concentration of 0.9%. This value is consistent with the spectroscopic measurement obtained using our device. The main source of error is the uncertainty in the strength of the absorption signal due to the wavenumber spacing issues.
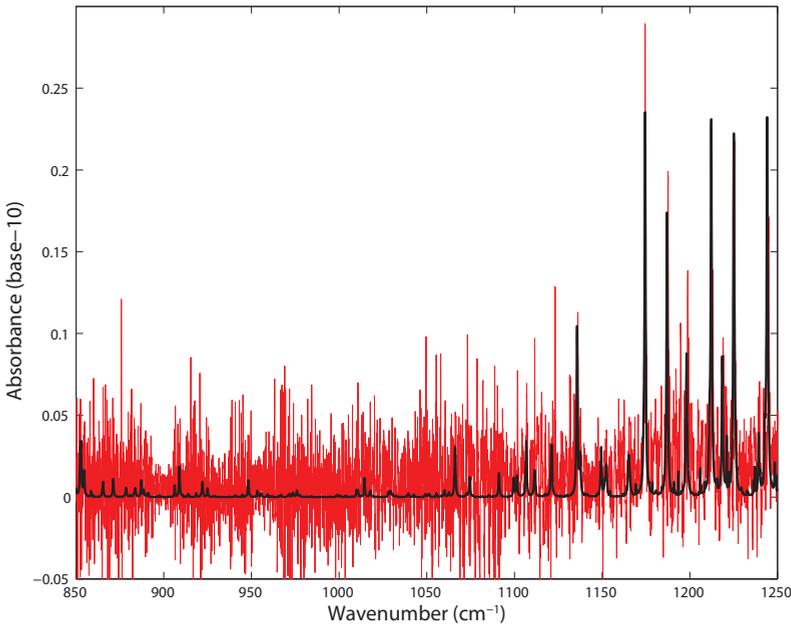
Figure 3.17: Lab air sample at 1001.0 mbar, at a temperature of 23.0 °C and a relative humidity of 31.0%. In black is depicted the spectrum of water from HITRAN database with a concentration of 1 %.

3.6.3. BREATH SAMPLE

Finally we measured a breath sample from a healthy volunteer from our research group. The pressure inside the multipass cell was 900 mbar and there was a continuous flow of the sample at 500 milliliter normal per minute (ml_n/min) [101]. As seen in Figure 3.18, we could resolve the absorption lines of CO_2 with a concentration of 3% and water at 2%. An interesting feature is the shift upwards of the base line in the region between 1170 and 1250 cm^{-1} as shown by the red curve in the inset of Figure 3.18. In this region

there is absorption from water and acetone. The black line in the inset of Figure 3.18 shows the water absorption and the 500 ppb$_v$ of acetone, giving a better match to the red curve corresponding to the measured spectrum. This concentration of acetone is in the normal range for a healthy person [102].

This measurement demonstrates the capability of the system to detect the presence of several molecules in a single sample. Using the overall profile allows to make an accurate and fast initial molecular identification of molecules with high concentrations. By examining closely the baseline of the spectrum we can identify and observe molecules with broad and smooth absorption profile. Their presence shifts the spectrum baseline and allows us to detect sub-ppb$_v$ concentrations.
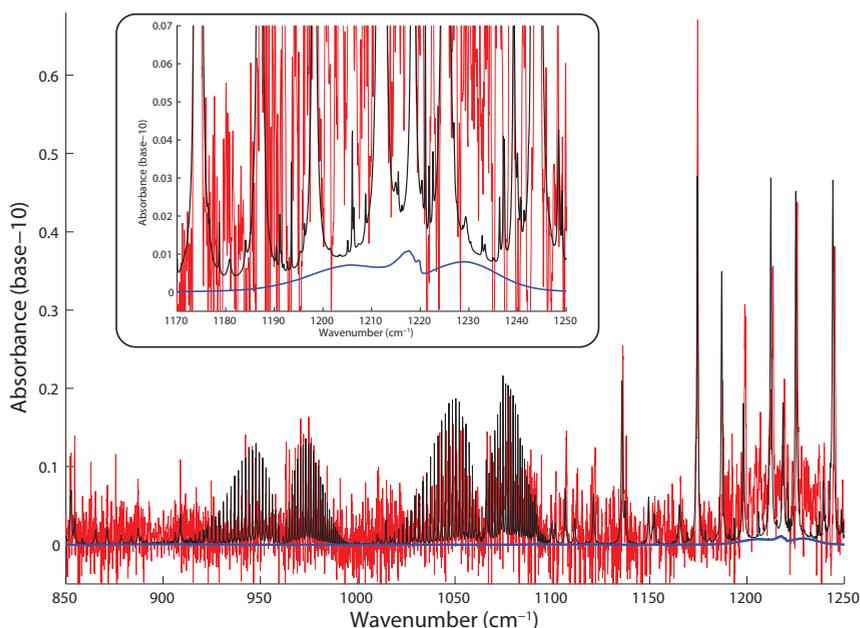
Figure 3.18: Breath sample at 900 mbar and flowing at 500 ml$_n$/min. In black is depicted the theoretical spectrum with 3% of CO$_2$ and 2% of water from HITRAN database and 500 ppb$_v$ of acetone from the PNNL database. In blue is the spectrum of acetone with a concentration of 500 ppb$_v$ taken from PNNL database.

3.7. CONCLUSIONS

In this Chapter we present a broadband spectroscopic gas analysis device which covers the molecular fingerprint region between 850 and 1250 cm^{-1}. It complements the direct absorption spectroscopic systems available in the 1250 - 3000 cm^{-1} region. At the same time broadens the spectral bandwidth covered by one single setup. This device enables the study of complex gas mixtures because the robustness of the multipass cell guarantees a stable interaction length over the full spectral scan. For applications requiring sensitivities in the ppm$_v$ and hundreds of ppb$_v$ levels this system is ideal to quickly esti-

mate the composition of a gas sample without any pre-treatment. The sensitivity characterization gave us a $NEAS$ of 2.99 $\times 10^{-7} cm^{-1} Hz^{-1/2}$. Furthermore, the sensitivity can be increased by incrementing the delay introduced by the multipass cell and measuring the absorption and the normalizing signal in one single detector [97].

The most important feature of the system is the fast identification of molecules with a broad and smooth spectral profile as shown with the acetone signature in the breath sample measurement. This shows that our system is reliable for applications such as the detection of molecular markers in complex gas mixtures like breath. Specially for cases where the concentration levels are in the order of hundreds of ppb_v or higher.

3

4

DATA ACQUISITION AND PROCESSING

The Devil is in the details

One of the big challenges in any experimental work is to define the best way to capture and save the data produced by the experiment. Even more difficult is to select the relevant information in order to speed up the data analysis without sacrificing important features. In this chapter we start by explaining the data acquisition used in our spectroscopic device. We continue describing the data analysis performed to get the absorption spectra. In particular we present and compare two methods used to correlate the information obtained from independent wavenumber scans. This is a critical step in the data analysis in order to get a reliable absorbance spectrum. Then we explain the process used to calibrate the measured absorbance spectra to a global reference: the absorbance spectra contained in the molecular databases. Using the calibrated spectra we make a comparison between the two correlations methods employed; we highlight their advantages and disadvantages and emphasize the situations where each one is more relevant. The chapter ends by giving a general description of theory behind the multiline fitting applied to determine the concentration of the gas sample components.

4.1. Data Acquisition

Let us start recalling the general structure of the experimental setup. It consists of two Quantum Cascade Lasers (QCLs), a multipass cell and two Mercury Cadmium Telluride (MCT) detectors, as depicted in Figure 4.1. The normalization detector (LaserScope unit, BlockEngineering) is used to measure the intensity of the QCLs immediately after the light is emitted. This allows us to record any intensity fluctuation produced by the lasing process or the wavenumber scanning mechanism described in Sections 3.2.1 and 3.2.2. The main detector (VIGO System) measures the absorption spectrum of the gas sample placed inside the multipass cell.
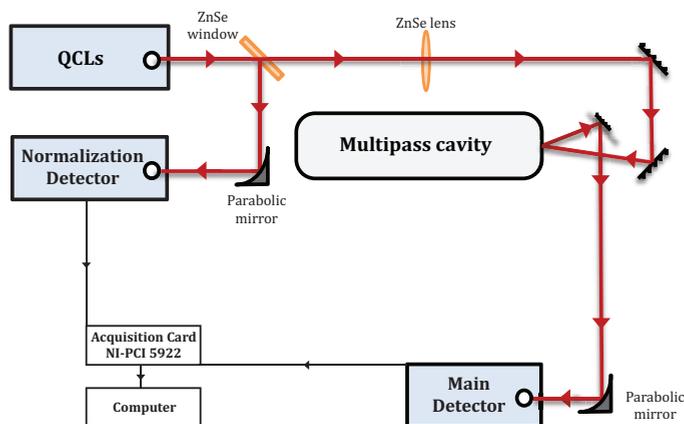
Figure 4.1: Experimental setup scheme. Only the main components of the setup are presented.

The emission of the QCLs is pulsed. The pulses have a Full Width at Half Maximum (FWHM) of 185 ns and a pulse to pulse separation of 5600 ns. In Chapter 3 we reported that the QCLs perform a wavenumber scan in 5.3 seconds, but this does not take into

account three stabilization stages performed during a wavenumber scan. We will explore these stabilization stages in Section 4.1.3. Including the stabilization stages, the total time needed for a full wavenumber scan is 6.5 seconds.

Dividing the duration of the wavenumber scan by the pulse to pulse separation we obtained the number of pulses emitted during a full wavenumber scan. This result gives us more than 1 million pulses for each single wavenumber scan. In Section 3.3.2 we digitized single pulses using a RIGOL DS1102E oscilloscope with a sampling rate of 1 Gigasample per second (GSa/s) [94]. If we use this oscilloscope to digitize the full wavenumber scan we obtain around 6.5×10^9 sample points. But this translates into a file size of more than 100 Gigabytes (GB) for each wavenumber scan for a single detector, which in turn increases the time needed to save the information. Fortunately, we do not need to record the full shape of each pulse because we only need to obtain the peak intensity of the pulses. Furthermore, the determination of the peak intensity is simplified because the pulses have an intensity plateau as can be observed in Figure 4.2a. We only need to measure one point in the plateau to determine the peak intensity, which relaxes the constraints for the digitization.

The data is digitized using a NI-PCI 5922 acquisition card from National Intruments [103]. It allows us to simultaneously capture the signals of the main and normalization detectors. The main signal is taken directly from the amplifier of the VIGO detector and the normalization signal is recorded using an output port of the LaserScope unit. The acquisition card is a two channel 24-bit flexible resolution digitizer. It can work at a maximum acquisition rate of 15 MegaSamples per second (MSa/s) with a 16-bit resolution. After analyzing the data obtained with different acquisition rates, we opted to use a 10 MSa/s rate with a resolution of 18-bits. This settings allow us to detect the peak intensity of the pulses and reduce the burden in hardware and software. With this sampling rate we acquire a data point each 100 ns, enabling us to detect at least one point from the intensity plateau of each pulse. In Figure 4.2 a pulse detected using the normalization detector is presented. In subfigure 4.2a a RIGOL oscilloscope was used and in subfigure 4.2b the NI-PCI 5922 acquisition card was employed. By comparing the two signals we can clearly observe that we manage to capture the peak of the pulse without the need to use the maximum acquisition rate of the NI card.

4

(a)   (b)
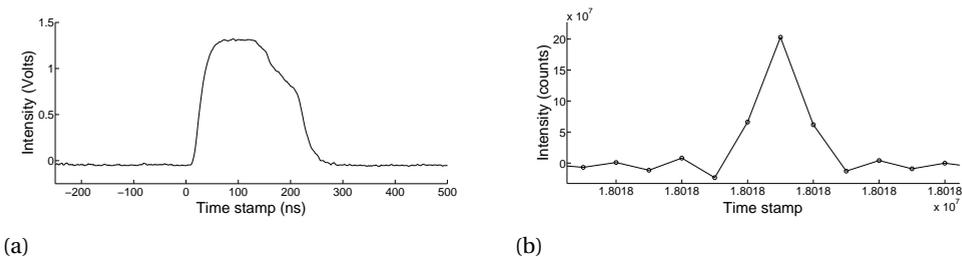
Figure 4.2: QCL pulse digitized using (a) a RIGOL oscilloscope with an acquisition rate of 1 GSa/s and (b) a NI-PCI 5922 acquisition card with an acquisition rate of 10 MSa/s

4.1.1. CONFIGURATION OF THE DATA ACQUISITION

To manage the LaserScope unit and the NI acquisition card we developed a LabVIEW program. It allows us to define when the QCLs emission starts, to record the signal from the main and normalization detectors simultaneously and to save the information in the most convenient location. The communication with the LaserScope unit is possible thanks to a remote control status available in its software. It allows to send a limited number of commands to the LaserScope unit, which are listed in Appendix B. For our purpose we only used the command *scan* which permits us to start a wavenumber scan at a specific moment.

The acquisition card has a built-in memory to store the data. For our application this memory is too small to record all the information produced during a wavenumber scan. When the card memory is full it overwrites the information, for this reason we use a Virtual Instrument (VI) to fetch the data to the computer hard disk. To guarantee that all the information is recorded, we start the acquisition and fetching before the wavenumber scan starts and we keep the acquisition running for a time window larger than the 6.5 seconds a wavenumber scan takes. In our experience an acquisition time of 8 seconds is enough.

The data acquisition process consists of the following steps:

1. Initialization of the communications with the LaserScope unit and the NI card.
2. Setting the number of wavenumber scans and the parameters for the NI card.
3. Define the file name and path for the information storage.
4. Create the files to save the data.
5. Command the NI card to start the data acquisition.
6. Send a trigger signal to initiate the fetching VI, if the trigger fails we stop the acquisition and start all over again.
7. Tell the LaserScope to start the wavenumber scan.
8. After 8 seconds the acquisition stops and we wait until the fetching VI finishes saving all the information still remaining in the memory card.
9. Check if all the wavenumber scans were performed.
10. Close all the communications.

In Appendix C we present a flow chart depicting the full acquisition process.

The parameters given to the acquisition card at the beginning of the LabVIEW program should be chosen properly to ensure an accurate data acquisition. Although we have optimized the values of some of these parameters, such as the acquisition rate, the interface of our LabVIEW program gives the freedom to change them to further develop the system. In Table 4.1 we present each parameter and its description.

From the parameters it is interesting to note that we can choose to use only one channel for the acquisition, this allows us to check the performance of a single detector. Another interesting feature is the maximum points per fetch, it indicates that the fetching VI saves the information by blocks. This is an effective way to release memory in the acquisition card to continue the data acquisition. More detailed information of the data acquisition is available in the work of Z. Hou [88].

Table 4.1: Parameters of the NI acquisition card

Parameter	Description
Channel Name	Channels used for the data acquisition. Allows the use of 1 or 2 channels for the acquisition.
Min. Sample Rate	Acquisition rate used during the data sampling. 10 MSa/s is used in this work.
Vertical Range	Defines the intensity range of the detection. As low as possible to increase the digitizing resolution. Large enough to avoid saturation.
Max points per fetch	Number of points to save during each fetch
Acq. time (Sec)	Time the acquisition is active
Number of Measurements	Number of laser scans

4.1.2. PEAK AND VALLEY DETECTION

A file containing all the data points recorded during the data acquisition and a time stamp taken with respect to the start of the acquisition has a size of around 1 GB. But most of the points correspond to the spacing between pulses as can be seen in Figure 4.3. We can reduce the files size and the time needed to save the information by recording only the peak intensity of each pulse and the level of the background between consecutive pulses. This idea allows us to reduce the size of each file to around 40 MB and to reduce the saving time by a factor of 6.
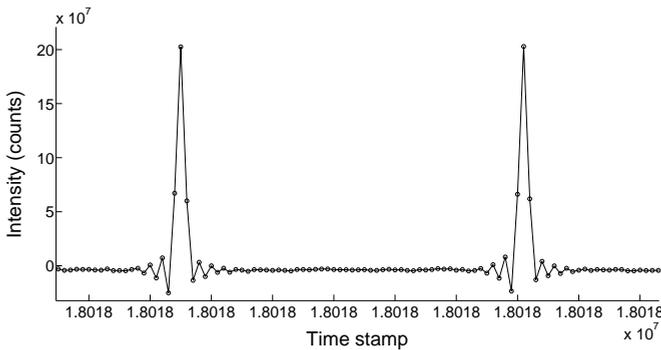
Figure 4.3: Two consecutive pulses recorded by the normalization detector. The circles show the individual data points recorded. 56 points are necessary to capture the 5600 ns period between pulses.

We did not limit ourselves to save only the peak of the pulses because the reverse bias applied to the MCT detectors changes the background level from which each pulse peak should be measured, as shown in Figure 4.4. This figure shows the background changes recorded by the normalization detector. The negative values recorded are the result of the reverse bias apply to the MCT detectors. It is worth mentioning that the background variation does not change dramatically around a single pulse, but it becomes relevant in the long time scale.
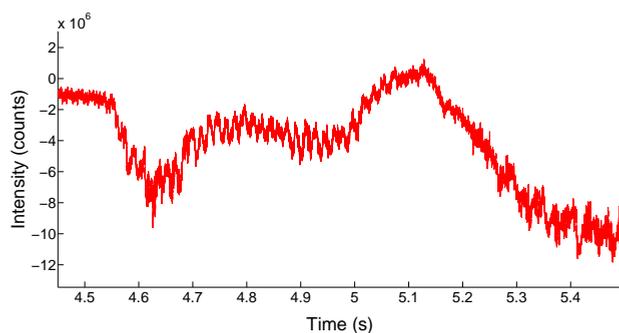
Figure 4.4: Background of the normalization signal. The change in level is clear over a long time range.

The peaks of the pulses are directly retrieved using a peak detection virtual instrument (pd-VI) from LabVIEW [104]. This pd-VI finds the location of the peaks in the original data and performs a least-square quadratic fit to determine their intensity. For each peak the pd-VI fits a parabola using neighboring points. The number of neighboring points used for the fit is defined in the width parameter of the pd-VI. LabVIEW needs at least three neighboring points to perform the fitting and it suggests to use a width value smaller than one half of the FWHM of the pulse. Once the pd-VI has performed the quadratic fit, it checks if the fitted parabola meets the criteria of a maximum. Then the pd-VI compares the maximum of the parabola with a chosen threshold. This last step permits us to pick up only the values corresponding to actual peaks and reduces the appearance of false positives.

Because the intensity and location of the peaks are determined using a quadratic fit, it is important to compare the calculated values with the values of the measurements to determine their accuracy. In Table 4.2 we present the error in intensity and location of the fitted peaks for different widths.

Table 4.2: Intensity and location accuracy of the peaks using the LabView VI

Width (# of points)	Intensity error (%)	Location error (%)
3	0.39	1.80
4	2.12	0.98
5	22.82	1.96

The intensity error retrieve by the pd-VI increases dramatically with the increment of the number of points used for the fitting. With a larger number of points for the fitting the pv-VI can reduce the residuals of the parabola fitting. However, the maximum of the parabola moves away from the value of the experimental peaks, which is the information we want to determine. From the results it is clear that three points give the best accuracy. The absorbance uncertainty induced by the 0.39% intensity error can be determined in the same way we calculated the noise of our setup in Section 3.5.1. We apply the Beer-Lambert law [79], equation 2.27 repeated below, assuming the signals $I(\widetilde{\nu})$ and $I_0(\widetilde{\nu})$ differ by 0.39% at each wavenumber $\widetilde{\nu}$,

$$A(\widetilde{v}) = -\log_{10}\left(\frac{I(\widetilde{v})}{I_0(\widetilde{v})}\right). \qquad\qquad \text{(2.27 revisited)}$$

The result gives us an uncertainty in the absorbance of 0.002, which is one tenth of the noise level associated with our setup, as calculated in Section 3.5.1. Therefore, the contribution of the intensity error is already included in the noise level of the system.

The uncertainty in the location of the peaks is not a problem because the separation between the QCL pulses is large enough to avoid overlapping. Because the pulses are equally spaced an error of 1.80% can be corrected easily, as we will show in Section 4.1.3.

In the case of the background level we use a different approach because the background is not defined by local minima. As shown in Figure 4.5, the background is a flat region with intensity variations due to noise. The region closest to the pulses is noisier than the central region between the pulses; for this reason we use this central part to determine the background level. First, we use the position of two adjacent peaks to calculate the location of the central point of the background; then we average the intensity of 23 points around this central point to get a background with less noise. From now on we will refer to the background points calculated this way as valleys.
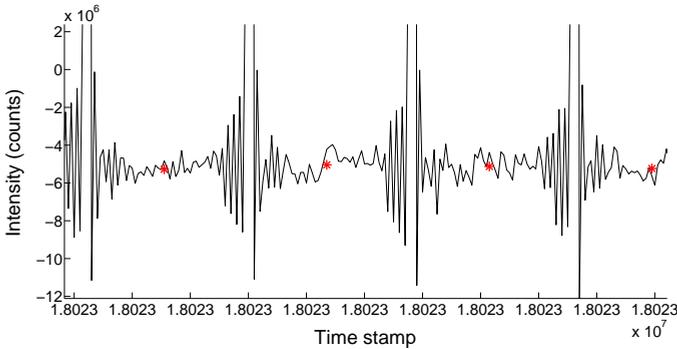
Figure 4.5: Background determination. The red asterisks correspond to the calculated valleys of the signal.

In this acquisition scheme only the values of the peaks and valleys are saved, reducing the size of the files to around 40 MB. During each wavenumber scan four files are saved, two corresponding to the peaks and valleys recorded by the main detector and two for the peaks and valleys measured by the normalization detector.

In Figure 4.6 the peaks and valleys captured by the normalization detector during a wavenumber scan are presented. The figure contains one peak and one valley for each pulse emitted by the QCLs, therefore it contains more than a million peaks and valleys. The horizontal axis shows the 8 seconds during which the acquisition card was active. From this signal we can obtain valuable information. The first thing we want to highlight is the absence of signal at the beginning and end part of the acquisition time frame. During these moments the only signal is the noise from the detectors because the QCLs are turned off. This absence of signal demonstrates that the threshold parameter chosen for the pd-VI is high enough to eliminate the noise. Furthermore, the presence of these two regions guarantees that all the laser information is collected. If the threshold level is

poorly chosen we see a low noise signal in these regions with an intensity several orders of magnitude lower than the intensity of the QCLs signal. During the signal checkup described in Section 4.1.3 we use this feature to determine the start and end of a full wavenumber scan and to remove any noisy signal at the beginning and end of the data.
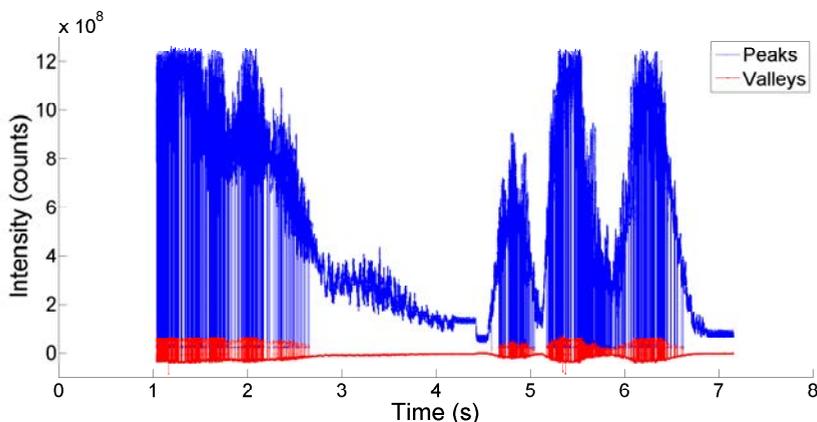
Figure 4.6: Peaks and valleys recorded by the normalization detector during a full wavenumber scan.

The second characteristic we observe in Figure 4.6 is the presence of regions where the values of the peaks and the valleys are very similar. In Figure 4.7 we zoom into one of this regions and observe that in indeed the peaks and valleys values are very similar and even in some cases they overlap. This happens when the peak detection VI detects more peaks than actual pulses. The origin of this lies in the small width parameter provided to the pd-VI. We prefer to keep these false peaks and valleys rather than sacrifice the intensity accuracy because we can easily identify the false data and remove it. In Section 4.1.3 we will explain in detail the data processing applied to remove the false peaks and valleys.
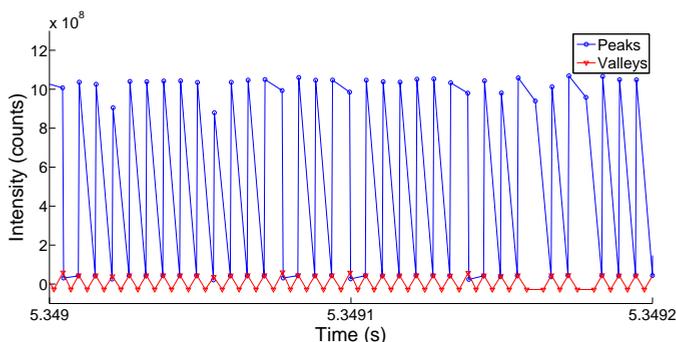
Figure 4.7: Zoom into the peaks and alleys signal recorded by the normalization detector during a full wavenumber scan.

4.1.3. SIGNAL CHECKUP AND DATA CORRECTION

From the recorded peaks and valleys we can get the intensity profile of the normalization and main signals. We use a MATLAB program to produce the corresponding intensity profiles. The first step to obtained these profiles is to remove any noise signal at the beginning and end of the recorded data. This is done simply by looking if there is any abrupt intensity change at the beginning or end of the peak values and removing the data corresponding to noise. Then, we remove any false peaks appearing in the remaining data using the periodicity of the laser pulses. The separation between two consecutive pulses is well established at 5600 ns. Therefore any point whose separation with its neighbors is smaller than 5600 ns is discarded.

The acquisition scheme can also miss recording some peak pulses. Thereupon we also look for regions where the spacing between points is larger than 1.5 times the spacing period of 5600 ns. If we find such regions we add a Not-a-Number (NaN) element to the peak and valley files as a marker to point the missing information.

The last step is to compute the difference between the peaks and valleys to get the intensity profile. Figure 4.8 presents a plot of the intensity profile obtained. From this plot we observe some interesting features of the QCLs behavior. First of all, we can clearly identify the moment when the laser emission switches between the QCLs in the middle of the figure. Other interesting features are the stabilization regions that appear at the beginning and end of each QCL scan. From the profile of these regions we deduct that the piezo element of each QCL performs a scan over a narrow bandwidth to calibrate its position at the beginning and end of each QCL scan.
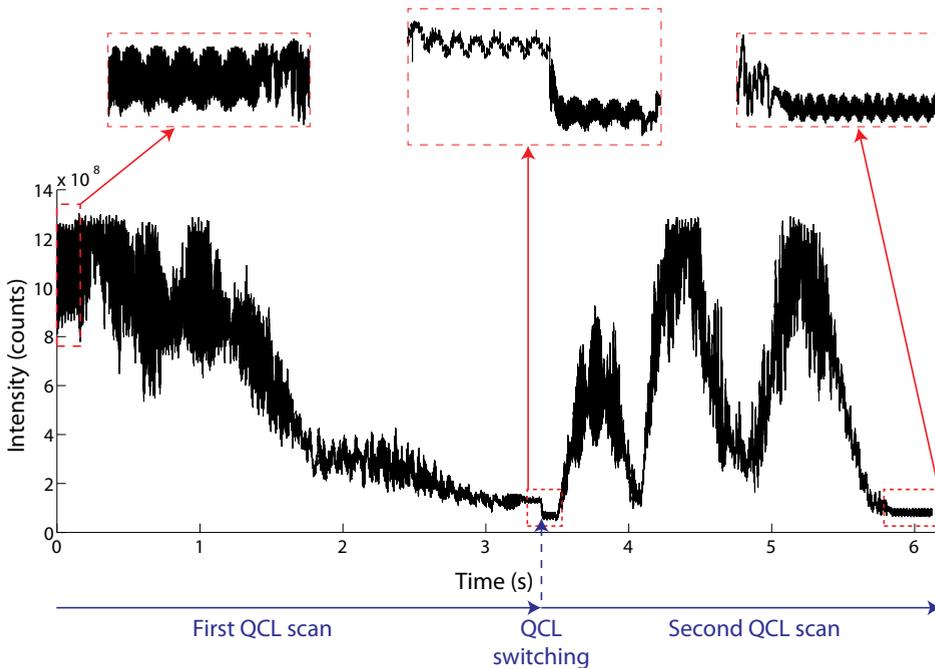
Figure 4.8: QCLs intensity profile obtained with the normalization detector.

Another interesting aspect of the laser scan is visible when we take a closer look into the intensity profile. In Figure 4.9 we zoom into a region where one of the QCLs is scanning. In this figure we can see two types of regions: one where the intensity changes are smooth, marked with a white background, and one where the intensity changes are abrupt, marked in light gray. This behavior by zones is typical of semiconductor lasers whose wavenumber emission is selected using an external cavity. The sudden changes are known as mode hops and originate from the competition between emission modes determined by the multiple Fabry-Perot cavities formed by the end-facets of the laser chip and the optical elements of the external cavity[78, 105]. We will denote the regions where the intensity changes are smooth as mode zones and the regions with abrupt changes as mode-hop zones.
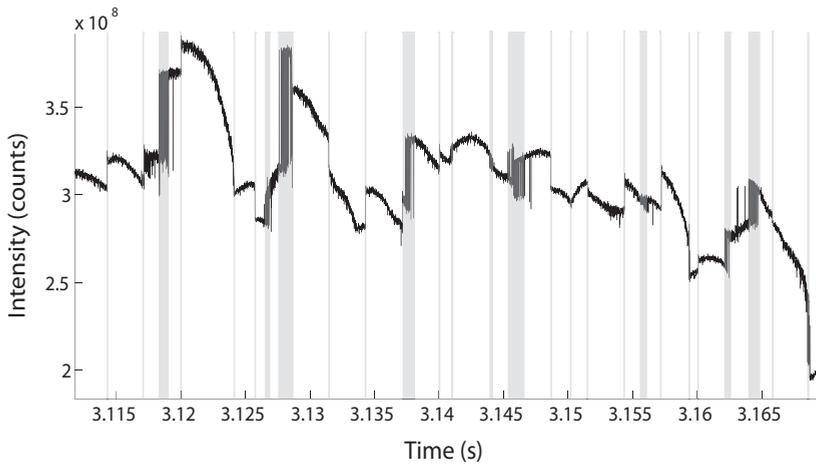
Figure 4.9: Mode zones (white) and mode-hop zones (gray) produced by the tuning of the wavenumber using an external cavity.

Until this point we have described how to check and correct the data obtained from each individual wavenumber scan. In order to analyze a gas sample we need to perform at least two wavenumber scans, one with the multipass cell empty and one with the gas sample in the cell. Furthermore we need to perform more wavenumber scans with the same conditions inside the multipass cell to reduce the noise level intrinsic to the QCLs emission. However when we compare the intensity profiles obtained from two consecutive wavenumber scans of the QCLs we observe a mismatch phenomenon in the signals. In Figure 4.10 we can observe this phenomenon for 3 consecutive signals taken with the normalization detector. We can clearly see that they do not overlap. Furthermore, the signals are compressed or stretched differently for different parts of the signal. This difference in the repeatability has its origin in the nature of the scan; mainly in the limitations of the piezo elements to reproduce its movement with enough accuracy as we will show in Section 4.2.1.
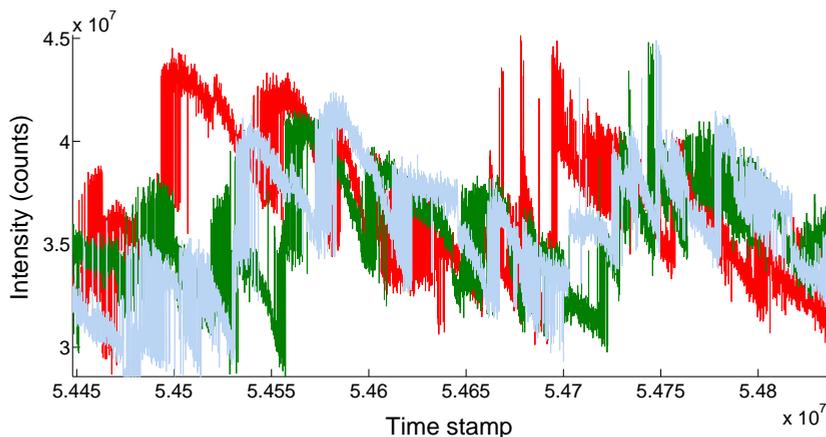
Figure 4.10: Mismatch in the intensity profile of three consecutive wavenumber scans.

So far all the wavenumber scans are independent from each other. In order to compare them we need to correlate the different wavenumber scans using a common reference, only then we can aspire to get any accurate spectroscopic information. During the development of this work we used two approaches to tackle this problem. One method uses selected points of the spectrum whose wavenumbers are provided by the LaserScope unit. The second method applies a warping algorithm to correct the mismatch between the different wavenumber scans. In the following sections 4.2 and 4.3 we explore in detail these two methods.

4.2. SCAN CORRELATION USING SELECTED WAVENUMBERS

The first method we use to correct the mismatch between different wavenumber scans is inspired by the method employed in the proprietary software of the LaserScope unit. The unit was originally developed to analyze materials with a strong absorbance signature using only one detector. For this purpose the LaserScope unit complies perfectly. However, the performance is not sensitive enough when the absorbance of the samples are considerable low, as in our case. Furthermore, the LaserScope method can not be implemented directly using two detectors. For clarity we will first explain the method used by the LaserScope unit and then we will give the details of our approach.

4.2.1. LASERSCOPE METHOD

The intensity profiles reported by the LaserScope unit always contain the same wavenumber values, which we call the selected wavenumbers. These selected wavenumbers are equally spaced by exactly 0.2 cm^{-1}, which is lower than the . This is possible thanks to a data post-processing applied by the software of the LaserScope unit.

During a wavenumber scan the LaserScope unit writes a temporary file containing the voltage applied to the scanning piezo, a wavenumber assigned to each voltage value and the intensity detected by the normalization detector. It saves a separate file for each

QCL. From this information we infer that the LaserScope software must contain a calibration file to correlate the voltage applied to the piezo and the assigned wavenumber. Unfortunately we do not have access to the calibration file and therefore we do no know how accurate this assignment is.

When we plot the assigned wavenumbers against the applied voltage we obtain the curve shown in subfigure 4.11a. We recover the same curve for different QCL scans, but the exact values of the voltage and the wavenumber are slightly different as can be seen in the zoom shown in subfigure 4.11b. In this subfigure we present the results obtained from ten independent QCL scans. The spread of the points demonstrates the inability of the LaserScope to repeat exactly the same wavenumber scan every time. However, all the points are aligned independently of their respective wavenumber scan, confirming that the assigned wavenumbers are obtained from a calibration curve. Moreover, the points of different wavenumber scans tend to cluster around the selected wavenumbers.

(a)   (b)
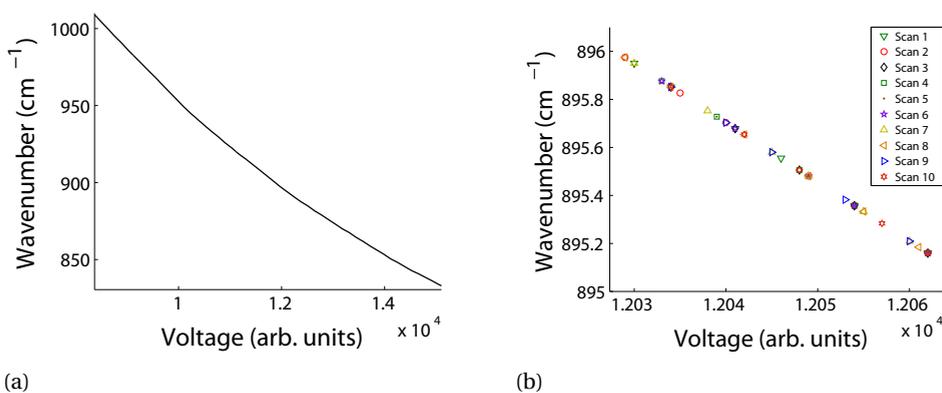
Figure 4.11: Wavenumber dependance on the voltage applied to the piezo scanner.

The LaserScope unit processes the data of each scan to establish the intensity of the selected wavenumbers. Because we do not have access to the details of the data processing we analyze the available information to discover the steps implemented by the LaserScope unit. In Figure 4.12 we present three sets of data, they all correspond to the same wavenumber scan. The blue dots correspond to the raw data saved by the LaserScope unit while the red asterisks show the result of the LaserScope data processing. In contrast to the raw data, the processed data is smoother and its points are equally separated by exactly 0.2 cm^{-1}. From this information we can only assume that a smoothing and an interpolation was applied to the original information. We tried to recreate the data processing by smoothing the raw data with a moving average of 7 points [89] and using a linear interpolation. The black markers in Figure 4.12 show that our results are close to the values obtained with the LaserScope unit data processing; but without knowing the full details of the LaserScope data processing we can not say more about the accuracy of this method.
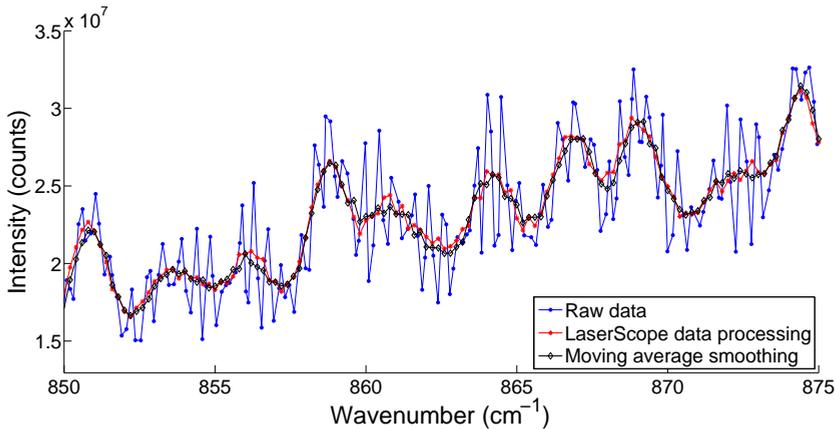
Figure 4.12: LaserScope method. Blue dots correspond to the raw intensity profile, in red is shown the result of the data processing and the black markers are the result of a moving average smoothing and the interpolation applied to the raw data.

Figure 4.13 presents the intensity profiles of ten independent wavenumber scans obtained with the LaserScope data processing.
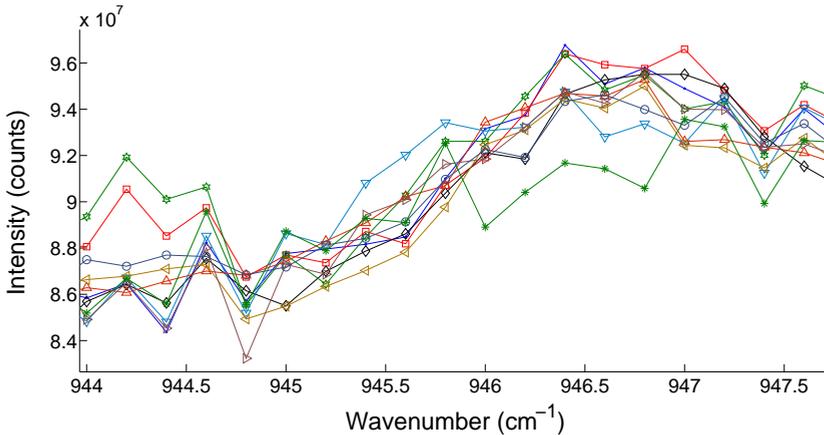
Figure 4.13: Intensity profiles obtained with the LaserScope method.

The final intensity profiles reported by the LaserScope unit only contain the average of all the wavenumber scans, but do not give any information of the associated uncertainty. As we mentioned before, this method is adequate for the purpose the LaserScope unit was originally developed for. We can not apply it directly to our system because we need two detectors, which has been proven to increase the accuracy of gas detection, specially for low concentrations [106, 107]. In practice we can apply the same concept to our signals by selecting the wavenumbers, as we describe below.

4.2.2. Implementation using two detectors

The LaserScope method can be used to correlate the intensity profiles obtained using the main and normalization detectors of our experimental setup. In order to apply the LaserScope method we need to implement a series of steps in order to get the correct result.

1. First, we find the temporal separation between the data points saved in the temporary files of the LaserScope unit.

2. Second, we determine the temporal coordinates of the data points of the temporary files.

3. In third place, we choose a separation of 0.05 cm^{-1} for the selected wavenumbers.

4. Then, we find the temporal coordinates of the selected wavenumbers by interpolating the data points of the temporary files.

5. We determine the intensity of the selected wavenumbers by averaging the intensity of 40 pulses around the temporal coordinates of the selected wavenumbers.

6. Using the same selected wavenumbers, we repeat the same method for independent scans to correlate them.

7. The last step is to improve the correlation building a reference profile.

In the next paragraphs we present in detail how to implement these steps.

Temporal separation of the data points of the temporary files.

The signals obtained with our NI acquisition card only contain the intensity of the pulses and the time coordinate when they where detected. It does not contained any information regarding the wavenumber of the pulses. We can use the temporary files of the LaserScope unit to determine the wavenumber of each pulse. However, in order to relate the information of the temporary files and the pulses captured by the detectors we need to determine the time coordinates of the data points contained in the temporary files. For this reason we first look for the temporal separation of the data points of the temporary files.

The temporary files are erased after each QCL scan, for this reason we use a program to copy them as soon as they are created. We must keep in mind that the LaserScope creates one temporary file for each QCL, so in for each complete wavenumber scan we obtained two temporary files. The first thing we note is a series of data points at the beginning and end of each temporary files for which the wavenumber information is not assigned. The number of data points in these regions varies for each wavenumber scan, but it always matches the number of mode zones appearing in the stabilization regions discussed in Section 4.1.3. Therefore we can conclude that the LaserScope unit saves a data point for each mode zone of the stabilization region. In Figure 4.14 we present the intensity profile of the first QCL and a zoom into the stabilization regions where we can observe the mode zones.
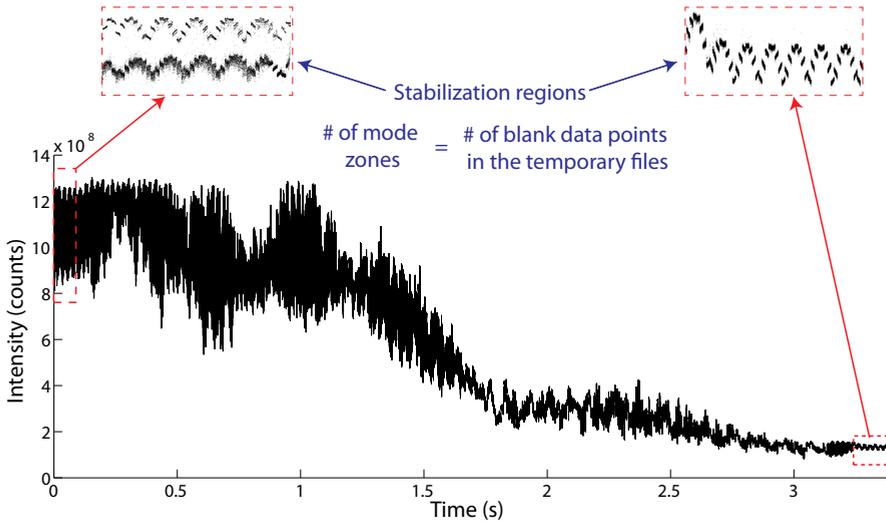
Figure 4.14: Intensity profile of the first QCL obtained with the normalization detector. The stabilization regions are highlighted.

The fact that for each mode zone there is a data point in the temporary files allows us to determine their temporal separation. In particular because the mode zones of the stabilization regions have a defined temporal width as can be seen in Figure 4.15. In this figure we look closely at one of the stabilization regions of the first QCL. Each mode zone in the stabilization region is composed of 511 pulses. Using the 5600 ns separation between pulses we get a value of 2.8616 ms for the width of the mode zones. Since the data points of the temporary files are equally separated the width of the mode zones is the temporal separation of the data points. We observe exactly the same behaviour in all the stabilization regions.
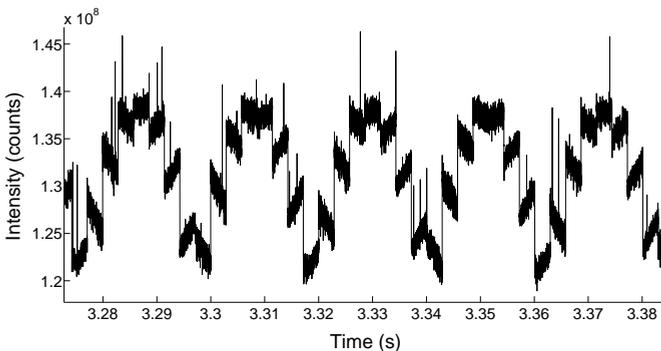
Figure 4.15: Intensity profile of a QCL stabilization region. The temporal width of the mode zones is constant at 2.8616 ms.

DETERMINATION OF THE TEMPORAL COORDINATES OF THE SELECTED WAVENUMBERS.
Once we know the temporal separation of the data points stored in the temporary files we can determine when they were recorded by the NI acquisition card. This way we get the temporal coordinates and the wavenumbers of the data points saved in the temporary files. Using this information we can make an interpolation to find the temporal coordinate of some selected wavenumbers. Furthermore, as the detectors record many more pulses than the LaserScope software we can look for the temporal coordinate of new selected wavenumbers with a separation smaller than the 0.2 cm^{-1} used in the LaserScope software. We decide to use selected wavenumbers with a separation of 0.05 cm^{-1} to improve the intensity profiles.

INTENSITY DETERMINATION
The last step to produce the intensity profile is to calculate the average intensity of 40 points around the new selected wavenumbers. This reduces the intensity variations of the selected wavenumbers that sit in mode-hop zones, where the intensity changes are large and fast. We are able to correlate independent wavenumber scans by using the same selected wavenumbers. In Figure 4.16 we present the result of this correlation process applied to ten independent wavenumber scans. The results show an intensity distribution qualitatively similar to the one obtained from the temporary files of the LaserScope unit. The main difference is the smoothness of the new intensity profiles, which results from using selected wavenumbers with a smaller separation.
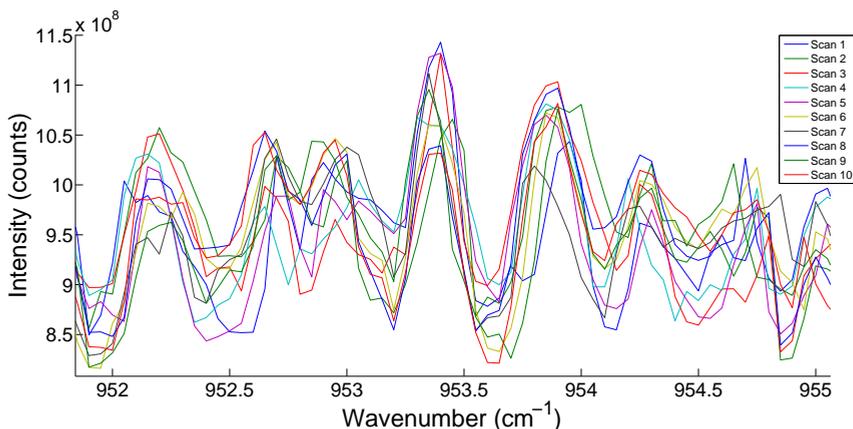
Figure 4.16: Intensity profiles obtained with selected wavenumbers separated by 0.05 cm^{-1}. Data obtained with the normalization detector.

This procedure can be applied simultaneously to the normalization and main signal using the same time coordinates. The reason lies in the fact that for each pulse detected in the normalization signal there is a corresponding pulse in the main signal. It is here when it becomes clear the importance of compensating any missing pulse in any of the signals with a Not-a-Number (NaN) element to conserve the one to one relationship, as mentioned in Section 4.1.3.

CORRELATION IMPROVEMENT

Although the intensity profiles obtained so far show a good match between independent wavenumber scans, the remaining mismatch is limiting the capacity of the measurements to perform accurate gas analysis. This mismatch can be appreciated in the wavenumber position of the maxima and minima of the intensity profiles presented in Figure 4.16. They do not occur at the same horizontal positions. To correct this error we apply a correlation improvement.

First we define a reference profile to which we can adjust all the wavenumber scans. The average of the normalization signals provides a reliable reference because the intensity variation between different normalization signals only depends on the changes in the QCLs emissions. The main signals are not adequate to build a reference because their intensity profile is determined by the gases present inside the multipass cell.

In Figure 4.17 we illustrate the wavenumber correction applied to one of the normalization signals. After getting the reference profile, plotted with a dashed curve, we find its local extrema, shown in black boxes. Then the local extrema of the normalization signal are locally shifted to the closest extrema of the reference profile, taking care that each minimum of the normalization signal is moved to a minimum in the reference profile. The data between the extrema of the normalization signal is re-sampled using linear interpolation.
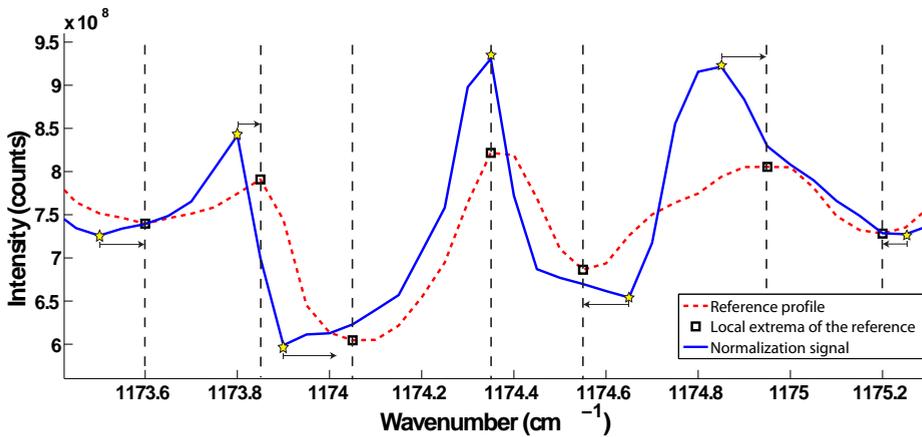
Figure 4.17: Wavenumber correction using the average of the normalization signals as base.

Thanks to the one to one relation between the data in the normalization and main signals of a single scan, the transformations that we apply to correct the wavenumbers of the normalization signal can be immediately used to correct the corresponding main signal.

We repeat this method for each pair of normalization and main signals. The result of this correlation improvement is presented in Figure 4.18. Each subfigure shows the intensity profile of ten wavenumber scans, five from a reference measurement with the multipass cell empty and five from a gas measurement. The upper plots show the uncorrected profiles from the normalization and main signals, while the lower ones show the

corrected profiles. One characteristic visible in the corrected profiles of the main signals (Figure 4.18d) is the difference between the reference measurements and the gas measurements in the region between 1174 and 1175 cm^{-1}. It corresponds to an absorption line of water present in the gas sample. This difference does not appear in the corrected normalization signals (Figure 4.18c). As mentioned before, the presence of these absorption features makes the main signals unsuitable to build a reference profile because they are determined by the gases present in the multipass cell. For the normalization signals this is not the case and therefore they are ideal to perfom the correlation correction even when the conditions inside the multipass cell are different.

(a)

(b)

(c)

(d)
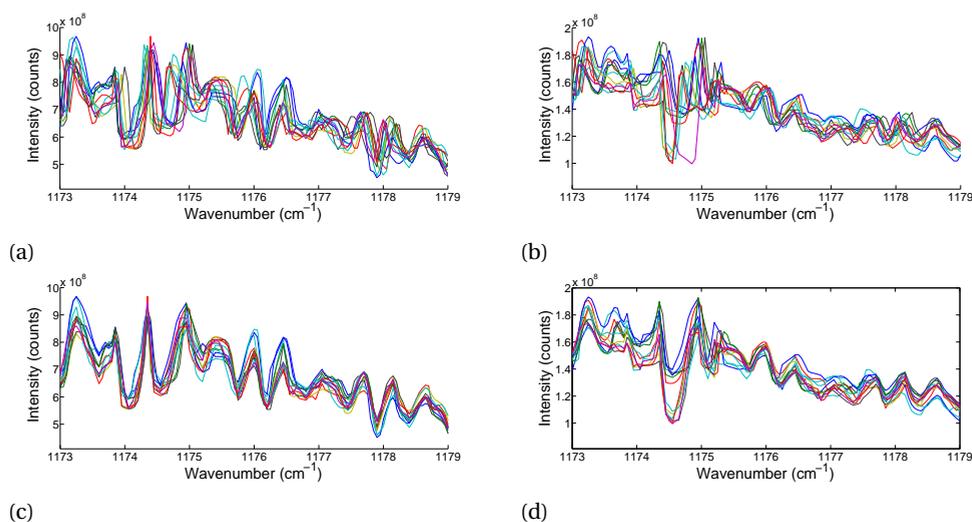
Figure 4.18: Correlation improvement. (a) Uncorrected normalization signals. (b) Uncorrected main signals. (c) Corrected normalization signals. (d) Corrected main signals.

Using the corrected intensity profiles we can build a reference and measurement spectra as described in Section 3.5 and get the absorbance spectrum of the gas sample. In Figure 4.19 we present the absorbance spectrum of a breath sample constructed using this method. We can clearly observe the absorption signature of water and CO_2.

The sharp absorption lines of water and CO_2 allow us to evaluate the wavenumber calibration obtained from this method. There is a discrepancy in the wavenumbers when we compared the absorption lines of CO_2 appearing in the measurement of Figure 4.19 and the values reported in the HITRAN database. In some places the discrepancy is as large as 1 cm^{-1}, which is in accordance with the specifications of the LaserScope software [82]. The origin of the error in the wavenumber calibration comes directly from the calibration file used by the LaserScope software to assigned wavenumber values to the data points of the temporary files. In Section 4.4 we explain how to correct the error in the wavenumber calibration using directly the information of the molecular databases.
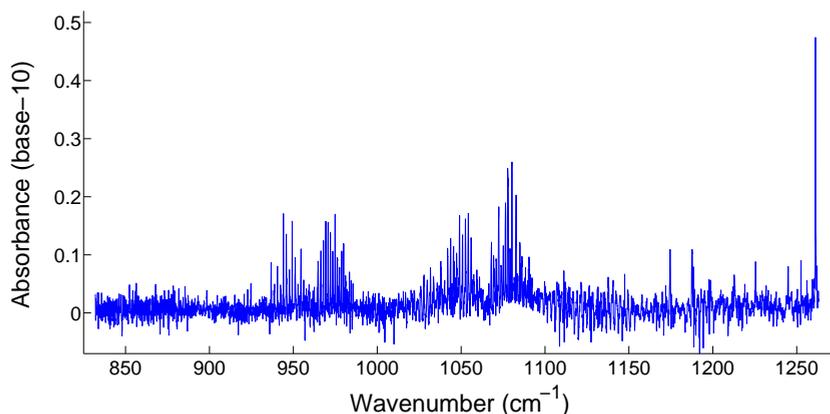
Figure 4.19: Absorbance spectrum of a breath sample built using the selected wavenumber correlation and the wavenumber correction.

4.3. SCAN CORRELATION USING SEMIPARAMETRIC TIME WARPING

A second approach is to correct the mismatch between different scans using the warping ideas originally developed to compare signals in speech recognition. In this context Bloemberg [108] defines warping as the shift, stretch, and compression applied to a signal, either completely or piecewise, that permits to match its features to a reference signal. The warping methods proved to be very useful when used to preprocess the spectra and chromatographs obtained from nuclear magnetic resonances (NMR), gas chromatography–mass spectrometry (GC–MS) and liquid chromatography – mass spectrometry (LC–MS) in order to determine the composition of the samples under study.

The work of Bloemberg [108, 109] gives an extensive and clear review of the different warping algorithms used to correct spectroscopic data, such as parametric time warping (PTW), semiparametric time warping (STW), dynamic time warping, correlation optimized warping, fuzzy warping, peak alignment by means of a genetic algorithm, peak alignment by beam search, peak alignment by fast Fourier transform, recursive alignment by fast Fourier transform, recursive segment-wise peak alignment and interval correlation-shifting. He discusses the advantages and disadvantages of each method. For our study we decided to use semiparametric time warping because it corrects the full signal, respects the intensity of the signal and provides an accurate matching between signals.

For clarity we will first give a general description of the theory behind STW and then we will present the implementation of the method in our system.

4.3.1. THEORY

The semiparametric time warping method (STW) was developed by Eilers [110]. The main idea of STW is to align an intensity profile $x_i = x(t_i)$ to a reference profile $y_i = y(t_i)$ by modifying its horizontal coordinate t_i using a warping function $w(t_i)$. The horizontal coordinate t_i can correspond to time, wavenumber or simply the indices of the profile points. The goal of the algorithm is to find the warping function $w(t_i)$ that minimizes the difference between the reference profile y_i and the warped profile obtained once the horizontal coordinate is resampled $x(w(t_i))$.

$$S = \sum_{i \in H} \left[y_i - x(w(t_i)) \right]^2 \tag{4.1}$$

here H indicates the set of indices i for which $x(w(t_i))$ can be computed.

In STW the warping function $w(t_i)$ consists of a series of B-splines constructed from polynomial pieces joined at specific t_i values. It has the form

$$w(t_i) = \sum_{j=1}^{n} b_j B_j(t_i) \tag{4.2}$$

where n corresponds to the number of B-splines, B_j is the jth B-spline computed at the points t_i and b_j is its coefficient. The appropriate warping function $w(t_i)$ is found by optimizing the warping coefficients b_j until equation 4.1 is minimized. In Reference 111 the reader can find a very clear description of B-splines and their properties.

If an approximation $\widetilde{b}$ to the coefficient vector b exists, the corresponding $\widetilde{w}(t)$ [112] is

$$\widetilde{w}(t) = \sum_{j=1}^{n} \widetilde{b}_j B_j(t) \tag{4.3}$$

The difference between any $w(t)$ and $\widetilde{w}(t)$ is

$$\Delta w(t) = \widetilde{w}(t) - w(t) = \sum_{j=1}^{n} \widetilde{b}_j B_j(t) - \sum_{j=1}^{n} b_j B_j(t) =$$
$$\sum_{j=1}^{n} \left[\widetilde{b}_j - b_j \right] B_j(t) = \sum_{j=1}^{n} \Delta b_j B_j(t) \tag{4.4}$$

Using the Taylor expansion for $x(w(t))$ and equation 4.4

$$x(w(t)) \approx x(\widetilde{w}(t)) + x'(\widetilde{w}(t))\Delta w(t) = x(\widetilde{w}(t)) + \sum_{j=1}^{n} \Delta b_j x'(\widetilde{w}(t))B_j(t) \tag{4.5}$$

By substituting equation 4.5 in equation 4.1 we obtain the relation

$$S = \sum_{i \in H} \left[y_i - x(\widetilde{w}(t_i)) - \sum_{j=1}^{n} \Delta b_j x'(\widetilde{w}(t_i))B_j(t_i) \right]^2 . \tag{4.6}$$

From this expression we obtain the correction of the warping coefficients Δb_j by the regression of the residuals $y_i - x(\widetilde{w}(t_i))$ on the set of basis functions $x'(\widetilde{w}(t_i))B_j(t_i)$.

However, it has been demonstrated that if the number of B-splines is too high, the warping function will induce more variations than the acceptable ones based on the original profile. To restrict the flexibility of the warping function, the STW method introduces a penalty term λR. The main idea of this penalty term is to control the roughness of the warped profile by also minimizing the difference between adjacent values $x(w(t_{i-1}))$ and $x(w(t_i))$. With the inclusion of this penalty term the function to minimize becomes

$$Q = S - \lambda R = S - \lambda \sum_{i \in H} [x(w(t_i)) - x(w(t_{i-1}))]^2 \tag{4.7}$$

Eilers et. al. [111] demonstrate the feasibility of defining the penalty term based on the kth-order finite differences of the coefficients of adjacent B-splines, $\Delta^k b_j$. These finite differences are defined as $\Delta b_j = b_j - b_{j-1}$, $\Delta^2 b_j = \Delta(\Delta b_j) = b_j - 2b_{j-1} + b_{j-2},...$ $\Delta^k b_j = \Delta(\Delta^{k-1} b_j)$. The penalty term proposed by Eilers has the form

$$\lambda R = \lambda \sum_{j=k+1}^{n} (\Delta^k b_j)^2 + \kappa. \tag{4.8}$$

The parameter λ allows to control the smoothness of the fit. If the parameter λ is big the warping function will approaches a parametric model. Eilers includes an extra parameter, κ, to control the flexibility in the correction of shifts between B-splines. Increases in the κ parameter reduces this flexibility. By substituting equation 4.8 in equation 4.7 we get the function used in the STW method to perform the warping.

$$Q = \sum_{i \in H} \left[y_i - x(\widetilde{w}(t_i)) - \sum_{j=1}^{n} \Delta b_j x'(\widetilde{w}(t_i))B_j(t_i) \right]^2 + \lambda \sum_{j=k+1}^{n} (\Delta^k b_j)^2 + \kappa \tag{4.9}$$

In practice, the warping coefficients and the penalty parameter are optimized iteratively. First, the warping coefficient b_j is optimized using a high $\lambda (= 10^{10})$, starting from $\Delta b_j = 0$, then the parameter λ is decreased by a factor of 10 as suggested by Nederkassel et. al. [110] and the warping coefficients b_j are re-optimized. This process continues until the value of equation 4.9 is lower than a predefined threshold.

When the STW algorithm is used to correct the spectra obtained from NMR, GC–MS or LC–MS it is necessary to apply a baseline removal and to smooth the data in order to get an adequate warping function. In our case we will not perform a baseline removal because we would loose important information related to the QCLs emission. We only apply a very specific smoothing method, the details are presented in the implementation section 4.3.2.

4.3.2. Implementation

The STW algorithm allows us to correct the mismatch between wavenumber scans without the need of selecting specific points as was done in Section 4.2.2. In this sense it is more powerful than the scan correlation using selected wavenumbers described in Section 4.2, because it uses all the information collected during the data acquisition. In

order to get a scan profile that we can use to produce an absorbance spectrum we need to follow the steps listed below:

1. First, we need to smooth the signals of the main and normalization detectors.

2. Then, we remove the stabilization regions at the beginning and end of the signals.

3. We proceed to find the warping function $w(t_i)$ for the normalization signals.

4. Finally, we apply the warping function to the main signal.

Smoothing and correction of the signals

To apply the STW method we first need to smooth the signals obtained with the main and normalization detectors. The raw signals have rapid variations in the mode-hop zones that limit the applicability of the STW method. Using the local maxima of the signal we build a profile envelope to smooth the signals. We opt to use this envelope because it contains intensities which were emitted by the laser. This way we preserve the experimental information. By selecting the envelopes, we are directly choosing a specific mode among all the modes which are competing in the mode-hop zones. In this sense, we assume that the mode with the highest intensity is the one emitting the wavenumbers of the mode-hop zone. We can not simply omit the mode-hop zones because they represent a considerable part of the full scan.

The bioinformatics toolbox of MATLAB has a function that is ideal to build the envelope profile of our signals. It is called *msbackadj* and it constructs the smooth profile of a signal using local extrema. The only thing to keep in mind is that the *msbackadj* function uses the local minima of the signal instead of the local maxima to get the profile. This is not a problem because all the information contained in our signals has positive values, therefore we can transform all the local maxima into local minima by multiplying the whole signal by -1. We also use the *msbackadj* function because the resulting smooth profile has the same number of data points as the raw signal. An important consequence is the preservation of the one to one relation of the normalization and main signals when the smoothing is applied to each signal independently.

The *msbackadj* function first estimates the smooth profile inside small signal windows of the same size. The windows are constructed by shifting their beginning and end points in such way that there is always an overlap between adjacent windows. For our data the windows consist of 5000 points to smooth the sharp variations from the mode-hop zones and the shift consists of 200 points. After this, the *msbackadj* function uses a spline approximation to regress the estimated profile to the local minima inside the windows.

To finally get the envelope of the signals we multiply the smooth profile obtained with the *msbackadj* function by -1 once more. The intensity of some points in the envelope is slightly different from their intensity in the original signal. This mainly happens in the mode-hop zones because of the fast variations in these zones. However the intensity differences are smaller than the fluctuations in the mode zones and therefore we can neglect them. In Figure 4.20 we have plotted the intensity profile acquired with the normalization signal and its corresponding envelope.

Once we have removed the fast variations in the mode-hop zones we get rid of the stabilization zones at the beginning and end of the QCL emission described in Section
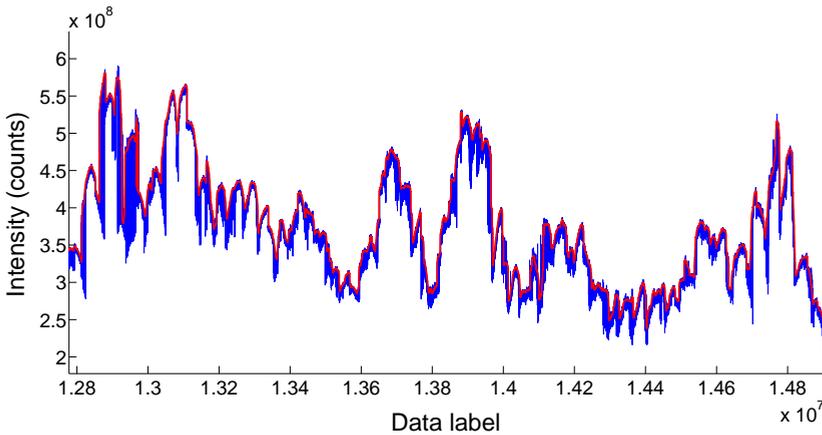
Figure 4.20: Intensity profile of a normalization signal and its envelope.

4.1.3. We do this because the stabilization zones have different lengths for each wavenumber scan which greatly affect the performance of the STW method. The knowledge of the envelope signal simplifies this task because it shows the periodic pattern of the stabilization regions and also uncovers the presence of specific features that the laser emission presents when it starts and finishes a scan. We use these features as reference points. By removing the same amount of points in the envelope of the normalization and main signal we respect the relation between their points.

APPLICATION OF THE STW METHOD

With the envelope we can proceed to apply the STW algorithm. We use the envelope of one of the normalization signals directly as the reference profile and adjust the rest of the signals using the STW algorithm. In fact we apply the algorithm five times using a different starting value of the penalty parameter λ: $10^{10}, 10^8, 10^6, 10^4$ and 10^2. Each time we use the result of the previous warping and we correct the main signal using the intermediate warping functions. The penalty parameter κ is kept constant at 100.

Thanks to the one to one relation between the normalization and main signals we only need to apply the STW method to the normalization signals. Once we get the right warping function $w(t_i)$ for the normalization signal we can directly apply it to correct the main signal. In Figure 4.21 ten different normalization signals and the corresponding main signals are shown. The upper subfigures correspond to the original data and the lower ones show the result after applying the STW correlation method. After the correction we see a good overlap between the different signals, this is specially clear in the normalization signals of subfigure 4.21c. For the main signals this method performs an excellent correction for the sharp peak in the left side of the intensity profile subfigure 4.21d). Specially taking into account that the warping signal used to correct the main signals was obtained from the normalization signals.

(a)   (b)

(c)   (d)
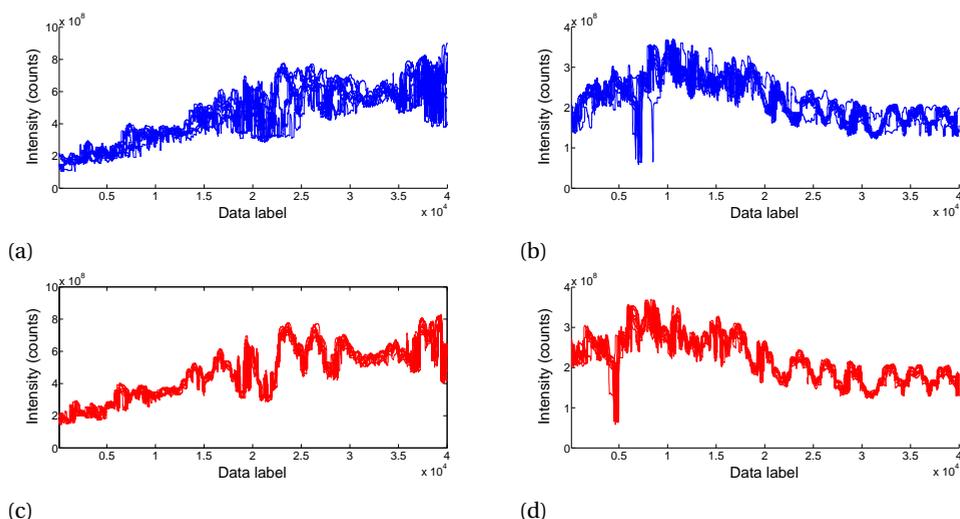
Figure 4.21: Correction using the STW method. (a) Uncorrected normalization signals. (b) Uncorrected main signals. (c) Corrected normalization signals. (d) Corrected main signals.

ABSORBANCE SPECTRUM BUILDING

We use the STW method to correct the signals from a reference measurement with the multipass cell empty and a breath sample measurement. Using the formalism discussed in Section 3.5 we obtain directly the absorbance spectrum. The resulting spectrum is presented in Figure 4.22. We only plot the result of the QCL emitting in the region between 1010 and 1250 cm^{-1} for clarity.
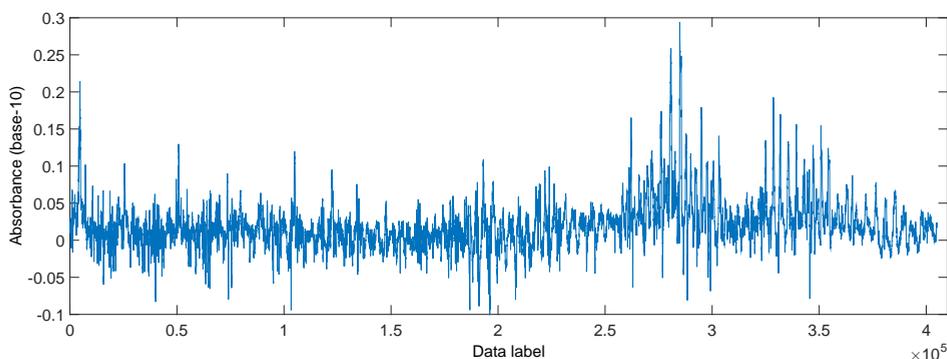
Figure 4.22: Absorbance spectrum of a breath sample in the wavenumber region between 1010 cm^{-1} and 1250 cm^{-1} obtained using the STW method. the horizontal axis is not calibrated to the corresponding wavenumbers.

From this spectrum we can clearly see an important feature missing in the result obtained from the STW method: the horizontal axis does not show the corresponding

wavenumbers. In the STW correlation method we never relate to the wavenumber emitted by the QCL. The wavenumber calibration has to be done using known absorption features in the absorbance spectrum to compare them to the spectra available in the molecular databases. This represents the main drawback of the STW method presented. The details of the wavenumber calibration are given in detail in Section 4.4.

4.4. Wavenumber calibration

The absorbance spectra obtained from the two scan correlation methods show the absorption features of the gas sample under study with different degrees of details. But both methods lack an essential characteristic to be useful in the determination of the gas composition and the concentrations: the resulting spectra are not calibrated against a global wavenumber reference. Now, we proceed to explain how to calibrate the specta using the absorption signature of known molecules present in the sample and the information contained in the molecular databases.

4.4.1. Wavenumber calibration for the scan correlation using selected wavenumbers

In the case of the scan correlation using selected wavenumbers method, the assigned wavenumbers come from the calibration provided by the LaserScope system. However, when we plot the absorbance spectrum of a breath sample and compare it with the CO_2 spectrum from the HITRAN database, we observe a clear mismatch between them, as shown in the upper plot of Figure 4.23.
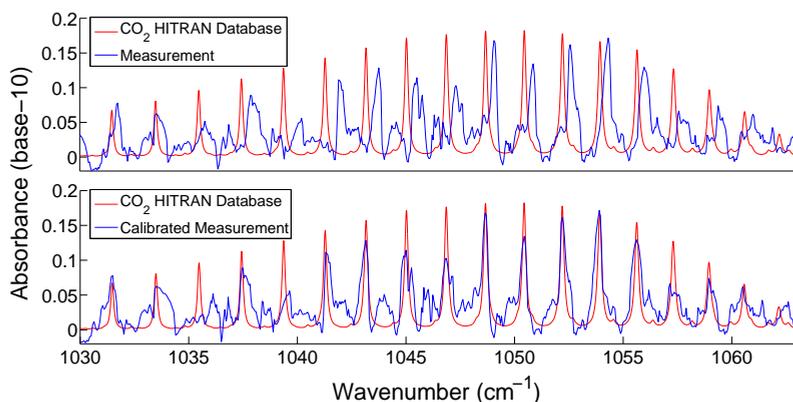
Figure 4.23: Absorbance spectrum obtained with the scan correlation using selected wavenumbers method. The upper plot shows a wavenumber mismatch when compared to the CO_2 absorbance of the HITRAN database. The lower plot shows the calibration applied to the data.

The mismatch is relatively small and we can easily correct the displacement using the maximum of the absorption peaks. In the upper plot of Figure 4.23 is clear that each peak in the measurement is close to its corresponding peak in the spectrum database, therefore we simply displace the peaks in the measurement to the correct wavenum-

ber and apply a linear interpolation between them. In the lower plot of Figure 4.23 we present the result of this process.

4.4.2. WAVENUMBER CALIBRATION FOR THE SCAN CORRELATION USING STW
In the case of the scan correlation using the STW method there is not even an attempt to calibrate the wavenumber axis. This method only aims at correcting the mismatch between the signals. It does not pretend to make any wavenumber calibration and is purely based on the intensity profile of the signals. To perform a successful wavenumber calibration we first need to make a rough calibration of the wavenumbers. In this first step we can use the most prominent features of the measured spectrum, such as the sharp absorption lines of water or the envelope of the P and R branches of the rovibrational spectrum of CO_2. In the following paragraphs we described in detail the calibration applied to the wavenumber region between 1030 cm^{-1} and 1090 cm^{-1}, the method described can be applied to any other region given the presence of prominent absorption features corresponding to well known molecules.

In Figure 4.24 we present the rough alignment applied using the envelope of CO_2. For this alignment we first use the MATLAB function *msbackadj* to define an envelope of the measurement, which is presented in the upper plot of Figure 4.24. Then we compare this envelope with the envelope of the spectrum of CO_2 provided in the HITRAN database, shown in the lower plot of Figure 4.24. From this comparison we conclude that the central minimum of the envelope of the measurement corresponds to the minimum between the P and R branches of CO_2 at 1064 cm^{-1}.
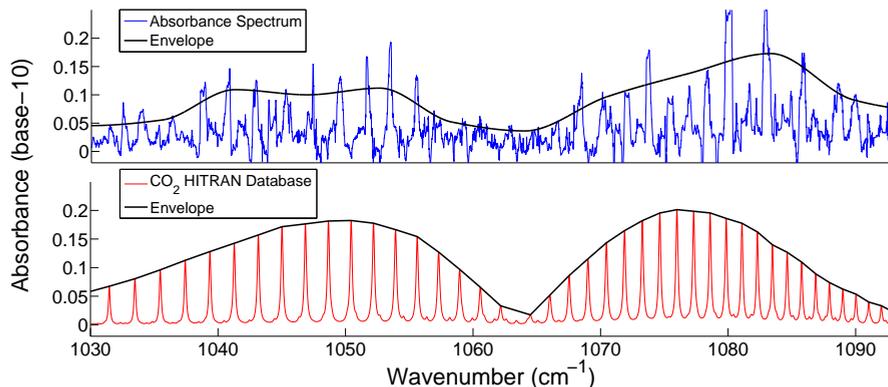
Figure 4.24: Rough calibration of the absorbance spectrum obtained with the scan correlation using the STW method. The upper plot shows the data measurement and its envelope. The lower plot shows the CO_2 absorbance of the HITRAN database and its envelope.

To further improve the wavenumber calibration we use the sharp absorption lines of the measured spectrum. In this case the mismatch between the measurement and the database information is not as small as in the case of the spectrum obtained from the scan correlation using selected wavenumbers. Therefore we cannot just move the peaks in the measurement to the closest peak in the database. Furthermore, the mea-

surement spectrum has more data points than the spectrum obtained from the selected wavenumbers method, which makes it more difficult to identify the peaks corresponding to the CO_2 absorption.

We start to enhance the identification of the CO_2 absorption peaks. Since we only want to highlight the CO_2 absorption peaks we transform any negative absorbance value to zero because their origin lies in the noise of our spectroscopic system. Then, we square all the absorbance values. This way we get a clear identification of the absorption peaks and keep a direct relation with the original data. Then we identify three of the highest peaks of the squared spectrum and we look for the closest peaks in the CO_2 spectrum database. We use three peaks in case one of the peaks is missing. To finish the calibration we look for the right wavenumber for the rest of the peaks of the squared spectrum. For this last step we take into account the equidistant position between adjacent peaks of the rovibrational spectrum of CO_2, this helps us to avoid any over-correction induced by the absence of an absorption peak in the squared spectrum. In Figure 4.25 we show the rough wavenumber calibration obtained using the envelope of the absorbance spectrum and the result of the calibration improvement done using the squared spectrum.
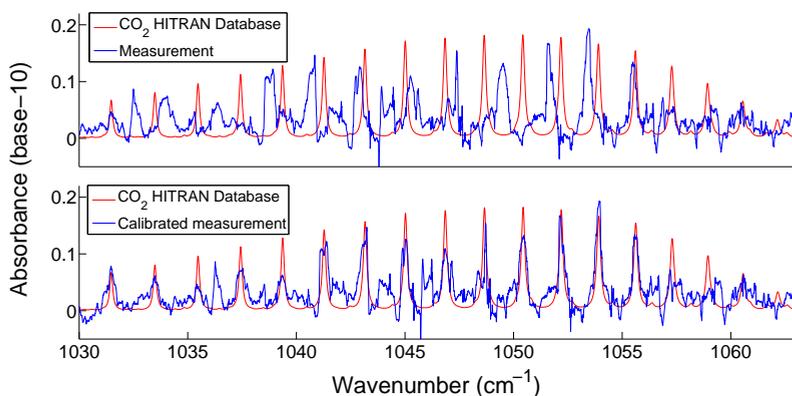
Figure 4.25: Absorbance spectrum obtained with the scan correlation using the STW method. The upper plot shows a wavenumber mismatch when compared to the CO_2 absorbance of the HITRAN database. The lower plot shows the calibration applied to the data.

4.4.3. COMPARISON OF RESULTS

The wavenumber calibration is essentially the same for both scan correlation methods. They need a first rough calibration which in the case of the scan correlation using selected wavenumbers is directly done during the correlation, while for the scan correlation using the STW method we need to analyze specific features to obtain the first calibration. Only then we can apply a more refined calibration using the absorption features of known molecules present in the measured sample. In Figure 4.26 we show the result of the wavenumber calibration for both correlation methods. To get these calibrated absorbance spectra we used the same data set.
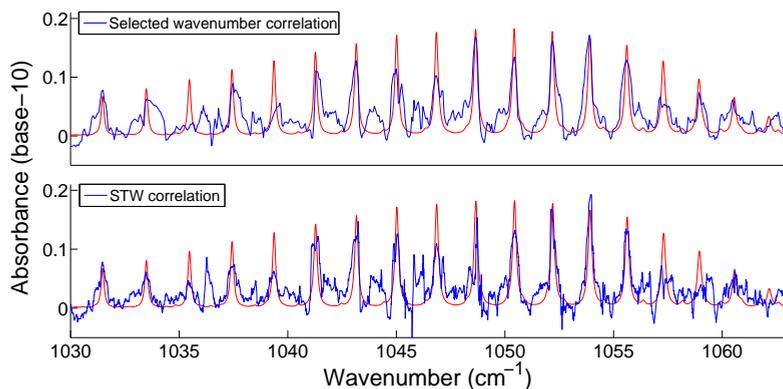
Figure 4.26: Wavenumber calibration of the absorbance spectrum obtained with the scan correlation using selected wavenumbers (upper subfigure) and the STW method (lower subfigure). The red plots show the CO_2 absorbance from the HITRAN database.

From Figure 4.26 it is clear that the STW method preserves more information because it uses more data points and it does not apply any smoothing of the original data. One of the advantages of this is a more reliable determination of the height of sharp peaks as can be observed for the CO_2 peak measured around 1054 cm^{-1}. A disadvantage of preserving all the information is the difficulty in differentiating between actual absorption signatures of molecules present in low concentration and noise.

The wavenumber correlation using the selected wavenumbers method is faster because it uses less data and requires less computational steps. This is particularly useful for applications where we only use the sharp details of the spectrum to calibrate the wavenumber and we want to determine the concentration of molecules with broad smooth absorption profiles, like acetone. In Chapter 5 we will present a good example of this implementation.

In both correlation methods the results show that the wavenumber adjustment applied is only adequate when there are clear sharp absorption lines in the measured spectrum, otherwise it is difficult to perform an accurate calibration. Also, the accuracy of the calibration depends on the number of absorption lines present in the calibrated region. For instance, the calibration is more accurate in the region where the CO_2 lines are present than in the region where the water lines lie, because CO_2 has more lines and they are more densely packed. However, by combining as many lines from different molecules present in the measurement is possible to obtain a good wavenumber calibration.

In the end a good wavenumber calibration is necessary to accurately determine the concentration of the molecular components of a gas sample. When the molecules under investigation have sharp absorption features it is better to use the STW method. If we are interested in the absorption of molecules with broad smooth absorption profiles we can save time using the wavenumber correlation using selected wavenumber method.

4.5. MULTILINE FITTING AND CONCENTRATION DETERMINATION

Once we have calibrated the wavenumber axis of a measured absorbance spectrum, $A_{meas}(\widetilde{\nu})$, we can proceed to determine the concentration level of the different molecules present in the measured gas sample. For gas samples containing a single molecular species with a broad smooth absorption profile and with a low noise level, we can accurately determine the molecular concentration, C, performing a non-linear least square multiline fitting. This type of fitting takes into account all the contribution of the different wavenumbers of the measurement, $\widetilde{\nu}$ [113]. For this analysis it is necessary to model the absorbance, $A(\widetilde{\nu})$, of the desired molecule. From the Beer-Lambert law we know the absorbance is the product of the concentration, the molar absorption coefficient, $\epsilon(\widetilde{\nu})$, and the interaction distance inside the multipass cell, l.

$$A(\widetilde{\nu}) = \epsilon(\widetilde{\nu})Cl. \qquad \text{(2.29 revisited)}$$

The molar absorption coefficient is available in the different molecular databases. We just need to use the molar absorption coefficient corresponding to the same pressure and temperature conditions of the experiment. In Section 3.3 we determined the interaction distance to be $l = 54.36$ m. With this information we look for the concentration value that minimizes the function

$$F(C) = \sum_{i=1}^{k} [A_{meas}(\widetilde{\nu}_i) - C\epsilon(\widetilde{\nu}_i)l]^2. \qquad (4.10)$$

Here k is the number of points in the absorbance spectrum. This method is highly reliable even if the noise level is considerable. In Figure 4.27 we show the result of using the multiline fitting to a gas sample with 2.6 ppm$_v$ of acetone in a matrix of Nitrogen. We used 2100 points to get the fitting, which shows a very good agreement with the experimental data.
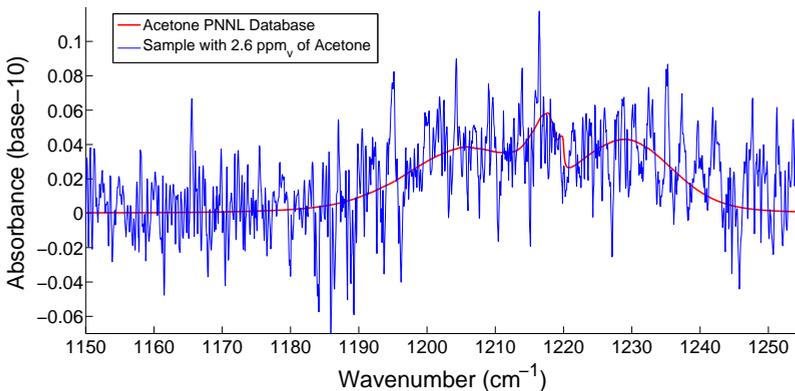
Figure 4.27: Absorbance spectrum of a gas sample with 2.6 ppm$_v$ of acetone.

This approach can be extended to n molecules taking into account that for diluted mixtures, like breath, the total absorbance is the sum of the individual absorbances $A_1(\widetilde{\nu}_i), A_2(\widetilde{\nu}_i), ..., A_n(\widetilde{\nu}_i)$ of each of the molecular species present in the sample. Therefore $A_{total}(\widetilde{\nu}_i) = \sum_{j=1}^{n} A_j(\widetilde{\nu}_i) = \sum_{j=1}^{n} C_j \epsilon_j(\widetilde{\nu}_i) l$. Then for each molecule we need to find the concentration C_j that minimizes the generalized function

$$F(C_1, ..., C_n) = \sum_{i=1}^{k} \left[A(\widetilde{\nu}_i) - l \sum_{j=1}^{n} C_j \epsilon_j(\widetilde{\nu}_i) \right]^2 \tag{4.11}$$

In Figure 4.28 we present a breath sample with 1.6 ppm$_v$ of acetone and 4% water in a matrix of Nitrogen. To obtain this figure we use the correlation using the selected wavenumbers method. This is the main reason that the water peaks do not match correctly with the peaks from the HITRAN database. Another possible reason is the presence of unknown molecules in low concentrations. However the acetone spectrum matches quite well the baseline of the spectrum, giving a reliable measurement of its concentration.
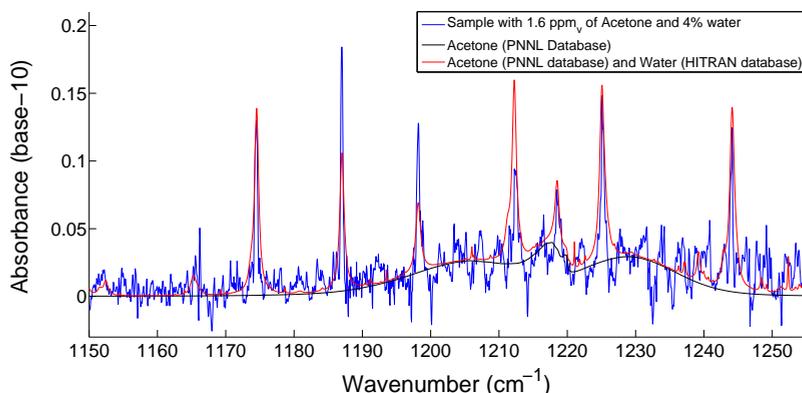
Figure 4.28: Absorbance spectrum of a breath sample with 1.6 ppm$_v$ of acetone and 4% water.

4.6. CONCLUSIONS

In this chapter we explain in detail the data acquisition and analysis system implemented in our experimental setup. We justify the different decisions taken in order to get reliable measurements. For this we collected the information of two independent MCT detectors to monitor the intensity behavior of our two QCLs. The use of two MCT detectors forced us to implement a more elaborate acquisition system than the one provided by the original LaserScope unit. Our acquisition system allows us to record more experimental information by collecting information from each one of the pulses emitted by the two QCLs. By measuring the peak intensity of the pulses we were able to reduce the burden on the acquisition system and to reduce the size of the saved data files to around 40 MB. The acquisition system can be further improved by coupling the frequency ac-

quisition of the MCT detectors to the peak of the pulses. This reduced the number of data points recorded by the NI acquisition card. Another improvement can be done by optimizing the writing process of the data files, at this moment we write directly on txt files, but we can use a different type of files to speed up the saving process.

Another important point discussed in this chapter is the verification and correction of the data. We do not only collect the information from the pulses but verify that we collect all the relevant information. At this point we also clean up the data produced by errors in the determination of the peaks and mark any missing data from the pulses. With this information we are able to create the intensity profile of the normalization and main signals.

From the intensity profiles we obtain important information about the operation of the QCLs. First of all, we identify the presence of different stabilization regions of a wavenumber scan. We also observe the presence of mode zones and mode-hop zones in the intensity profile. But without doubt the most important discovery lies in the fact that independent wavenumber scans are uncorrelated and warped. For this reason we explore two methods to correlate the wavenumber scans: the scan correlation using selected wavenumbers and the scan correlation using STW.

Each correlation method has its advantages and drawbacks. The scan correlation using selected wavenumbers is faster and ideal to determine the concentration of molecules with a broad smooth absorption profile. But it misses the small details of the profiles because it only uses part of the information, which makes it unsuitable to establish the concentration of molecules whose absorption profile consist of sharp lines. The scan correlation using STW preserves most of the experimental data and reproduces more details of the absorption spectra of the molecules. For this reason this second method can be used to determine the concentration of molecules with all type of absorption profiles. The presence of more details in the absorption spectrum by itself is a challenge because it is more difficult to differentiate between the signature of molecules with sharp absorption features in low concentration and the noise of the signals. Additionally this correlation method is slower because of the different computational steps.

4

5

EXHALED ACETONE IN TYPE 1 DIABETES

For people could close their eyes to greatness, to horrors,
to beauty, and their ears to melodies or deceiving words.
But they couldn't escape scent.
For scent was a brother of breath.
(...) He who ruled scent ruled the hearts of men.

Patrick Süskind, Perfume. The Story of a Murderer

Quantum cascade laser (QCL) spectroscopy has a very promising future in the study of diabetes and diabetic ketoacidosis. In this Chapter we study the acetone concentration exhaled in the breath of three Type 1 Diabetes patients (two minors and one adult) and one healthy volunteer using the QCL spectroscopic system developed in our lab. Using the acetone signature between 1150 and 1250 cm^{-1} and a multiline fitting method the concentration variations in the order of parts per billion by volume (ppm_v) were measured. Blood glucose and ketone concentrations in blood measurements were performed simultaneously to study their relation with acetone in exhaled breath. We focus in personalized studies to better understand the role of acetone in diabetes. For each volunteer we performed a series of measurements over a period of time, including overnight fastings of 11 ± 1 hour and during ketosis - hyperglycemia events for the minors. Our results highlight the importance of performing personalized studies because the response of the minors to the presence of ketosis was consistent but unique for each individual. Also, our results emphasize the need of performing more studies with Type 1 Diabetes (T1D) minors, since the acetone concentration in the breath of the minors differs, with respect to those reported in literature, which are based on adults.

5.1. Introduction

Diabetes is one of the major challenges for modern society because of its increasing occurrence. In 2013 the International Diabetes Federation estimated that 382 million people had diabetes worldwide and 46% were undiagnosed [115]. Furthermore, the Diabetes Programme of the World Health Organization projects that by 2030 diabetes will be the 7th leading cause of death [116]. In view of the rapid progression of the diabetes epidemic, the scientific and medical communities around the world are making great efforts to understand the causes and find more efficient ways to detect and control the disease. Although blood-based diagnosis and monitoring methods are robust and reliable, they are uncomfortable and invasive; especially for diabetic patients who have to monitor their glucose and ketone levels constantly. For this reason the search for alternative methods is a very active field. On one hand, research efforts aim to develop cheap, fast, reliable and accurate diagnosis and monitoring techniques which are also comfortable and non-invasive [117, 118]. On the other hand, a more through study is necessary to fully understand the mechanisms involved in diabetes. Our particular research falls in this area, we focus in the development of personalized studies and the use of spectroscopy to identify the relevant molecules for which specific sensors should be developed.

One of the most promising solutions to achieve the requirements of modern diagnosis methods is the analysis of exhaled human breath. It is based on the identification of molecular biomarkers present in breath whose production is linked to specific metabolic processes. The ability to detect variations in the biomarkers' concentrations can be used to identify changes in the metabolism of a person, providing valuable information of its health status. A good example is the exacerbation of nitric oxide production as a result of airway inflammation, which is currently a standard method to diagnose and monitor asthmatic patients [19]. For most of the biomarkers the concentration changes are in the order of parts per billion by volume (ppb_v) and parts per trillion by volume (ppt_v), which are now detectable thanks to the latest advances in molecular spectroscopy. The

data analysis and the available databases have also improved the accuracy of molecular identification in complex gas mixtures. Furthermore, there are robust and portable prototypes that show the possibility of miniaturization in the coming decades [119]. Due to its potential, the development of exhaled breath analysis is reaching technological and commercial importance [120].

In particular, acetone is a very attractive molecule to study diabetes because its volatility makes it easy to detect in breath. Also its relation with improperly managed insulin levels is well known and the metabolic pathways that produce acetone are well-studied [121]. Even though studies of acetone concentrations in exhaled breath started in the 1950's using Gas Chromatography and Mass Spectroscopy [122], the stringent requirements made these methods impractical for many years. In the last 20 years the field has revived thanks to the improvements in the processes of Gas Chromatography and Mass Spectrometry, the development of new chemical detection methods [119, 123, 124] the introduction of the e-nose [125, 126] and the development of suitable laser spectroscopy techniques [127, 128]. To date there have been less than 50 independent studies analyzing the acetone concentrations in exhaled breath in different conditions and with different methods [121]. So far, these studies have narrowed down the acetone concentrations in exhaled breath for healthy people to a range between 0.39 and 1.09 ppm$_v$ [121]. Broader standardized studies are necessary to define a more robust acetone concentration range in exhaled breath for healthy and diabetic individuals and to be able to use the acetone concentrations in exhaled breath for diagnostic purposes.

In the specific case of acetone Type 1 Diabetes (T1D) there are only 6 studies that explicitly report acetone concentration measurements in exhaled breath, to the best of our knowledge [129–134]. Despite the small number of studies performed important conclusions have been drawn. All the studies have measured an elevated mean acetone concentration in the exhaled breath of T1D patients. These studies have mainly concentrated on adults, but T1D is more common in children under 15 years making it highly relevant to investigate the acetone concentration for this age group. By 2013 the global estimations indicated that 497,100 children have T1D worldwide [115]. Moreover, it is demonstrated that elevated levels of acetone in exhaled breath are associated with the occurrence of ketosis [135, 136]. When ketosis manifests in combination with hyperglycemia [137] and acidosis [138] the presence of diabetic ketoacidosis is evident. Diabetic ketoacidosis is a serious life-threating complication in which toxic molecules are produced and accumulated in the body. Diabetic ketoacidosis is a leading cause of morbidity and mortality in children with T1D [115]. According to the International Diabetes Federation only in 2013 more than 79,000 children developed T1D and up to 80% of them already had diabetic ketoacidosis when they were diagnosed with diabetes [115]. Hence, it is important to develop fast non-invasive diagnostic tools for ketosis and diabetic ketoacidosis. Such diagnostic appliances and techniques can be included in the standard general medical examinations and provide an early warning when chronic ketosis is detected, which in turn will help to detect diabetes at an early stage.

A promising tool to diagnose ketosis and diabetic ketoacidosis is the detection of elevated acetone concentrations in exhaled breath using direct quantum cascade laser (QCL) spectroscopy [81, 128]. Employing mid-infrared QCLs, fixed multipass cells and Mercury Cadmium Telluride (MCT) mid-infrared detectors it is possible to detect ace-

5

tone concentrations in the ppb_v level. The identification of acetone in breath is possible thanks to the recently developed QCLs emitting in the absorption spectral region between 1150 and 1250 cm^{-1}, where acetone has a strong absorption signature. State-of-the-art fixed multipass cells provide a robust and reliable tool to detect sub-ppm_v concentrations. Moreover, since the QCLs and the MCT detectors operate at room temperature the measurements can be performed in flexible environmental conditions. Our approach is very versatile since the exhaled breath samples do not require any preprocessing or preparation. Furthermore, the relatively small size of the components makes QCL spectroscopy a strong candidate for future portable devices that can be used to complement well-establish blood test and eventually can be sent to regions where blood test cannot be performed easily due to lack of laboratories, sanitary conditions or cultural views. Combined with other techniques, such as the e-nose, it can eventually lead to the development of accurate, non-invasive hand-held devices to monitor diabetic ketoacidosis and diabetes at home.

5.2. MATERIALS AND METHODS

In this research we measured the blood glucose levels and the acetone concentrations in exhaled breath of one healthy adult volunteer as control and 3 volunteers with T1D, one adult and two 14 year old minors. In all cases the volunteers and/or their parents granted their informed consent and we followed a strict confidentiality protocol. All the T1D volunteers have been diagnosed with diabetes for at least one year and they keep a strict control of their blood glucose level. Both T1D minors wear insulin dispensers while the T1D adult administers the insulin manually. To confirm the health status of the control volunteer a high-performance liquid chromatography measurement was performed by a certified clinical laboratory, the results gave a 5.3 % level of glycated hemoglobin A1c, confirming the absence of diabetes [139].

The measurements were taken in two different conditions: after an overnight fasting and during four events in which the T1D minors suffered hyperglycemia and ketosis simultaneously. The overnight fasting measurements were used to establish the reference blood glucose level and acetone concentration in exhaled breath of all the volunteers. The measurements were taken during school/working days when the volunteers follow a fixed routine. In this routine the overnight fasting period is 11 ± 1 hours. The hyperglycemia - ketosis events were caused by a malfunction of the insulin dispensers. These events were identified because the T1D minors showed mild symptoms of ketosis [140]. In all cases the blood glucose levels were measured using a Glucose meter (Bayer Contour Link) and a breath sample from a single exhalation was collected in a 1 liter sample bag (SKC FlexFoil PLUS) to measure the acetone concentration in exhaled breath. The breath sample was transported to the laboratory and measured within 6 hours. During the hyperglycemia - ketosis episodes the ketone levels were also measured using a ketone-meter (Medi Sense Precision Xceed). In these episodes the T1D minors received insulin to restore the blood glucose level and acetone concentration in exhaled breath immediately after the measurements and breath sample were collected.

The acetone concentrations in exhaled breath were determined using the QCL spectroscopy setup schematically shown in Figure 5.1. The setup has a Noise Equivalent Absorbance Sensitivity of 2.99×10^{-7} cm^{-1} $Hz^{-1/2}$ and its capability to detect sub-ppm_v

concentrations of acetone has been demonstrated [81]. The details of the setup are available in Chapter 3 and Reference 81. The setup incorporates a QCL source operated at room temperature (LaserScope, Block Engineering Inc., USA), a home-built multipass cell and two thermo-electric cooled Mercury Cadmium Telluride (MCT) detectors (LaserScope, Block Engineering Inc. and PVI-4TE-10.6-0.3x0.3-TO8-BaF2, VIGO System S. A., Poland). A set of mass flow controllers (F-201CV-1K0-RAD-22-V, Bronkhorst High-Tech B.V., The Netherlands), a pressure controller (P-702CV-1K1A-RAD-22-V, Bronkhorst High-Tech B.V., The Netherlands) and a diaphragm vacuum pump (MD 1, Vacuumbrand GMBH + CO KG, Germany) are used to control the injection and evacuation of the gas samples from the multipass cell.
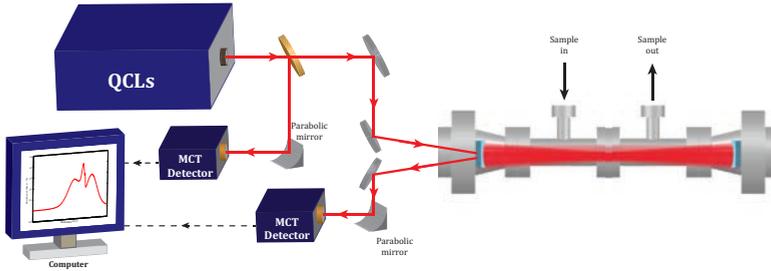
Figure 5.1: Schematic of the experimental setup. Red lines correspond to the laser beam path, solid black arrows show the sample flow and the dashed lines indicate electrical connections.

The QCL scans the wavenumber ($\widetilde{\nu}_i$) region between 850 and 1250 cm^{-1} with a wavenumber resolution of 0.1 cm^{-1}. Here the subscript i accounts for the discreet nature of the scan. The multipass cell allows to obtain an effective interaction distance of $l = 54.36$ meters between the gas samples and the laser in a volume of only 0.6 liters. The signal acquired by the MCT detectors is processed using the scan correlation using selected wavenumbers method described in Section 4.2 to obtain a normalized transmission spectrum of the multipass cell. For our measurements we first record a reference transmission spectrum when the multipass cell is empty, $I_0(\widetilde{\nu}_i)$, then we fill the cell with a gas sample until the pressure inside the cell reaches 900 mbar, immediately after we measure the new transmission spectrum of the cell, $I(\widetilde{\nu}_i)$. With these two spectra we obtained the absorbance $A(\widetilde{\nu}_i)$ of the gas samples using the Beer-Lambert law discussed in Section 2.2, equation 2.27 repeated below [79].

$$A(\widetilde{\nu}_i) = -\log_{10}\left(\frac{I(\widetilde{\nu}_i)}{I_0(\widetilde{\nu}_i)}\right) \qquad \text{(2.27 revisited)}$$

For this study we focused in the bandwidth between 1150 and 1250 cm^{-1} where acetone has a broad absorption signature. Since water presents strong absorption features in this region [81] we used the information available in the High Resolution Transmission (HITRAN) database [72] to model the absorbance of water, $A_w(\widetilde{\nu}_i)$, for a concentration of 1% at the same experimental conditions of temperature, pressure and interaction distance. We also use the sharp absorption water lines in this region to calibrate the wavenumber of our system. We compare the position of the water peaks in our measured

absorbance, $A(\widetilde{v}_i)$, with those in the absorbance of water, $A_w(\widetilde{v}_i)$. Then we correct the wavenumber position by linearly interpolating the points between adjacent water peaks. The calibrated absorbance contains k discrete wavenumbers $\widetilde{v}_i$ between 1150 and 1250 cm^{-1} separated by 0.1 cm^{-1}.

After the calibration we apply the nonlinear least-squares multiline fitting of Section 4.5 to estimate the water and acetone concentrations C_w and C_a [113]. Using the molar absorption coefficient of acetone, $\epsilon_a(\widetilde{v}_i)$, provided by the Pacific Northwest National Laboratory (PNNL) database [80], the molar absorption coefficient of water, $\epsilon_w(\widetilde{v}_i)$ from the HITRAN database and an interaction distance of $l = 54.36$ m we minimize the function

$$F(C_w, C_a) = \sum_{i=1}^{k} [A(\widetilde{v}_i) - l(C_w \epsilon_w(\widetilde{v}_i) + C_a \epsilon_a(\widetilde{v}_i)]^2. \tag{5.1}$$

We use an initial water concentration of 1% and an initial acetone concentration of 1 ppm$_v$.

Using the estimated C_w and C_a concentrations we fit a Voigt profile to each water feature and remove them from the spectrum. Then we apply again a multiline fitting with one single variable C_a to get a better determination of the acetone concentration [113]. From this last multiline fitting we obtained an acetone concentration detection limit of 0.05 ppm$_v$ (1σ), which confirms the suitability of our setup to measure the acetone concentrations in exhaled breath levels. To report the standard deviation we use the 3σ value. Figure 5.2 shows the acetone multiline fitting of a measurement with a concentration of 0.46 ppm$_v$. The residuals of the fitting are plotted in the lower plot and show a background without structure, this demonstrates that the wavenumber calibration and the removal of the water peaks are adequate for this data analysis [113].
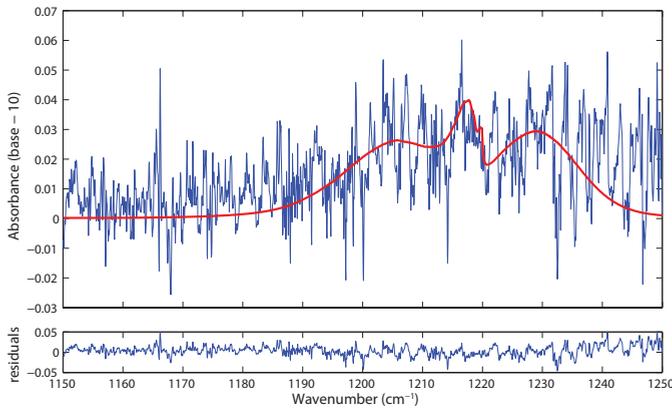
Figure 5.2: Absorbance of acetone with 0.46 ppm$_v$ concentration after the water lines were removed. The uncertainty is associated to the acetone concentration detection limit of 0.05 ppm$_v$ (1σ). The red line shows the absorbance profile provided by the PNNL database, its strength was found using the multiline fitting routine. The absence of structure in the residuals corroborate that the water removal and the wavenumber calibration are performed correctly.

After each measurement the multipass cell is cleaned by flushing zero-air once (Purity grade 5.5, Linde, The Netherlands) and nitrogen twice (Purity grade 6, Linde, The Netherlands) . The protocol to clean the sample bags includes flushing them with zero-air once and nitrogen twice, then they are heated overnight at 60 degrees Celsius and finally they are flushed again twice, first with zero-air and then with nitrogen. We have observed that these cleaning procedures eliminate the acetone residuals below the sensitivity detection of our system.

To calibrate the acetone concentrations measured with our setup we used calibrated gas mixtures provided by the Dutch metrology Institute (VSL). We used 4 different concentrations of 0.34, 0.68, 1.54 and 3 ppm_v in a matrix of Nitrogen. For each concentration we performed 10 measurements. The results are shown in Figure 5.3, where the circles correspond to individual measurements, the continuous line shows the linear fit of the calibration and the dotted lines indicate the 95% confidence interval.
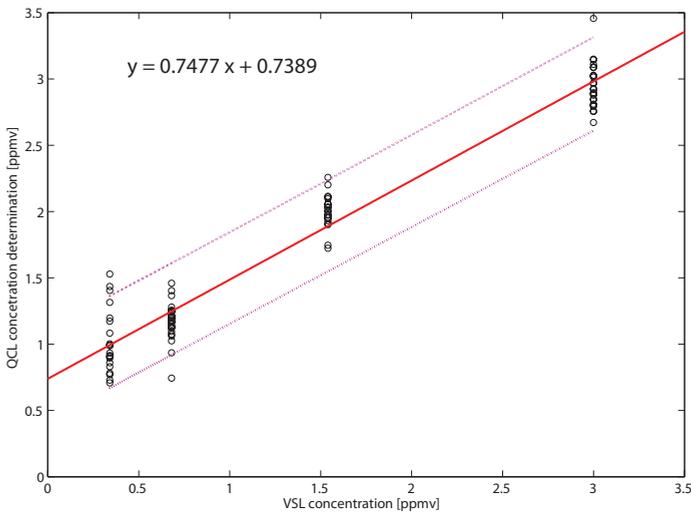
Figure 5.3: Calibration curve of the QCL spectrometer using acetone mixtures provided by the Dutch Metrology Institute. The equation corresponds to the linear fit of the calibration.

5.3. Results and Discussion

5.3.1. Overnight fasting measurements

The results of the overnight fasting measurements for each volunteer are presented in Figure 5.4. We use the same axis range for all the volunteers to show how the measurement distribution is different for each individual. The dashed vertical lines in the plots indicate the overnight fasting blood glucose goal range recommended by the American Diabetes Association for adolescents between 13 and 19 years old with T1D and diabetic adults, this range goes from 4.99 to 8.32 mmol/l (90 - 150 mg/dL) [139].
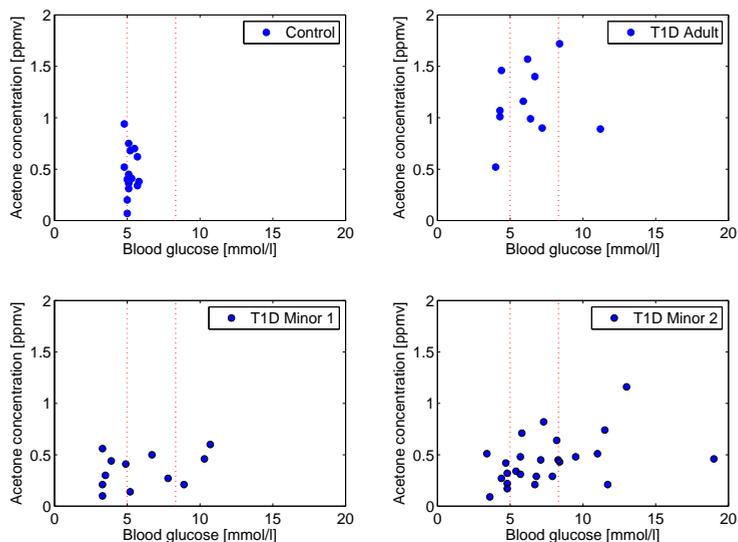
Figure 5.4: Overnight fasting measurements of the volunteers. Each dot corresponds to an independent measurement. The dashed vertical lines indicate the overnight fasting blood glucose goal range recommended by the American Diabetes Association. [139]

From the distributions we clearly observe how the blood glucose of the control agglutinates around the mean value of 5.2 mmol/l while for all the T1D volunteers the blood glucose values spread over the full range. In the case of the T1D Minor 2 there is even a measurement with a blood glucose level of 19 mmol/l, which is more than twice the upper limit of the goal range; although in that specific case the volunteer did not present symptoms of ketosis [140] we decided to measured the ketone concentration and did not detect any ketones. In Table 5.1 we summarized the mean value and the standard deviation of the blood glucose and acetone concentrations for each volunteer.

Table 5.1: Overnight fasting measurements

	# Measurements	**Blood Glucose** [mmol/l]		**Acetone concentration** [ppm_v]	
		Average	std	Average	std
Control	17	5.2	0.3	0.47	0.21
T1D Adult	12	6.4	2.1	1.17	0.34
T1D Minor 1	14	6.0	2.8	0.35	0.17
T1D Minor 2	28	7.6	3.5	0.44	0.24

In Figure 5.5 we have plotted the mean value and the standard deviation of the acetone concentration for each volunteer. The region between the dashed horizontal lines corresponds to the acetone concentrations in exhaled breath range of healthy people,

between 0.39 and 1.09 ppm$_v$ [121]. From this figure we can clearly see that the acetone concentrations of the T1D adult are in the upper level of the healthy range, which corroborate that adult T1D patients tend to have higher acetone concentrations in exhaled breath [129–134]. Regarding the acetone concentration from the T1D Minors, the values are in the lower range of the healthy range. This is a new finding because, as we mentioned in the introduction, previous works have focused on adults. To the best of our knowledge only the study of Reference 129 explicitly reports acetone measurements in two T1D Minors. Looking closely at their data and ours we observe that the acetone concentration of the T1D Minors are lower than the levels measured in T1D adults, this suggest that the acetone levels of T1D minors may differ considerably from those measured in T1D adults.
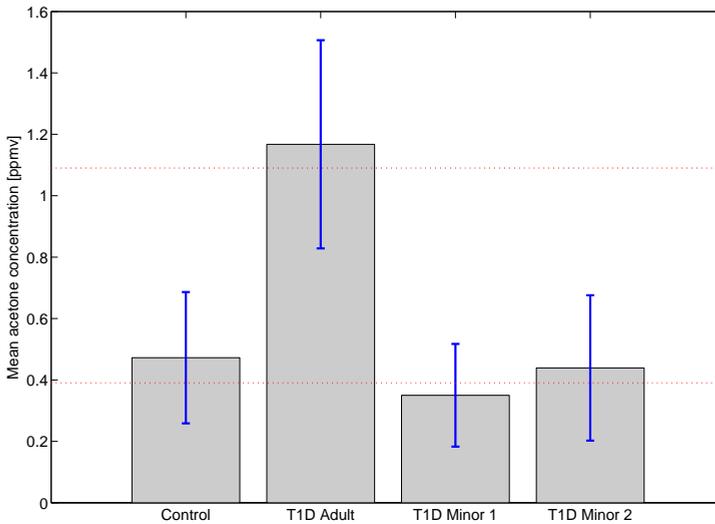
Figure 5.5: Mean value and standard deviation of the acetone concentration of the volunteers. The dashed horizontal lines show the range of the acetone concentrations in exhaled breath for healthy people from 0.39 to 1.09 ppm$_v$

5.3.2. HYPERGLYCEMIA - KETOSIS MEASUREMENTS

The results of the measurements of the hyperglycemia - ketosis events are summarized in Table 5.2. In all the events the blood glucose, the ketone levels in blood and the exhaled acetone in breath were measured as soon as the events were detected. In the particular case of event E4 from the T1D Minor 2 we had the opportunity to take a second measurement one hour after the occurrence of the event and after insulin had been administered to the volunteer, at this point the ketosis symptoms had disappeared and the ketone level shows that ketosis was under control.

Table 5.2: Hyperglycemia - ketosis measurements

Volunteer	Event	Blood Glucose [mmol/l]	Ketone level	Acetone concentration [ppm$_V$]
T1D Minor 1	E1	24.7	0.1	0.24
	E2	21.1	0.1	0.14
T1D Minor 2	E3	29.0	1.2	1.47
	E4	22.2	1.0	2.55
		29.7*	0.1*	2.81*

*Measurements taken one hour after the event happened.

Between the two T1D Minor volunteers we observe a clear difference in the ketone levels when they experienced a hyperglycemia - ketosis events. The Minor 1 has a lower ketone concentration in blood than Minor 2, but the glucose levels in all cases are above 20.0 mmol/l, which is more than three times the average level obtained from the overnight fasting measurements. Since ketosis in T1D patients appears when the body cannot process glucose due to the lack of insulin [115], this suggest that a threshold in the blood glucose level exists around 20.0 mmol/l where the hyperglycemia - ketosis events start. In Figure 5.6 are plotted the overnight fasting measurements and the measurements of the hyperglycemia - ketosis events for each T1D Minor. The overnight fasting results are enclosed in green and the red ovals encloses the information of the hyperglycemia - ketosis events. In the case of T1D Minor 1 the only observable difference is the increase in the blood glucose, showing that for hyperglycemia - ketosis events where the ketone levels remain low the acetone levels in breath does not change considerably. This is consistent with the observation that acetone diffusion from plasma to the lungs strongly depends on the acetone concentration present in blood [141].
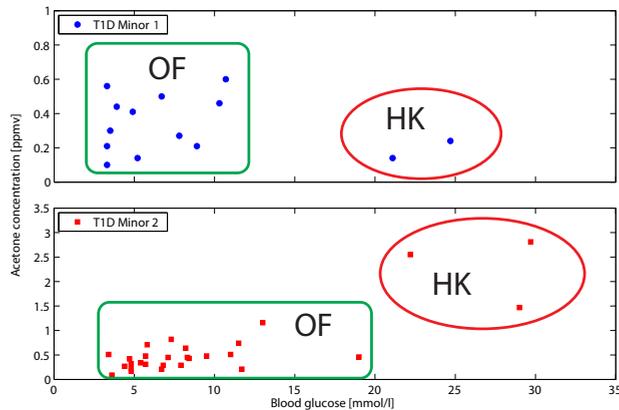
Figure 5.6: OF: overnight fasting. HK: hyperglycemia - ketosis events

For the T1D Minor 2 there is a clear difference between the overnight fasting condition and the hyperglycemia - ketosis events. The blood glucose levels and acetone concentrations in exhaled breath are definitely higher in the presence of hyperglycemia - ketosis. In particular for this volunteer the acetone concentrations in exhaled breath is more than twice the average acetone concentrations in exhaled breath obtained from the overnight fasting measurements. An interesting observation for this volunteer is that although in the overnight fasting measurements the blood glucose level spreads over a wide range it does not overlap with the blood glucose levels at which ketosis occurs, in the same way the acetone concentrations in exhaled breath levels are clearly separated. This is different to the behavior observed for Minor 1 for whom only the glucose level are well separated, highlighting the need of performing more personalized studies in the future.

For the case of the event E4 we observed that in the second measurement when the ketosis had been controlled the blood glucose and acetone concentrations in exhaled breath levels were higher than in the first measurement, see Table 5.2. For the acetone concentrations it confirms that the elimination of acetone through breath is a slow process which is responsible for the presence of acetone in breath long after ketosis has been controlled [142]. This observation shows the suitability of our system to study the elimination of acetone through the lungs without the need of any sample pre-processing [141], which can help to better understand the body reaction when insuline is administered to control a ketosis event.

5.4. CONCLUSIONS

Our research corroborates that stable T1D adult patients tend to have acetone concentrations in exhaled breath above the range for healthy individuals [121] but it also suggests that this is not necessarily true for T1D minors. It becomes highly relevant to perform more studies focused on children under 15 years old because the majority of T1D patients are in this age range. Our measurements in the presence of hyperglycemia and ketosis illustrate the importance of performing personalized studies because each T1D Minor showed a different response during the events. Also, the variability we found in the blood glucose levels and the acetone concentrations in exhaled breath during the overnight fasting measurements support the importance of performing more personalized studies or a through clinical study with several participants. By monitoring each patient over a period of several days we can gain important information about the human metabolism during the overnight fasting, which in turn can help to improve the monitoring and control of ketosis and diabetes in T1D patients.

One of the main limitations in detecting ketosis using exhaled breath is the slow diffusion of acetone from the blood to the lungs, making the detection of acetone in breath only possible when the concentrations in blood are elevated. This can be use to detect chronic ketosis where the levels of acetone in blood remain relatively high during a prolonged period of time as demonstrated by diverse studies [135, 136]. In this case we would expect a closer relation between the levels of acetone in exhaled breath and the acetone concentration in blood. In this case our method would be ideal given its non-invasiveness but further studies and measurements are necessary to fully validate this hypothesis.

Our approach has potential as a routine test in the general medical examinations to detect ketosis in children because it is non-invasive, it is very robust and it does not require any sample preparation. Such test can be performed more often than the blood tests to diagnose ketosis and diabetes and could serve as a reference to perform more complex medical examinations when ketosis is detected. This can help to diagnose T1D earlier and to decrease the morbidity and mortality of children caused by diabetes ketoacidosis. Particularly in low income countries this approach will help to reduce the diagnosis costs and improve the quality of life of T1D patients.

Our present device scans the full range between 850 and 1250 cm^{-1}, but for acetone a narrower scan between 1150 and 1250 cm^{-1} is enough. Therefore, it is possible to develop a faster and cheaper device able to detect sub-ppm$_v$ concentrations of acetone. Furthermore, the sensitivity can be improved by increasing the interaction length inside the multipass cell and using one single detector to produced the normalized transmission spectrum [81]. The future of diagnostic methods based on acetone concentrations in exhaled breath is highly promising. By combining our setup with recently develop methods to exclusively trap acetone [124] we can increase the sensitivity of our system and boost the performance of laser-based devices for breath analysis.

5

6

BREATH ANALYSIS IN PEDIATRICS

Adonis REYES-REYES

Esther VAN MASTRIGT

*But magic is really only the utilization
of the entire spectrum of the senses.*

Michael Scott, The Alchemyst

The identification and measurement of specific endogenous biomarkers is a promising tool for diagnosis and monitoring in the breath research field. This chapter is dedicated to explore the applicability of our quantum cascade laser (QCL) spectroscopic system in the study of respiratory diseases. The main purpose is to identify statistical differences between the absorption profile of the exhaled breath of three groups: children with asthma, children with cystic fibrosis (CF) and healthy children. By studying the distribution of the absorption values for each wavenumber we found spectroscopic regions where the differences between the different groups are significant. Using these spectroscopic regions we were able to list a series of suspect molecules which in future studies should be investigated in depth to determine their influence in asthma and CF. Finally, we also performed a pilot study to determine the standard concentration of acetone and methane in the breath of healthy children.

6.1. INTRODUCTION

For many years the search of endogenous biomarkers associated to specific diseases has been a very active field of research in medicine. An important idea to keep in mind is that within the scope of our research, biomarkers are molecules that are produced by the body in normal circumstances; only when a substantial change happens in the metabolism the concentration of the molecules is modified significantly [144, 145].

In the last decades, the measurement of volatile organic compounds (VOCs) as biomarkers in exhaled breath has raised great interest as a potential tool for diagnosis and management of respiratory diseases. Especially in pediatric respiratory medicine, exhaled breath profiling is attractive to assess the presence and degree of airway inflammation, to distinguish disease phenotype and to enable targeted and individualized treatment. Hydrocarbons are potential biomarkers for airway inflammation, as they derive from lipid peroxidation. This can be induced by e.g. reactive oxygen species during inflammatory reactions in the airways. Previous studies on VOC profiling in children with asthma and CF have identified hydrocarbons as the most discriminating VOCs between patients with and without asthma or CF. Hydrocarbons were helpful in the diagnosis of asthma and CF, and predicted clinical deterioration [146–150]. However, there is not yet a clear knowledge of the standard range of concentrations of the different VOCs in the exhaled breath of healthy individuals. Only for a very small group of VOCS a standardization has been performed. This information is even scarcer or non- existent when dealing with children.

Currently, gas chromatography coupled to mass spectrometry (GC-MS) is the gold standard for the identification of VOCs. Disadvantages of this technique include the need for sample pre-concentration, the fragmentation of molecules and the necessity for expensive equipment and highly trained personnel. An alternative is the electronic nose (eNose), an easy and fast technique which allows online real-time analysis of the complete spectrum of VOCs, but it cannot identify separate molecular components [151]. In this work we used a QCL spectroscopic technique to detect VOC profiles in exhaled breath [81]. Our QCL spectroscopic system allows us to study the absorption spectra of exhaled breath in the wavenumber range between 850 and 1250 cm^{-1}. This range is important because it covers a considerable span of the fingerprint absorption spectra of many hydrocarbons. Moreover, our technique enables the detection and identification

of whole molecules within minutes, and it has the potential for miniaturization, thereby combining the advantages of the GC-MS and eNose methods. In the present study, we explored the clinical applicability of QCL spectroscopy by determining the repeatability and comparing exhaled breath samples from healthy children with samples from children with either stable asthma or stable CF lung disease. These two conditions have different types of chronic airway inflammation as a prominent feature. Furthermore, we also study the concentrations of acetone and methane in the exhaled breath of healthy children to determine the standard concentrations.

6.2. BREATH PROFILES OF HEALTHY CHILDREN AND CHILDREN WITH ASTHMA AND CF

For this research we study the exhaled breath of 89 minors with ages ranging between 6 and 18 years old. They comprise 3 different groups: a control group with 35 healthy children, an asthmatic group with 39 children with stable asthma and a CF group conformed of 15 children with stable CF.

The control group was recruited from a primary school in Rotterdam, The Netherlands. The asthma and CF groups were recruited at the outpatient clinic KinderHaven of the Harbour Hospital in Rotterdam and the Sophia Children's Hospital of the Erasmus Medical Center of Rotterdam, between September 2013 and February 2014. In all cases we obtained written informed consent from the parents or caretakers and from children of 12 years and older. This study was approved by the Medical Ethical Committees of the Erasmus Medical Center - Sophia Children's Hospital and of the Harbour Hospital of Rotterdam.

To include the volunteers in the control group they had to fulfill the following criteria: they should have a normal forced expiratory volume in 1 sec (FEV1) [152] and a negative result from the questionnaire for asthma and allergy designed by the International Study of Asthma and Allergies in Childhood [153].

In the case of the asthmatic group, the individuals had a medical diagnosis of asthma and they should had a well controlled status at the time of the study according to the criteria established by the Global Initiative for Asthma [154]. A well controlled status is defined when the patient uses a stable dose of inhaled corticosteroids, needs short acting ß2 agonists less than 3 times per week, has minimal daytime symptoms (< 3 times per week), does not have any limitations of activity, has not had any exacerbation in the previous 3 months and has a normal lung function. A normal long function is defined when the FEV1 is higher than 80% of the FEV1 predicted by the equations published by the Global Lung Initiative of 2012 [155] or when the FEV1 of the volunteer is higher than 80% of its personal best. The genetic tendency to develop allergic diseases (atopy) of the individuals in the asthmatic group was determined when the radioallergosorbent test (RAST) was positive or when the subject had a positive skin-prick test for at least one aeroallergen [156].

The criteria to include the CF patients in the study are: they had been previously diagnosed with CF with a sweat test or two CF specific DNA mutations and they were not treated for exacerbations with oral or intravenous antibiotics in the 6 weeks previous to the study.

6

6.2.1. STUDY DESIGN

We collected two breath samples with a difference of 30 minutes for all the groups. In the case of the control group we also investigate the short and long term reproducibility by collecting two extra breath samples, after 24 hours and 1 week from the time the first samples were taken.

For all groups we performed spirometry measurements (Masterscreen, CareFusion, Germany) to determine the FEV1 level of the subjects the same day the first breath samples were collected [152]. For the control and asthmatic group we also measured the fractional exhaled nitric oxide score (FeNO) of the volunteers [19] during the first day of breath sample collection (NIOX MINO NO analyzer, Aerocrine, Sweden). FeNO is a biomarker of eosinophilic airway inflammation. Both, the spirometry and the FeNO measurements were conducted according to the guidelines established by the American Thoracic Society (ATS) and the European Respiratory Society (ERS) [19, 152].

Table 6.1 summarizes the clinical characteristics of the three groups.

Table 6.1: Clinical characteristics of the subjects. The data is presented as the median (with interquartile range). Abbreviations: M = male. F = female, NA = not applicable.

	Control	Asthma	CF
Subjects (n)	35	39	15
Age [years]	9.9 (9.0-10.7)	10.9 (8.7-14.0)	11.0 (7.7-14.5)
Gender (M:F)	15:20	26:13	4:11
Height [cm]	141.5 (133.0-149.0)	150.0 (133.4-160.5) n=38	145.2 (129.3-164.9)
Weight [kg]	33.0 (27.5-37.0)	42.9 (35.8-52.8) n=38	36.4 (27.4-59.1)
FEV1 [% predicted]*	94.6 (83.0-106.0) n=32	90.2 (84.7-99.8) n=33	95.7 (83.9-105.6)
FeNO	15.0 (11.0-18.0) n=33	25.0 (12.0-39.0) n=35	NA
Atopy (yes:no)	NA	31:8	NA
Asthma control test score	NA	22.0 (18.5-23.0) n=29	NA

* % predicted according to the Global Lung Initiative equations of 2012 [155]   .

The exhaled breath samples were collected during a single, slow exhalation from total lung capacity after 5 seconds of breath holding. The children exhaled via a disposable mouthpiece and one way valve into a 3 liter re-usable Tedlar bag (231-03, SKC Inc., USA). Eating and exercise were not allowed within one hour before sampling. The bags were transported the same day to the laboratory where the samples were analyzed immediately.

The cleaning protocol of the Tedlar bags is based on the recommendation method 422 of the Environmental Protection Agency of USA from 1991 [157]. First the bags are flushed once with synthetic air with a purity grade of 5.5 (Linde, The Netherlands), then they are flushed three times with nitrogen with a purity grade of 6 (Linde, The Netherlands). Afterward they are heated overnight at 60 °C and finally they are flushed once

with synthetic air and once with nitrogen.

6.2.2. DETAILS OF THE SPECTROSCOPIC MEASUREMENTS

The exhaled breath samples were analyzed using the broadband quantum cascade laser spectroscopy system described in Chapter 3 and Reference 81. The setup consists of a LaserScope unit (Block Engineering Inc., Massachusetts, USA), a home-built multipass cell of 0.6 liters and two infrared mercury cadmium telluride (MCT) detectors (Laser-Scope, Block Engineering Inc. and PVI-4TE-10.6-0.3x0.3-TO8-BaF2, VIGO System S. A., Poland) as shown in Figure 6.1.
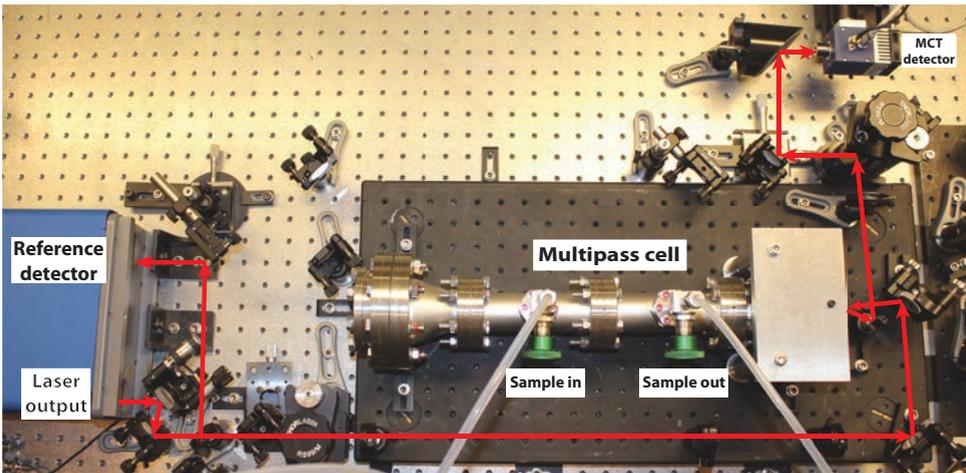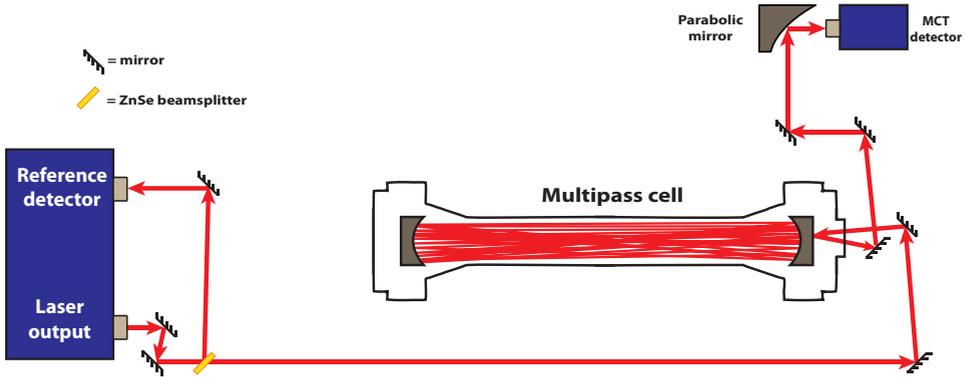
Figure 6.1: Spectroscopic setup.

The LaserScope unit contains a quantum cascade laser which scans the wavenumber region between 850 and 1250 cm^{-1}. The laser is directed to the multipass cell in which two astigmatic mirrors (Aerodyne Research Inc., USA) ensured multiple reflections, confining the laser to travel for 54.36 meters inside the volume of the multipass cell. When the laser exits the multipass cell its intensity profile is measured using a MCT detector.

To measure the absorption profile of the breath samples we first record the laser signal with the multipass cell empty, I_0. For this we reduce the pressure inside the cell to less than 0.1 mbar using a pressure controller (P-702CV-1K1A-RAD-22-V, Bronkhorst High-Tech BV, The Netherlands) and a diaphragm vacuum pump (MD 1, Vacuumbrand GMBH + CO KG, Germany). Subsequently, a breath sample is injected in the multipass cell using a set of mass flow controllers (F-201CV-1K0-RAD-22-V, Bronkhorst High-Tech BV, The Netherlands). This way we ensure a continuous flow of 500 milliliters per minute normal (ml_n/min) until the internal pressure of the multipass cell reaches 900 mbar. At this point we reduce the flow to 50 ml_n/min and let the internal pressure to stabilize before taking the gas measurement, I. After each measurement we clean the multipass cell and the tubing by flushing them once with synthetic air and twice with nitrogen. In each flushing step we empty the cell to less than 10 mbar. In all these measurements the breath sample passes through a Nafion dryer (MD-070-24S-4, Perma Pure LLC, USA) before entering the multipass cell to reduce the water content.

We use the data acquisition discussed in Chapter 4 to record the reference and gas measurements. We use the scan correlation by point selection method described in Section 4.2 to correlate the 10 different scans of the measurements because it is the fastest of the correlation methods developed in this thesis and because it was the most mature at the time of this study. The absorption profiles of the breath samples were obtained using directly the Beer-Lambert law in the form of equation 2.27, which we reproduce below for clarity.

$$A(v) = -\log_{10}\left(\frac{I(\widetilde{v})}{I_0(\widetilde{v})}\right). \qquad \text{(2.27 revisited)}$$

here the dependence on the wavenumber, $\widetilde{v}$, is explicit after the scan correlation. The wavenumber resolution of the final absorption spectra is 0.05 cm^{-1}.

6.2.3. DATA ANALYSIS

The data analysis aims to study the repeatability of the absorption profile of the breath samples and to find significant differences between the study groups. First we prepared the data to improve our analysis. We observed that for each wavenumber point the distribution of the absorbance values of the samples is approximately Gaussian with similar median and variances. To increase the sensitivity of any subsequent statistical analysis we applied a shifted modified Z-score [158–160] to the absorbance distribution of each wavenumber point. To calculate the Z-score we subtracted the median of the absorbance distribution from the distribution itself and divided the result by the median absolute deviation (MAD) [160]. Then we shifted the result by adding the global median of all the sample spectra. All the statistical calculations performed in this data analysis were done using R version 3.1.2.

REPEATABILITY OF EXHALED BREATH PROFILES

To study the repeatability we used Spearman's correlation coefficient, ρ [161], to compare the 1st and 2nd exhaled breath samples in each patient group and to compare the 1st with the 2nd, 3rd and 4th exhaled breath sample in the control group.

We found a poor overall correlation in all the comparisons as can be seen in Table 6.2 with a maximum value of 0.46 for the control group.

Table 6.2: Repeatability between different measurements of each study group. Abbreviations: n = number of subjects used for the calculation, NA = not applicable.

	Spearman's correlation (ρ)		
	1st vs 2nd measurement	1st vs 3rd measurement	1st vs 4th measurement
Control group	0.46 (n=35)	0.39 (n=34)	0.36 (n=34)
Asthma group	0.43 (n=31)	NA	NA
CF group	0.43 (n=15)	NA	NA

The agreement between the 1st and 2nd exhaled breath samples of each group was assessed using a Bland and Altman plot [162, 163]. The results showed a moderate agreement when individual wavenumber points were assessed as demonstrated in Figure 6.2. In this figure we show the result obtained from the 1st and 2nd exhaled breath sample profiles of the 35 healthy children. Each dot represents a single wavenumber point. The vertical axis depicts the difference between the profiles and the horizontal axis corresponds to the average of the 1st and the 2nd exhaled breath profiles. In this specific case the mean difference of the profiles has a value of -0.03 and the 95% limits of agreement have the values 0.60 and -0.66.
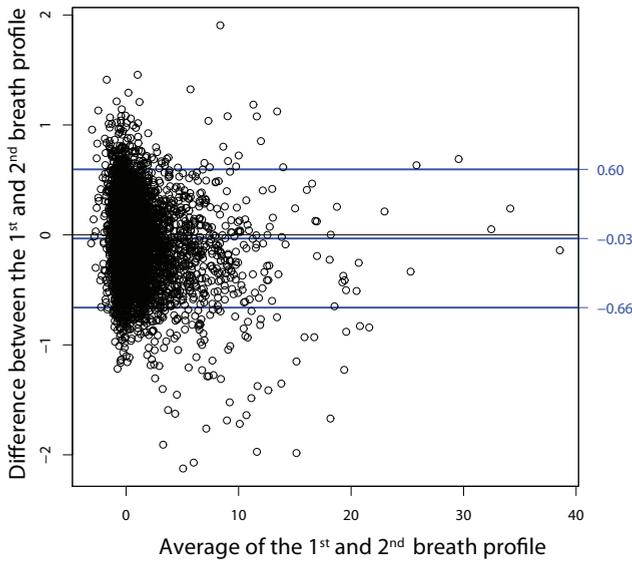
Figure 6.2: Bland and Altman plot between the 1st and 2nd exhaled breath sample profile of the control group.

6.2.4. RELEVANT DIFFERENCES BETWEEN STUDY GROUPS

One of the objectives of this study is to identify possible difference between the study groups. For this reason we performed two statistical analysis: a principal component analysis (PCA) to look for differences using the complete spectral information of the breath samples and the analysis of the calculated probability, p-value, to identify differences between the study groups for each wavenumber point.

PRINCIPAL COMPONENT ANALYSIS

PCA is a method for exploratory data analysis. It transforms a set of observations of possibly correlated variables into a set of linearly uncorrelated variables, which are known as the principal components. The result of the transformation is an orthogonal basis in the directions of the largest variances. The first principal component (PC1) corresponds to the highest possible variance and each succeeding principal component (PC2, PC3, ...) corresponds to the next highest variance. [164]

In our study the PCA was performed on the mean of the 1st and 2nd measurement of each volunteer. In Figure 6.3 we present the scatterplots depicting the separation between the members of the control group and the members of the asthmatic and CF groups using two principal components.
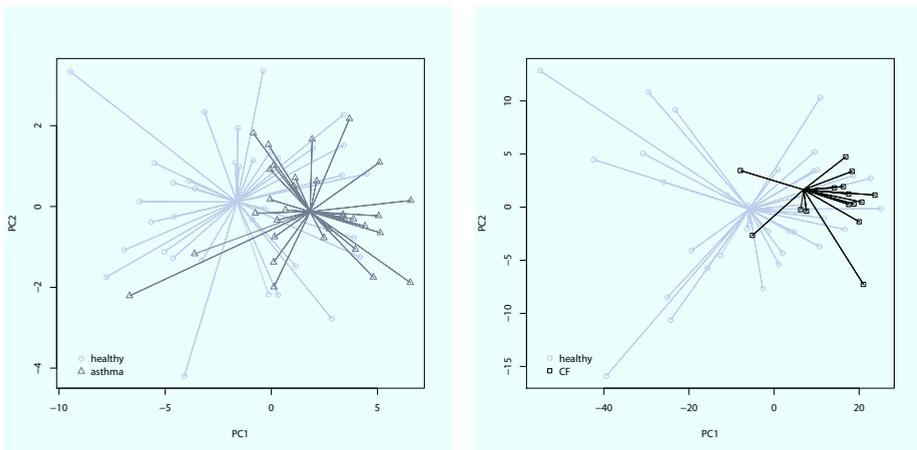
Figure 6.3: Principal component analysis. The dots correspond to the values of the 1st and 2nd PCAs and the intersection of the lines shows the average of the PCAs. The subfigure on the left shows the results of comparing the control and asthmatic group. The subfigure on the right corresponds to the comparison between the control and the CF groups.

The PCA analysis uses all the information of each absorption spectrum as a complete set. For this reason we do not observe a clear and definite separation between the different study groups in Figure 6.3.

P-VALUES ANALYSIS

A different approach is to look for significant differences among the study groups for each independent wavenumber point. This is done by estimating the p-value for each wavenumber [161].

A p-value is defined as the probability to obtain a specific result under the assumption that a specific hypothesis is fulfilled, known as the null hypothesis [161, 165]. In our particular case the null hypothesis is that the absorption profile of the patients (asthma or CF groups) is the same as the absorption profile of the healthy volunteers. Therefore, for each wavenumber the distribution of the absorbance values of the control group and the corresponding to the patient groups should be similar. By fitting a normal distribution to each wavenumber of the patient and control groups we can calculate their corresponding means and standard deviations. If the null hypothesis is true, then the mean and standard deviation of the patient group should be exactly the same as the ones of the control group. If they are different, then we calculate the probability that the absorbance values of the patient group get the values they actually have, assuming that those values follow the distribution of the control group, which is simply the null hypothesis.

When the p-value is lower than a chosen threshold, it indicates that the patient group is significantly different with respect to the control group. However this criteria alone is not enough to define if there is a significant difference at the corresponding wavenumbers because we are looking for differences in the absorption profiles caused by molecules whose absorption fingerprints cover several consecutive wavenumbers. For this reason we look for wavenumbers regions where at least 10 consecutive points have a p-value of 0.05 or lower, which we will called significantly different regions. This choice of p-value follows the traditional threshold level of 5% [166]. By choosing a spectroscopic region of 0.5 cm^{-1} we increase the probability that the significant difference are the result of changes in the concentration of one or more molecular species. To calculate the p-values we employed the BioConductor package limma (version 3.22.1) [167]. Briefly, limma uses linear models combined with methods for variance stabilization and post-hoc multiple correction

In Figure 6.4 we present the p-values obtained from the comparison between the asthma and the control groups. We can observe the presence of many wavenumbers where the p-value is lower than the threshold. But in most of the cases they are single points or very narrow wavenumber regions. In this case we found two significantly different regions in the 1181.80 – 1182.55 cm^{-1} and 1261.40 – 1262.05 cm^{-1} sections.
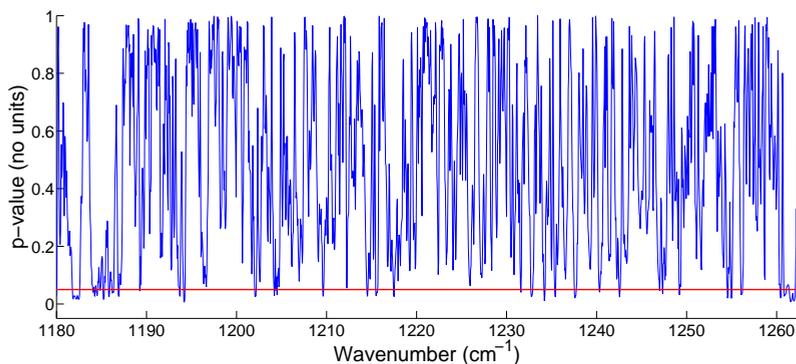
Figure 6.4: P-values obtained by comparing the control group and the asthma group. The red line shows the 5% threshold level.

The p-values corresponding to the comparison between the CF and control groups are shown in Figure 6.5. In this case we only found one significantly different region in the 1260.70 – 1261.65 cm^{-1} range.
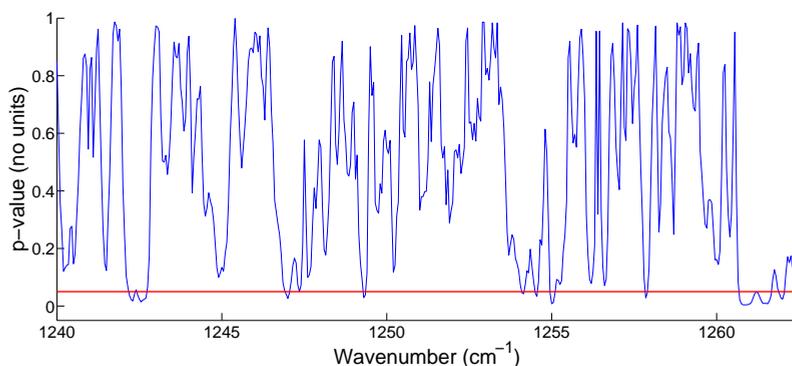
Figure 6.5: P-values obtained by comparing the control group and the CF group. The red line shows the 5% threshold level.

A closer look to the p-values of the CF and control groups reveals a wavenumber range, 1242.20 - 1242.65 cm^{-1}, that should be considered as a significantly different region because only the p-value at 1242.40 cm^{-1} has a slightly higher p-value of 0.056 as shown in Figure 6.6. Since we are looking for significant difference caused by molecules with absorption fingerprints over wavenumber regions rather than single wavenumber points, we will also use this region to look for molecules responsible for the differences.
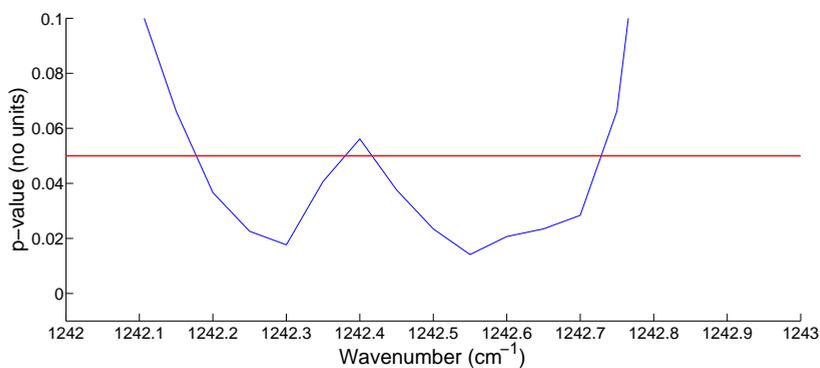
Figure 6.6: P-values from the control and CF group comparison in the 1242.20 - 1242.65 cm^{-1} wavenumber. The red line shows the 5% threshold level.

In Section 6.2.5 we will compare the results obtained using a p-value threshold of 5.6% with the strict case where only regions with p-values smaller than 5% are taken into account.

6.2.5. Suspect molecules

Once we know the significantly different regions we can search in the Pacific Northwest National Laboratory (PNNL) database [80] the molecules with a relevant fingerprint absorption in such regions. But we most choose carefully the parameters that determine the strength of their absorption. From the acetone study reported in Section 5 and Reference 114 we know that our setup can detect molecular concentrations changes in the order of hundreds of ppb$_v$ levels and even tens of ppb$_v$. For this reason we chose to look for molecules that produce an absorption signal in the significantly different regions when their concentration is higher than 10 ppb$_v$ and when the interaction distance with the laser is the same as the 54.36 meters of the multipass cell.

Using these constraints we found the molecules included in the PNNL database that may be responsible for the difference between the control group and the asthma and CF groups. This study is only exploratory because the presence of the molecules was not confirmed with a independent technique. The full list of molecules is presented in Appendix D. In Table 6.3 we summarize the number of suspect molecules found using our methodology and the number of molecules that have been reported in previous works [168].

Table 6.3: Comparison of the number of molecules found using two different p-value thresholds.

No. of molecules	Healthy vs Asthma	Healthy vs CF 5% p-value threshold	5.6% p-value threshold
Suspected	82	42	58
Previously reported	23	12	14

For the specific case of the control vs CF group comparison we include the results obtained using a p-value threshold of 5% and those obtained with 5.6%. There is a difference of 16 suspected molecules in these results and among those molecules there are two previously reported molecules. Moreover, seven of the previously reported molecules have a signature in both wavenumber regions considered with the 5.6% threshold. This make them interesting candidates for further studies. The last analysis highlights the importance of understanding the origin of the significant differences. If we define our selection based exclusively on the traditional 5% threshold level, we would exclude a relevant region and missed some interesting molecules.

6.2.6. Discussion

Our results showed a poor repeatability of the complete exhaled breath profiles, both on the short term and on the long term. However, we also showed a relatively good agreement of individual wavenumber points in the first and second breath sample. A previous study, in which 3 breath samples were collected with a 2 minutes interval, also showed high coefficients of variation for VOC profiles when compared to individual molecules [169]. The limited repeatability might be explained by the fact that we investigated a large spectroscopic range comprising more than 8000 individual wavenumber points. We consider plausible that the high variability of exhaled breath content could be the result of multiple daily life factors, like eating, drinking, exercise and the variable composi-

tion of ambient air. Surprisingly, the long term repeatability in the healthy children was similar to the short term repeatability. This suggests that VOCs reflecting the pulmonary condition may be relatively stable over time, but they only comprise a small amount of the complete profile.

Previous studies have shown that exhaled breath VOC profiles measured with GC-MS and the eNose could discriminate healthy controls from children with asthma or CF with a relatively high sensitivity and specificity and with a limited intra-individual variability [147, 148, 170–172]. We are not able to discriminate healthy children from children with stable asthma or stable CF. The limited separation between the groups in the PCA suggests that the majority of the wavenumber points and spectra representing VOCs did not differ between healthy and diseased children. Previous studies were inconsistent, reporting different sets of best-discriminative VOCs for specific diagnoses, and external validation was often lacking [147–150, 170–173].

Our original aim was to detect and study hydrocarbons present in exhaled breath. Surprisingly, the suspect molecules found in this study include carboxylic acids,esters and ethers, which can derive from hydrocarbons. A recent review showed that these type molecules represent around 20% of all molecules found in exhaled breath samples [168]. When we focused on molecules which have been described before in exhaled breath, we noticed overlap between those in asthma and CF. This is not unexpected, as CF is characterized by chronic neutrophilic airway inflammation, and varying degrees of neutrophilic inflammation have been observed in childhood asthma as well [174].

The main limitations of our study and earlier breath profiling studies include lack of standardization of sample collection and analysis between studies, thereby hampering the overall comparability. For instance, the effect of expiratory flow remains unclear although current data do not suggest that flow control is needed [172, 175–177]. Another worry is contamination from external sources. We cannot rule out contamination of exhaled breath samples by disposable plastic components (mouthpieces, tubes and valves). As we used the same type of disposables in all children, this would not affect the differences between groups. In order to avoid contamination we cleaned the multipass cell after each measurement, and used the laser signal of the empty cell immediately after performing the cleaning procedure as a reference for the following measurement. Finally, our QCL spectroscopic setup has a sensitivity in the order of parts per billion by volume (ppb_v) for some molecules, which is less than the parts per trillion by volume (ppt_v) sensitivity of other laser-based techniques such as cavity ringdown spectroscopy [178]. To what extent higher sensitivity would result in better results remains to be shown.

6.3. Standard concentrations of acetone and methane in healthy children

One of the advantages of using a broadband QCL spectroscopic system is the amount of information collected from each single breath sample. By looking into specific regions of the recorded spectra we can obtain relevant information of the content of the samples. And we can also asses the performance of the system. In particular, an in-depth analysis allowed us to investigate the concentration of known molecules such as acetone and

methane. We studied the concentration of acetone and methane in the exhaled breath of healthy children using the absorption spectra of 11 individuals from the control group.

In Chapter 5 we pointed out the lack of information regarding the standard concentration of acetone in children. We obtained important information in this regard by looking into the specific fingerprint of acetone in the 1150 and 1250 cm^{-1} region.

To further explore the possibilities and limitations of our QCL spectroscopic setup we look into the spectral fingerprint of methane, between 1258 and 1262.5 cm^{-1}. We chose methane because the studies of methane concentration in the exhaled breath of healthy children are more abundant [179, 180].

6.3.1. CONCENTRATION DETERMINATION METHODS

Acetone and methane present two types of absorption spectra as shown in Figure 6.7. Acetone has a broad smooth profile covering 100 cm^{-1} while methane presents a profile with sharp absorption lines, as shown by the 3.5 cm^{-1} plotted in the figure. For this reason we have to be careful when analyzing and calibrating the absorption spectrum of the breath samples.

(a)

(b)
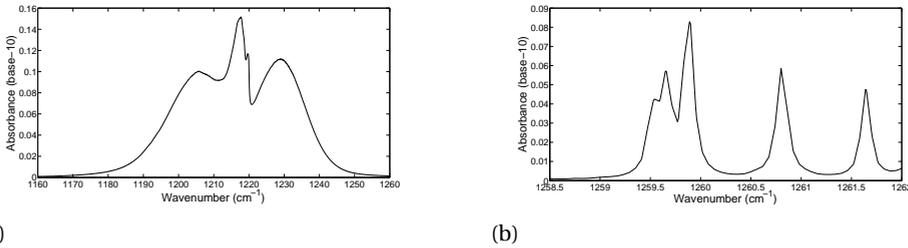
Figure 6.7: Absorption spectrum of (a) acetone and (b) methane. The spectrum of acetone is smooth and covers a broad wavenumber range covering 100 cm^{-1}. The spectrum of methane consist of sharp absorption lines as point out by the narrow wavenumber range of 3.5 cm^{-1} presented.

ACETONE

For acetone we determined its concentration in exhaled breath using the same method applied in our study of type 1 diabetes of Chapter 5 and Reference 114. As a reminder, this method uses the scan correlation using selected wavenumbers discussed in Section 4.2 and the wavenumber calibration mentioned in Section 4.4.1. In this case we used a spectral resolution of 0.05 cm^{-1}.

The acetone concentration determination is done with the multiline fitting method. As described in Section 4.5, this method finds the concentration values C_j, with $j = 1, 2, ..., n$, that minimizes the function

$$F(C_1, ..., C_n) = \sum_{i=1}^{k} \left[A(\widetilde{v}_i) - l \sum_{j=1}^{n} C_j \epsilon_j(\widetilde{v}_i) \right]^2 \qquad \text{(4.11 revisited)}$$

with $A(\widetilde{v}_i)$ the absorbance of the measured spectrum, $\epsilon_1(\widetilde{v}_i), \epsilon_2(\widetilde{v}_i), ..., \epsilon_n(\widetilde{v}_i)$ corresponding to the molar absorption coefficient of each molecular species present in the sample,

$l = 54.36$ meters is the interaction distance in the multipass cell and k denotes the number of wavenumber points in the absorbance spectrum. In this specific case we only use the information of acetone from the PPNL database and the water molar absorption coefficient provided in the High Resolution Transmission (HITRAN) database [72]. The resulting concentrations are corrected using the calibration curve obtained with the calibrated gas mixtures provided by the Dutch metrology Institute (VSL) which was presented in Section 5.2.

METHANE

In the case of methane we need a better wavenumber resolution to study its sharp absorption peaks. For this reason we used the scan correlation using semiparametric time warping of Section 4.3 and its designated wavenumber calibration method, described in Section 4.4.2. In this case we used water absorption peaks at 1244.15 cm^{-1} and 1260.35 cm^{-1} as initial references for the wavenumber calibration.

To improve the calibration we apply an additional step in the region between 1258 and 1262.5 cm^{-1} where methane has its absorption signature. We use the water absorption peak centered at 1260.35 cm^{-1} as reference because its profile covers three narrow peaks of methane. First we identify the absorption peak in the measurement corresponding to the water reference peak, in other words, the nearest peak in the measurement to 1260.35 cm^{-1}. Then we obtain the baseline of the peak using a region of 1.5 cm^{-1} width around the maximum of the peak. This baseline is calculated using the *msbackadj* function from the bioinformatics toolbox of MATLAB. Afterward, we apply a Gaussian fit to the baseline to take into account only the contribution of water and reduce the influence of methane or noise. We also apply a Gaussian fit to the water reference peak with a width of 1.5 cm^{-1}. We compare the parameters of both Gaussian fits and use them to displace and compress or expand the wavenumber axis of the measurement. In Figure 6.8 we present an absorption measurement before and after the calibration improvement. Here we can appreciate how the absorption spectrum is shifted to the left and how its width is expanded. The effect of this last expansion is observed by looking how the dips at 1259.65 cm^{-1} and 1261.26 cm^{-1} in the upper subfigure are moved to 1259.30 cm^{-1} and 1261.46 cm^{-1} respectively in the lower subfigure.
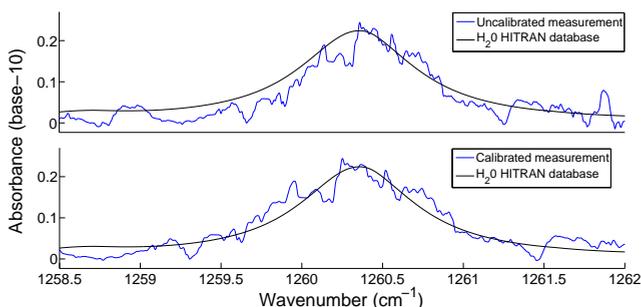
Figure 6.8: Absorption measurement before and after improving the wavenumber calibration using an absorption line of water.

The result of the wavenumber calibration improvement is shown in Figure 6.9. The wavenumber resolution of the final absorption spectrum is 0.01 cm^{-1}, which allows us to study the details of the methane spectrum.
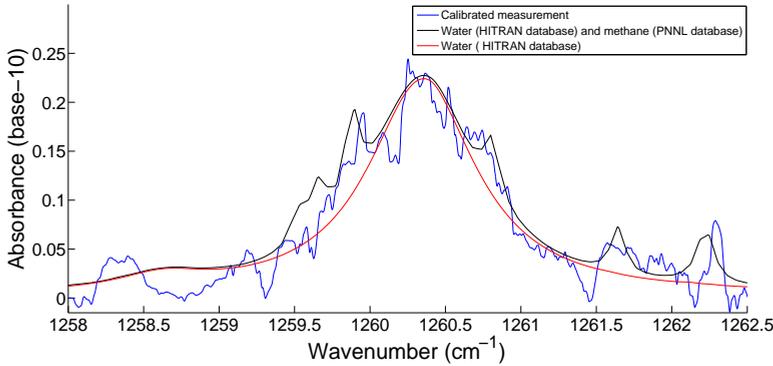
Figure 6.9: Absorption spectrum depicting the overlapping lines of water and methane.

To determine the concentration of methane we used again the multiline fitting of Section 4.5, using only the molar absorption coefficients of water and methane for the fitting. For methane we did not have access to calibrated samples from the Dutch metrology institute and therefore we did not apply any correction to our measurements, nevertheless we obtain interesting information that we will discuss in Section 6.3.2 below.

6.3.2. RESULTS AND DISCUSSION
The results of our concentration measurements are depicted in Table 6.4.

Table 6.4: Mean and standard deviation of acetone and methane in healthy children.

	Concentration [ppm$_v$]	
	average	standar deviation
Acetone	0.27	0.19
Methane	0.05	0.03

The first interesting feature we observe is the average and standard deviation of the acetone concentration. The values are below the standard range for healthy adults between 0.39 and 1.09 ppm$_v$ [121]. This result is similar to the findings of our study of acetone in the exhaled breath of children with type 1 diabetes under stable conditions (Chapter 5 and Reference 114). Both results support our thesis that the standard concentrations of acetone in the exhaled breath of children differ from those of adults and therefore more studies should be performed.

In the case of methane, we observe rather low concentration levels. This directly points out the need to establish a calibration curve to correct the concentration values.

Nevertheless, the small average value and the small concentration distribution shown by the standard deviation is consistent with previous studies [180]. These studies reported that in the age range we studied only 10% to 20% of the individuals present methane concentrations higher than 1 ppm$_v$ in their exhaled breath, known as producers. According to the distribution of our results the children in our study are non-producers and their concentrations would fall below 1 ppm$_v$, but this needs to be corroborated with an independent measurement. Another interesting point to take into consideration is the comparison with respect to the methane levels in the environment, which according to previous reports can present values as high as 2 ppm$_v$ [181, 182]. But this value changes from place to place and should be measured directly.

6.4. CONCLUSIONS

We developed a QCL spectroscopic technique that covers a broad wavenumber region. It allows the detection of VOC profiles in exhaled breath samples [81]. Previous studies using laser-based spectroscopy methods have focused on narrow wavenumber regions (10-20 cm^{-1}), covering the absorption signature of only one or two specific molecules [183–185].

We used this QCL spectroscopic system to study the exhaled breath of healthy children and children with stable asthma or stable CF. We found an overall poor repeatability and agreement on the short and long term of the complete exhaled breath profiles. However, we were able to identified wavenumber regions where the profiles of the healthy children are significantly different from the profiles of children with asthma or CF. We used these significantly different regions and the information from the PNNL and HITRAN molecular databases to identify VOCs that can be used to discriminate between healthy children and children with asthma or CF. The identification of VOCs is important to understand underlying inflammatory processes, which may be useful to develop new and/or personalized treatments.

An interesting result was obtained from the p-value analysis used in the comparison between the control and CF groups. It showed that a clear and deep understanding of the origin of the differences between groups is crucial in order to correctly identify the relevant VOCs. In this case by applying a more flexible threshold p-value of 5.6 % we included a second significantly different region, which otherwise would have been neglected. The inclusion of this extra significantly different region provided more information regarding the possible discriminating VOCs. Because of the exploratory nature of our approach we must highlight the need of additional validation and attempts to improve the repeatability before further clinical studies are undertaken.

The study of acetone and methane demonstrated the potential of our technique to explore the presence of specific molecules by focusing on particular regions of the spectrum covered by our QCL spectroscopic system. For acetone, the concentration results corroborate the need to performed more dedicated studies to establish the standard concentrations of acetone in the exhaled breath of children. In the case of methane, the results highlighted the need of defining a calibration curve for our system. Their distribution does not show a broad scattering of values, which suggests consistency with the concentrations reported in previous studies, but a more in depth study is necessary.

The advantages of our QCL method are the ability to detect whole molecules instead

of fragmented molecules, as in the case of GC-MS, the relatively fast and easy analysis and the potential for miniaturization. They make QCL spectroscopy attractive to investigate the diagnostic and prognostic potential of VOCs in exhaled breath.

6

7

VALORISATION

There are more things in Heaven and Earth, Horatio,
than are dreamt of in your philosophy.

William Shakespeare, Hamlet

Two centuries ago we did not know what air was. Now we know the components of air and their role in the atmosphere. Furthermore we are fully aware that the quality of life depends on the quality of the air we breathe. Nowadays, the advances in the production of light sources emitting in the mid-infrared are changing the way we study the composition of air, giving way to new methods and techniques to perform breath analysis. Specially the production of quantum cascade lasers (QCLs) is growing fast and it is presenting its potential in the gas and breath analysis applications in a decisive way. We are experiencing the beginning of a new revolution in healthcare that will reach its maturity in the next decade when QCL will become cheaper and more reliable.

The value of any research is not only in what it demonstrates but on the possibilities it opens. In this chapter we explore the potential offer by the combination of QCL spectroscopy with different methods developed for other gas analysis techniques. We aim to draw a roadmap of the possibilities taking as reference the recent achievements obtained in mid infrared spectroscopy using sample pre-concentration methods, molecular filtering, and hollow waveguides (HWGs). In the final part of this Chapter we make an evaluation of the social and economic impact that breath analysis methods will have in the near future by studying two successful examples: the breathalyzer and the fractional exhaled nitric oxide (FeNO) analyzers.

7.1. Pre-concentration of the gas samples

Trace gas analysis using laser-based methods is without doubts governed by the Beer - Lambert law and the absorbance $A(\widetilde{v})$

$$A(\widetilde{v}) = \epsilon(\widetilde{v})Cl. \qquad\qquad \text{(2.29 revisited)}$$

Until now most of the techniques using mid-infrared sources have focused on increasing the absorbance by choosing either the wavenumber, $\widetilde{v}$, that corresponds to the highest molar absorption coefficient, $\epsilon(\widetilde{v})$, or by augmenting the interaction distance, l, as much as possible. But the proper handling of the concentration, C, can provide a very powerful tool to increase the sensitivity of laser-based methods for trace gas analysis and specifically for exhaled breath analysis.

In fact, it has already been demonstrated that pre-concentrating the gas samples boosts the sensitivity of QCL absorption spectroscopic devices. Already in the 1990s Bergamaschi et. al. studied the methane (CH_4) concentration in the atmosphere [186]. They used a pre-concentration step that uses charcoal at 77 K to trap the CH_4. The concentration was determined with a tunable diode laser absorption spectrometer (TDLAS) that had two Pb-salt diode lasers to scan the absorption signature of CH_4 around 3007 cm^{-1} [187]. The main disadvantage of this system relays in the fact that it needed liquid nitrogen to operate the charcoal traps and the diode lasers, which makes it costly and reduces its use for on-field applications.

More recently a pre-concentration unit originally developed for gas chromatography - mass spectrometry (GC-MS) systems [188] was adapted to work with QCL absorption spectroscopy systems [189]. It has permitted the detection of nitrous oxide (N_2O) in the atmosphere [190–193]. The QCL spectrometer is similar to the one presented in this thesis, the main difference is that it only scans the wavenumber region corresponding to

the absorption lines of N_2O. Moreover, this system made possible to detect, distinguish and measure different isotopes of N_2O. The pre-concentration system combines an adsorbent and a low temperature refrigeration unit [188]. It allows to increase 200 times the concentration of N_2O to get a sample with 60 ppm_v, which is in accordance with the 320 ppb_v currently reported in ambient air using other techniques.

In a similar way Eyer et. al. [194] used a pre-concentration unit to increase the CH_4 concentration by a factor of up to 500. They also used a QCL absorption spectrometer. They were able to detect and differentiate two CH_4 isotopes. The results reported concentrations levels of 750 ppm_v in the concentrated sample, which corresponds to an un-concentrated sample of around 1.5 ppm_v. Furthermore, the reported precision on the concentration sample is 0.1‰ for the δC-isotope and 0.5‰ for the δD-CH_4 isotope, with an average sampling of 10 minutes. The use of a different adsorbent in this pre-concentration device has been suggested to measure carbon monoxide and other volatile organic compounds (VOCs) [189].

The previous examples show the feasibility of improving the sensitivity of laser-based trace gas detection with the use of adequate pre-concentration methods. They are attractive and useful for monitoring the atmosphere and for applications where gas sample volumes of 5 liters or more are available. However, this requirement makes them impractical for applications where the sample volume is limited, like in breath analysis. In these cases is necessary to look for more efficient pre-concentration techniques. The desired pre-concentration methods have to fulfill some basic requirements to be reliable:

- Efficient trapping of the molecular species under investigation.

- Filter non-desirable molecules to avoid interference from other constituents of the gas sample.

- High and easy recovery of the desired molecular species.

- Chemical neutrality of the pre-concentration trap to avoid concentration changes.

In the chemical sensing field there are already good candidates to fabricate the desired pre-concentration units [195]. The basic element for these units is the fabrication of nanomaterials whose surface can be functionalized to trap specific molecules. Among the potential candidates are the silicon-doped nanostructured WO_3 films developed by Righettoni et. al. for acetone detection [124]. Originally these films were developed to measure the acetone concentration using the change in the resistivity of the films induced by the presence of acetone. They demonstrated sensitivities as low as 20 ppb_v with response and recovery times lower than 12 seconds. With some extra research these films can be used to produce an efficient pre-concentration unit that can be coupled to a laser-based spectrometer. Even more, by using the resistivity properties of the films we can produce a system with two methods to determine the concentration, which will provide another level of confidence to the measurements.

A second example of the nanomaterials used to produce films that can be used for the fabrication of pre-concentration units is polyaniline nanoparticles deposited on a silver electrode using inkjet-printing [196]. These films are highly selective for ammonia

detection. They allow to detect concentration as low as 40 ppb$_v$ by measuring their electrochemical response. In this case the films were conceived as disposable and therefore the recovery of ammonia has not been investigated, but it should be possible to implement a recovery method that permits the use of these films in pre-concentration units for laser-based spectroscopy.

Other examples of materials that could be used in pre-concentration units are the nanotubes and nanowire arrays. Their potential for chemical sensing is proven [197–199]. Moreover the functionalization of their surface allows a high selectivity to bind to very specific molecules [200]. An interesting example for trace gas analysis in this field is the nanowire arrays on plastic substrate developed by McAlpine et. al. [201]. Its electrochemical response permits the detection of Nitrogen dioxide (NO_2) in concentrations of 20 ppb$_v$.

To have an idea of how the pre-concentration will improve the detection sensitivity in exhaled breath we present the expected absorbance of eight molecular species in Figure 7.1 . The signals correspond to the reported average concentrations in breath multiplied by a concentration factor. We use the same interaction distance of 54.36 meters of the multipass cell used in this thesis to get an idea of the concentration enhancement needed to get a meaningful signal. Furthermore, the concentration factors are chosen in order to get absorbance levels well above the 0.02 noise level of our system. This way we can ensure good signal to noise ratios and reliable signals.
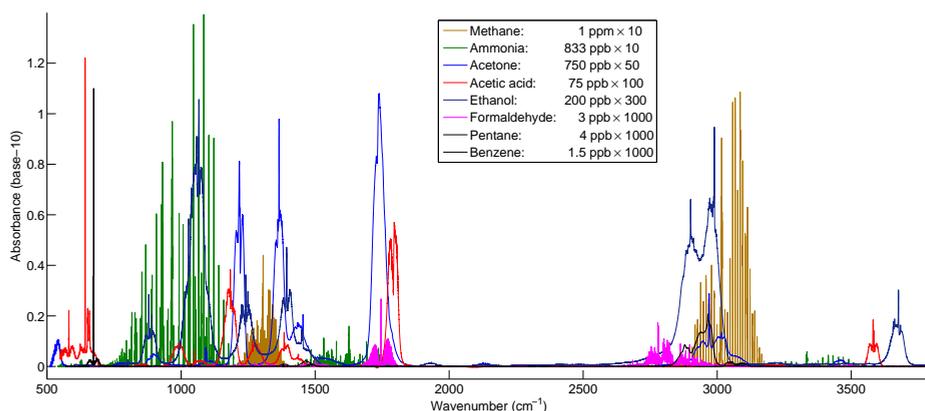
Figure 7.1: Absorbance signature of eight molecules reported in breath after being pre-concentrated. The signal strength corresponds to their average concentration multiply by an adequate enhancement factor. An interaction distance of 54.36 meters is assumed.

The results presented in Figure 7.1 simply confirm that for molecules with low ppb concentrations the enhancement needs to be in the order of 1000. This constraint is even more stringent when detecting molecular concentrations in the ppt level. However, we must remember that by focusing on a very specific molecule we can target smaller wavenumber regions and lower the noise level of the spectroscopic system. This way we can reduce the enhancement factor needed to get a meaningful absorbance. In the Intro-

duction chapter 1 we already mentioned some techniques that can be used. Among them are the differential optical absorption spectroscopy (DOAS) [30], differential absorption lidar (DIAL) [31] and the tunable diode laser absorption spectroscopy (TDLAS) with wavenumber and frequency modulation [34].

7.2. HOLLOW MID-INFRARED FIBERS

The current fabrication methods allow the production of QCLs and mid-infrared detectors chips with dimensions of less than 10 mm [202, 203]. These components are reaching a state of maturity in the search for miniaturized systems. A key element missing is the miniaturization of the gas cell where the samples are placed. Until now the laser-based spectrometers use multipass cells and resonant cavities to increase the interaction distance between the laser light and the gas samples. But these gas cells are bulky. An excellent approach to reduce their size is to use hollow waveguides (HWGs). They offer a way to confined the laser light in thin channels where small volumes of the gas sample are placed. So far the available HWGs for mid-infrared have been mainly used to deliver high powers from CO_2 lasers for industrial and surgical applications [204]. But their use is being extend to gas detection to the point that we can already find patents of systems that used HWGs as micro gas cells[205].

An ideal HWG for gas analysis must provide freedom to transmit the appropriate wavenumber, have a large bore, allow for a path length in the order of few meters and be suitable to be placed in small devices, for example by coiling it. Currently, we can find working HWG cells transmitting in all the range from 1 to 25 μm [204, 206], with internal diameters of 1 mm and volumes lower than 100 microliters [207]. They allow interactions distances of up to 4 meters [208]. This relatively small interaction distance is not a limitation because in the HWGs the laser fills all the volume of the cell, which increases the photon density inside the cell and rises the probability of interaction with the molecules [204, 209]. As an example, Charlton et. al. [204] reported an effective increase in the sensitivity of their absorption system by a factor of 1.2 when compared to the results obtained with a 3 meter multipass cell.

In table 7.1 we present a summary of the gas analysis applications that have used HWGs solely as micro gas cells. We can observe some important achievements from the information presented in this table. First of all, we observe detection capabilities in the order of ppb when coherent sources like CO_2 lasers and QCLs are used.

We also observe a variety of available fibers depending of the substance under study. For example, for the detection of CH_4 and n-C_4H_{10} is better to use HWG internally coated with silver (Ag) and zinc sulfide (Zns). For the detection of nitrogen dioxide (NO_2) and sulfur dioxide (SO_2) is preferable to use silicon dioxide (SiO_2) HWGs. While for CO_2 we can use sapphire fiber or silica HWGs internally coated with silver and silver iodide (Ag/AgI). In particular the Ag/AgI-coated HWGs offer transmission over a broadband range (3 - 15 cm^{-1}) with a loss of only 0.15 dB m^{-1} [210, 211] and it has shown a high immunity to chemical degradation [207].

7

Table 7.1: ID: internal diameter, LOD: calculated limit of detection. FTIR: Fourier Transform Infrared Spectroscopy

	Detection method	Detected molecules	Concentration	ID [mm]	Length [m]	Characteristics
Sato et. al. [212]	Incoherent source Direct absorption	CH_4 n-C_4H_{10}	0.1 - 1.9 % 0.1 - 1.3 %	1	1	ZnS/Ag coated HWG
		NO_2 SO_2		1	0.25	SiO_2 HWG
Kozodoy et. al. [213]	Incoherent source FTIR	CO_2	200 ppm	0.7	0.23	Ag/AgI-coated
Saggese et. al. [214]	Incoherent source FTIR	CO_2	20%	1.06	0.15	Sapphire fiber
Micheels et. al. [215]	Incoherent source FTIR	BTEX * Trichloroethylene	> 20 ppm_v > 25 ppm_v	2	0.9	Ag/AgI-coated 1 meter bend
Saito et. al. [216]	CO_2 laser Direct absorption	NH_3	5 ppm	1.5	1	
Worrell et. al. [217]	CO_2 laser Direct absorption	Ethylene Trichorethylene 2-Furaldehyde	LOD: 86 ppb LOD: 4.5 ppm LOD: 1.9 ppm	1	0.9 2 2	Straight _____ Coiled
Hvozdara et. al. [218]	QCL FTIR	Ethene	250 ppm	1.13	0.424	Response time ~ 100 ms
Young et. al. [219]	QCL FTIR	Dichloromethane Ethyl chloride Trichloromethane	LOD: 7 ppm LOD: 4 ppb LOD: 11 ppb			Ag/AgI-coated
Charlton et. al. [208] [220]	QCL Direct absorption	Ethyl chloride	LOD: 500 ppb_v LOD: 30 ppb_v	0.7	4 1	Ag/AgI-coated Photonic band-gap
Gayraud et. al. [221]	OPO ** FTIR	CH_4	0.1 %	0.045 (core)	0.8	Photonic band-gap

* BTEX: Benzene, toluene, ethylbenzene, and xylenes

** OPO: Optical parametric oscillator

Also in table 7.1 we can observe how the use of different technologies improves the detection achievable with HWGs. Specifically there is a direct improvement in the ethyl chloride concentrations measured by Charlton et. al. using photonic band-gap structures [208] compared to the values obtained with Ag/AgI coated HWGs [220]. Using photonic band-gaps they managed to lower the limit of detection from 500 ppb$_v$ to 30 ppb$_v$ and at the same time they reduced the interaction length from 4 meters to only 1 meter. The use of photonic band-gaps can be applied to improve the detection of other molecules. Their structure allows to select better the wavenumber transmitted through the HWGs, increasing the molecular selectivity.

HWGs have been also used to study the carbon isotope ratio (^{13}C/^{12}C) in small samples [209]. In combination with a direct absorption spectrometer they permit the identification of ^{13}C in concentrations of 10 ppm in a gas sample containing CO_2 concentrations in the range of 400–750 ppm. The same approach can be use to study the isotope composition of other gases such as CO_2, CO, O_2, SO_2, SO, SF_6 and CH_3Cl [207].

The fast response of HWGs to concentration changes offers a unique characteristic for real time studies. The small volume of HWGs reduces their response time to flow changes making it faster than bulk multipass cells [204]. Thompson et. al. [222] compared the response time of a HWG with a length of 55 cm and a multipass cell with a path-length of 3 meters. The response time of the HWG is one order of magnitude faster than the corresponding to the multipass cell. This proves the advantage of using HWGs in real time applications. Some of the fields where this feature become highly attractive include process monitoring, surveillance and combustion processes.

We still need to optimize many aspects of the HWGs before their use can reach its maturity. One of the most important aspects under study is the losses introduce by bends in the HWGs. These losses have been under investigation since HWGs were introduced [210, 211, 213, 217, 220, 223–226]. It is crucial to minimize these losses in order to incorporate HWGs in portable devices. Another characteristics under study are the relation between the internal diameter of the HWGs and the mode quality [207]. Particularly important is to find adequate filtering methods to reduce the noise introduced by the higher order modes of the HWGs [207, 211, 227]. Also under investigation are the effects and improvements that tapered HWGs can introduce in the their transmission [207]. These are just some of the challenges posed by the use of HWGs as gas cells.

In terms of size reduction and integration to portable devices an interesting development was achieved with the introduction of the substrate-integrated hollow waveguide (iHWG). iHWGs are layered structures integrated into a solid-state substrate material that work as waveguides. Their characteristics allow to obtain interaction path lengths of more than 22 cm in a bulk system of only 75 mm × 50 mm × 12 mm [228]. In combination with Fourier transform infrared (FTIR) systems iHWGs have been used to measure ozone in air in real time, obtaining a limit of detection of 3.5 ppm [229].

In conclusion, HWGs are showing a great potential to become the so much needed miniaturized gas cells of the future.

7.3. HYBRID SYSTEMS

Just as GC, MS - based methods have shown great advances thanks to their combination with other chemical and electrochemical techniques, the best performance of laser-

based gas analysis systems is expected from the combination with different techniques. In the previous section we already observe how the combination of QCLs and HWGs permits the detection of low concentrations of molecules. In this section we would like to present some examples where the combination of different filtering, concentration and analysis methods has been used to produce trace gas analysis setups with higher selectivity and sensitivity.

One type of setups use preconcentration units to increase the concentration of the chemical under study and HWGs as miniaturized gas cells. Pogodina et. al. [230] preconcentrated ethene samples using a sorption tube sampling. They use thermal desorption to introduce the sample into a HWG with an internal diameter of 0.7 mm and a length of 47 cm. Using FTIR spectroscopy they obtained a limit of detection of 1 ppb. With a similar setup Young et. al. [231] obtained limits of detection in the order of ppbs for Benzene, Toluene, and Xylenes from environmental air. Perez et. al. [232] achieved a limit of detection of 32 ppb for isoprene from exhaled breath by coupling a compact preconcentrator column, an iHWG and a FTIR spectrometer. A disadvantage of all these systems is that they can not operate in real time.

A different and highly interesting approach is to use HWGs directly as concentrators. As an example, Yang et. al. [233] coated the inner surface of HWGs with a hydrophobic film to trap chlorinated aromatic compounds. Using an FTIR spectrometer they got a limit of detection of 500 ppb.

Another example where the combination of techniques has improved the selectivity and sensitivity of gas analysis is found in the work of Wu et. al. [234]. They used of a GC column to separate the components of a gas sample before using a QCL-HWG spectrometer to measure the concentration of specific compounds. In a variation of the system it is possible to include a Pt-Cu catalytic oxidation module before the HWG to convert any hydrocarbon compounds into CO_2 and water, by measuring the CO_2 content it is possible to indirectly determine the concentration of the hydrocarbons.

Another example of hybrid systems using a filter before the HWG setup is the work developed by de Melas et. al. [235]. They used a membrane sampling device to analyze different chemicals extracted from aqueous solutions by injecting a carrier gas. The setup allows the detection of ppb_v levels of chlorinated hydrocarbons and ppm_v levels of 1,4-dioxane in a continuous measurement. The same system was employed to measure concentrations of benzene, toluene, xylenes isomers and chloroform, in which case a sensititvity in the ppb_v level is reported [236].

A very important application where hybrid QCL-HWG setups can greatly contribute is the detection of explosives. Hildenbrand et. al. [237] showed that is possible to detect TNR and TATP explosives using a desorption unit to extract traces from the containers and membranes used to collect the samples. They achieved a limit of detection of 0.43 ppm using a QCL-HWG spectrometer. This represents the ability to detect 2 ng of TNT and less than 1 ng of TATP in the gas phase with a HWG of only 50 cm.

HWGs also make possible to measure the concentration of chemical species with weak infrared absorptions by converting them into a chemical with a stronger absorption signal in a small volume. In their work, da Silveira Petruci used UV-assisted conversion to transform hydrogen sulfide (H_2S) into sulfur dioxide (SO_2) [238]. The resulting SO_2 is placed in a iHWG and measured with FTIR spectroscopy. The results show a limit

of detection of 3 ppm for H_2S, with a response time of less than 60 seconds.

The previous examples illustrate the world of possibilities opened by the combination of different filtering, concentration and analysis techniques. This is just the starting point of a new revolution in trace gas analysis that will continue growing with the inclusion of methods already used in other areas such as FTIR, GC-MS and nano-fabrication. We can expect to see the first portable and highly sensitive gas analysis devices in the market in the next 10 years.

7.4. SOCIAL AND ECONOMIC IMPACT

During the last century advances in medicine increased life expectancy globally. Along with this increment the healthcare sector has experienced an increasing pressure to find more efficient and economic ways to detect, monitor and treat diseases. For this reason the healthcare sector is one of the fastest growing markets because new technologies are having a huge societal and economical impact. Specifically, the optical technologies applied to healthcare are having one of the highest growth rates. In 2007 they had an estimated market value of €23 billion and it was expected to grow at an 8% compound annual growth rate (CAGR)[1] to reach €43 billion in 2015 [240, 241]. Assuming the same CAGR, we can expect a market value of €63 billion by 2020.

Among the different optical devices for healthcare, the photonic devices working in the extended infrared region (1 - 1000 μm) are particularly attractive because of the information they can provide. For this reason the European technology platform Photonics21 established the development of high performance and affordable photonic devices in the extended infrared region as one of the major photonics needs in the multiannual strategic roadmap 2014 – 2020 [241]. In particular, the development of devices for the mid infrared is experiencing a continuous growth. In 2012 the overall market for mid infrared sensors got a value of $789 million and it is projected to reach $7 billion by 2019 [242].

At the same time the interest in breath analysis is also growing quickly. Breath analysis has already shown huge potential in the identification of different health conditions: from alcohol-impairment [25], the diagnosis of gastroenteric diseases[243] like *Helicobacter pylori* infection [244], the diagnosis of airway inflammation, asthma and cystic fibrosis by measuring the exhaled nitric oxide [19], the control of ketogenic diets of epileptic patients [136], to the first attempts to diagnose diabetes [114, 121], gastric cancer [245] and lung cancer using exhaled breath [246–249]. In 2014 the overall market of breath analyzers, which included all available technologies, was estimated at $479.3 million and the economic forecast places its value at $1,539.1 million by 2019 [250].

In the next paragraphs we present two examples to highlight the influence breath analysis studies are having in the society: the fractional exhaled nitric oxide (FeNO) devices and the breathalyzer. They are practical examples where exhaled breath analysis have become an important player for the economy.

---

[1]CAGR is the mean annual growth rate of an investment over a specified period of time longer than one year [239].

FRACTIONAL EXHALED NITRIC OXIDE ANALYZERS

Among the respiratory conditions, asthma is one of the major chronic respiratory disease in the world [116]. For this reason the interest in asthma research is growing. According to the estimations of the 2014 global asthma report as many as 334 million people in the world suffer from asthma [251]. Furthermore, asthma is the most common chronic disease in children according to the World health organization [252–254]. In 2013 the global atlas of asthma estimated that 15 million of DALYs are being lost each year [255]. Worldwide there is not yet an estimation of the total cost of asthma to the society, but we can have a picture by looking at the numbers reported in the 2014 global asthma report. In 2007 the direct cost in the United States of America was estimated at US$56 billion. In Europe the estimated total cost reached €19.3 billion in 2011 from people within 15 to 64 years old [251]. The indirect costs of asthma mainly comes from the reduction in productivity of workers suffering from asthma. The result of many studies indicate that the burden of the indirect cost is as large as the direct cost.

The previous reasons are the driving force to find new, more efficient methods to diagnose and control asthma. In particular fractional exhaled nitric oxide (FeNO) detection is a very interesting candidate. However commercial FeNO devices are relatively new. With their actual degree of maturity FeNO devices are complementary to other well-established medical procedures, not a replacement. Although the benefits of FeNO analyzers have been observed and reported [256] their use is under a strict scrutiny because the information available is still limited. But the potential of FeNO analysis is being recognized to the point that the American thoraxic society has already established the guidelines to interpret FeNO levels in patients with asthma.

The novelty of FeNO devices have fueled an intense debate over their cost efficiency compared to other diagnosis methods. The latest assessment of the national institute for health research of the UK [257] concluded that FeNO devices have the potential to be cost effective in asthma monitoring and management, but the available published data is insufficient to draw a final conclusion. The use of FeNO detection alone is not seen as a viable candidate for asthma detection and monitoring at this stage. The economic potential of FeNO setups appears when they are used in combination with other diagnostic methods. Particularly promising is the use of FeNO monitoring in combination with bronchodilator reversibility as the dominant diagnosis test.

In terms of economic growth, the reports of the main supplier of FeNO devices showed a revenue of £10.3 millions in the period between June 2015 and March 2016, which represents a 32% growth with respect to the same period in the previous year [258]. This indicates the growing interest in FeNO technology. Furthermore, the implementation of alternative techniques, such as optical spectroscopy, can help to reduce the cost of FeNO devices and improve their accuracy [259].

7

BREATHALYZER

The breathalyzer is an electrochemical sensor that allows to measure the blood alcohol concentration (BAC) through the exhaled breath [25]. Over the years it has helped authorities to reduce the number of road accidents by identifying alcohol impaired drivers. The global status report on alcohol and health 2014 of the World health organization shows a strong evidence of the effectiveness of controlling the BAC levels to avoid drink-

driving [260]. The establishment of low BAC limits (0.02% to 0.05%) and their enforcement through sobriety checkpoints and random breath testing[2] has reduced the traffic crashes by 20% [261, 262]. Furthermore, the same measures have helped to decrease the alcohol consumption among drivers [263].

In the European Union the '2006 Report of alcohol in Europe' confirmed a reduction in traffic accidents of 20% [264]. This represents a reduction of between 161 to 460 disability adjusted life years (DALYs) [3] per million people per year, with a cost of €43 to €62 per 100 people annually [265]. This report also estimated the impact of implementing unrestricted breath testing. The results indicate that random breath testing would help to avoid 111,000 years of disability and premature death at a cost of €233 million per year.

In the United States the results of implementing BAC levels for drivers decreased the fatality rate by 27%[4] between 2005 and 2014 as shown by the report on traffic safety of 2014 [266].

Although there has been a considerable reduction in the alcohol related traffic fatalities, the number of accidents involving alcohol impaired drivers still accounts for a 25% of all accidents in Europe [267] and 31% in the United States [266]. Furthermore, globally the implementation of alcohol laws to reduce drink-driving is not optimal. The global status report on road safety 2015 of the World health organization reported that 176 countries out of 180 have a national drink–driving law [268]. But only 134 of these laws are based on BAC limits and only 121 countries use random breath testing.

The effectiveness shown by alcohol breath testing to reduce traffic accidents is extending its use globally. Every year more countries are adjusting their national regulations in order to extend the implementation of low BAC levels and enforce them using random breath testing. This explains the present growth of the market value of the breathalyzers, which in 2011 reached $284.6 million and it is expected to increase to $3.2 billion in 2018 [269]. For this reasons the breathalyzer is a fine example of the impact breath analysis can have in the economy and more importantly in society.

7

---

[2]Sobriety checkpoints are defined as checkpoints or roadblocks established by the police on public roadways to control for drink–driving. Random breath testing is defined as the situation in which any driver can be stopped by the police at any time to test the breath for alcohol consumption. [260]

[3]A DALY is a measure of the health of a person living with a disease. It combines the number of years of life lost due to premature death and the years of life lost due to disability [264].

[4]This fatality rate is taken per 100 million vehicle miles traveled.

8

GENERAL CONCLUSIONS

Every real story is a never ending story.

Michael Ende, The Neverending Story

We have presented the results of a research project that took up the challenge to design, build, test, validate and deploy a trace gas detection system for human breath analysis based on quantum cascade lasers (QCLs). The primary goal of this research was to study the absorption profile of exhaled breath over a broadband range, between 850 and 1250 cm^{-1}. The implemented spectroscopic setup allowed us to study the difference between the absorption profile of healthy and ill persons. In the following paragraphs we will wrap up the main lessons learned during the development of this research.

SETUP AND DATA PROCESSING

Broadband QCL spectroscopy was the best approach for the exploratory nature of our research. It is the only optical method that allows to study the complex composition of exhaled breath. In Chapter 1 we mentioned that other studies have shown high sensitivities in the detection of low molecular concentrations, but they have always focused on one or two specific molecules by scanning a narrow wavenumber regions (10-20 cm^{-1}) [183–185]. The interest of developing optical spectroscopic systems that can analyze more complex gas mixtures is seen as an alternative to more established techniques like gas chromatography and mass spectrometry. Optical methods do not necessarily represent a replacement to these established methods, as mentioned in Chapter 7 they can complement them.

A solid understanding of the physical origin of the absorption spectrum of molecules was essential to the development of this work. For this reason we devoted Chapter 2 to present the main aspects of the theory of gas detection. This placed the foundations for the design of the experimental setup, from the choice of the QCLs, the definition of the parameters of the multipass cell, to the selection of the appropriate Mercury Cadmium Telluride (MCT) detectors.

The chosen QCLs cover a wavenumber region where many of molecules present in human breath absorb. Unfortunately, due to the time constraints of the project we could only get QCLs within an enclosed system. Despite the unavailability to modify the mechanical, electronic or thermal controls of the QCLs we were able to make an exhaustive characterization of their emission. This characterization showed the presence of mode-hop in the QCLs' emission and their inability to accurately repeat consecutive wavenumber scans. These are the main two reasons why most of the trace gas research using QCLs is devoted to narrow wavenumber bands. But we demonstrated that by monitoring the QCL emission and using appropriate analytical methods is possible to correlate independent scans.

The two correlations methods presented in this thesis permit to correct the warping between independent scans. The degree of correlation obtained lowered the noise equivalent absorption sensitivity (NEAS) of the system to 2.99 $\times 10^{-7}$cm^{-1}Hz$^{-1/2}$, allowing to obtain sensitivities in the ppm$_v$ and tens of ppb$_v$ levels. Each correlation methods implemented present different advantages and drawbacks. The scan correlation using selected wavenumbers permits to quickly identify molecules with broad smooth absorption profiles, but at the same time is limited to this type of molecules because it does not capture the fine details of the spectra. On the other hand the scan correlation using semiparametric time warping (STW) provides a more detailed absorption spectrum including the sharp absorption lines but it has heavier computational requirements that make it slower.

8

The choice of the multipass cell with astigmatic mirrors used in the gas analysis setup is also based on the broadband wavenumber under investigation. The coating of the mirrors provides the reflectivity and robustness needed to guarantee a stable interaction distance over the full spectral bandwidth. This is not the case for resonant cavities whose ability to provide interaction distances in the order of kilometers is restricted to very narrow bandwidths due to the high specificity of their mirrors' coating. Additionally, the reflection geometry provided by the astigmatic mirrors permits to fold the interaction distance on a space with a reasonable small volume of 0.6 liters, which is ideal for the limited sample volume available for breath analysis. At the same time, the reflection geometry allows the laser to cover a higher portion of the multipass cell's volume. However, as discussed in Chapter 7 the multipass cells are still too bulky to use the systems in a generalize clinical practice. For this reason we encourage to investigate the potential of hollow waveguides (HWGs) in the development of small and portable breath analyzers.

In terms of the MCT detectors used in the setup, we can only acknowledge their excellent performance, sensitivity and reliability. The use of thermoelectric cooling systems have helped them to be usable at room temperature. MCT detectors have reached a degree of maturity that will ensure their use in many of the QCL and mid infrared future applications.

A major challenge in our project was the identification of the different molecular components of a gas mixture. This problem in itself is the focus of a full research field known as chemometrics. The scan correlations methods detailed in Sections 4.3 and 4.2 are the starting point in the process to determine the molecular composition of a gas sample. In a second phase we calibrated the experimental spectra to a global reference. For this we use the the absorption features of known molecules like CO_2 and water. The results of this calibration are satisfactory, specially when we look at the overlap obtained from the sharp absorption features of CO_2 presented in Section 4.4.3. The final step addressed the identification and quantification of molecular components whose absorption signature overlaps. In this case we use the multiline fitting method explained in Section 4.5. We were very successful in this regards as was demonstrated with our measurement of acetone and methane in exhaled breath presented in Chapters 5 and 6. We could analyze the sharp absorption lines of water on top of a the broad smooth spectrum of acetone and we also could analyze two overlapping molecules with sharp absorption lines, in this case water and methane.

EXHALED BREATH TESTS

From the results of the clinical test performed in children with asthma and cystic fibrosis (CF) and reported in Chapter 6 is clear that a setup with more sensitivity is desired. This is particularly clear from the overall poor repeatability and agreement found in the the short and long term profile comparison. Nevertheless, the significant differences observed between the healthy group and the asthma and CF groups show the possibilities offered by QCL spectroscopy. With the current degree of sensitivity we found a considerable number of suspect molecules that may play a role in asthma and CF and should be further investigated. For asthma we found 82 suspect molecules out of which 23 have been previously reported. For CF we found 58 suspected molecules with 14 of them previously reported. Particularly interesting is to mention the fact that for CF we choose

a higher p-value of 5.6 % to include a second significant region that would have been neglected if we would have limited ourselves to the more conservative limit of 5 %.

The reliability of our systems is further corroborated by the results obtained from the studies of acetone and methane concentration in breath. In both cases the resulting concentration regions matched the results previously reported. The acetone case is particularly interesting because the number of studies in children is limited and our results from two different studies point out that the acetone levels in minors is lower than in healthy adults. This was observed in the study performed in a healthy group of 11 individuals and in the study practiced with 2 children with type 1 diabetes (T1D). The natural suggestion is to perform a broader study in healthy children to establish the standard acetone concentration range.

The measurement of acetone in breath can be immediately applied in the detection of chronic ketosis. As mentioned in Chapter 5 when the acetone levels in the blood remain relatively high during a prolong period of time, the slow diffusion of acetone to the lungs increases the acetone concentration in breath. This way we could detect chronic ketosis in a non-invasive and reliable way. This can mark a breakthrough in the diagnosis of T1D in children because a great number of T1D patients present chronic ketosis for a long period, some times even for years, before they are diagnose with diabetes. In this sense QCL spectroscopy and its future developments can provide a fast method to assess ketosis and become a routine test in medical examinations. It can be a reference point to perform more complex medical test when ketosis is detected.

The findings of our different studies in exhaled breath in children strongly support the idea that personalized healthcare test will help to improve the diagnosis and monitoring of diseases. In this landscape QCL spectroscopy is a strong candidate for the development of cheaper, faster and reliable point of care devices.

RECOMMENDATIONS

The QCL setup and the data analysis presented have the capability to detect concentrations in the ppb$_v$ level and to discriminate molecules with overlapping absorption signatures. There are a series of practical steps that can be implemented to increment the sensitivity of the system.

First of all, the interaction distance in the multipass cell can be increased by changing the angle of the input laser beam. During the construction and test of the multipass cell we achieve configurations of up to 100 meters. We did not use such configurations because we tried to maintain an adequate intensity in the output beam and because they are less robust to vibrations and misalignment. With the current knowledge of the system we are confident that the intensity obtained with a longer interaction distance will be enough to perform the absorption measurements. This increase in the interaction distance will double the strength of the absorption profiles, but we must be aware that it will not decrease the noise level.

The noise level of the system and its sensitivity can be improved by using one single MCT detector to capture the absorption and the normalizing signals of the QCLs [97]. As mentioned in Chapter 3 this will reduces the electronic delay of the signals and it should also reduce the noise contribution due to the use of two independent MCT detectors. A more complex improvement to reduce the noise level of the system will be to measure all

8

the energy carried by each QCL pulse. So far, we limited ourselves to measure the peak power of the pulses and depended on a peak detection algorithm. By integrating the energy of the pulses we would avoid the appearance of false pulses, reduce the uncertainty in the measured power and obtain cleaner QCL signals. This particular approach is not trivial because it requires the implementation of fast detection electronics to capture the profile of each pulse, to have a large amount of memory available to temporarily store the information of each pulse and perform the integration of the pulse's energy. It also requires a fast data transmission to write the information in the final data files. Without these elements we would loose valuable information in each wavenumber scan.

Probably the most important element in the quest to boost the sensitivity of our setup lies in the improvement of the correlation between independent scans. Only by reducing the uncertainty introduced by the warping of the scans we can clean the resulting absorption profiles and improved the concentration determination. The acquisition of cleaner QCL signals by integrating individual pulses will enhanced the performance of the scan correlation methods described in Chapter 4, specially for the study of broadband wavenumber regions.

Our results on acetone and methane detection showed that depending on the application sometimes is better to use a smaller wavenumber bandwidth. In principle this should be possible with the current QCLs by tuning the system to the central wavenumber of the absorption region of interest and then applying a modulation to the current supply. With this we should be able to scan the emission of the QCLs over a small wavenumber region without mode hops. However, this approach should be throughout investigated because the current QCL system uses the same power supply to feed all its electronics, including the lasers, the thermoelectric coolers and the piezo elements.

Finally, we would encourage to explore the possibilities presented in the valorisation Chapter 7. Particularly, we believe some of the pre-concentration techniques can be implemented succinctly. Most of the elements required to concentrate the samples are commercially available and are compatible with the multipass cell and the gas system already implemented. The implementation of HWGs is also possible but it requires a more in depth investigation.

A WORD ON MULTIDISCIPLINARY RESEARCH

One of the big lessons learned from this research project is that a period of 4-5 years is only enough to build the foundation for a more complete, long reaching project. Even though we were able to perform some clinical trials, the main suggestion is to increase the time a project of this magnitude runs. A realistic suggestion is to increase the time span of the whole project to at least 8 years. With one quarter of the time devoted exclusively to the design and construction of the setup, one quarter used to develop the data analysis, another quarter for the calibration and validation of the system and the last quarter employed in the clinical test. This way we can get a more extensive understanding of each stage of the research and address the different issues encountered in a more efficient manner.

Another important lesson learned from this project is that the success of our research was only possible thanks to a multidisciplinary collaboration. This type of collaborations faces many challenges, but particularly important is the language barrier between different fields. The establishment of a common language is essential to define of a proper

8

agenda and to keep the project running. And maybe this was the most challenging step in the early stages of our particular collaboration, which fortunately we were able to over-come.

REFERENCES

[1] M. W. Sigrist, *Introduction to environmental sensing,* in *Air monitoring by spectroscopic techniques. Chemical analysis series, Vol. 127,* edited by M. W. Sigrist (John Wiley & Sons, Inc., 1994) Chap. 1, pp. 1–26.

[2] E. Dlugokencky and P. Tans, *Trends in atmospheric carbon dioxide* (NOAA/ESRL) http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html. [Accessed: 04-02-2016].

[3] *Mosby's Medical Dictionary*, 9th ed. (Elsevier, St. Louis, MO, 2013) p. 154.

[4] IPCC 2014, *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*, edited by Core Writing Team, R. K. Pachauri, and L. A. Meyer (IPCC, Geneva, Switzerland, 2014).

[5] F. Szabadvary, *History of analytical chemistry* (Pergamon Press, 1966) p. 420.

[6] S. N. Ketkar, *1. Introduction to gas analysis: past and future,* in *Trace analysis of specialty and electronic gases*, edited by W. M. Geiger and M. W. Raynor (Wiley, Hoboken, New Jersey, 2013) Chap. 1, pp. 1–20.

[7] J. V. Hinshaw, *A century of chromatography — Gas analysis in the first 50 years,* LCGC Europe , 1 (2003).

[8] S. P. Burg and E. A. Burg, *Role of ethylene in fruit ripening,* Plant Physiology **37**, 179 (1962).

[9] I. J. Church and A. L. Parsons, *Modified atmosphere packaging technology: A review,* Journal of the Science of Food and Agriculture **67**, 143 (1995).

[10] H. H. Funke, B. L. Grissom, C. E. McGrew, and M. W. Raynor, *Techniques for the measurement of trace moisture in high-purity electronic specialty gases,* Review of Scientific Instruments **74**, 3909 (2003).

[11] J. Feng, R. Clement, and M. Raynor, *Characterization of high-purity arsine and gallium arsenide epilayers grown by MOCVD,* Journal of Crystal Growth **310**, 4780 (2008).

[12] K. C. Cossel, F. Adler, K. A. Bertness, M. J. Thorpe, J. Feng, M. W. Raynor, and J. Ye, *Analysis of trace impurities in semiconductor gas via cavity-enhanced direct frequency comb spectroscopy,* Applied Physics B: Lasers and Optics **100**, 917 (2010).

[13] P. Kolla, *Gas-chromatography, liquid-chromatography and ion chromatography adapted to the trace analysis of explosives,* Journal of Chromatography A **674**, 309 (1994).

[14] J. Yinon, *Trace analysis of explosives in water by gas chromatography-mass spectrometry with a temperature-programmed injector,* Journal of Chromatography A **742**, 205 (1996).

[15] P. A. Smith, C. J. Lepage, M. Lukacs, N. Martin, A. Shufutinsky, and P. B. Savage, *Field-portable gas chromatography with transmission quadrupole and cylindrical ion trap mass spectrometric detection: Chromatographic retention index data and ion/molecule interactions for chemical warfare agent identification,* International Journal of Mass Spectrometry **295**, 113 (2010).

[16] M. C. Phillips, M. S. Taubman, B. E. Bernacki, B. D. Cannon, R. D. Stahl, J. T. Schiffern, and T. L. Myers, *Real-time trace gas sensing of fluorocarbons using a swept-wavelength external cavity quantum cascade laser,* Analyst **139**, 2047 (2014).

[17] T. Yuanyuan, L. Wenqing, K. Ruifeng, L. Jianguo, H. Yabai, Z. Yujun, X. Zhenyu, R. Jun, and G. Hui, *Measurements of NO and CO in Shanghai urban atmosphere by using quantum cascade lasers,* Optics Express **19**, 20224 (2011).

[18] TexAQS II Rapid Science Synthesis Team, *Final Rapid Science Synthesis Report : Findings from the Second Texas Air Quality Study (TexAQS II)* (Texas Commission on Environmental Quality, 2007).

[19] R. A. Dweik, P. B. Boggs, S. C. Erzurum, C. G. Irvin, M. W. Leigh, J. O. Lundberg, A.-C. Olin, A. L. Plummer, and D. R. Taylor, *An official ATS clinical practice guideline: Interpretation of exhaled nitric oxide levels (FeNO) for clinical applications,* American Journal of Respiratory and Critical Care Medicine **184**, 602 (2011).

[20] IUPAC., *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). ,* (1997), xML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook.

[21] PerkinElmer, Inc., *Sensitivity Versus Detection Limit. Technical note.* (2016), https://www.perkinelmer.com/lab-solutions/resources/docs/AASensitivityvsDetectionLimitTechNote.pdf [Accessed: 01-12-2016].

[22] H. Hartkamp, J. Rottmann, and M. Schmitz, *Trace gas analysis using thermoanalytical methods,* Fresenius Journal of Analytical Chemistry **337**, 729 (1990).

[23] P. R. Griffiths and J. A. de Haseth, *Fourier transform infrared spectroscopy,* 2nd ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).

[24] T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner, eds., *Handbook of machine olfaction: Electronic nose technology* (Wiley-VCH, 2003).

[25] R. F. Borkenstein and H. W. Smith, *The breathalyzer and its applications,* Medicine, Science and the Law **2**, 13 (1961).

[26] A. T. James and A. J. P. Martin, *Gas-liquid partition chromatography. A technique for the analysis of volatile materials,* Analyst **77**, 915 (1952).

[27] J. H. Gross, *Mass spectrometry*, 2nd ed. (Springer, Berlin, Heidelberg, 2011).

[28] C. Wang and P. Sahay, *Breath analysis using laser spectroscopic techniques: Breath biomarkers, spectral fingerprints, and detection limits.* Sensors **9**, 8230 (2009).

[29] E. Hecht, *Optics*, 3rd ed. (Addison-Wesley, USA), 1998).

[30] U. Platt, *Differential optical absorption spectroscopy (DOAS),* in *Air monitoring by spectroscopic techniques. Chemical analysis series, Vol. 127,* edited by M. W. Sigrist (John Wiley & Sons, Inc., 1994) Chap. 2, pp. 27–84.

[31] S. Svanberg, *Differential absorption lidar (DIAL),* in *Air monitoring by spectroscopic techniques. Chemical analysis series, Vol. 127,* edited by M. W. Sigrist (John Wiley & Sons, Inc., 1994) Chap. 3, pp. 85–164.

[32] P. Kluczynski, J. Gustafsson, Å. M. Lindberg, and O. Axner, *Wavelength modulation absorption spectrometry — an extensive scrutiny of the generation of signals,* Spectrochimica Acta Part B: Atomic Spectroscopy **56**, 1277 (2001).

[33] G. C. Bjorklund, M. D. Levenson, W. Lenth, and C. Ortiz, *Frequency modulation (FM) spectroscopy,* Applied Physics B Photophysics and Laser Chemistry **32**, 145 (1983).

[34] F. Tittel, D. Richter, and A. Fried, *Mid-infrared laser applications in spectroscopy,* in *Solid-state mid-infrared laser sources,* edited by I. T. Sorokina and K. L. Vodopyanov (Springer-Verlag, Berlin, Heidelberg, 2003) pp. 445–516.

[35] D. Hofstetter, M. Beck, J. Faist, M. Nägele, and M. W. Sigrist, *Photoacoustic spectroscopy with quantum cascade distributed-feedback lasers.* Optics letters **26**, 887 (2001).

[36] A. A. Kosterev, Y. A. Bakhirkin, R. F. Curl, and F. K. Tittel, *Quartz-enhanced photoacoustic spectroscopy.* Optics letters **27**, 1902 (2002).

[37] G. Berden and R. Engeln, eds., *Cavity ring-down spectroscopy: Techniques and applications* (Wiley, 2009).

[38] D. S. Baer, J. B. Paul, M. Gupta, and A. O'Keefe, *Sensitive absorption measurements in the near-infrared region using off-axis integrated-cavity-output spectroscopy,* Applied Physics B: Lasers and Optics **75**, 261 (2002).

[39] M. S. Taubman, R. M. Williams, T. L. Myers, F. Capasso, C. Gmachl, D. Sivco, A. J. Hutchinson, and A. Y. Cho, *Sub-doppler nice-ohms spectroscopy at 8.5 microns using a quantum cascade laser,* in *Conference on Lasers and Electro-Optics* (Optical Society of America, 2002) p. CFA2.

[40] M. J. Thorpe and J. Ye, *Cavity-enhanced direct frequency comb spectroscopy,* Applied Physics B **91**, 397 (2008).

[41] C. Fischer and M. W. Sigrist, *Mid-IR difference frequency generation,* in *Solid-State Mid-Infrared Laser Sources*, edited by I. T. Sorokina and K. L. Vodopyanov (Springer-Verlag, Berlin, Heidelberg, 2003) pp. 99–143.

[42] D. D. Arslanov, M. Spunei, J. Mandon, S. M. Cristescu, S. T. Persijn, and F. J. M. Harren, *Continuous-wave optical parametric oscillator based infrared spectroscopy for sensitive molecular gas sensing,* Laser & Photonics Reviews **7**, 188 (2013).

[43] G. Cerullo and S. De Silvestri, *Ultrafast optical parametric amplifiers,* Review of Scientific Instruments **74**, 1 (2003).

[44] T. Y. Chang, *Vibrational transition lasers,* in *CRC Handbook of laser science and technology. Vol. II. Gas lasers*, edited by M. J. Weber (CRC Press, Inc, Boca Raton, Florida, 1982) Chap. 3.2, p. 575.

[45] C. B. Hitz, J. J. Ewing, and J. Hecht, *Carbon dioxide and other vibrational lasers,* in *Introduction to laser technology* (John Wiley & Sons, Hoboken, 2012) 4th ed., pp. 239–247.

[46] S. Mirov, V. Fedorov, I. Moskalev, D. Martyshkin, and C. Kim, *Progress in Cr 2+ and Fe 2+ doped mid-IR laser materials,* Laser & Photonics Reviews **4**, 21 (2010).

[47] M. Tacke, *Lead–salt lasers,* Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **359**, 547 (2001).

[48] C. Sirtori and R. Teissier, *Quantum cascade lasers: Overview of basic principles of operation and state of the art,* in *Intersubband transitions in quantum structures*, edited by R. Paiella (McGraw-Hill, New York, 2006) Chap. 1, pp. 1–44.

[49] C. R. Petersen, U. Møller, I. Kubat, B. Zhou, S. Dupont, J. Ramsay, T. Benson, S. Sujecki, N. Abdel-Moneim, Z. Tang, D. Furniss, A. Seddon, and O. Bang, *Mid-infrared supercontinuum covering the 1.4 - 13.3 μm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre,* Nature Photonics **8**, 830 (2014).

[50] I. Vurgaftman, W. W. Bewley, C. L. Canedy, C. S. Kim, M. Kim, J. R. Lindle, C. D. Merritt, J. Abell, and J. R. Meyer, *Mid-ir type-ii interband cascade lasers,* IEEE Journal of Selected Topics in Quantum Electronics **17**, 1435 (2011).

[51] Z. Tian, L. Li, H. Ye, R. Q. Yang, T. D. Mishima, M. B. Santos, and M. B. Johnson, *Inas-based interband cascade lasers with emission wavelength at 10.4 μm,* Electronics Letters **48**, 113 (2012).

[52] F. K. Tittel, Y. Bakhirkin, A. a. Kosterev, and G. Wysocki, *Recent advances in trace gas detection using quantum and interband cascade lasers,* The Review of Laser Engineering **34**, 275 (2006).

[53] A. Hugi, R. Maulini, and J. Faist, *External cavity quantum cascade laser,* Semiconductor Science and Technology **25**, 083001 (2010).

[54] G. N. Rao and A. Karpf, *External cavity tunable quantum cascade lasers and their applications to trace gas monitoring,* Applied Optics **50**, A100 (2011).

[55] A. A. Kosterev and F. Tittel, *Chemical sensors based on quantum cascade lasers,* IEEE Journal of Quantum Electronics **38**, 582 (2002).

[56] Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, *Room temperature quantum cascade lasers with 27 % wall plug efficiency,* Applied Physics Letters **98**, 181102 (2011).

[57] R. Maulini, A. Lyakh, A. Tsekoun, R. Go, C. Pflügl, L. Diehl, F. Capasso, and C. K. N. Patel, *High power thermoelectrically cooled and uncooled quantum cascade lasers with optimized reflectivity facet coatings,* Applied Physics Letters **95**, 151112 (2009).

[58] P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, J. B. Khurgin, Y. Dikmelik, X. Wang, J.-Y. Fan, and C. F. Gmachl, *Highly power-efficient quantum cascade lasers,* Nature Photonics **4**, 95 (2010).

[59] A. Hugi, G. Villares, S. Blaser, H. C. Liu, and J. Faist, *Mid-infrared frequency comb based on a quantum cascade laser.* Nature **492**, 229 (2012).

[60] Q. Y. Lu, M. Razeghi, S. Slivken, N. Bandyopadhyay, Y. Bai, W. J. Zhou, M. Chen, D. Heydari, A. Haddadi, R. McClintock, M. Amanti, and C. Sirtori, *High power frequency comb based on mid-infrared quantum cascade laser at $\lambda \sim 9\ \mu m$,* Applied Physics Letters **106** (2015).

[61] P. Atkins and J. D. Paula, *Atkins' physical chemistry*, 8th ed. (Oxford University Press, 2006) pp. 783–827.

[62] W. Demtroder, *Molecular Physics* (Wiley, 2008).

[63] B. H. Bransden and C. J. Joachain, *Physics of atoms and molecules* (Longman, New York, 1983).

[64] G. H. Herzberg, *Molecular spectra and molecular structure. I. Spectra of diatomic molecules* (Krieger Publishing Company, Florida, 1991).

[65] G. H. Herzberg, *Molecular spectra and molecular structure. II. Infrared and Raman spectra of polyatomic molecules* (Krieger Publishing Company, Florida, 1991).

[66] G. H. Herzberg, *Molecular spectra and molecular structure. III. Electronic spectra and electronic structure of polyatomic molecules* (Krieger Publishing Company, Florida, 1991).

[67] National Institute of Standards and Technology, *CODATA Internationally recommended 2014 values of the fundamental physical constants*, http://physics.nist.gov/cuu/Constants/index.html . [Accessed: 10-05-2016].

[68] M. Born and R. Oppenheimer, *Zur quantentheorie der molekeln,* Annalen der Physik **20**, 457 (1927).

[69] M. Born, R. Oppenheimer, and S.M. Blinder (Translator), *On the quantum theory of molecules. Translation from the German*, http://www.ulb.ac.be/cpm/people/scientists/bsutclif/bornopn_corr.pdf [Accessed: 24-09-2015].

[70] A. Reyes-Reyes, *Towards a trace gas detection setup using a VIPA spectrometer*, Master thesis, Delft University of Technology (2010).

[71] H. Goldstein, C. Poole, and J. Safko, *Classical mechanics*, 3rd ed. (Addison-Wesley, 2001).

[72] L. S. Rothman, I. E. Gordon, A. Barbe, D. C. Benner, P. F. Bernath, M. Birk, V. Boudon, L. R. Brown, A. Campargue, J. P. Champion, K. Chance, L. H. Coudert, V. Dana, V. M. Devi, S. Fally, J. M. Flaud, R. R. Gamache, A. Goldman, D. Jacquemart, I. Kleiner, N. Lacome, W. J. Lafferty, J. Y. Mandin, S. T. Massie, S. N. Mikhailenko, C. E. Miller, N. Moazzen-Ahmadi, O. V. Naumenko, A. V. Nikitin, J. Orphal, V. I. Perevalov, A. Perrin, A. Predoi-Cross, C. P. Rinsland, M. Rotger, M. Šimečková, M. A. H. Smith, K. Sung, S. A. Tashkun, J. Tennyson, R. A. Toth, A. C. Vandaele, and J. Vander Auwera, *The HITRAN 2008 molecular spectroscopic database,* Journal of Quantitative Spectroscopy and Radiative Transfer **110**, 533 (2009).

[73] M. Diem, *Modern vibrational spectroscopy and micro-spectroscopy: Theory, instrumentation and biomedical applications* (John Wiley & Sons, 2015) p. 28.

[74] G. Graner, *What can molecular spectroscopy tell us about hot bands?* Revista Mexicana de Astronomia y Astrofisica **23**, 79 (1992).

[75] L. S. Rothman, R. L. Hawkins, R. B. Wattson, and R. R. Gamache, *Energy levels, intensities, and linewidths of atmospheric carbon dioxide bands,* Journal of Quantitative Spectroscopy and Radiative Transfer **48**, 537 (1992).

[76] T. Shimanouchi, *Molecular Vibrational Frequencies,* in *NIST Chemistry WebBook, NIST Standard Reference Database Number 69,* edited by P. J. Linstrom and W. G. Mallard (National Institute of Standards and Technology, Gaithersburg MD, 20899) http://webbook.nist.gov , [Accessed: 07-10-2015].

[77] R. A. Toth, L. R. Brown, C. E. Miller, V. Malathy Devi, and D. C. Benner, *Spectroscopic database of CO_2 line parameters 4300 - 7000 cm^{-1},* Journal of Quantitative Spectroscopy and Radiative Transfer **109**, 906 (2008).

[78] A. Reyes-Reyes, *Espectroscopia de polarizacion en vapor de Rubidio Utilizando un diodo laser con cavidad extendida.*, Bsc. thesis, Facultad de Ciencias, Universidad Nacional Autonoma de Mexico, Mexico City, Mexico (2008).

[79] S. E. Braslavsky, *Glossary of terms used in photochemistry, 3rd edition (IUPAC Recommendations 2006),* Pure and Applied Chemistry **79**, 293 (2007).

[80] S. W. Sharpe, T. J. Johnson, R. L. Sams, P. M. Chu, G. C. Rhoderick, and P. A. Johnson, *Gas-phase databases for quantitative infrared spectroscopy.* Applied spectroscopy **58**, 1452 (2004).

[81] A. Reyes-Reyes, Z. Hou, E. van Mastrigt, R. C. Horsten, J. C. de Jongste, M. W. Pijnenburg, H. P. Urbach, and N. Bhattacharya, *Multicomponent gas analysis using broadband quantum cascade laser spectroscopy,* Optics Express **22**, 18299 (2014).

[82] Block Engineering, *LaserScope user manual,* (2012).

[83] J. B. McManus, P. L. Kebabian, and M. S. Zahniser, *Astigmatic mirror multipass absorption cells for long-path-length spectroscopy,* Applied Optics **34**, 3336 (1995).

[84] G. O. Nelson, *Gas mixtures; preparation and control* (Lewis Publisher, Inc., 1992) pp. 110 –116.

[85] B. Lax, in *Proceedings of the international symposium on quantum electronics*, edited by C. H. Townes (Columbia Univ Press, New York, 1960) p. 428.

[86] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, *Quantum cascade laser,* Science **264**, 553 (1994).

[87] M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior, *Continuous wave operation of a mid-infrared semiconductor laser at room temperature,* Science **295**, 301 (2002).

[88] H. Zhe, *Breath analysis using infrared spectroscopy*, Master thesis, Delft University of Technology (2013).

[89] V. Docter, *Spectroscopic measurements using a quantum cascade laser*, Master thesis, Delft University of Technology (2013).

[90] D. Herriott, H. Kogelnik, and R. Kompfner, *Off-axis paths in spherical mirror interferometers,* Applied Optics **3**, 523 (1964).

[91] J. Altmann, R. Baumgart, and C. Weitkamp, *Two-mirror multipass absorption cell.* Applied optics **20**, 995 (1981).

[92] J. A. Silver, *Simple dense-pattern optical multipass cells.* Applied optics **44**, 6545 (2005).

[93] L.-y. Hao, S. Qiang, G.-r. Wu, L. Qi, D. Feng, Q.-s. Zhu, and Z. Hong, *Cylindrical mirror multipass Lissajous system for laser photoacoustic spectroscopy,* Review of Scientific Instruments **73**, 2079 (2002).

[94] RIGOL Technologies, Inc., *DS1000E, DS1000D Series. Data Sheet,* (2014), http://www.rigolna.com/products/digital-oscilloscopes/ds1000e/ds1102e/ [Accessed: 21-10-2015].

[95] A. Rogalski, *HgCdTe infrared detector material: history, status and outlook,* Reports on Progress in Physics **68**, 2267 (2005).

[96] Block Engineering, *MCT IR Detector Module,* (2013), http://www.blockeng.com/products/mct.pdf [Accessed: 21-10-2015].

[97] D. D. Nelson, J. H. Shorter, J. B. McManus, and M. S. Zahniser, *Sub-part-per-billion detection of nitric oxide in air using a thermoelectrically cooled mid-infrared quantum cascade laser spectrometer,* Applied Physics B **75**, 343 (2002).

[98] E. J. Moyer, D. S. Sayres, G. S. Engel, J. M. St Clair, F. N. Keutsch, N. T. Allen, J. H. Kroll, and J. G. Anderson, *Design considerations in high-sensitivity off-axis integrated cavity output spectroscopy,* Applied Physics B **92**, 467 (2008).

[99] M. S. Taubman, T. L. Myers, B. D. Cannon, R. M. Williams, and J. F. Schultz, *Ultra-trace chemical sensing with long-wave infrared cavity-enhanced spectroscopic sensors,* February (Pacific Northwest National Laboratory, Richland, Washington, 2003).

[100] A. L. Buck, *New equations for computing vapor pressure and enhancement factor,* Journal of Applied Meteorology **20**, 1527 (1981).

[101] Mass Flow ONLINE B.V., *SCCM, ln/min or other Mass Flow units, where do they stand for?* (2014), http://www.massflow-online.com/faqs/what-do-lnmin-lsmin-slm-and-sccm-stand-for/ [Accessed: 01-04-2014].

[102] C. Deng, J. Zhang, X. Yu, W. Zhang, and X. Zhang, *Determination of acetone in human breath by gas chromatography-mass spectrometry and solid-phase microextraction with on-fiber derivatization.* Journal of chromatography. B, Analytical technologies in the biomedical and life sciences **810**, 269 (2004).

[103] National Instruments, *NI PXI / PCI-5152 Specifications,* (2015), http://sine.ni.com/nips/cds/view/p/lang/en/nid/202657 [Accessed: 21-10-2015].

[104] National Instruments, *Peak Detector VI,* (2012), http://zone.ni.com/reference/en-XX/help/371361J-01/lvanls/peak_detector/ [Accessed: 21-10-2015].

[105] D. C. Christopher, *Laser and Electro-Optics. Fundamentals and Engineering* (Cambridge University Press, 1996).

[106] A. A. Kosterev, F. K. Tittel, R. Köhler, C. Gmachl, F. Capasso, D. L. Sivco, A. Y. Cho, S. Wehe, and M. G. Allen, *Thermoelectrically cooled quantum-cascade-laser-based sensor for the continuous monitoring of ambient atmospheric carbon monoxide,* Applied Optics **41**, 1169 (2002).

[107] A. A. Kosterev, R. F. Curl, F. K. Tittel, R. Köhler, C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, *Transportable automated ammonia sensor based on a pulsed thermoelectrically cooled quantum-cascade distributed feedback laser,* Applied Optics **41**, 573 (2002).

[108] T. G. Bloemberg, *New Warping and Multivariate Analysis Methodologies for Complex Proteomic and Metabolomic Data,* Ph.D. thesis, Radboud Universiteit Nijmegen (2015).

[109] T. G. Bloemberg, J. Gerretzen, A. Lunshof, R. Wehrens, and L. M. C. Buydens, *Warping methods for spectroscopic and chromatographic signal alignment: a tutorial.* Analytica Chimica Acta **781**, 14 (2013).

[110] A. M. van Nederkassel, M. Daszykowski, P. H. C. Eilers, and Y. V. Heyden, *A comparison of three algorithms for chromatograms alignment.* Journal of Chromatography. A **1118**, 199 (2006).

[111] P. H. C. Eilers and B. D. Marx, *Flexible smoothing with B-splines and penalties,* Statistical Science **11**, 89 (1996).

[112] P. H. C. Eilers, *Parametric time warping,* Analytical Chemistry **76**, 404 (2004).

[113] F. Adler, P. Masłowski, A. Foltynowicz, K. C. Cossel, T. C. Briles, I. Hartl, and J. Ye, *Mid-infrared Fourier transform spectroscopy with a broadband frequency comb,* Opt. Express **18**, 21861 (2010).

[114] A. Reyes-Reyes, R. C. Horsten, H. P. Urbach, and N. Bhattacharya, *Study of the exhaled acetone in type 1 diabetes using quantum cascade laser spectroscopy.* Analytical chemistry **87**, 507 (2015).

[115] *IDF diabetes atlas*, 6th ed. (International Diabetes Federation, Brussels, 2013).

[116] *WHO | Global status report on noncommunicable diseases 2010* (World Health Organization, Geneva, 2011).

[117] I. Harman-Boehm, A. Gal, A. M. Raykhman, J. D. Zahn, E. Naidis, and Y. Mayzel, *Noninvasive glucose monitoring: a novel approach.* Journal of Diabetes Science and Technology **3**, 253 (2009).

[118] N. Ozana, N. Arbel, Y. Beiderman, V. Mico, M. Sanz, J. Garcia, A. Anand, B. Javidi, Y. Epstein, and Z. Zalevsky, *Improved noncontact optical sensor for detection of glucose concentration and indication of dehydration level.* Biomedical Optics Express **5**, 1926 (2014).

[119] M. Righettoni and A. Tricoli, *Toward portable breath acetone analysis for diabetes detection,* Journal of Breath Research **5**, 37109 (2011).

[120] *Toshiba Press release. Toshiba Develops Breath Analyzer for Medical Applications* (Toshiba Japan, 2014) available from: http://www.toshiba.co.jp/about/press/date.htm [Accessed date: 18-03-2014].

[121] Z. Wang and C. Wang, *Is breath acetone a biomarker of diabetes? A historical review on breath acetone measurements,* Journal of Breath Research **7**, 37109 (2013).

[122] M. J. Henderson, B. A. Karge, and G. A. Wrenshall, *Acetone in the breath: A study of acetone exhalation in diabetic and nondiabetic human subjects,* Diabetes **1**, 188 (1952).

[123] M. Fleischer, E. Simon, E. Rumpel, H. Ulmer, M. Harbeck, M. Wandel, C. Fietzek, U. Weimar, and H. Meixner, *Detection of volatile compounds correlated to human diseases through breath analysis with chemical sensors,* Sensors and Actuators B **83**, 245 (2002).

[124] M. Righettoni, A. Tricoli, and S. E. Pratsinis, *Si:WO$_3$ sensors for highly selective detection of acetone for easy diagnosis of diabetes by breath analysis,* Analytical Chemistry **82**, 3581 (2010).

[125] J. W. Gardner, H. W. Shin, and E. L. Hines, *An electronic nose system to diagnose illness,* Sensors and Actuators B **70**, 19 (2000).

[126] P. Wang, Y. Tan, H. Xie, and F. Shen, *A novel method for diabetes diagnosis based on electronic nose.* Biosensors & Bioelectronics **12**, 1031 (1997).

[127] G. Hancock, C. E. Langley, R. Peverall, G. a. D. Ritchie, and D. Taylor, *Laser-based method and sample handling protocol for measuring breath acetone,* Analytical Chemistry **86**, 5838 (2014).

[128] L. Ciaffoni, G. Hancock, J. J. Harrison, J.-P. H. van Helden, C. E. Langley, R. Peverall, G. a. D. Ritchie, and S. Wood, *Demonstration of a mid-infrared cavity enhanced absorption spectrometer for breath acetone detection.* Analytical Chemistry **85**, 846 (2013).

[129] C. Wang, A. Mbi, and M. Shepherd, *A study on breath acetone in diabetic patients using a cavity ringdown breath analyzer: Exploring correlations of breath acetone with blood glucose and glycohemoglobin A1C,* IEEE Sensors Journal **10**, 54 (2010).

[130] G. Rooth and S. Östenson, *Acetone in alveolar air, and the control of diabetes,* The Lancet **288**, 1102 (1966).

[131] C. N. Tassopoulos, D. Barnett, and T. Russell Fraser, *Breath-acetone and blood sugar measurements in diabetes,* The Lancet **293**, 1282 (1969).

[132] C. Turner, C. Walton, S. Hoashi, and M. Evans, *Breath acetone concentration decreases with blood glucose concentration in type I diabetes mellitus patients during hypoglycaemic clamps,* Journal of Breath Research **3**, 46004 (2009).

[133] D. Guo, D. Zhang, L. Zhang, and G. Lu, *Non-invasive blood glucose monitoring for diabetics by means of breath signal analysis,* Sensors and Actuators B: Chemical **173**, 106 (2012).

[134] T. D. C. Minh, S. R. Oliver, J. Ngo, R. Flores, J. Midyett, S. Meinardi, M. K. Carlson, F. S. Rowland, D. R. Blake, and P. R. Galassetti, *Noninvasive measurement of plasma glucose from exhaled breath in healthy and type 1 diabetic subjects,* American Journal of Physiology. Endocrinology and Metabolism. **300**, E1166 (2011).

[135] *The Management of Diabetic Ketoacidosis in Adults,* 2nd ed. (Joint British Diabetes Societies Inpatient Care Group, 2013).

[136] K. Musa-Veloso, S. S. Likhodii, and S. C. Cunnane, *Breath acetone is a reliable indicator of ketosis in adults consuming ketogenic meals,* The American Journal of Clinical Nutrition **76**, 65 (2002).

[137] *Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia. Report of a WHO/IDF consultation.* (WHO/IDF, Geneva, 2006).

[138] T. M. Wallace and D. R. Matthews, *Recent advances in the monitoring and management of diabetic ketoacidosis,* The Quarterly Journal of Medicine. **97**, 773 (2004).

[139] American Diabetes Association, *Standards of medical care in diabetes – 2014.* Diabetes Care **37**, S14 (2014).

[140] G. L. Blackburn, J. C. Phillips, and S. Morreale, *Physician's guide to popular low-carbohydrate weight-loss diets.* Cleveland Clinic Journal of Medicine **68**, 761 (2001).

[141] E. M. P. Widmark, *Studies in the acetone concentration in blood, urine, and alveolar air. III: The elimination of acetone through the lungs,* Biochemical Journal **14**, 379 (1920).

[142] M. J. Sulway and J. M. Malins, *Acetone in diabetic ketoacidosis,* The Lancet **296**, 736 (1970).

[143] E. van Mastrigt, A. Reyes-Reyes, K. Brand, N. Bhattacharya, H. P. Urbach, A. P. Stubbs, J. C. de Jongste, and M. W. Pijnenburg, *Exhaled breath profiling using broadband quantum cascade laser-based spectroscopy in healthy children and children with asthma and cystic fibrosis,* Journal of Breath Research **10**, 26003 (2016).

[144] W. Miekisch, J. K. Schubert, and G. F. E. Noeldge-schomburg, *Diagnostic potential of breath analysis — focus on volatile organic compounds,* Clinica Chimica Acta **347**, 25 (2004).

[145] A. Smolinska, E. M. M. Klaassen, J. W. Dallinga, K. D. G. van de Kant, Q. Jobsis, E. J. C. Moonen, O. C. P. van Schayck, E. Dompeling, and F. J. van Schooten, *Profiling of volatile organic compounds in exhaled breath as a strategy to find early predictive signatures of asthma in children,* PLoS ONE **9**, e95668 (2014).

[146] M. Barker, M. Hengst, J. Schmid, H. J. Buers, B. Mittermaier, D. Klemp, and R. Koppmann, *Volatile organic compounds in the exhaled breath of young patients with cystic fibrosis,* European Respiratory Journal **27**, 929 (2006).

[147] M. Caldeira, A. S. Barros, M. J. Bilelo, A. Parada, J. S. Câmara, and S. M. Rocha, *Profiling allergic asthma volatile metabolic patterns using a headspace-solid phase microextraction/gas chromatography based methodology.* Journal of chromatography A **1218**, 3771 (2011).

[148] M. Caldeira, R. Perestrelo, a. S. Barros, M. J. Bilelo, a. Morête, J. S. Câmara, and S. M. Rocha, *Allergic asthma exhaled breath metabolome: A challenge for comprehensive two-dimensional gas chromatography,* Journal of Chromatography A **1254**, 87 (2012).

[149] K. D. G. van de Kant, J. J. B. N. van Berkel, Q. Jöbsis, V. Lima Passos, E. M. M. Klaassen, L. van der Sande, O. C. P. van Schayck, J. C. de Jongste, F. J. van Schooten, E. Derks, E. Dompeling, and J. W. Dallinga, *Exhaled breath profiling in diagnosing wheezy preschool children,* European Respiratory Journal **41**, 183 (2013).

[150] J. W. Dallinga, C. M. H. H. T. Robroeks, J. J. B. N. van Berkel, E. J. C. Moonen, R. W. L. Godschalk, Q. Jöbsis, E. Dompeling, E. F. M. Wouters, and F. J. van Schooten, *Volatile organic compounds in exhaled breath as a diagnostic tool for asthma in children.* Clinical & Experimental Allergy **40**, 68 (2010).

[151] E. van Mastrigt, J. C. de Jongste, and M. W. Pijnenburg, *The analysis of volatile organic compounds in exhaled breath and biomarkers in exhaled breath condensate in children - clinical tools or scientific toys?* Clinical & Experimental Allergy , 1170 (2014).

[152] M. R. Miller, J. Hankinson, V. Brusasco, F. Burgos, R. Casaburi, a. Coates, R. Crapo, P. Enright, C. P. M. van der Grinten, P. Gustafsson, R. Jensen, D. C. Johnson, N. MacIntrye, R. McKay, D. Navajas, O. F. Pedersen, R. Pellegrino, G. Viegi, and J. Wagner, *Standardisation of spirometry,* European Respiratory Journal **26**, 319 (2005).

[153] M. I. Asher, U. Keil, H. R. Anderson, R. Beasley, J. Crane, F. Martinez, E. A. Mitchell, N. Pearce, B. Sibbald, and A. W. Stewart, *International study of asthma and allergies in childhood (ISAAC): Rationale and methods.* European Respiratory Journal **8**, 483 (1995).

[154] Global Initiative for Asthma, *Global strategy for asthma management and prevention. 2015 update* (2015) available from: http://www.ginasthma.org.

[155] P. H. Quanjer, D. J. Brazzale, P. W. Boros, and J. J. Pretto, *Implications of adopting the Global Lungs Initiative 2012 all-age reference equations for spirometry,* European Respiratory Journal **42**, 1046 (2013).

[156] W. Nystad, *Daycare attendance, asthma and atopy,* Annals of Medicine **32**, 390 (2000).

[157] California environmental protection agency. Air Resources Board, *Method 422. Determination of volatile organic compounds in emissions from stationary sources,* (1991), available from: http://www.arb.ca.gov/testmeth/vol3/meth422.pdf.

[158] NIST/SEMATECH, *Detection of outliers,* in *e-Handbook of Statistical Methods,* http://www.itl.nist.gov/div898/handbook/eda/section3/eda35h.htm [Accessed: 14-01-2016].

[159] B. Iglewicz and D. Hoaglin, *How to Detect and Handle Outliers,* Volume 16 of the ASQC basic references in quality control: Statistical techniques (ASQC Quality Press, Milwaukee, Wis., 1993).

[160] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, *Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median,* Journal of Experimental Social Psychology **49**, 764 (2013).

[161] M. R. Spiegel, J. J. Schiller, and R. A. Srinivasan, *Schaum's outline of probability and statistics*, 3rd ed. (McGraw-Hill, 2009).

[162] D. G. Altman and J. M. Bland, *Measurement in medicine: The analysis of method comparison studies,* The Statistician **32**, 307 (1983).

[163] J. M. Bland and D. G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet **1**, 307 (1986).

[164] I. T. Jolliffe, *Principal component analysis*, 2nd ed., Springer Series in Statistics (Springer-Verlag, New York, 2002).

[165] O. Grouwstra, *Breath analysis by quantum cascade laser spectroscopy*, Master thesis, Delft University of Technology (2015).

[166] Y. Benjamini and Y. Hochberg, *Controlling the false discovery rate: A practical and powerful approach to multiple testing,* Journal of the Royal Statistical Society. Series B **57**, 289 (1995).

[167] R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. H. Yang, and J. Zhang, *Bioconductor: Open software development for computational biology and bioinformatics.* Genome Biology **5**, R80 (2004).

[168] B. de Lacy Costello, A. Amann, H. Al-Kateb, C. Flynn, W. Filipiak, T. Khalid, D. Osborne, and N. M. Ratcliffe, *A review of the volatiles from the healthy human body.* Journal of Breath Research **8**, 014001 (2014).

[169] C. Phillips, N. Mac Parthaláin, Y. Syed, D. Deganello, T. Claypole, and K. Lewis, *Short-term intra-subject variation in exhaled volatile organic compounds (VOCs) in COPD patients and healthy controls and its effect on disease classification.* Metabolites **4**, 300 (2014).

[170] T. Paff, M. P. van der Schee, J. M. A. Daniels, G. Pals, P. E. Postmus, P. J. Sterk, and E. G. Haarman, *Exhaled molecular profiles in the assessment of cystic fibrosis and primary ciliary dyskinesia,* Journal of Cystic Fibrosis **12**, 454 (2013).

[171] C. M. Robroeks, J. J. van Berkel, Q. Jöbsis, F. J. van Schooten, J. W. Dallinga, E. F. Wouters, and E. Dompeling, *Exhaled volatile organic compounds predict exacerbations of childhood asthma in a 1-year prospective study,* European Respiratory Journal **42**, 98 (2013).

[172] C. M. H. H. T. Robroeks, J. J. B. N. van Berkel, J. W. Dallinga, Q. Jöbsis, L. J. I. Zimmermann, H. J. E. Hendriks, M. F. M. Wouters, C. P. M. van Der Grinten, K. D. G. van de Kant, F. J. van Schooten, and E. Dompeling, *Metabolomics of volatile organic compounds in cystic fibrosis patients and controls,* Pediatric Research **68**, 75 (2010).

[173] N. Fens, A. H. Zwinderman, M. P. van der Schee, S. B. de Nijs, E. Dijkers, A. C. Roldaan, D. Cheung, E. H. Bel, and P. J. Sterk, *Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma.* American Journal of Respiratory and Critical Care Medicine **180**, 1076 (2009).

[174] L. Fleming, L. Tsartsali, N. Wilson, N. Regamey, and A. Bush, *Sputum inflammatory phenotypes are not stable in children with asthma,* Thorax **67**, 675 (2012).

[175] S. Dragonieri, R. Schot, B. J. Mertens, S. Le Cessie, S. A. Gauw, A. Spanevello, O. Resta, N. P. Willard, T. J. Vink, K. F. Rabe, E. H. Bel, and P. J. Sterk, *An electronic nose in the discrimination of patients with asthma and controls,* Journal of Allergy and Clinical Immunology **120**, 856 (2007).

[176] Z. Lazar, N. Fens, J. van der Maten, M. P. van der Schee, A. H. Wagener, S. B. de Nijs, E. Dijkers, and P. J. Sterk, *Electronic nose breathprints are independent of acute changes in airway caliber in asthma,* Sensors **10**, 9127 (2010).

[177] J. J. B. N. van Berkel, J. W. Dalling, G. M. Moeller, R. W. L. Godschalk, E. Moonen, E. F. M. Wouters, and F. J. van Schooten, *Development of accurate classification method based on the analysis of volatile organic compounds from human exhaled air,* Journal of Chromatography B **861**, 101 (2008).

[178] M. H. P. Murtz, *Cavity ring-down spectroscopy for medical applications,* in *Cavity Ring-Down Spectroscopy: Techniques and Applications,* edited by G. Berden and R. Engeln (Wiley, Wiley, 2009) pp. 213–236.

[179] J. H. Bond, R. R. Engel, and M. D. Levitt, *Factors influencing pulmonary methane excretion in man. An indirect method of studying the in situ metabolism of the methane-producing colonic bacteria.* The Journal of Experimental Medicine **133**, 572 (1971).

[180] Y. Peled, T. Gilat, E. Liberman, and Y. Bujanover, *The development of methane production in childhood and adolescence,* Journal of Pediatric Gastroenterology and Nutrition **4**, 575 (1985).

[181] K. Dryahina, D. Smith, and P. Spanel, *Quantification of methane in humid air and exhaled breath using selected ion flow tube mass spectrometry,* Rapid Ccommunications in Mass Spectrometry **24**, 1296 (2010).

[182] B. P. J. de Lacy Costello, M. Ledochowski, and N. M. Ratcliffe, *The importance of methane breath testing: A review.* Journal of Breath Research **7**, 024001 (2013).

[183] M. R. McCurdy, Y. Bakhirkin, G. Wysocki, R. Lewicki, and F. K. Tittel, *Recent advances of laser-spectroscopy-based techniques for applications in breath analysis,* Journal of Breath Research **1**, 14001 (2007).

[184] F. M. Schmidt, O. Vaittinen, M. Metsälä, M. Lehto, C. Forsblom, P. H. Groop, and L. Halonen, *Ammonia in breath and emitted from skin,* Journal of Breath Research **7**, 017109 (2013).

[185] J. Mandon, M. Högman, P. J. F. M. Merkus, J. van Amsterdam, F. J. M. Harren, and S. M. Cristescu, *Exhaled nitric oxide monitoring by quantum cascade laser: comparison with chemiluminescent and electrochemical sensors.* Journal of Biomedical Optics **17**, 017003 (2012).

[186] P. Bergamaschi, C. A. M. Brenninkmeijer, M. Hahn, T. Röckmann, D. H. D. H. Scharffe, P. J. Crutzen, N. F. Elansky, I. B. Belikov, N. B. A. Trivett, and D. E. J. Worthy, *Isotope analysis based source identification for atmospheric CH_4 and CO sampled across Russia using the trans-siberian railroad,* Journal of Geophysical Research **103**, 8227 (1998).

[187] P. Bergamaschi, M. Schupp, and G. W. Harris, *High-precision direct measurements of $^{13}CH_4/^{12}CH_4$ and $^{12}CH_3D/^{12}CH_4$ ratios in atmospheric methane sources by means of a long-path tunable diode laser absorption spectrometer,* Applied Optics **33**, 7704 (1994).

[188] B. R. Miller, R. F. Weiss, P. K. Salameh, T. Tanhua, B. R. Greally, J. Mühle, and P. G. Simmonds, *Medusa: a sample preconcentration and GC/MS detector system for in situ measurements of atmospheric trace halocarbons, hydrocarbons, and sulfur compounds.* Analytical Chemistry **80**, 1536 (2008).

[189] J. Mohn, C. Guggenheim, B. Tuzson, M. K. Vollmer, S. Toyoda, N. Yoshida, and L. Emmenegger, *A liquid nitrogen-free preconcentration unit for measurements of ambient N2O isotopomers by QCLAS,* Atmospheric Measurement Techniques **3**, 609 (2010).

[190] J. Mohn, B. Tuzson, A. Manninen, N. Yoshida, S. Toyoda, W. A. Brand, and L. Emmenegger, *Site selective real-time measurements of atmospheric N_2O isotopomers by laser spectroscopy,* Atmospheric Measurement Techniques **5**, 1601 (2012).

[191] J. Mohn, C. Steinlin, L. Merbold, L. Emmenegger, and F. Hagedorn, *N2O emissions and source processes in snow-covered soils in the Swiss Alps,* Isotopes in Environmental and Health Studies **49**, 520 (2013).

[192] J. Mohn, B. Wolf, S. Toyoda, C. T. Lin, M. C. Liang, N. Brüggemann, H. Wissel, A. E. Steiker, J. Dyckmans, L. Szwec, N. E. Ostrom, K. L. Casciotti, M. Forbes, A. Giesemann, R. Well, R. R. Doucett, C. T. Yarnes, A. R. Ridley, J. Kaiser, and N. Yoshida, *Interlaboratory assessment of nitrous oxide isotopomer analysis by isotope ratio mass spectrometry and laser spectroscopy: Current status and perspectives,* Rapid Communications in Mass Spectrometry **28**, 1995 (2014).

[193] B. Wolf, L. Merbold, C. Decock, B. Tuzson, E. Harris, J. Six, L. Emmenegger, and J. Mohn, *First on-line isotopic characterization of N2O above intensively managed grassland,* Biogeosciences **12**, 2517 (2015).

[194] S. Eyer, B. Tuzson, M. E. Popa, C. van der Veen, T. Röckmann, M. Rothe, W. A. Brand, R. Fisher, D. Lowry, E. G. Nisbet, M. S. Brennwald, E. Harris, C. Zellweger, L. Emmenegger, H. Fischer, and J. Mohn, *Real-time analysis of $\delta 13C$- and δD-CH_4*

in ambient air with laser spectroscopy: method development and first intercomparison results, Atmospheric Measurement Techniques **9**, 263 (2016).

[195] U. Tisch and H. Haick, *Chemical sensors for breath gas analysis: The latest developments at the Breath Analysis Summit 2013,* Journal of Breath Research **8**, 027103 (2014).

[196] T. Hibbard, K. Crowley, and A. J. Killard, *Direct measurement of ammonia in simulated human breath using an inkjet-printed polyaniline nanoparticle sensor.* Analytica Chimica Acta **779**, 56 (2013).

[197] J. T. W. Yeow and Y. Wang, *A review of carbon nanotubes-based gas sensors,* Journal of Sensors **2009** (2009).

[198] R. M. Penner, *Chemical Sensing with Nanowires,* Annual Review of Analytical Chemistry **5**, 461 (2012).

[199] J. Zhang, J. Guo, H. Xu, and B. Cao, *Reactive-template fabrication of porous SnO_2 nanotubes and their remarkable gas-sensing performance,* ACS Applied Materials & Interfaces **5**, 7893 (2013).

[200] J. Guo, J. Zhang, M. Zhu, D. Ju, H. Xu, and B. Cao, *High-performance gas sensor based on ZnO nanowires functionalized by Au nanoparticles,* Sensors and Actuators B **199**, 339 (2014).

[201] M. C. McAlpine, H. Ahmad, D. Wang, and J. R. Heath, *Highly ordered nanowire arrays on plastic substrates for ultrasensitive flexible chemical sensors,* Nature Materials **6**, 379 (2007).

[202] Thorlabs, Inc, *Quantum cascade lasers (QCLs) & Interband cascade lasers (ICLs), 3.42 - 11.00 μm* (2016) https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id =6932 [Accessed: 21-06-2016].

[203] VIGO System S. A., *Detector TO8 - XTE. Technical Drawing* (2015) http://www.vigo.com.pl/pub/File/application_notes/Technical%20Drawings/ TO8.pdf [Accessed: 21-06-2016].

[204] C. M. Charlton, B. T. Thompson, and B. Mizaikoff, *Hollow waveguide infrared spectroscopy and sensing,* in *Frontiers in chemical sensors: Novel principles and techniques*, edited by G. Orellana and M. C. Moreno-Bondi (Springer Berlin Heidelberg, Berlin, 2005) pp. 133–167.

[205] M. L. Alexander, J. F. Kelly, R. L. Sams, J. J. Moran, and M. K. Newburn, *Capillary absorption spectrometer and process for isotopic analysis of small samples,* (2014), US Patent 20140220700 A1.

[206] J. M. Lopez-Higuera, *Introduction to fibre optic sensing technology,* in *Handbook of optical fibre sensing technology*, edited by J. M. Lopez-Higuera (Wiley, 2002) pp. 1–22.

[207] J. F. Kelly, R. L. Sams, T. A. Blake, and J. M. Kriesel, *Further developments of capillary absorption spectrometers using small hollow-waveguide fibers,* Proceedings of SPIE **8993**, 89931O (2013).

[208] C. Charlton, F. de Melas, A. Inberg, N. Croitoru, and B. Mizaikoff, *Hollow-waveguide gas sensing with room-temperature quantum cascade lasers,* IEE Proceedings - Optoelectronics **150**, 306 (2003).

[209] J. F. Kelly, R. L. Sams, T. A. Blake, M. Newburn, J. Moran, M. L. Alexander, and H. Kreuzer, *A capillary absorption spectrometer for stable carbon isotope ratio (13C/12C) analysis in very small samples,* Review of Scientific Instruments **83** (2012).

[210] T. Abel, J. Hirsch, and J. A. Harrington, *Hollow glass waveguides for broadband infrared transmission,* Optics Letters **19**, 1034 (1994).

[211] C. M. Bledt, J. A. Harrington, and J. M. Kriesel, *Loss and modal properties of Ag/AgI hollow glass waveguides,* Applied Optics **51**, 3114 (2012).

[212] S. Sato, M. Saito, and M. Miyagi, *Infrared hollow waveguides for capillary flow cells,* Applied Spectroscopy **47**, 1665 (1993).

[213] R. L. Kozodoy, R. H. Micheels, and J. A. Harrington, *Small-bore hollow waveguide infrared absorption cells for gas sensing,* Applied Spectroscopy **50**, 415 (1996).

[214] S. J. Saggese, J. A. Harrington, and G. H. Sigel, *Attenuation of incoherent infrared radiation in hollow sapphire and silica waveguides,* Optics Letters **16**, 27 (1991).

[215] R. H. Micheels, K. Richardson, D. J. Haan, and J. A. Harrington, *FTIR-based instrument employing a coiled hollow waveguide cell for rapid field analysis of volatile organic compounds,* Proceedings of SPIE **3540**, 66 (1999).

[216] Y. Saito, T. Kanaya, A. Nomura, and T. Kano, *Experimental trial of a hollow-core waveguide used as an absorption cell for concentration measurement of NH_3 gas with a CO_2 laser,* Optics Letters **18**, 2150 (1993).

[217] C. A. Worrell and N. A. Gallen, *Trace-level detection of gases and vapours with mid-infra-red hollow waveguides,* Journal of Physics D: Applied Physics **30**, 1984 (1997).

[218] L. Hvozdara, S. Gianordoli, G. Strasser, W. Schrenk, K. Unterrainer, E. Gornik, C. S. S. S. Murthy, M. Kraft, V. Pustogow, B. Mizaikoff, A. Inberg, and N. Croitoru, *Spectroscopy in the gas phase with GaAs/AlGaAs quantum-cascade lasers,* Applied Optics **39**, 6926 (2000).

[219] C. Young, S. S. Kim, Y. Luzinova, M. Weida, D. Arnone, E. Takeuchi, T. Day, and B. Mizaikoff, *External cavity widely tunable quantum cascade laser based hollow waveguide gas sensors for multianalyte detection,* Sensors and Actuators B **140**, 24 (2009).

[220] C. Charlton, B. Temelkuran, G. Dellemann, and B. Mizaikoff, *Midinfrared sensors meet nanotechnology: Trace gas sensing with quantum cascade lasers inside photonic band-gap hollow waveguides,* Applied Physics Letters **86**, 194102 (2005).

[221] N. Gayraud, b. W. Kornaszewski, J. M. Stone, J. C. Knight, D. T. Reid, D. P. Hand, and W. N. MacPherson, *Mid-infrared gas sensing using a photonic bandgap fiber,* Applied Optics **47**, 1269 (2008).

[222] B. T. Thompson, A. Inberg, N. Croitoru, and B. Mizaikoff, *Characterization of a mid-infrared hollow waveguide gas cell for the analysis of carbon monoxide and nitric oxide,* Applied Spectroscopy **60**, 266 (2006).

[223] M. Miyagi, K. Harada, and S. Kawakami, *Wave propagation and attenuation in the general class of circular hollow waveguides with uniform curvature,* IEEE Transactions on Microwave Theory and Techniques **32**, 513 (1984).

[224] Y. Matsuura and M. Miyagi, *Bending losses and beam profiles of zinc selenide-coated silver waveguides for carbon dioxide laser light,* Applied Optics **31**, 6441 (1992).

[225] D. J. Haan and J. A. Harrington, *Hollow waveguides for gas sensing and near-IR applications,* Proceedings of SPIE **3596**, 43 (1999).

[226] J. M. Kriesel, N. Gat, B. E. Bernacki, R. L. Erikson, B. D. Cannon, T. L. Myers, C. M. Bledt, and J. A. Harrington, *Hollow core fiber optics for mid-wave and long-wave infrared spectroscopy,* Proceedings of SPIE **8018**, 80180V (2011).

[227] J. M. Kriesel, G. M. Hagglund, N. Gat, V. Spagnolo, and P. Patimisco, *Spatial mode filtering of mid-infrared (mid-IR) laser beams with hollow core fiber optics,* Proceedings of SPIE **8993**, 89930V (2013).

[228] A. Wilk, J. C. Carter, M. Chrisp, A. M. Manuel, P. Mirkarimi, J. B. Alameda, and B. Mizaikoff, *Substrate-integrated hollow waveguides: A new level of integration in mid-infrared gas sensing,* Analytical Chemistry **85**, 11205 (2013).

[229] J. F. da Silveira Petruci, P. R. Fortes, V. Kokoric, A. Wilk, I. M. Raimundo Jr, A. Alves Cardoso, and B. Mizaikoff, *Real-time monitoring of ozone in air using substrate-integrated hollow waveguide mid-infrared sensors,* Scientific Reports **3**, 3174 (2013).

[230] O. A. Pogodina, V. V. Pustogov, F. de Melas, C. Haberhauer-Troyer, E. Rosenberg, H. Puxbaum, A. Inberg, N. Croitoru, and B. Mizaikoff, *Combination of sorption tube sampling and thermal desorption with hollow waveguide FT-IR spectroscopy for atmospheric trace gas analysis: Determination of atmospheric ethene at the lower ppb level,* Analytical Chemistry **76**, 464 (2004).

[231] C. R. Young, N. Menegazzo, A. E. Riley, C. H. Brons, F. P. Disanzo, J. L. Givens, J. L. Martin, M. M. Disko, and B. Mizaikoff, *Infrared hollow waveguide sensors for simultaneous gas phase detection of benzene, toluene, and xylenes in field environments,* Analytical Chemistry **83**, 6141 (2011).

[232] D. Perez-Guaita, V. Kokoric, A. Wilk, S. Garrigues, and B. Mizaikoff, *Towards the determination of isoprene in human breath using substrate-integrated hollow waveguide mid-infrared sensors,* Journal of Breath Research **8**, 26003 (2014).

[233] J. Yang and J. W. Her, *Development of a hollow waveguide sampler for detection of chlorinated aromatic compounds in soils,* Analytical Chemistry **72**, 878 (2000).

[234] S. Wu, A. Deev, M. Haught, and Y. Tang, *Hollow waveguide quantum cascade laser spectrometer as an online microliter sensor for gas chromatography,* Journal of chromatography A **1188**, 327 (2008).

[235] F. de Melas, V. V. Pustogov, N. Croitoru, and B. Mizaikoff, *Development and optimization of a mid-infrared hollow waveguide gas sensor combined with a supported capillary membrane sampler,* Applied Spectroscopy **57**, 600 (2003).

[236] F. de Melas, V. V. Pustogov, D. K. Wolcott, D. C. Olson, A. Inberg, N. Croitoru, and B. Mizaikoff, *Combination of a mid-infrared hollow waveguide gas sensor with a supported capillary membrane sampler for the detection of organic compounds in water,* International Journal of Environmental Analytical Chemistry **83**, 573 (2003).

[237] J. Hildenbrand, J. Herbst, J. Wöllenstein, and A. Lambrecht, *Explosive detection using infrared laser spectroscopy,* Proceedings of SPIE **7222**, 72220B (2009).

[238] J. da Silveira Petruci, P. R. Fortes, V. Kokoric, A. Wilk, I. M. Raimundo, A. A. Cardoso, and B. Mizaikoff, *Monitoring of hydrogen sulfide via substrate-integrated hollow waveguide mid-infrared sensors in real-time,* Analyst **139**, 198 (2014).

[239] *Compound Annual Growth Rate - CAGR* (Investopedia, LLC) http://www.investopedia.com/terms/c/cagr.asp [Accessed: 26-04-2016].

[240] E. T. P. Photonics21, *Strategic Research Agenda Lighting the way ahead* (2010).

[241] European Technology Platform Photonics21, *Towards 2020 – Photonics driving economic growth in Europe* (Brussels, 2013).

[242] E. T. Curtiss and S. Eustis, *Mid ir sensors: Market shares, strategies, and forecasts, worldwide, 2013 to 2019* (WinterGreen Research, Inc., Lexington, Mass., 2013).

[243] K. H. Kim, S. A. Jahan, and E. Kabir, *A review of breath analysis for diagnosis of human health,* Trends in Analytical Chemistry **33**, 1 (2012).

[244] J. P. Gisbert and J. M. Pajares, *Review article: ^{13}C-urea breath test in the diagnosis of Helicobacter pylori infection – A critical review.* Alimentary Pharmacology & Therapeutics **20**, 1001 (2004).

[245] Z. q. Xu, Y. Y. Broza, R. Ionsecu, U. Tisch, L. Ding, H. Liu, Q. Song, Y.-y. Pan, F.-x. Xiong, K.-s. Gu, G.-p. Sun, Z.-d. Chen, M. Leja, and H. Haick, *A nanomaterial-based breath test for distinguishing gastric cancer from benign gastric conditions,* British Journal of Cancer **108**, 941 (2013).

[246] M. Phillips, R. N. Cataneo, A. R. C. Cummin, A. J. Gagliardi, K. Gleeson, J. Greenberg, R. A. Maxfield, and W. N. Rom, *Detection of lung cancer with volatile markers in the breath,* Chest **123**, 2115 (2003).

[247] H. Yu and P. Wang, *SPME-GC analysis of alkanes and aromatic hydrocarbons in human breath to screen lung cancer,* IFMBE Proceedings **19**, 396 (2008).

[248] S. Kischkel, W. Miekisch, A. Sawacki, E. M. Straker, P. Trefz, A. Amann, and J. K. Schubert, *Breath biomarkers for lung cancer detection and assessment of smoking related effects – confounding variables, influence of normalization and statistical algorithms.* Clinica Chimica Acta **411**, 1637 (2010).

[249] J. Rudnicka, T. Kowalkowski, T. Ligor, and B. Buszewski, *Determination of volatile organic compounds as biomarkers of lung cancer by SPME-GC-TOF/MS and chemometrics.* Journal of Chromatography B **879**, 3360 (2011).

[250] *Breath analyzers market by technology (fuel cell, semiconductor, infrared, crystal), by application (alcohol, drug abuse, tuberculosis, asthma, cancer), by end user (law enforcement, enterprises, individuals) - Analysis & global forecast to 2019* (Markets and Markets, 2015).

[251] Global Asthma Network, *The global asthma report 2014*, Vol. 5 (Auckland, New Zealand, 2014).

[252] Global Initiative for Asthma, *Global strategy for asthma management and prevention 2016* (2016) available from: http://www.ginasthma.org.

[253] World Health Organization, *WHO | Asthma* (2015) http://www.who.int/ respiratory/asthma/en/ [Accessed: 25-04-2016].

[254] World Health Organization, *Global surveillance, prevention and control of chronic respiratory diseases: A comprehensive approach.*, edited by J. Bousquet and N. Khaltaev (Geneva, 2007).

[255] C. A. Akdis and I. Agache, eds., *Global atlas of asthma* (European Academy of Allergy and Clinical Immunology, 2013).

[256] J. Berg and P. Lindgren, *Economic evaluation of FENO measurement in diagnosis and 1-year management of asthma in Germany,* Respiratory Medicine **102**, 219 (2008).

[257] S. E. Harnan, P. Tappenden, M. Essat, T. Gomersall, J. Minton, R. Wong, I. Pavord, M. Everard, and R. Lawson, *Measurement of exhaled nitric oxide concentration in asthma: a systematic review and economic evaluation of NIOX MINO, NIOX VERO and NObreath,* Health Technology Assessment **19**, 1 (2015).

[258] Circassia, *Preliminary Results Presentation,* (2016), Available from: http://www.circassia.com/media/presentations/ [Accessed: 09-04-2016].

[259] D. Conteduca, F. Dell'Olio, C. Ciminelli, and M. N. Armenise, *New miniaturized exhaled nitric oxide sensor based on a high Q/V mid-infrared 1D photonic crystal cavity,* Applied Optics **54**, 2208 (2015).

[260] W. H. Organization, *Global status report on alcohol and health* (Geneva, 2014).

[261] C. Peek-Asa, *The effect of random alcohol screening in reducing motor vehicle crash injuries,* American Journal of Preventive Medicine **16**, 57 (1999).

[262] R. W. Elder, R. a. Shults, D. a. Sleet, J. L. Nichols, S. Zaza, and R. S. Thompson, *Effectiveness of sobriety checkpoints for reducing alcohol-involved crashes,* Traffic Injury Prevention **3**, 266 (2002).

[263] T. F. Babor, *Alcohol: No ordinary commodity - A summary of the second edition,* Addiction **105**, 769 (2010).

[264] P. Anderson and B. Baumberg, *Alcohol in Europe* (Institute of Alcohol Studies, London, 2006).

[265] D. Chisholm, J. Rehm, M. V. Ommeren, and M. Monteiro, *Reducing the global burden of hazardous alcohol use: A comparative cost-effectiveness analysis.* Journal of Studies on Alcohol **65**, 782 (2004).

[266] NHTSA's National Center for Statistics and Analysis, *Traffic Safety Facts 2014 data: alcohol-impaired driving.* (Washington, DC, 2015).

[267] SafetyNet, *Alcohol* (2009) http://ec.europa.eu/transport/road_safety/specialist/knowledge/pdf/alcohol.pdf [Accessed: 08-04-2016].

[268] World Health Organization, *Global status report on road safety 2015* (Geneva, 2015).

[269] E. T. Curtiss and Susan Eustis, *Breath analyzer market shares, strategies, and forecasts, worldwide, 2012 to 2018* (WinterGreen Research, Inc, Lexington, Mass., 2012).

APPENDICES

A. MULTIPASS CELL COMPONENTS

The multipass cell uses commercially available components and custom - made parts to hold the astigmatic mirrors in place. Figure A.1 shows the components of the multipass cell before assembly.

Figure A.1: Multipass cell components.

The commercial vacuum components used in the multipass cell are listed in Table A.1.

Table A.1: Commercial components used in the multipass cell.

Part number	Component	Quantity
212-100N	Bored Flange, Non-Rotatable with bore .938", DN25CF	1
450-212-100Z	Zero Length Reducer Flange DN64CF/25CF	1
2NRC-450-275	Conical Reducer Nipple 1 Flange Rotatable DN64CF/DN40CF	2
3TR-150-075	Reducer Tee, 2-Flanges Rotatable DN40CF/DN19CF	2
450-275-150Z	Zero Length Reducer Flange DN64CF/40CF	1
212-000N	Blank Flange, Non-Rotatable DN25CF	1
133-STV-1/4SWAG	Flange with up to air valve 1/4" Swagelok, DN19CF	2

In Figure A.2 we present a picture of the astigmatic mirrors used in the multipass cell. The hole in the center of the input mirror is used to couple the laser light inside the cavity.
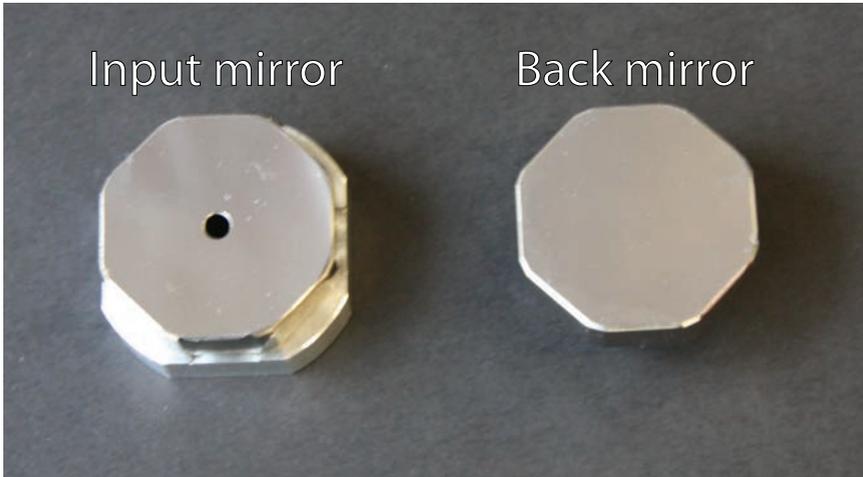
Figure A.2: Astigmatic mirrors.

We fabricated supports for each of the mirrors. In Figure A.3 we show a picture of the supports.
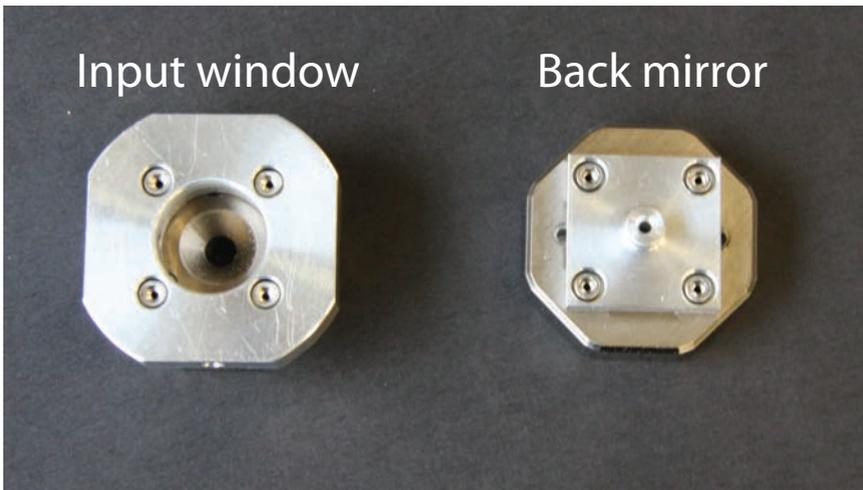
Figure A.3: Mirrors support attached to the rear of the astigmatic mirrors.

The details of the supports are presented in Figures A.4 and A.5.
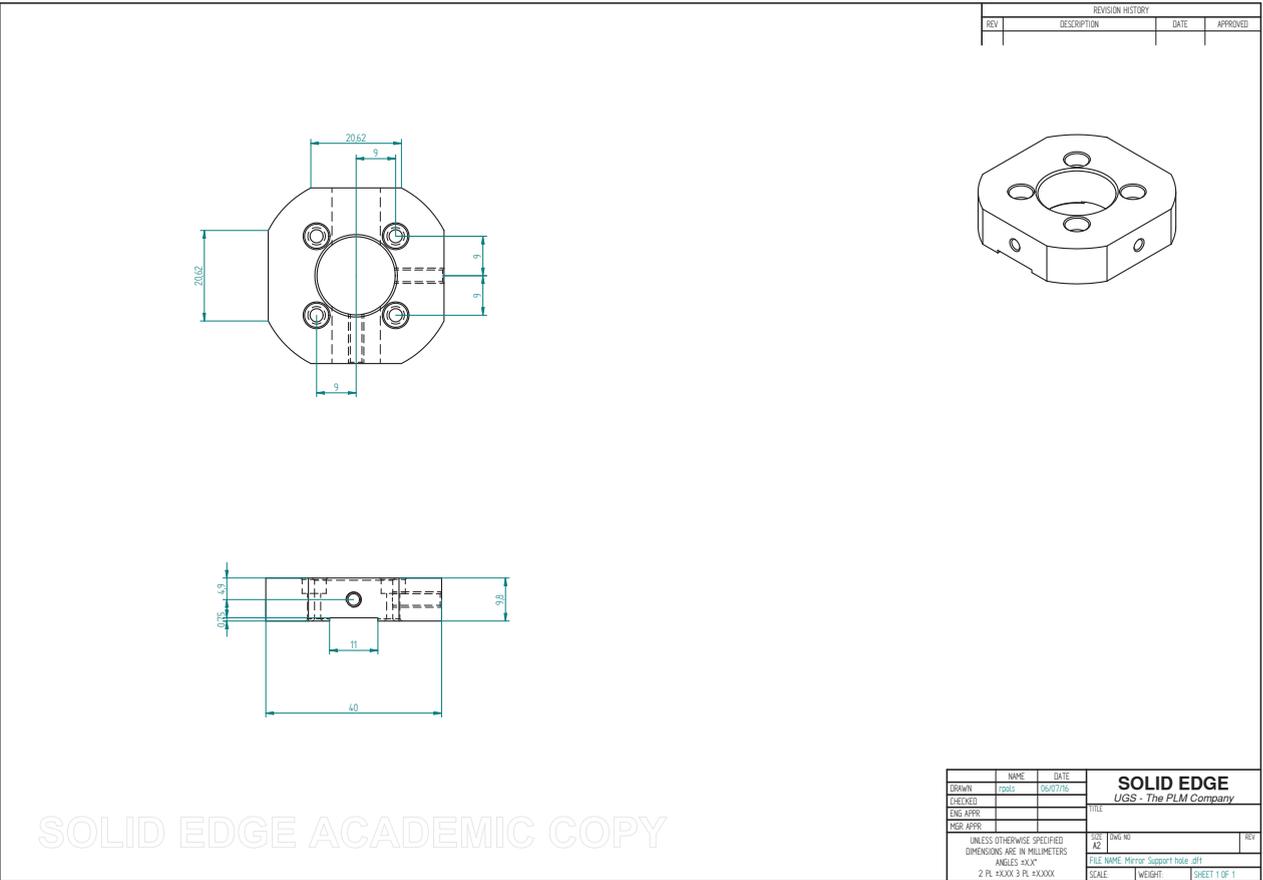
Figure A.4: Details of the support of the mirror.

REVISION HISTORY

REV	DESCRIPTION	DATE	APPROVED

	NAME	DATE
DRAWN	rpols	06/07/16
CHECKED		
ENG APPR		
MGR APPR		

SOLID EDGE
UGS - The PLM Company

TITLE

SIZE A2	DWG NO	REV

FILE NAME: Mirror Support .dft

| SCALE: | WEIGHT: | SHEET 1 OF 1 |

UNLESS OTHERWISE SPECIFIED
DIMENSIONS ARE IN MILLIMETERS
ANGLES ±X.X°
2 PL ±X.XX 3 PL ±X.XXX
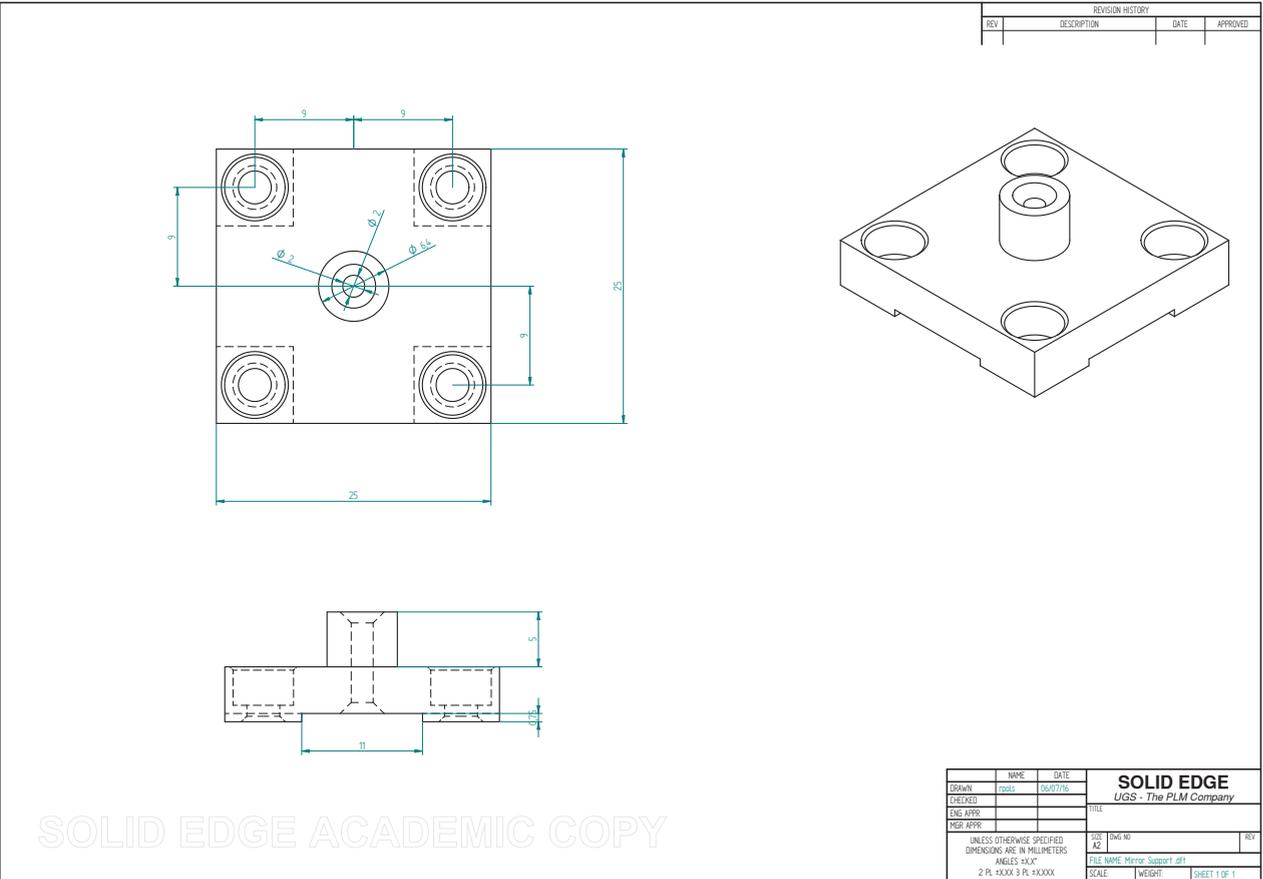
Custom mounts were built to hold the mirrors in place. The input mirror does not required any adjustment, we just have to keep a clear aperture for the introduction of the laser light. In Figure A.6 we show a picture of the mount

Figure A.6: Picture of the input mirror holder.

In Figure A.7 we can observe the position of the mirror holder and the mirror mount before they are joined and secured.

Figure A.7: Picture of the input mirror holder.

The details of the mount are presented in Figure A.8
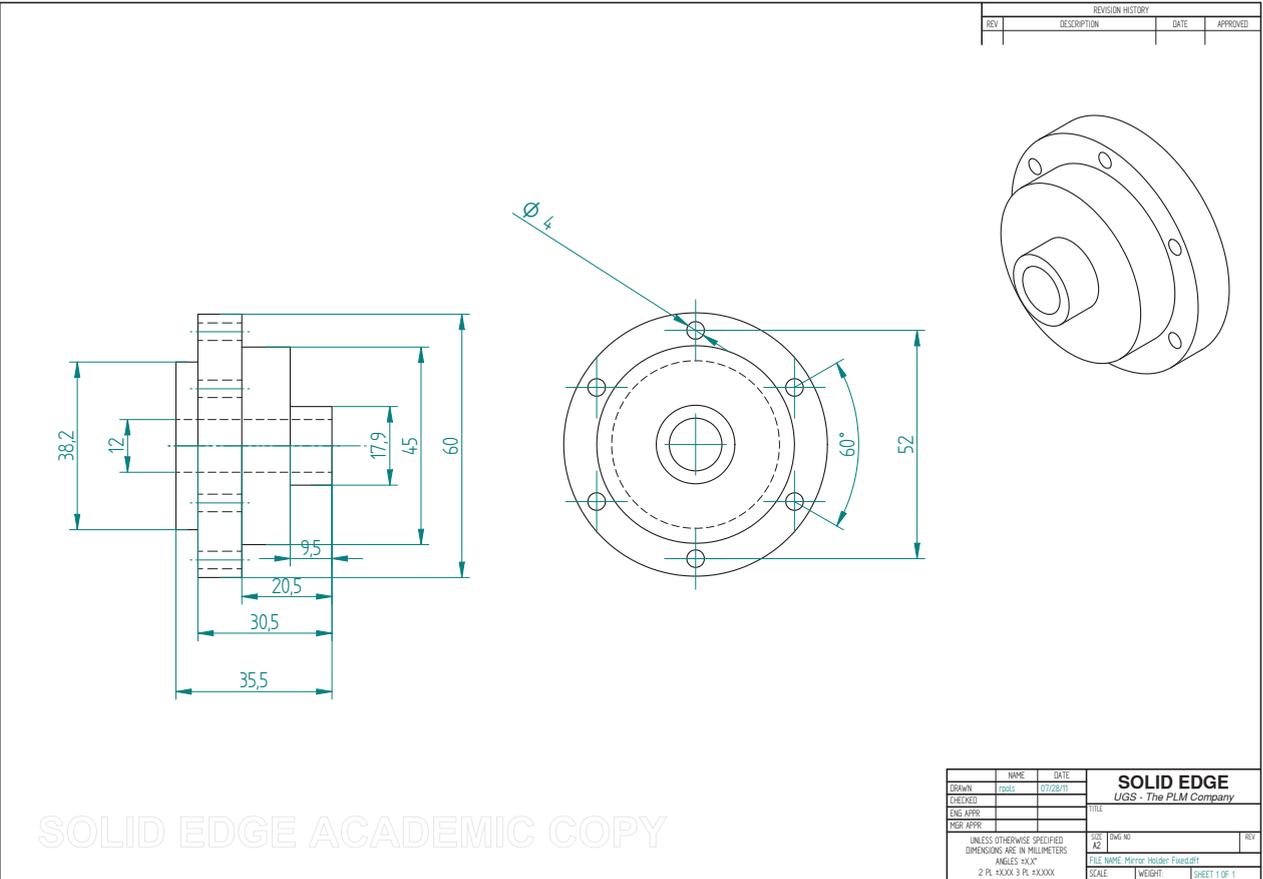
Figure A.8: Details of the support of the mirror.

The mount of the back mirror includes a commercial tilt platform (Newport, U50-PL1) to align the multipass cell mirrors. The platform is encase on a custom made rotational mount that allows to adjust the relative angular position of the principal axis of the astigmatic mirrors. The rotational mount consist of three parts: a reducer, the rotation stage that encases the Newport platform and a lock. In Figure A.9 we show a picture of the component of the rotational mount.

Figure A.9: Picture of the input mirror holder.

In Figure A.10 we can see how the rotational mount should be assembled.

Figure A.10: Picture of the input mirror holder.

The details of the three parts of the rotational mount are presented in Figures A.11, A.12 and A.13
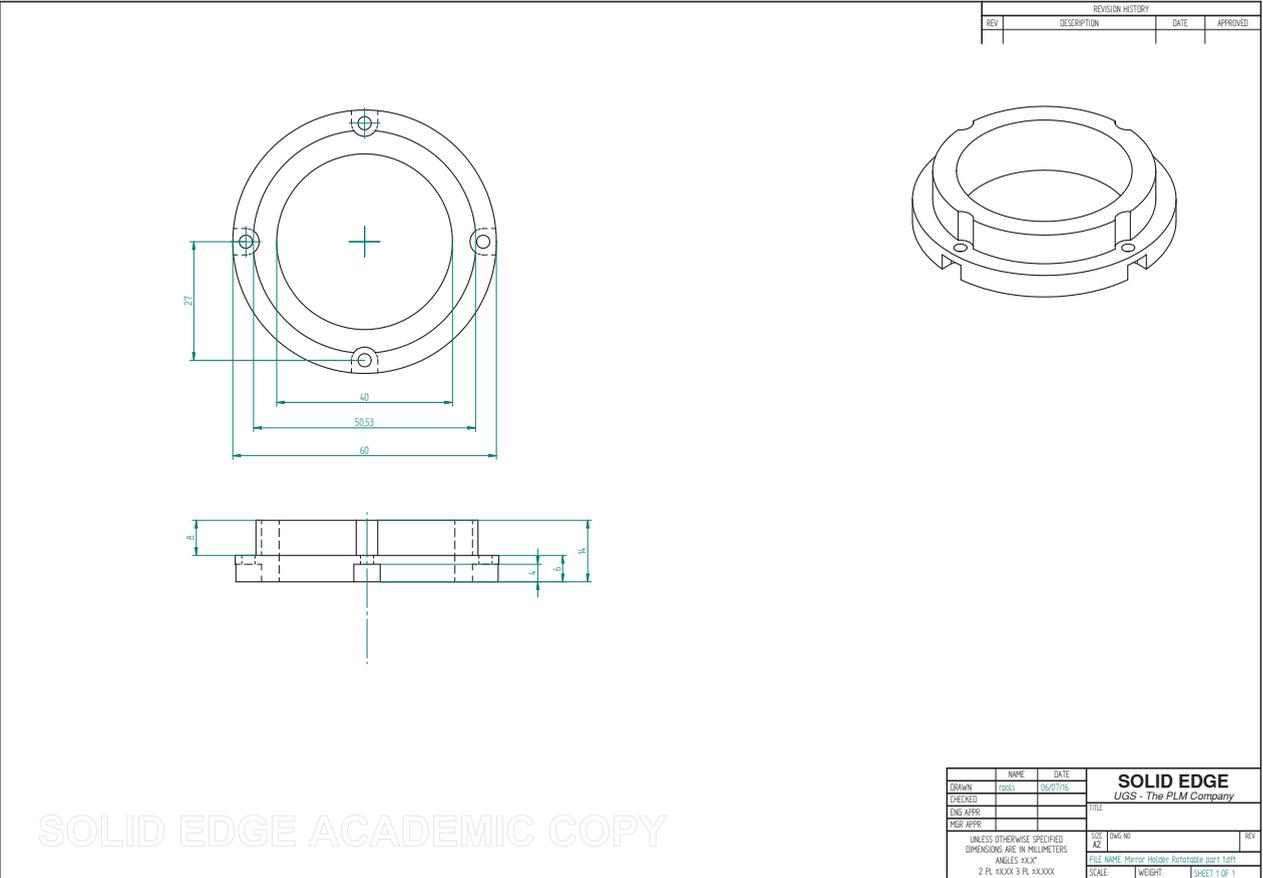
Figure A.11: Details of the reducer of the rotational mount.

	NAME	DATE	SOLID EDGE
DRAWN	rpals	06/07/16	UGS - The PLM Company
CHECKED			
ENG APPR			
MGR APPR			

REVISION HISTORY

REV	DESCRIPTION	DATE	APPROVED

UNLESS OTHERWISE SPECIFIED
DIMENSIONS ARE IN MILLIMETERS
ANGLES ±X.X°
2 PL ±X.XX 3 PL ±X.XXX

TITLE

SIZE A2   DWG NO   REV

FILE NAME: Mirror Holder Rotatable part 1.dft

SCALE: WEIGHT: SHEET 1 OF 1

Figure A.12: Details of the support that holds the Newport platform in the rotational mount.

	NAME	DATE	**SOLID EDGE**		
DRAWN	rpals	06/07/16	UGS - The PLM Company		
CHECKED			TITLE		
ENG APPR					
MGR APPR					

REVISION HISTORY

REV	DESCRIPTION	DATE	APPROVED

UNLESS OTHERWISE SPECIFIED
DIMENSIONS ARE IN MILLIMETERS
ANGLES ±X.X°
2 PL ±X.XXX 3 PL ±X.XXXX

SIZE A2 | DWG NO | REV

FILE NAME: Mirror Holder Rotatable part 2.dft

SCALE: | WEIGHT: | SHEET 1 OF 1

Figure A.13: Details of the lock of the rotational mount.

At the input of the multipass cell we place a Zinc Selenide (ZnSe) window with anti-reflective coating from 7 to 12 μm (Thorlabs WG71050-G) to let the laser light in. The ZnSe window is hold using a custom - made window flange. In Figure A.14 we show the window flange and the commercial components used for the input port of the multipass cell

Figure A.14: Flange holding for the ZnSe input window.

In Figure A.15 we show the ZnSe window mounted in the multipass cell.
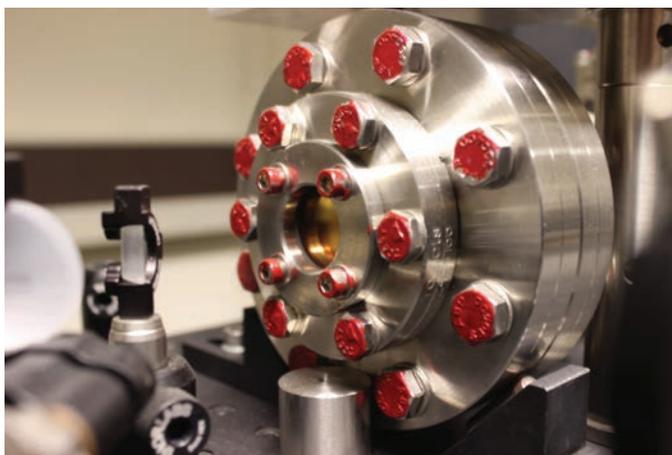
Figure A.15: ZnSe input window mounted in the multipass cell.

The specifications of the window flange built in our lab are depicted in Figure A.16.
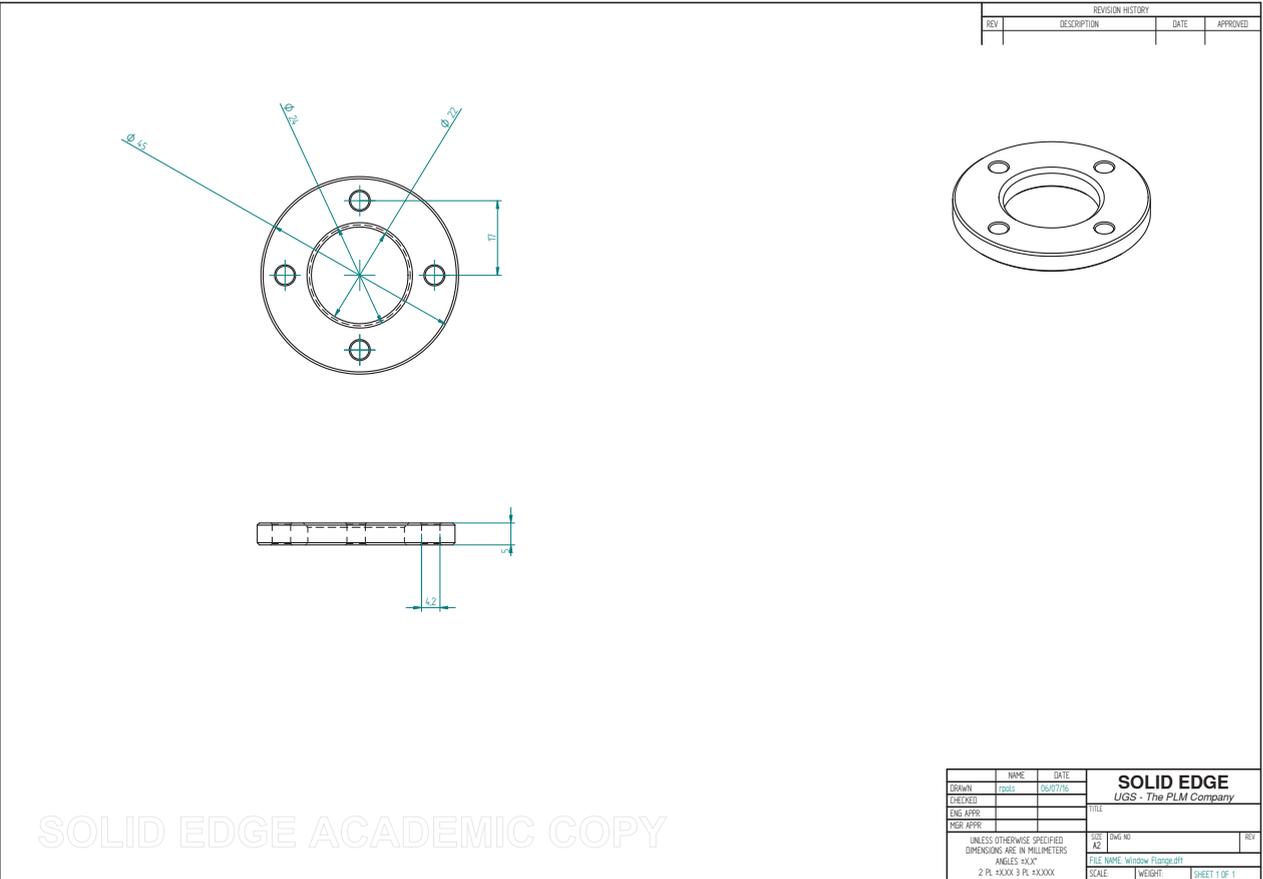
Figure A.16: Flange holding the ZnSe input window.

	REVISION HISTORY		
REV	DESCRIPTION	DATE	APPROVED

	NAME	DATE
DRAWN	rpals	06/07/16
CHECKED		
ENG APPR		
MGR APPR		

SOLID EDGE
UGS - The PLM Company

TITLE

UNLESS OTHERWISE SPECIFIED
DIMENSIONS ARE IN MILLIMETERS
ANGLES ±X.X°
2 PL ±X.XX 3 PL ±X.XXX

SIZE A2	DWG NO		REV
	FILE NAME: Window Flange.dft		
SCALE:	WEIGHT:	SHEET 1 OF 1	

B. LASERSCOPE COMMANDS FOR REMOTE CONTROL

The Laser Scope is set to a remote control mode and is controlled by LabVIEW. The connection address is 127.0.0.1 and the remote port is 2055. After the connection is built, we can directly send commands from LabVIEW to control the laser. The commands and their descriptions are listed in table B.1.

Table B.1: Laser Scope commands

Command	Description
Background <coadds>	Acquire background with the specified number of coadds.
On <wavenum>	Turn on laser at <wavenum>. (ex. "On 1020", "On 1011.2")
Off	Turn off laser
Ptr <0,1>	Visible Pointer on/off. (ex. "Ptr 1", "Ptr 0")
Range	Display available wavenumber range.
Scan <coadds> [NOREFBG]	Scan with the specified number of coadds. Optionally specify NOREFBG to acquire single beam only. (ex. "Scan 1", "Scan 1 NOREFBG")
Sync-Offset <offset in ns>	Sync-Out Offset in multiples of 16 ns (ex. Sync-Offset 16", "Sync-Offset 32)
Trigger <0,1>	Trigger <0=internal, 1=external>. (ex. "Trigger 1", "Trigger 0")
Version	Display version info
Help or ?	Show this help text

C. FLOW CHART OF THE DATA ACQUISITION PROCESS

The Flow chart chart below depicts the full data acquisition process implemented to capture the data recorded by the MCT detectors of our system. It was implemented using LabVIEW.
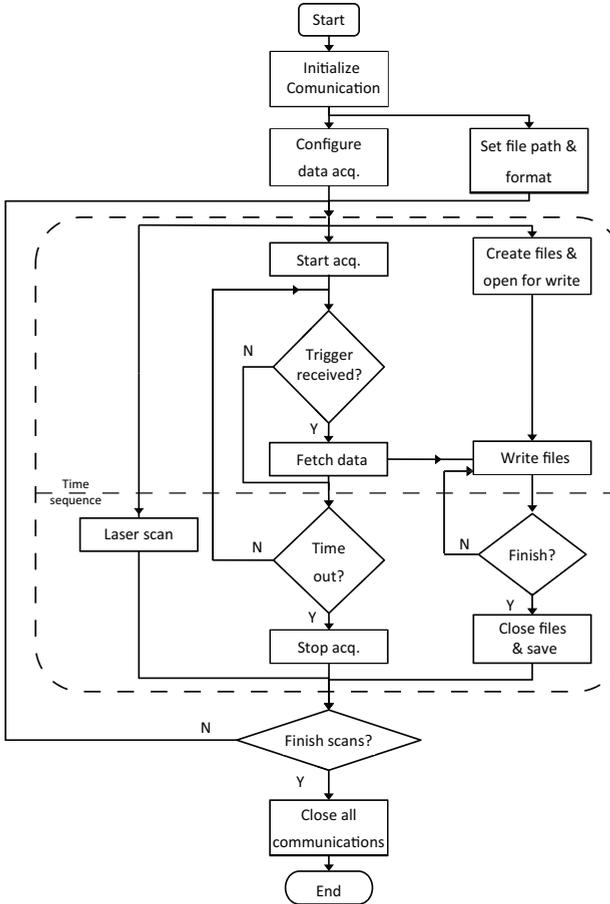
Figure C.1: Flow chart of the LabVIEW data acquisition program.

D. SUSPECT MOLECULES

In Chapter 6 we found the regions where the exhaled breath profile of healthy children, children with asthma and children with cystic fibrosis presents significant differences. In Tables D.1 and D.2 we present the molecules with absorption in these regions.

Table D.1: Molecules with absorption in the significantly different regions found from comparing the exhaled breath of healthy children and children with asthma. * VOCs that have been identified before in exhaled breath.

Compounds		Wavenumber region (cm^{-1})	
CAS Number	Name	1181.80 - 1182.55	1261.40 - 1262.05
354-25-6	1-Chloro-1,1,2,2-tetrafluoroethane (R124A)	✓	✓
75-68-3	1,1,1-Chlorodifluoroethane (Freon-142B)	✓	
421-50-1	1,1,1-Trifluoroacetone	✓	
420-46-2	1,1,1-Trifluoroethane (R143A)		✓
431-89-0	1,1,1,2,3,3,3-Heptafluoropropane (HFC227EA)		✓
124-73-2	1,2-Dibomotetrafluoroethane (Freon-114B2)	✓	
106-93-4	1,2-Dibromoethane (EDB)	✓	
354-23-4	1,2-Dichloro-1,1,2-trifluoroethane (F132A)	✓	
123-91-1 *	1,4-Dioxane		✓
75-88-7	2-Chloro-1,1,1-trifluoroethane (R133A)		✓
2837-89-0	2-Chlroro-1,1,1,2-tetrafluoroethane (HCFC-124)	✓	
111-15-9 *	2-Ethoxyethyl acetate		✓
138495-42-8	2H, 3H-Perfluoropentane	✓	✓
306-83-2	2,2-Dichloro-1,1,1-trifluoroethane (Freon-123)	✓	
76-11-9	2,2-Difluorotetrachloroethane (F112A)	✓	
75-89-8	2,2,2-Trifluoroethanol	✓	✓
25256-77-4	2,2,4-Trimethyl-1,3-pentanediol isobutyrate (Texanol)	✓	✓

Table D.1: (continued)

CAS Number	Name	1181.80 - 1182.55	1261.40 - 1262.05
64-19-7 *	Acetic acid	✓	
75-86-5	Acetone cyanohydrin	✓	
123-86-4 *	Butyl acetate		✓
107-92-6 *	Butyric acid	✓	
353-50-4	Carbonyl fluoride		✓
107-30-2	Chloromethyl methyl ether	✓	
76-15-3	Chloropentafluoroethane (R115)	✓	
1189-71-5	Chlorosulfonyl isocyanate (CSI)	✓	
75-09-2 *	Dichloromethane		✓
108-20-3	Diisopropyl ether	✓	
108-18-9	Diisopropylamine	✓	
616-38-6 *	Dimethyl carbonate		✓
115-10-6 *	Dimethyl ether	✓	
79-44-7	Dimethylcarbamoyl chloride		✓
34590-94-8 *	Dipropylene glycol methyl ether	✓	✓
141-78-6 *	Ethyl acetate		✓
140-88-5 *	Ethyl acrylate	✓	✓
105-54-4 *	Ethyl butyrate	✓	✓
541-41-3	Ethyl chloroformate	✓	
109-94-4 *	Ethyl formate	✓	
383-63-1	Ethyl trifluoroacetate	✓	
76-13-1	Freon-113 (1,1,2-Trichlorortrifluoroethane)	✓	
76-14-2	Freon-114 (1,2-dichlorortetrafluoroethane)	✓	✓
811-97-2	Freon-134a (1,1,1,2-tetrafluoroethane)	✓	
76-19-7	Freon-218 (octafluoropropane)		✓
115-25-3	Freon-C318 (octafluorocyclobutane)		✓
684-16-2	Hexafluoroacetone		✓
76-16-4	Hexafluoroethane (Freon-116)		✓
381-10-5	Hexafluoroisobutylene	✓	✓
116-15-4	Hexafluoropropene	✓	

Table D.1: (continued)

Compounds		Wavenumber region (cm^{-1})	
CAS Number	Name	1181.80 - 1182.55	1261.40 - 1262.05
110-19-0 *	Isobutyl acetate		✓
108-21-4 *	Isopropyl acetate		✓
108-39-4 *	m-Cresol	✓	
124-63-0	Methanesulfonyl chloride	✓	
79-20-9 *	Methyl acetate		✓
96-33-3	Methyl acrylate	✓	✓
93-58-3	Methyl benzoate	✓	✓
107-31-3 *	Methyl formate	✓	
547-63-7	Methyl isobutyrate	✓	✓
80-62-6 *	Methyl methacrylate	✓	
598-98-1	Methyl pivalate	✓	
554-12-1 *	Methyl propionate	✓	
119-36-8	Methyl salicylate	✓	✓
79-22-1	Methylchloroformate (MCF)	✓	
	Methyldichlorodisilanes (mixed isomers)		✓
78-93-3 *	Methylethyl ketone	✓	
628-63-7	n-Amyl acetate		✓
97-66-7	N,N-Diethylaniline		✓
123-63-7	Paraldehyde	✓	
355-25-9	Perfluorobutane	✓	✓
382-21-8	Perfluoroisobutylene (PFIB)	✓	
624-65-7	Propargyl chloride		✓
79-09-4 *	Propionic acid (and some dimer)	✓	
109-06-4	Propyl acetate		✓
7791-25-5	Sulfuryl chloride	✓	
2699-79-8	Sulfuryl fluoride		✓
75-73-0	Tetrafluoromethane		✓
994-05-8	tert-Amyl methyl ether	✓	
79-01-6 *	Trichloroethylene	✓	
75-05-1	Trifluoroacetic acid	✓	✓
407-25-0	Trifluoroacetic anhydride	✓	✓
354-32-5	Trifluoroacetyl chloride	✓	✓
75-46-7	Trifluoromethane (Freon-23)	✓	
373-80-8	Trifluoromethylsulfur pentafluoride		✓
334-99-6	Trifluoronitrosomethane	✓	✓

Table D.2: Molecules with absorption in the significantly different regions found from comparing the exhaled breath of healthy children and children with cystic fibrosis. * VOCs that have been identified before in exhaled breath.

Compounds		Wavenumber region (cm^{-1})	
CAS Number	Name	1242.20 - 1242.65	1260.70 - 1261.65
354-25-6	1-Chloro-1,1,2,2-tetrafluoroethane (R124A)	✓	✓
421-50-1	1,1,1-Trifluoroacetone	✓	
420-46-2	1,1,1-Trifluoroethane (R143A)	✓	✓
431-89-0	1,1,1,2,3,3,3-Heptafluoropropane (HFC227EA)	✓	✓
124-73-2	1,2-Dibomotetrafluoroethane (Freon-114B2)	✓	
354-23-4	1,2-Dichloro-1,1,2-trifluoroethane (F132A)	✓	
123-91-1 *	1,4-Dioxane		✓
75-26-3	2-Bromopropane	✓	
75-88-7	2-Chloro-1,1,1-trifluoroethane (R133A)		✓
111-15-9 *	2-Ethoxyethyl acetate	✓	✓
138495-42-8	2H, 3H-Perfluoropentane	✓	✓
75-89-8	2,2,2-Trifluoroethanol		✓
25256-77-4	2,2,4-Trimethyl-1,3-pentanediol isobutyrate (Texanol)	✓	✓
108-24-7 *	Acetic anhydride	✓	
123-86-4 *	Butyl acetate	✓	✓
353-50-4	Carbonyl fluoride	✓	✓
76-15-3	Chloropentafluoroethane (R115)	✓	
75-43-4	Dichlorofluoromethane (Freon-21)	✓	
75-09-2 *	Dichloromethane		✓
674-82-8	Diketene	✓	
616-38-6 *	Dimethyl carbonate		✓
79-44-7	Dimethylcarbamoyl chloride		✓
34590-94-8 *	Dipropylene glycol methyl ether		✓
141-78-6 *	Ethyl acetate	✓	✓

Table D.2: (continued)

Compounds		Wavenumber region (cm^{-1})	
CAS Number	Name	1242.20 - 1242.65	1260.70 - 1261.65
140-88-5 *	Ethyl acrylate		✓
105-54-4 *	Ethyl butyrate	✓	✓
383-63-1	Ethyl trifluoroacetate	✓	
462-06-6	Fluorobenzene	✓	
14523-98-9	Formic acid, dimer	✓	
76-19-7	Freon-218 (octafluoropropane)	✓	✓
115-25-3	Freon-C318 (octafluorocyclobutane)	✓	✓
684-16-2	Hexafluoroacetone	✓	✓
76-16-4	Hexafluoroethane (Freon-116)	✓	✓
381-10-5	Hexafluoroisobutylene	✓	✓
110-19-0 *	Isobutyl acetate	✓	✓
78-59-1	Isophorone	✓	
108-21-4 *	Isopropyl acetate	✓	✓
79-20-9 *	Methyl acetate	✓	✓
93-58-3	Methyl benzoate	✓	✓
547-63-7	Methyl isobutyrate		✓
554-12-1 *	Methyl propionate	✓	
119-36-8	Methyl salicylate	✓	✓
	Methyldichlorodisilanes (mixed isomers)		✓
628-63-7	n-Amyl acetate	✓	✓
97-66-7	N,N-Diethylaniline		✓
7782-77-6	Nitrous acid	✓	
355-25-9	Perfluorobutane	✓	✓
382-21-8	Perfluoroisobutylene (PFIB)	✓	
624-65-7	Propargyl chloride		✓
109-06-4	Propyl acetate	✓	✓
2699-79-8	Sulfuryl fluoride		✓
75-73-0	Tetrafluoromethane		✓
75-05-1	Trifluoroacetic acid	✓	✓
407-25-0	Trifluoroacetic anhydride	✓	✓
354-32-5	Trifluoroacetyl chloride	✓	✓
373-80-8	Trifluoromethylsulfur pentafluoride	✓	✓
334-99-6	Trifluoronitrosomethane	✓	✓
108-15-4	Vinayl acetate	✓	

Acknowledgments

Life is a wonderful trip in which we are always accompanied by amazing fellow travelers. The PhD has been a small part of this trip full of excitement and incredible experiences. It was only possible to the help, company and support of many amazing friends and colleagues to whom I want to express my gratitude. First of all, I want to thank Prof. Paul H. Urbach for allowing me to be part of the Optics Research Group and give me the freedom to experience the different aspects of a research project. I am sincerely grateful to Dr. Nandini Bhattacharya, my supervisor during my PhD and Master studies. Nandini, over the years you have been a mentor, an example, a guide and overall a friend on whom I can rely professionally and personally. I big word of gratitude is for the members of the graduation committee for reviewing my thesis and for their invaluable feedback. Prof. J. Dankelman, Dr. O. Vaittinen, Dr. F. J. M. Harren and Prof. L. J. van Vliet, thanks to you I learned a bit more. Among the committee members I will always be in debt with Prof. P. H. C. Eilers for providing me the code to calibrate the data, thanks to you I discover the fascinating world of time warping. There is a very special team I will always be grateful to: Prof. Johan de Jongste, Dr. Marielle Pijnenburg and Dr. Esther van Mastrigt from the Erasmus Medical Center. I will always remember with a big smile our meetings and discussions. With you I learned to break the barrier between scientific fields, to dream of the possibilities without losing the connection with reality and, above all, to see the human side of medicine. Thank you doctors.

A huge thank you is for Dr. Stefan Persijn, who was one of the creators of this project and who was always willing to help and support me. A phone call or a short ride to get material was never too much for you to provide assistance.

To the magicians of the Biostatistics group of the Erasmus Medical center, Karl Brand and Andrew Stubbs, thank you for the help with the analysis of the experimental data. Without you it would have been difficult to get the real value of this research project.

I dedicate this work to the brave volunteers and patients who participated in this research. You are the heroes of this story and some of the most courageous people I have ever met.

I will always be in debt to the support team of the Optics Research Group, Roland Horsten, Thim Zuidwijk, Rob Pols and Yvonne van Aalst, you all showed me what creativity can do. I will always treasure your friendship.

A very special thanks goes to all the students who were brave enough to take me as supervisor for their Master graduation projects and internships. Timothy Lucas, Zhe Hou, Vincent Docter, Chiara Graziosi, Ton Verhoeven, Olav Grouwstra, Maxim Spreij, Alberto Maran, Alonso Millan, Adeline Remoue, Virginie Poncéry and Shan Du, I owe you more than I will ever be able to return you back. Thanks to you I learned the art of teaching and guiding.

During the years I was part of the Optics team in Delft I had the opportunity to meet amazing people. To all of them I am thankful for their teachings and great coffee mo-

ments. From Prof. Joseph Braat, Prof. Paul Planken, Silvania Pereira, Florian Bociort and Omar El Gawhary I learned many important scientific and social lessons. I will always be grateful with Aurele Adams for the fun moments and encouraging words. To Peter Somers, I want to thank you for your help with the Dutch summary of this thesis, I will always admire your readiness to help others. I thank Mounir Zeitouny, who was at the beginning of this project, for helping me to put this project up and running. When I arrived in Delft there was a group who showed me what the Optica spirit is. Thanks for that warm welcome, Man Xu, Sven van Haver, Morris Cui, Pascal van Grol, Olaf Janssen, Maarten van Turnhout, Reshmi Chakkitakandy, Joseph Knab, Aura Nugrowati, Gopakumar Ramakrishnan, Alberto da Costa Assafrao. You can be sure that although sometimes the spirit changes the roots are always there.

Like me, many others joined and left the team over the years and everyone contributed to enrich the atmosphere of the group. As one of us in the group used to say: everyone has its own weirdness. Lei Wei, Katsiaryna Ushakova, Mahsa Nemati, Sarathi Roy, Matthias Strauch, Edgar Rojas Gonzalez, Hamed Ahmadpanahi, Gerward Weppelman, Priya Dwivedi, Yifeng Shao, Fellipe Grillo Paternella, Diego Guamialama, Anna Ezerskaia, Alessandro Polo, Thomas Liebig, Axel Wiegmann, Marco Mout, Liesbeth Dingemans, Gyllion Loozen, Rik Starmans, Peter Petrik, Jeffrey Meisner, Nick Schilder, Nick Spook, Sander Konijnenberg, Daniel Nacimento Duplat, Ying Tang, Jörn Zimmerling, Daniel Asoubar, Quincy Berg, Hugo Bordier, Luiz Couto Correa Pinto Filho, Radek Smid thank you all for this special weirdness that made the days more fun and interesting.

I will always keep in my hearth the great moments I spent in the office, sometimes talking science, sometimes talking non-sense. Thanks Nitish Kumar and Wouter Westerveld for being the best office mates I could have ever had. I will always be a proud member of the ANW department inside the TNW building. Although we never shared office, I also had a great amount of non-sense fun with Andreas Hansel, Luca Cisotto and Gopika Ramananda. Thank you for the continuous friendship, let it last forever.

A person to whom I will always be grateful for his guidance is Mr. Harry Linders. Thank you Harry for helping me to find my way and for teaching me that life has many available roads.

Outside the University I had received the support and love of many good friends. The space is too short to name you all, but be sure I always will be thankful for sharing this adventure with me. I specially want to thank my Dutch-Mexican familie, Antonio Peralta and Julieta Pelaez. Thanks for opening your hearth and home for a couple of friends looking for their place in this world. To Monique Belien and Jose Luis Almodovar, the amazing friends who I do not need to see and meet every week but who have been always there for a drink, a word and a hug. To the meditation enthusiast: Carmen (thanks for showing me that there are many paths), Pavel, Kasia, Bernat, I hope one day we can share again a beer while in a silent retreat. To the SoSalsa! friends, you taught me that wherever I go there will always be place for a dance and time for a good song, thanks Vivian, Jan, Hendrik, Aurinke, Vicky, Remko, Ricardo. To the TU Delft student ambassadors and specially to Ina Dijstelbloem and Patricia Carrion, you are making a big difference by focusing in the core element of the university: the humans who study there. To Dr. Maurijn van der Zee for reminding me that research is not only about personal curiosity but helping to shape a better future for the coming generations.

Anyone who meet me will discover that I am proud Mexican and a happy Latino. My stay in Delft would not have been as fun without the Mexican gang, for whom the nights are always too short and who are always ready to give the best of them: Hugo Cueto (I will never be able to copy your napping style), Omar Jimenez (man, Risk is good for you!), Zeus Guevara, Diana Lucatero, Behtzua Elizondo, Perla Cruz, Fabian Carranza, Ileana Hernandez, Antonio Jarquin, Antonio Ramos, Hugo Cruz, Jorge Maxil, Edmundo Molina, Mar Palmeros, Alberto Nol, Diego Alatorre, Ivan Piña, Alejandro Luna and family, Oziel Garcia, David Calvillo, Gabriel Vazquez, Alondra Torres and all the people who help to establish the Mexicanos en TU Delft association and the Latin American Student Association (LATITUD).

The final words of gratitude are in my mother tongue, Spanish, in which everything sounds more delicious and colorful. Gracias familia por haberme ayudado en los momentos felices y en los difíciles, por haberme enseñado a soñar y a volar, por haberme mostrado que en la vida sólo hay un objetivo: ser feliz y ayudar a otros a ser felices. Mama, Betito, Ada, yo nunca he estado solo porque en cada paso que doy ustedes están conmigo, hoy quiero asegurarles que yo siempre estaré con ustedes. Tania, mi gran compañera de aventuras, gracias por tu paciencia y amor, por enseñarme que la vida es cambio y que nunca debemos dejar de crecer y jugar.

Let's keep playing!

ABOUT THE AUTHOR

Adonis Reyes Reyes was born on February 28th 1983 in Atlixco, Puebla, Mexico. He obtained his BSc degree from the National University of Mexico (UNAM) in 2008. During his BSc thesis he worked in the implementation and stabilization of Magneto-Optical Traps (MOT) to cool and trap Rubidium atoms. For his BSc thesis he was awarded the Prometeo Award for Experimental Physics by the Alvarez-Cisneros/Castillo-Garza Foundation and the Physics Institute of UNAM.

He joined the dual degree Erasmus Mundus Master OpSciTech program (2008 - 2010) where he obtained his MSc. degree in Optics, Matter and Plasma from the Institutd'Optique, Graduate School, Paris, France, and his MSc. degree in Applied Physics from Delft University of Technology, The Netherlands. During his master thesis he started working on spectroscopic techniques to detect trace gases in human breath.

He continued working on human breath analysis thanks to the support of the Dutch Foundation for Fundamental Research on Matter (FOM). During his PhD he used quantum cascade laser spectroscopy to study the composition of breath and to determine possible biomarkers associated with asthma, cystic fibrosis and diabetes. Currently, Adonis is working as a consultant of TMC in The Netherlands, where he enjoys the freedom and challenges of being an employeneur. Specifically, he holds the position of Invention disclosure coordinator at IMEC, Leuven, where he is combining his passion for science and technology with his interest in the valorisation of ideas.

In his free time, Adonis enjoys bouldering, swimming, dancing and enjoying a good meal and drink with friends.

LIST OF PUBLICATIONS

Journal publications

- **Reyes-Reyes A.**, Van Mastrigt E., Brand K., Stubbs A., De Jongste J. C., Pijnenburg M. W., Urbach H. P. and Bhattacharya N., *Exhaled breath profiling in children with asthma and cystic fibrosis: a new broadband quantum cascade laser spectroscopy method*, In preparation.

- Van Mastrigt E, **Reyes-Reyes A.**, Brand K., Bhattacharya N., Urbach H. P., Stubbs A., De Jongste J. C. and Pijnenburg M. W., *Exhaled breath profiling using broadband quantum cascade laser-based spectroscopy in healthy children and children with asthma and cystic fibrosis*, Journal of Breath Research **10**, 026003 (2016).

- **Reyes-Reyes A.**, Horsten R. C., Urbach H. P. and Bhattacharya N., *Study of the exhaled acetone in type 1 diabetes using quantum cascade laser*, Analytical Chemistry **87**, 507 (2015).

- **Reyes-Reyes A.**, Hou Z., van Mastrigt E., Horsten R. C., de Jongste J. C., M. W. Pijnenburg M. W., Urbach H. P. and Bhattacharya N., *Multicomponent gas analysis using broadband quantum cascade laser spectroscopy*, Optics Express **22**, 18299 (2014).

Conference Proceedings

- **Reyes-Reyes A.**, Zeitouny M. G., van Mastrigt E., Persijn S. T., Bhattacharya N. and Urbach H. P., *Cavity-enhanced direct frequency comb spectroscopy*, Proc. SPIE 8011, 22nd Congress of the International Commission for Optics: Light for the Development of the World **8011**, 80112O (2011).

Conference Contributions

- **Reyes-Reyes A.**, van Mastrigt E., Pijnenburg M. W., de Jongste J. C., Bhattacharya N. and Urbach H. P. *Smelling diseases using quantum cascade laser spectroscopy. Shining light on human breath analysis*, Dutch Biophysics 2016. 3-4 October 2016, Veldhoven, The Netherlands.

- **Reyes-Reyes A.**, Hou Z., van Mastrigt E., Horsten R. C., de Jongste J. C., Pijnenburg M. W., Urbach H. P. and Bhattacharya N., *Quantum cascade laser spectroscopy for exhaled breath profiling*, Breath Analysis 2014. 6-9 July 2014, Torun, Poland.

- **Reyes-Reyes A.**, Van Mastrigt E., Brand K., S.T. Persijn, Stubbs A., De Jongste J. C., Pijnenburg M. W. H., Bhattacharya N. and Urbach H. P., *Smelling diseases using mid-infrared spectroscopy*, BioDay Delft. 10 June 2014, TU Delft, The Netherlands.

- **Reyes-Reyes A.**, Van Mastrigt E., S.T. Persijn, Bhattacharya N. and Urbach, *Mid-infrared spectroscopy of human breath*, Physics@FOM Veldhoven. 21 - 22 January, 2014, Veldhoven, The Netherlands.

179

- **Reyes-Reyes A.**, Zeitouny M. G., van Mastrigt E., Bhattacharya N., Persijn S. T. and Urbach H. P., *High resolution spectroscopy with a Frequency Comb Laser*, Summer School of Ultrafast Laser Science and Applications. Spanish network of ultrashort ultraintense laser. Consolider SAUUL. 10-15 June 2012, Menorca, Spain.

- **Reyes-Reyes A.**, Zeitouny M. G., van Mastrigt E., Persijn S. T., Bhattacharya N. and Urbach H. P., *High resolution and high sensitive spectroscopy for gas analysis*, 113th annual DGaO conference. TU Eindhoven. 29 May - 2 June 2012, Eindhoven, The Netherlands.

- **Reyes-Reyes A.**, Zeitouny M. G., van Mastrigt E., Persijn S. T., Bhattacharya N. and Urbach H. P., *VIPA and FTIR Spectroscopy for Gas Analysis*, Breath Analysis workshop. VSL. April 2011, Delft, The Netherlands.

- **Reyes-Reyes A.**, Zeitouny M. G., Bhattacharya N., Persijn S. T. and Urbach H. P., *Application of cavity-enhanced direct frequency comb spectroscopy in the detection of biomarkers in exhaled breath*, EOS Annual Meeting 2010. European Optical Society. 26 - 29 October 2010, Paris, France.

- **Reyes-Reyes A.**, Zeitouny M. G., Bhattacharya N., Persijn S. T. and Urbach H. P., *High resolution spectroscopy using a Virtually Imaged Phased Array and its application to human breath analysis*, Advanced Laser Technologies 2010 (ALT 10). 11-16 September, 2010, Egmond aan Zee, The Netherlands.

- **Reyes-Reyes A.**, Zeitouny M. G., van Mastrigt E., S.T. Persijn, Pijnenburg M. W. H., de Jongste J. C., Bhattacharya N. and Urbach H. P., *Interferometric applications with the Frequency Comb*, Latin Laser Lab Meeting, University of Salamanca, 5-6 March 2012, Salamanca, Spain.

- **Reyes-Reyes A.**, Bhattacharya N. and Urbach H. P., *Gas detection in the infrared using spectroscopic techniques*, Symposium on Photonic Integration and IOP-PD and STW-GTIP/SOS Meeting. TU Eindhoven. November 2011, Eindhoven, Netherlands.

Propositions

accompanying the dissertation

SHINING LIGHT ON HUMAN BREATH ANALYSIS
WITH QUANTUM CASCADE LASER SPECTROSCOPY

by

Adonis REYES REYES

1. In human breath research, if the molecular components of breath can be detected then it smells like trouble. (This thesis)

2. In the future, exhaled acetone measurements should be included among routine medical examinations because the detection of high concentrations of acetone will enable the early diagnosis of chronic ketosis and type 1 diabetes, improving the quality of life of patients.(Chapter 5)

3. The selection of the threshold for discriminating parameters, such as p-values, should take into account the experimental data and method because the use of only standard recipes can lead to the exclusion of valuable information. (Chapter 6)

4. The commercial breath analyzer will be a hybrid combining laser spectroscopy, chemistry and material science developments. (Chapter 7)

5. Curiosity and imagination are the fuel of science. Arrogance is its brake.

6. Art and science raise questions. Science provides answers. Art places statements.

7. Education is more than just literacy. There is a lot of uneducated people with academic degrees.

8. Mental health does not receive the attention it deserves because the brain is not consciously regarded as other human organs.

9. If humans fear what they do not understand, scientists are the bravest people in the world.

10. Human reproduction is defended as a fundamental right, but responsible parenting is not conceived as a basic obligation.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotor prof. dr. H. P. Urbach.

Stellingen

behorende bij het proefschrift

SHINING LIGHT ON HUMAN BREATH ANALYSIS
WITH QUANTUM CASCADE LASER SPECTROSCOPY

door

Adonis REYES REYES

1. Als in onderzoek naar menselijke adem de moleculaire componenten daarvan kunnen worden gedetecteerd, dan riekt dit naar problemen. (Dit proefschrift)

2. In de toekomst zouden metingen van uitgeademde aceton deel moeten uitmaken van routinematig medisch onderzoek, omdat de detectie van hoge concentraties aceton vroegtijdige diagnose van chronische ketose en type 1 diabetes mogelijk zal maken, waarmee de kwaliteit van leven van patiënten zal worden worden verbeterd. (Hoofdstuk 5)

3. Bij de keuze van drempelwaarden voor onderscheidende parameters, zoals p-waardes, zou rekening moeten worden gehouden met de experimentele gegevens en methode omdat beperking tot het gebruik van standaardrecepten kan leiden tot uitsluiting van waardevolle informatie. (Hoofdstuk 6)

4. De commerciële ademanalysator zal een hybride zijn waarin laserspectroscopie, en ontwikkelingen op het gebied van chemie en materiaalwetenschappen samengaan. (Hoofdstuk 7)

5. Nieuwsgierigheid en verbeeldingskracht vormen de brandstof voor wetenschap. Arrogantie vormt de rem.

6. Kunst en wetenschap roepen vragen op. Wetenschap levert antwoorden. Kunst doet uitspraken.

7. Opvoeding in de zin van opleiding betekent meer dan alleen geletterheid. Er zijn talloze onopgevoede mensen met een academische graad.

8. Geestelijke gezondheid krijgt niet de aandacht die het verdient omdat de hersenen minder strikt als orgaan worden gezien dan andere menselijke organen.

9. Als mensen vrezen wat zij niet kennen, zijn wetenschappers de dapperste mensen in de wereld.

10. De menselijke voortplanting wordt verdedigd als een fundamenteel recht, maar verantwoordelijk ouderschap wordt niet opgevat als een basisverplichting.

Deze stellingen worden opponeerbaar en verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotor prof. dr. H. P. Urbach.