Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

Karina Strøm

Mid-Infrared Laser Spectroscopy for Glucose Sensing

Master's thesis in Electronics Systems Design and Innovation
Supervisor: Prof. Astrid Aksnes
June 2019

NTNU
Norwegian University of
Science and Technology

Karina Strøm

Mid-Infrared Laser Spectroscopy for Glucose Sensing

Master's thesis in Electronics Systems Design and Innovation
Supervisor: Prof. Astrid Aksnes
June 2019

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Project description

Spectroscopy for qualitative and quantitative sensing is a rapidly expanding area of sensing that employs optical methods. Spectroscopy has traditionally been used for gas sensing and other industrial uses and is now extending into other fields such as sensing applications in medicine. In the Artificial Pancreas project (APT), an artificial pancreas where the diabetic user does not have to administer insulin or input information about meal times is the goal. The aim will be to improve the optical sensors that are being developed for glucose sensing within the APT project. In this master thesis, mid-Infrared (mid-IR) spectroscopy will be performed using a quantum cascade laser (QCL) to predict the level of glucose found in samples containing different analytes. The measurement setup for fiber-coupled configurations will be modified to improving the signal-to-noise ratio. This will be done by using different fiber sizes in the setup as well as utilizing an attenuated total reflection (ATR) prism which might be a better sensor design for a personal and portable sensor. Different interferents will be included in the measured samples, such as ketones and albumin if these are found to absorb at the mid-IR wavelengths covered by the QCL. If time, the signal processing and data analysis might be modified to better fit the specific setup. This might include adding more filtering options to the data analysis together with averaging and aligning the data in order to remove noise.

Abstract

An improved mid-infrared spectroscopy measurement method for detecting glucose is wanted in order to make a continuous glucose monitoring device with a longer lifetime in order to improve the quality of life for diabetics. Measurements were obtained through fibre-coupled transmission spectroscopy and attenuated total reflection (ATR) spectroscopy. Regression analysis was then performed on the measurements from both of the setups in order to predict the glucose concentrations in the measured samples. The samples that were measured were; pure glucose solution, pure β-Hydroxybutyric acid (BHA) solution and a more complex solution with glucose, lactate, urea and albumin dissolved in phosphate-buffered saline. These other analytes were included in the measurements in order to find the setups' glucose selectivity since they have overlapping absorption bands with glucose. However, they are also related to complications like sepsis/septic shock and renal failure and therefore it could be beneficial to predict these values as well. BHA was measured to see how well the setups would be able to predict the concentration in order to detect diabetic ketoacidosis. Through regression analysis the root-mean square error of prediction (RMSEP) and the root-mean square error of cross-validation (RMSECV) were found to be smaller for the measurements obtained with the fibre-coupled transmission setup. This lead to the conclusion that this setup was superior to the ATR setup. The fibre-coupled transmission setup was also able to predict the concentration of urea, lactate and albumin with high accuracy. The RMSEP/RMSECV values achieved for the glucose selectivity by the fibre-coupled transmission setup was very good, which indicates that the measurement method provides a robust and accurate glucose prediction. The analyte concentrations measured were down to physiological concentrations. The linear response obtained through regression analysis indicated that accurate measurements of glucose concentrations was possible even at low concentrations (< 100 mg/dl). The ATR setup was also able to predict the glucose selectivity and sensitivity as well as the albumin concentration with high accuracy. It was not able to predict the concentrations of urea and lactate due to a low signal-to-noise ratio. The spectra obtained with the setup showed that the signal was very low and improvements can be done to this setup in regards to increasing the signal strength.

Sammendrag

For å kunne lage en enhet for kontinuerlig glukosemåling som har lang levetid er det ønskelig med en forbedret spektroskopi målemetode for å kunne detektere glukose. En slik enhet vil kunne forbedre livskvaliteten til diabetikere. I denne avhandlingen ble målinger anskaffet ved bruk av fiber-koblet transmisjonsspektroskopi og dempet totalrefleksjons (ATR) spektroskopi. Regresjonsanalyse ble brukt på målingene for begge oppsettene for å kunne predikere glukose konsentrasjonene i de målte prøvene. Prøvene som ble målt var; ren glukoseløsning, ren β-hydroksybutyrat (BHA) løsning og en mer kompleks løsning med glukose, laktat, urea og albumin oppløst i fosfatbufret saltvann. Disse andre analyttene ble inkludert i målingene for å finne oppsettets glukose selektivitet da disse har overlappende absorpsjonsbånd med glukose. Det er også ønskelig å kunne predikere disse analyttene da de kan relateres til komplikasjoner som sepsis/septisk sjokk og nyresvikt. BHA ble målt for å se hvor godt oppsettene klarte å predikere konsentrasjonen og dermed kan bli brukt til å detektere diabetisk ketoacidose. Gjennom regresjonsanalysen ble det kvadratiske gjennomsnittets avvik av prediksjon (RMSEP) og det kvadratiske gjennomsnittets avvik av kryss-validering (RMSECV) funnet til å være mindre for målingene fra det fiber-koblede transmisjons oppsettet. Dette førte til konklusjonen om at dette oppsettet var bedre enn ATR oppsettet. Det fiber-koblede transmisjons oppsettet klarte også å predikere konsentrasjonene av urea, laktat og albumin med høy nøyaktighet. RMSEP/RMSECV verdiene som en oppnådde for glukose selektiviteten med dette oppsett var veldig gode, som indikerer at denne målingsmetoden kan sørge for robuste og nøyaktige glukose prediksjoner. Analytt konsentrasjonene var helt ned til fysiologiske konsentrasjoner. Den lineære responsen oppnådd gjennom regresjonsanalysen indikerer at nøyaktige målinger av glukosekonsentrasjonene var mulig, til og med for lave konsentrasjoner ($<$ 100 mg/dl). ATR oppsettet klarte også å predikere glukose selektiviteten og sensitiviteten samt albumin konsentrasjonen, med høy nøyaktighet. Dette oppsettet klarte ikke å predikere konsentrasjonene av urea og laktat på grunn av det lave signal-til-støy forholdet. Spektrene som en oppnådde med ATR oppsettet viste at signalet var veldig lavt og oppsettet kan dermed forbedres i sammenheng med å forsterke signalet.

Preface

This master thesis is submitted in partial fulfillment of the requirements for the degree of Master of Science in Electronics Systems Design and Innovation at the Department of Electronic Systems (IE) at the Norwegian University of Science and Technology (NTNU). The work presented here is a continuation of a one semester project performed in the Fall 2018, which consisted of a literature study on biomedical applications for mid-IR laser spectroscopy and the detection of the glucose selectivity and sensitivity in a free-space transmission setup. Part of the work is reproduced here in the theory chapter for completeness..

I would like to thank my supervisor Prof. Astrid Aksnes for guidance and assistance during this work. For the helpful discussions related to the work presented here and for the feedback related to the writing of this master thesis. Secondly I would like to my co-supervisor PhD candidate Ine Larsen Jernelv for all of the helpful discussions, for always taking her time to help me and guide me in my work and for sharing her experiences from the laser lab. I would also like to extend my thanks to the mechanical workshop at the IES department for the help they could provide in regards to the prism-holder and to the other users of the laser lab for helpful advice regarding my setup.

Table of Contents

Abbreviations

ATR = Attenuated total reflection
BGL = Blood glucose level
BHA = β-hydroxybutyric acid
CGM = Continuous glucose monitoring
DFB = Distributed feedback
EC = External Cavity
FFT = Fast Fourier Transform
FIR = Finite impulse response
FP = Fabry Pérot
FTIR = Fourier-transform infrared spectroscopy
GUI = Graphical user interface
IIR = Infinite impulse response
IR = Infrared
IRE = Internal reflecting element
ISF = Interstitial fluid
MA = Moving Average
Mid-IR = Mid Infrared
PBS = Phosphate-buffered saline
PC = Principal component
PCA = Principal Component Analysis
PCR = Principal Component Regression
PLSR = Partial Least Square Regression
QCL = Quantum Cascade Laser
RMSECV = Root-mean-square error of cross-validation
RMSEP = Root-mean-square error of prediction
SNR = Signal-to-noise ratio
TIR = Total internal reflection
UV = Ultraviolet
ZnS = Zinc sulfide
ZnSe = Zinc selenide

1 | Introduction

1.1 Biomedical Laser Spectroscopy

The mid-infrared (mid-IR) spectral region (4000-400 cm^{-1}) is of great importance in biomedical spectroscopy. Many of the biological molecules found in the body have specific absorption bands in this region. Through mid-IR spectroscopy it is thus possible to detect, identify and even quantify biological analytes. Mid-IR spectroscopy used for biomedical application provides an effective, fast and painless way of sensing and monitoring different biological analytes related to various diseases [1, 2]. Mid-IR can also be used for breath analysis which offers non-invasive, real-time and point-of-care and disease diagnostics with minimal risk and negligible discomfort for patients [3]. It has also been used in order to detect drugs in saliva [4] which makes mid-IR spectroscopy versatile within biomedical applications. The spectroscopy method does not rely on any reagents which result in biomedical devices with a long lifetime. The sample amount needed for quantification of biological analytes is also very small compared to other methods used today. However, the light sources required in mid-IR spectroscopy for biomedical applications are very expensive.

A disease that affects many in the world is diabetes. As many as 422 million adults have diabetes and it is the cause of 1.5 million deaths per year. [5] Due to the severity of the disease a lot of research related to mid-IR spectroscopy is concentrated around improving the devices used for glucose detection. Approximately 85 % of the world marked for biomedical sensors is comprised of blood glucose measurement sensors [6].

1.2 Diabetes Mellitus

Diabetes mellitus, commonly known as 'Diabetes', is a chronic disease related to abnormally high levels of glucose in the blood. The abnormal glucose level is a result of underproduction of the glucose regulating hormone insulin (type 1) or insulin resistance (type 2). A diabetic is identified as an individual whose blood glucose concentration does not decrease to less than 140 mg/dl within two hours of a meal [7]. It is important for diabetics to monitor their blood glucose level (BGL) frequently to prevent it from fluctuating outside normal range. Hypoglycemia occurs when the BGL is below 70 mg/dl. The opposite of this is hyperglycemia, which occurs when the BGL is above 180 mg/dl [8]. Severely lowered hypoglycemia can possibly lead to diabetic coma, which can be fatal if not treated properly. Elevated hyperglycemia can also lead to diabetic coma but long-term effects are the main problem related to elevated hyperglycemia. Both hypo- and hyperglycemia are unwanted, and the long-term effects of poorly regulated diabetes are many and can strongly reduce the quality of life and lifetime.

Having a high glucose level over a long period can lead to stiffening of blood vessels which may lead to a reduction in the supply of oxygen and nutrients to different tissues. This stiffening occurs both in larger blood vessels which leads to the heart, brain and the lower extremities, as well as smaller blood vessels that lead to the eyes and kidneys [9]. If the oxygen delivery is reduced to these organs it can lead to so-called ”long-term complications” e.g. damaged sight, kidney failure, and cardiovascular diseases [10]. A lot of the routine care around diabetic patients is therefore directed towards prevention of these kinds of complications. Another serious complication of diabetes is diabetic ketoacidosis. When the body isn’t able to produce enough insulin it will start to burn fatty acids and fat and the resulting product of this is ketones. If the level of ketones in the blood thus increases significantly the blood can become acidic, which leads to diabetic ketoacidosis [8]. The life expectancy of a diabetic can be prolonged if the diabetic is able to manage hypo- and hyperglycemia episodes and reduce the occurrence of these types of episodes, which one believes can be done with the use of continuous glucose monitoring devices [9].

1.3 Continuous Glucose Monitoring Devices

In order for diabetics to keep their BGL stable and to keep hypo- and hyperglycemia episodes to a minimum, they have to monitor their BGL daily. Different devices have been developed and can be used for self-monitoring of the BGL level and the most common is an electrochemical device that requires a finger prick to measure the level [11]. The method is both inconvenient and painful for the patient, which might lead to diabetics measuring their BGL less frequently than they should. This makes it more difficult to avoid episodes of hypo- and hyperglycemia and thus the risk of short- and long-term complications are not prevented. There is also no way to monitoring the levels during the night which would be ideal in order to know more about how the BGL behaves during the night. By monitoring the BGL using such a finger prick device the measurements occur sporadically and the levels cannot be measured continuously. To avoid hypo- and hyperglycemia episodes and prevent short- and long-term complications, measurements of the BGL in a continuous manner is beneficial. A device that does this will be referred to as a continuous glucose monitoring (CGM) device.

Electrochemical CGM devices have already been developed and can be found on the market today [12, 13]. These measure the glucose concentration very well and are available commercially. However, a drawback with these is the degradation of the chemical reagents and the biofouling of the sensor. Due to this, these types of CGM devices have to be replaced on a regular basis, approximately once per week [14]. Many of them also require daily finger pricks in order to calibrate the device, which is inconvenient and painful. Attention has therefore been drawn towards methods that use other ways for detecting the glucose levels, which does not degrade and does not need to be replaced as often as the electrochemical CGM devices. Spectroscopy is one method that many are investigating in relation to making a CGM device. This method does not rely on any reagents and is suited for a reagent-free quantification of glucose as well as other biological analytes in the body. If spectroscopy-based CGM devices can provide precise and robust measurements it will be an optimal replacement for the electrochemical CGM devices on the market today which would not need to be replaced as often.

The interest for both non-invasive and invasive optical CGM devices has increased over the last years and are being investigated by many groups which are working towards overcoming the different challenges related to the methods. The main optical technologies for glucose measurements are; Raman, near-IR and mid-IR spectroscopy. Challenges related to Raman spectroscopy are the small scattering cross section compared to the fluorescence cross section and long acquisition time. The challenges related to near-IR are strong water absorption and other biological analytes with absorbance similar to glucose. Challenges related to mid-IR are the small penetration depth and suitable mid-IR sources that are strong enough to overcome the strong water absorption. However, with the invention and development of the quantum cascade laser, which is a high emission power source, new avenues have opened for mid-IR spectroscopy [14].

1.4 Aim of This Master

In this master thesis mid-IR spectroscopy will be investigated, in relation to the detection of glucose. Attenuated total reflection (ATR) spectroscopy and fibre-based transmission spectroscopy will be tested in order to see if the use of mid-IR light to detect glucose can be improved. Both of the setups will be used together with a quantum cascade laser (QCL) which is a high emission powered mid-IR source. Samples containing different combinations of glucose, urea, lactate, albumin and the ketone β-hydroxybutyric acid (BHA) will be measured in order to compare the measurement methods. Urea, lactate and albumin are included in the measured samples to investigate if they interfere with the measurement and prediction of glucose concentration. Glucose and BHA are measured separately to see if it is possible to predict them with a small error since they are both directly related to diabetes. Regression analysis will be used to assess how well the analyte concentrations can be predicted.

1.5 Thesis Outline

Chapter 1 begins with an introduction of the master thesis and some motivation. Chapter 2 then presents an overview of infrared light in Section 2.1 and different spectroscopy methods in Section 2.2. Section 2.3 presents the theory on the high emission power mid-IR source, the QCL. Section 2.4 presents an overview of mid-IR spectroscopy research for glucose measurements and previous work related directly to the two measurement methods used in this master thesis. Theory about data regression analysis and signal processing is presented in Section 2.5 and the biological analytes that will be measured in this master thesis is presented in Section 2.6.

Chapter 3 describes the simulation method used to simulate different ATR prisms that are discussed in Section 3.1. The polishing method for the fibres used in the setups is presented in Section 3.2 and the setups are presented in Section 3.3. How the measured data is processed is presented in Section 3.4 and then it is presented in Section 3.5 how the data is analysed. In Section 3.6 the measured samples are presented with their concentrations.

The results are presented in Chapter 4 and Section 4.1 presents values related to the ATR spectroscopy. Section 4.2 presents the results regarding the absorption bands for urea and BHA. The simulation results are presented in Section 4.3 and Section 4.4 presents the results obtained in relation to the polishing methods and choice of fibres. The results obtained from measuring the samples with the two measurement methods are presented in Section 4.5. Chapter 5 includes the discussion concerning the results and Chapter 6 presents the conclusion of this master thesis.

2 | Theory

Parts of the theory section have been included from the project thesis for completeness with some additions and modifications.

2.1 Infrared Light

Since mid-infrared (mid-IR) is the wavelength range that will be utilized for measurements in this master thesis the following theory on IR spectroscopy is included to understand why mid-IR is desired for bio-medical sampling over the other IR regions. The main problems concerning the IR region and measurement of biological samples in the body will also be highlighted. The theory in the text below is based on the book *Infrared Spectroscopy: Fundamentals and Applications* by Barbara H. Stuart [15].

The IR wavelengths are found between microwaves and visible light in the electromagnetic spectrum. IR radiation has longer wavelengths and therefore lower energy than visible light. The IR region can be divided into three smaller regions; near-IR [13000-4000 cm^{-1}], mid-IR [4000-400 cm^{-1}] and far-IR [400-33 cm^{-1}], see Figure 2.1. Wavenumbers ($\bar{\nu}$) are used as a unit instead of wavelengths (λ) since this is more frequently used in spectroscopy. Wavenumbers are inversely proportional to wavelength ($\bar{\nu} = 1/\lambda$).

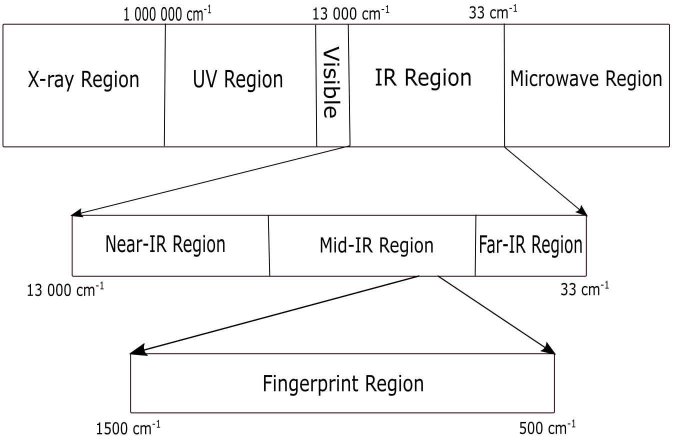


Figure 2.1: Electromagnetic spectrum showing how the IR region is divided into near-IR, mid-IR and far-IR. This work focuses on a region in the mid-IR range called the fingerprint region, between 1500 cm^{-1} and 500 cm^{-1}. Adapted from [15].

2.1.1 IR Absorption

When shining visible light or ultraviolet (UV) light onto a sample, the electrons will absorb the light and be excited to a higher energy state. For an electron to be excited, the energy received through absorption has to be higher than the energy of the first excited state, $E = h\nu > E_1$. E is the energy received through absorption, E_1 is the energy of the first excited state, h is Planck's constant and $\nu = c/\lambda$ where c is the speed of light and λ is the wavelength of the absorbed light. IR light is less energetic than both visible light and UV light, so the energy absorbed by electrons will not result in excitation to higher energy states. However, there is enough energy to excite vibrational states in a given molecule. The difference between absorption of UV/visible light and IR radiation is shown in Figure 2.2. Not all of the radiation can be absorbed by the molecule, only the radiation with frequencies that corresponds to the fundamental vibrational modes of the molecule. In addition, there is a selection rule saying that for a molecule to absorb the infrared radiation the electric dipole moment has to change during vibration. This means that not all of the fundamental vibrational modes are able to absorb energy even though the frequency corresponds to the mode. As a result of the selection rule, symmetrical molecules experience less infrared-active vibrations, compared to asymmetrical molecules. If two atoms, connected through a bond, are far from each other in the periodic table the absorption will be stronger and the bands in the IR-spectra will appear more intense.

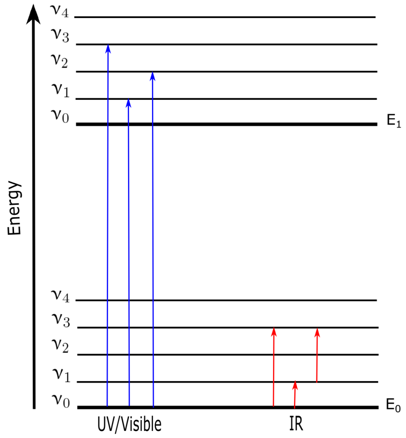


Figure 2.2: Energy diagram demonstrating the difference between absorption of UV/Visible light (excitation to a higher energy state) and IR radiation in molecules (vibrational excitation). E_n $n = 0, 1, 2, ..$ are the energy states and ν_n $n = 0, 1, 2, ...$ are the vibrational states.

2.1.2 Other Factors Affecting Absorption Bonds

In an IR spectrum it is not always possible to relate a specific peak directly to a vibrational excitation. In some cases, the peaks and structures that can be seen are a result of overtones, combination bands, Fermi resonance, hydrogen bonds or skeletal vibrations. Each fundamental vibrational mode has multiples that are called overtone bands which appears at $n\nu_0$, where ν_0 is the wavenumber of the fundamental mode and n is the index of the overtone band ($n = 2, 3, 4, ..$). These overtone bands, and other bands, might absorb at the same time in one molecule. If this is the case, the combined absorption band will appear at the wavenumber which is the sum of the two combined wavenumbers. If there is Fermi resonance in an IR spectrum the frequencies and intensities of two absorption bands that are close or overlapping will experience a shift. The bands are usually shifted away from each other and their intensities become more similar. Hydrogen bonds have the ability to influence the bond stiffness in a molecule, and by changing the bond stiffness the vibrational mode might be shifted in the spectrum. The effects of hydrogen bonds affect water in particular and often arise due to the water. Skeletal vibrations are vibrations that couple over a large part of a molecule. These vibrations result in patterns that are called fingerprints, which are distinct for different molecules. Skeletal vibrations are most commonly found in the mid-IR region in the area 1500-500 cm^{-1} [15].

In the near-IR region, the absorptions that can be observed are mainly due too overtones or combinations of the fundamental stretching bands which occur in a specific region. A common problem with using near-IR in spectroscopy is that the bands are weak in intensity and broad, and they often overlap. Use of the far-IR region is limited, but the region contains information about vibrations of molecules with heavy atoms, molecular torsions, and prism lattice vibrations [16]. The mid-IR region can be divided into four regions were the bands in three of them are groups regions and the last one contains the vibrational patterns called fingerprints, see Figure 2.3. Based on this the mid-IR region is often more useful for quantitative analysis than near-IR and far-IR.

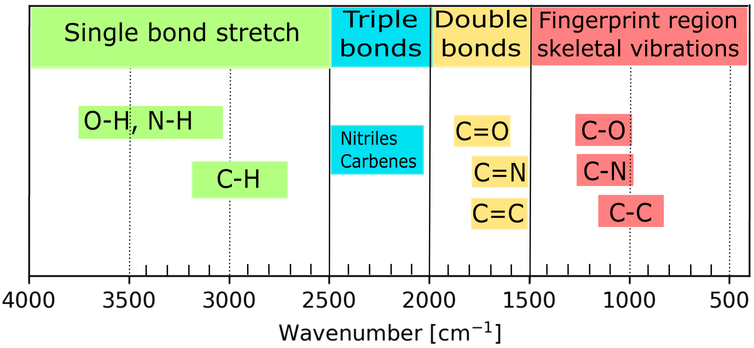


Figure 2.3: An overview of how the mid-IR region can be divided into four regions with their typical vibrations. The fingerprint region is the most interesting region in mid-IR since the fingerprints are distinct for the molecule as a whole.

2.1.3 Experimental Methods

Even though the mid-IR region is more suited for quantitative analysis, there have been some challenges with using this region for biomedical applications. The challenge with mid-IR spectroscopy has been the limited amount of high-energy sources, which are expensive and hard to produce. Thermal sources are suitable for bench-top applications, but the source does not provide high enough SNR for biomedical applications. In order to measure mid-IR radiation different instruments have been used. A dispersive spectrometer was the first instrument that obtained an infrared spectrum. Such a spectrometer can use a prism or gratings in order to disperse the light before it is transmitted through a monochromator. The light is then transmitted through the sample into the detector. It is very time-consuming to scan the entire wavelength range due to the monochromator, which removes most of the source light to achieve one wavelength at a time, and the instrument is not used that much anymore.

Another spectroscopy method that is more frequently used today is the Fourier-transform infrared (FTIR) spectroscopy. This type of setup or instrument employs an interferogram, which is generated when the radiation of two beams interfere. The interferogram results in a signal that is the function of the difference in pathlength between two beams. This signal is then transformed from distance to frequency using Fourier transformation. Different interferometers can be used in FTIR spectroscopy and the most common is to use a Michelson interferometer. Figure 2.4 shows a sketch of a basic setup for an FTIR spectrometer using a Michelson interferometer.

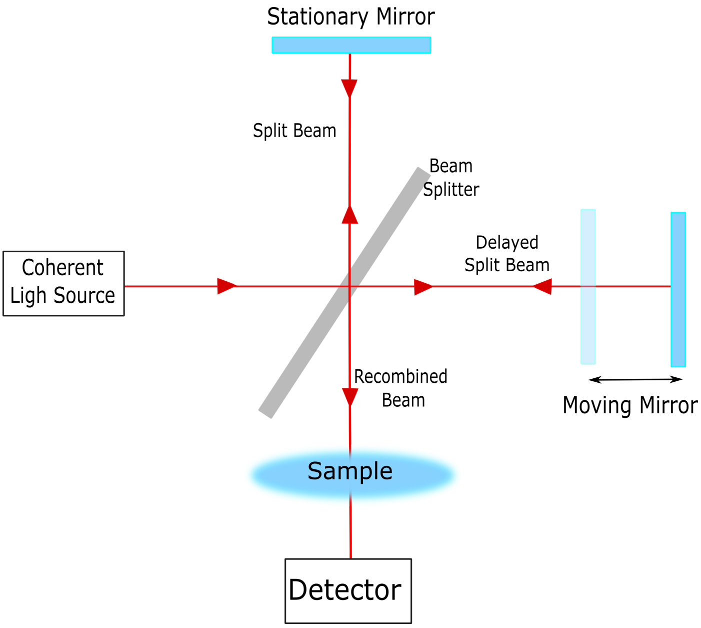


Figure 2.4: Basic set-up for an FTIR spectrometer, using a Michelson interferometer. Adapted from [17]

This type of interferometer consists of a beamsplitter that splits a coherent light source into two beams. One of the split beams travels towards a stationary mirror and the other towards a moving mirror which will be delayed. Both beams are reflected at the mirrors and returned to the beamsplitter and recombine while interfering. The recombined beam is then transmitted through the sample and into a detector. Only 50 % of the beam will be sent through the sample since the other 50 % will be sent back to the source due to the nature of the beamsplitter. The optical path difference which will be transformed by using the Fourier transformation is a result of the moving mirror. The sources, beamsplitters, and detectors used in FTIR spectroscopy depends on the wavelength. For mid-IR the most common source is a silicon carbide element which is heated to about 1200 K, KBr (Potassium bromide) is the most commonly used material for the beamsplitter and the detector mostly used is a DTGS (deuterated triglycine sulfate) pyroelectric detector. The source most commonly used is not a high-energy source and the resulting signal-to-noise ratio (SNR) have thereby been low. A need for higher-energy sources has thereby been required to improve the spectroscopy methods using mid-IR.

2.2 Spectroscopy

Within infrared spectroscopy there are two main methods of acquiring infrared spectra; transmission and reflectance. The following section is included in order to understand the difference between how the two methods obtain infrared spectra. The section is mainly based on the book *Infrared Spectroscopy: Fundamentals and Applications* by Barbara H. Stuart [15].

2.2.1 Transmission Methods

The transmission method is considered to be the simplest of the two since the setups usually consist of few components. In general the transmission setups mainly consist of a source, sample and detector, see Figure 2.5. The light may be transmitted through the sample via an optical fibre or through air, and is then detected on the other side of the sample by a detector. The light will be absorbed by analytes in the sample, and at the wavelengths where the analytes absorb the energy will be lower. This method can be used for measuring samples in liquid, solid or gaseous form. Liquids can be used directly and are contained in a cell or some sort of container. It is important that the windows in the cell are transparent to the infrared radiation. Gas samples are also measured and are usually contained in a chamber which the light will be transmitted through. Solid samples however are usually too thick to be used directly in the transmission method. Therefore solid samples normally have to be processed into thin discs or films, which will allow for transmission of infrared radiation through them.

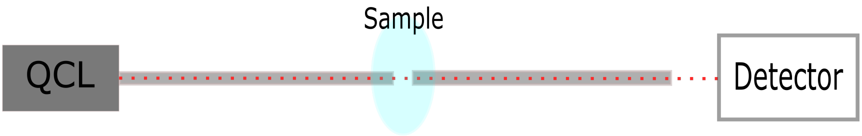


Figure 2.5: Sketch of fibre-coupled transmission i.e. where the light is transmitted to the sample through a fibre.

In a transmission setup there is a trade-off between the pathlength and the signal strength. When light is transmitted through the sample some of the light will be absorbed. If the pathlength is large then more of the light will be absorbed by the sample according to Beer-Lamberts Law, see Equation 2.1 [18]. In the equation A is the absorbance, I is the intensity of the sample measurement, I_0 is the intensity of the reference measurement, ε is the absorption coefficient, l is the pathlength and c is the concentration of the sample. If the pathlength is increased the absorbance will increase accordingly and the intensity will be reduced.

$$A = -log\frac{I}{I_0} = \varepsilon lc \tag{2.1}$$

2.2.2 Reflectance Methods

The second method for acquiring infrared spectra is the reflectance method. This method can further be divided into two categories; internal and external reflectance, see Figure 2.6. The external reflectance method is based on an infrared beam being directly reflected from a sample surface, this then measures e.g. specular or diffuse reflectance of the surface. Since the sample surface has to be reflective this method can not be used for all samples. The method that will be used in this master thesis is internal reflectance which can be utilized by using attenuated total reflectance (ATR) spectroscopy; this will be described below.

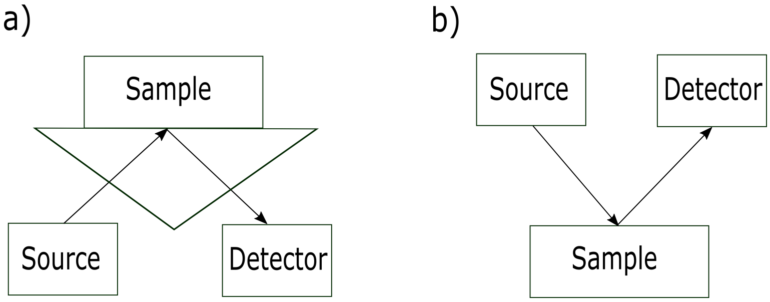


Figure 2.6: Sketch showing the difference between the two reflectance methods for spectroscopy. **a)** Internal reflection is where the reflection is induced by an object and then the light interacts with the sample. **b)** External reflectance is when the sample itself is reflective and induces the reflection at the surface.

ATR spectroscopy uses a prism or fibre made of high-index material and total internal reflection (TIR). Since a prism is used in this thesis, the method will be explained utilizing a prism. To undergo TIR, meaning that all of the light is reflected and none is refracted, the angle of incident radiation θ_i has to be greater than the critical angle, see Figure 2.7. The critical angle is defined as when the angle of refraction (θ_2) becomes 90° and Snell's law is reduced from Equation 2.2 to Equation 2.3 [19]. Here n_1 is the refractive index of the the prism and n_2 is the refractive index of the sample placed on top of the prism.

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{n_2}{n_1} \tag{2.2}$$

$$\sin(\theta_c) = \frac{n_2}{n_1}, n_1 \neq n_2 \tag{2.3}$$

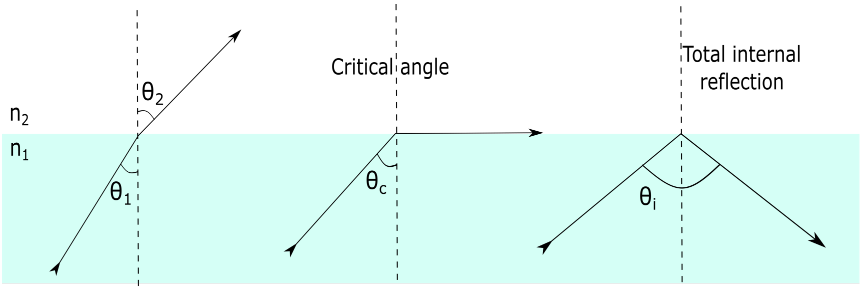


Figure 2.7: Schematics of Snell's law. When the angle of the incident ray (θ_i) is larger than the critical angle θ_c all of the light is reflected and TIR is achieved.

When the angle of incident light in the prism is larger than the critical angle ($\theta_i > \theta_c$) the beam undergoes total internal reflection. As a result of this an evanescent field will be induced which will penetrate out of the prism at the reflecting surface, see Figure 2.8. This evanescent field will then be able to interact with a sample placed on top of the prism and the beam will lose energy at the wavelengths where the specific analyte absorbs.

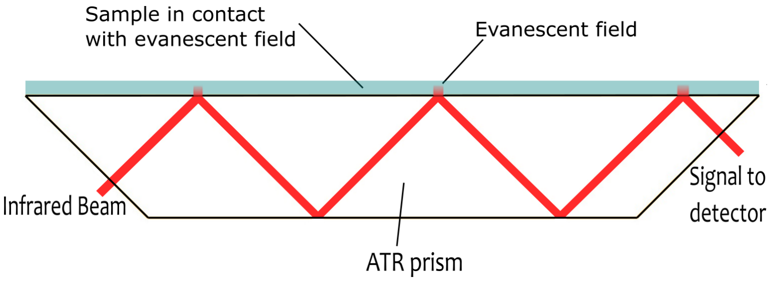


Figure 2.8: Sketch of a reflectance spectroscopy setup, utilizing an ATR prism. When TIR occur an evanescent field is induced which can penetrate out of the prism in order to interact with a sample placed on top of the prism.

The angle of the light sent into the prism will be refracted when the light is transmitted from air or fibre into the prism, see Figure 2.9. The angle of the refracted light within the prism is θ_2', which can be found by using Equations 2.4 - 2.6. In the equations $\theta_{i,f}$ is the angle of the light sent into the prism. The angle related to the critical angle in order to achieve TIR is $\theta_{i,p}$, which can be found using Equation 2.7. This angle has to be larger than the critical angle θ_c for TIR to occur. n_2 is the refractive index of the prism and n_1 is the refractive index of the surrounding medium from which the lights propagates from into the prism.

$$\theta_1 = \theta_{i,f} - 45^\circ \tag{2.4}$$

$$\theta_2 = \sin^{-1}\left(\frac{n_1}{n_2} sin\theta_1\right) \tag{2.5}$$

$$\theta_2' = 45^\circ + \theta_2 \tag{2.6}$$

$$\theta_{i,p} = 90^\circ - \theta_2' \tag{2.7}$$

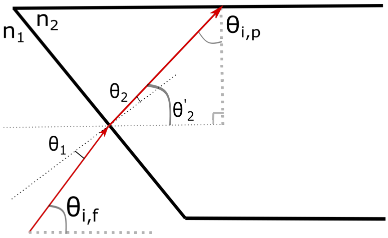


Figure 2.9: A ray trace through a prism where n_2 is the refractive index of the prism, n_1 is the refractive index of the medium the light goes from into the prism e.g. air or fibre. The refraction angle is θ_2', the angle of the incident light into the prism is $\theta_{i,f}$ and the incident light to the top surface of the prism is $\theta_{i,p}$.

It is desirable for the beam to undergo TIR multiple times within the prism. Each time the beam undergoes TIR, an evanescent field is induced at the surface which can interact with the sample. Due to this, more reflections within the prism leads to more evanescent fields induced at the surface of the prism which again leads to a larger interaction area with the sample. It is possible to measure both liquid and solid samples with ATR spectroscopy. Liquid samples are simply placed on top of the prism and solid samples are measured by pressing the sample on to the prism.

The penetration depth of the light into the sample in ATR spectroscopy depends on the wavelength λ, the refractive index of the prism n_2, the refractive index of the sample n_1 and the angle of the incident radiation $\theta_{i,f}$. To calculate the maximum and minimum penetration depth that can be achieved with a scan, from 1200 cm^{-1} to 905cm^{-1} (8.33 µm to 11.04 µm) Equation 2.8 is used [15].

$$d_p = \left(\frac{\lambda}{n_1}\right) / \left(2\pi[\sin\theta_{i,f} - \left(\frac{n_1}{n_2}\right)^2]\right)^{\frac{1}{2}} \tag{2.8}$$

The prisms used in ATR spectroscopy are made of materials with low solubility in water and with a high refractive index. The most commonly used materials are diamond, Zinc selenide (ZnSe) and Zinc sulfide (ZnS) [20]. The refractive index for both diamond and ZnSe is 2.4 and for ZnS the refractive index is 2.2. Diamond is highly resistant to scratching and abrasion but due to the high price the use of this material in a commercial product is limited. The drawback with ZnSe is that the material is toxic which does not makes it suitable for biomedical spectroscopy. ZnS is therefore the most suited material for the prism used in this thesis, despite of the lower refractive index.

2.2.3 Measurement Placement for a CGM Device

Biomedical spectroscopy can be used for both invasive and noninvasive glucose monitoring. If the glucose concentration would be measured directly through the skin it is only possible to measure it in the interstitial fluid (ISF). Due to strong water absorption the penetration depth of mid-IR in skin is short, between 50 - 100 μm [21]. The blood vessels are thus too far down in the skin for mid-IR radiation to reach them. ISF is the fluid which surrounds the cells in the body and the amount of different analytes in the ISF is proportional to the amount in the blood, with a delay of about 5-10 min [22].

The main problem with this approach is that the water absorption is very high in the mid-IR region [7] which makes it hard to detect absorption from other analytes in the ISF, see Figure 2.10. The strong water absorption can be clearly seen in this figure together with the weaker glucose absorption. Glucose absorbs at other wavelengths than the one marked in the figure (around 3000 cm^{-1} and 1400 cm^{-1}) but these are strongly overlapping with the water absorption peaks [14]. Water absorption is also found in the region from 900 to 1500 cm^{-1} but since it is low here, it is possible to detect other analytes. Previously the problem with water absorption meant that the light would not be able to penetrate deep enough into the body to interact with neither ISF or the blood vessels. However this problem was overcome by the introduction of stronger mid-IR sources.

Another approach to measure the glucose level is with an invasive measurement method. By placing the measuring device inside of the body the glucose level can be measured in the ISF or directly in the blood. The main problem about measuring analytes directly in the blood using mid-IR light is the strong absorption of hemoglobin. Similar to the strong water absorption this makes it difficult to measure and detect the absorption of other analytes, due to the low SNR and overlap between absorption bands.

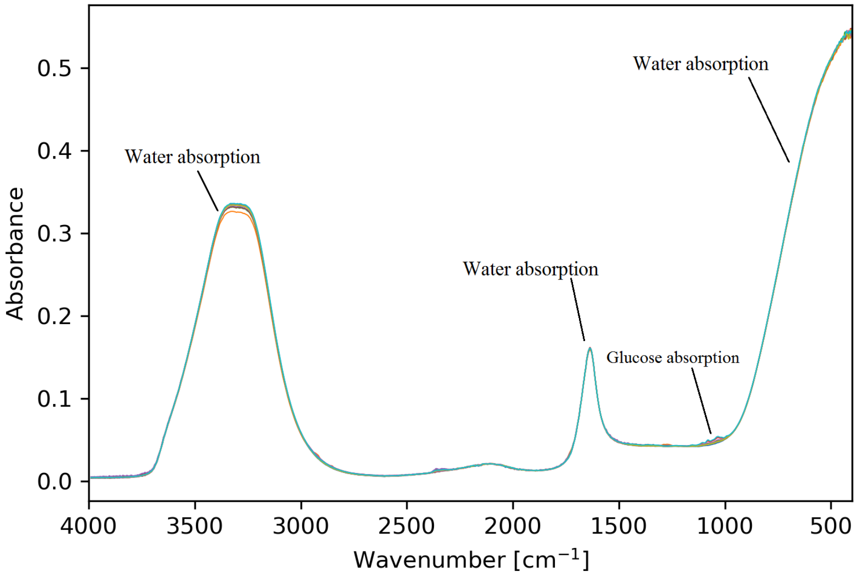


Figure 2.10: Measurements of samples with different glucose concentrations in water using FTIR spectroscopy showing how strong water absorption can affect the measurement of analytes with weaker absorption. Glucose absorption is found around 3000, 1400 and 1200-1000 cm^{-1}, only the glucose absorption at 1200-1000 cm^{-1} is marked in the figure.

2.3 Quantum Cascade Laser

In order to use mid-IR spectroscopy to measure biological samples, improvements regarding mid-IR light sources were necessary. The FTIR measurement method has mostly been used so far. Thermal light sources are commonly used as the light source in conventional FTIR spectrometers [21] and will here be used for comparisons to the newer and improved mid-IR light sources. FTIR has thus far not been suitable for non-invasive and invasive measurements of biological analytes in the human body. The spectral power density has not been strong enough to overcome the problem of high water absorption in the mid-IR regime which has resulted in a low SNR. Newer light sources such as quantum cascade lasers (QCL) offer an increased spectral power density by a factor of 10^4 and more [21]. This increase in intensity results in a higher SNR, and having higher emission power permits a longer path length for transmission measurements, even in the presence of water absorption. Single-wavelength QCLs have mainly been used for gas phase analysis, but when external cavity QCLs (EC-QCLs) were introduced the large spectral tuning range together with the high intensity made them more usable for liquid samples as well. It has also been shown that the acquisition time of an IR spectrum is shorter in QCL-based IR spectroscopy, compared to FTIR spectroscopy [21].

The main advantages of this improved mid-IR source are the wider tunability, stable operation at room temperature, as the first of its kind and the high emission power. The quantum cascade laser is the main component in the setup utilized in this thesis to measure glucose concentrations. Due to this, it is described in detail in order to get a better understanding of how the advantages over the standard semiconducting laser are achieved. The text is based on the book *Quantum Cascade Lasers* by Jérôme Faist [23] and an article with the same name written by Federico Capasso et. al. [24]. Both Faist and Capasso were part of the group that invented and first demonstrated the quantum cascade laser in 1994.

2.3.1 Interband and Intersubband Transitions

Standard semiconducting lasers are based on the recombination of electrons and holes when traveling over a given band gap (E_g). To change the wavelength for this kind of laser the material needs to be changed in order to obtain a different band gap, this will result in emission with a different wavelength. To get wavelengths in the mid-IR region using this scheme, the band gap would have to be reduced greatly. The operation of such a laser becomes more challenging because of the maximum operating temperature and the fact that stable temperature is required to avoid the thermal runaway effect and thermal recycling. Due to these thermal effects, it is not possible to achieve a semiconducting laser with wavelengths in the mid-IR region. The difference between QCLs and standard semiconducting lasers is that QCLs do not rely on the band gap for light emission, but rather achieve emission by an intersubband transition of electrons. The principle of the QCL is that laser photons are generated by electrons that undergo quantum jumps between energy levels [24]. The transition energy required here will then be the difference between the confinement energies of the individual electronic states [25]. The electrons jump between energy level E_n and E_{n-1}, which then creates a photon with frequency $\frac{(E_n - E_{n-1})}{h}$.

2.3.2 Quantum Cascade Laser Design

The quantum well structure consists of an ultra-thin layer of a material with a narrow band gap that is surrounded by layers of another material with a wider band gap in a varying manner, forming a superlattice. The active regions are structured into quantum wells to make the energy levels. Since the positions of the energy levels are not determined by the material but by the thickness of the layers, there is a greater possibility to tune the laser over a wide range of frequencies in the same heterostructured material [25]. The standard semiconducting laser has band transitions where the lasing occurs due to recombination of electrons and holes, which will remove the given electron from the conduction band. In the QCL however, one electron has the possibility of inducing more than one lasing photon. The electrons stay in the ”conduction band” after emission and have the ability to tunnel down to the next quantum well and energy level. This results in a cascade emission of photons, which is the reason for the name, quantum cascade laser. To get this emission, the active regions have to be alternated with doped electron injectors and it is important that a sufficient bias voltage is applied to the structure. The QCLs are able to achieve such a high power output due to the cascade emission of photons. Because the levels appear in the quantum wells, the well structure will be oblique, as shown in Figure 2.11.

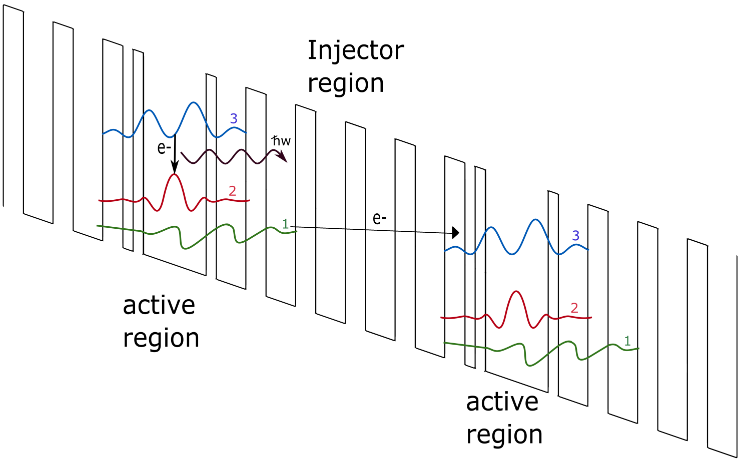


Figure 2.11: Schematic band diagram which illustrates two stages of a cascaded well structure of a QCL chip. Each stage consists of an active region and an injector region. When only three energy states are considered the radiative transition occurs between state 3 and 2 and the population inversion is maintained due to a rapid depopulation of the lower state 2 into state 1. The electron will after a lasing transition couple into the next region. Adapted from [23, 24].

The number of stages (i.e. the amount of active and injector regions) in the laser can be all from one to over 100, but it is common to have between 20 and 35 stages. Population inversion is required for lasing actions and if a 3-energy-layer quantum well is considered there has to be a population inversion between state 3 and state 2. The laser transition will be defined by the energy difference between these two states. To obtain population inversion the electron relaxation time between the two highest states, 3 and 2, has to be significantly longer than the lifetime of state 2. To fulfill this requirement the two lowest states should be separated by the energy of an optical phonon ($\approx$ 34 meV). The energy difference is thus larger between the two highest states than between the two lower states.

Electrons are injected into the highest state and what is important then is to prevent these electrons from escaping out to the right in Figure 2.11. An escape like that will reduce the population of the level, which will make it harder to achieve population inversion. The barrier heights should be between 0.7 and 0.8 eV in order to reduce electron leakage over the well barriers. The energy separation between the lower state of the laser transition and the injector ground state is also an important parameter for improving the lasers temperature performance.

2.3.3 Types of QCLs and Their Field of Applications

QCLs can be classified into groups based on what type of resonator design they use. The three resonators that are most commonly used in the lasers are; Fabry Pérot (FP), distributed feedback (DFB) and external cavity (EC) [21]. The active regions will be the same for all the different types so when there is a difference in the losses it is due to the different resonator designs. Of the three most common types the FP-QCL consists of the simplest resonator design as it is only based on the QCL chip having a high reflecting coating on the end facet of the laser ridge. The determining parameter in this type of QCL will be the cavity length. The reason for this is that the gain from the active material has to be sufficient and the distance between the mirrors has to lead to constructive interference. Over a wide spectral range there will be multimode emission as a result of the standing wave condition being fulfilled.

In spectroscopic applications when the sample is in gas phase, a laser with single mode emission is preferred over multimode emission. To achieve single mode the DFB-QCL can be used. This type selects a single mode within the gain from the active medium by having the Bragg grating on top of the QCL chip along the light propagation direction. The Bragg grating leads to losses in all other modes except the one it is designed for. The wavelength of such a laser can be tuned by changing the injector current and/or changing the operating temperature. By changing one or both of these parameters the effective refractive index of the laser chip material will change which result in a shift of the resonance wavelength of the Bragg grating. The tuning range for a DFB-QCL is approximately 5 cm^{-1}. This type of QCL is more suited for gas measurements as it has narrowband single-mode emission. It is possible to combine multiple DFB-QCLs into an array to get a broader spectral coverage [21].

For the EC design, an external diffraction grating is incorporated in order to get broadband spectral tuning. It is possible to get tuning ranges of up to several hundred wavenumbers by changing the angle of the diffraction grating relative to the QCL chip. This type of QCL can have emission both as pulsed and as a standard continuous wave. Both of these operations allow for very large spectral tuning range. However, when used in high-resolution gas-phase spectroscopy the pulsed and continuous wave configurations might lead to problems due to mode-hops that might occur during tuning. The mode-hops/jitter is a result of the grating not being completely accurate. The effect leads to the laser emission ”hopping” from one mode to another [21]. A figure showing an EC-QCL is shown in Figure 2.12, and the design of this type is included since the laser utilized in this thesis is such a laser.

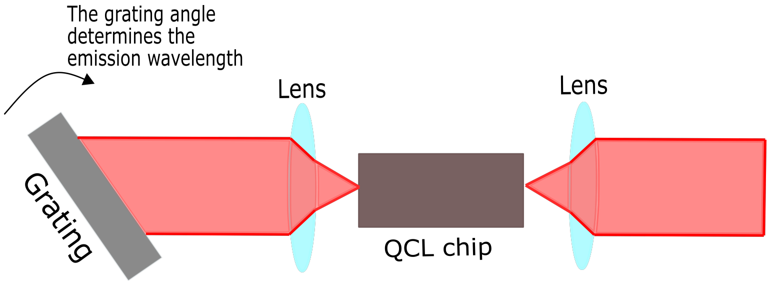


Figure 2.12: Schematic of an external cavity quantum cascade laser. This type of QCL enables broadband spectral tuning by changing the angle of the grating, which determines the emission wavelength. Adapted from [26].

2.4 Overview of Mid-IR Spectroscopy Research for Glucose Measurements

2.4.1 Other Research in General

Previous work will be described to get a clear image of what the obstacles are related to measuring biochemical analytes in the body using mid-IR. Based on previous work it is clear that measuring chemical analytes in the body using mid-IR is a difficult task due to the strong water and hemoglobin absorption in this spectral range. This has shown itself to be a problem not just for non-invasive measurement methods, but also for invasive measurements methods. Previous attempts of analyzing blood or ISF using mid-IR spectroscopy have mainly been done with the use of FTIR. Sources used in FTIR have not provided strong enough intensity to give a precise and robust measurement of analytes in blood, ISF, or water. In addition to this they are considered to be bulky, which makes them less suited for e.g. personal CGM devices [27]. Recent work have instead used the newer mid-IR light source, the QCL, which has higher emission power. A discussion follows on recent work in the field of mid-IR spectroscopy, mainly in relation to the measurement of glucose.

ATR Spectroscopy

Since strong water and hemoglobin absorption is a problem with mid-IR spectroscopy, different research groups have tried to improve the signal strength of the light transmitted into the body in non-invasive measurements. To achieve higher analyte sensitivity ATR prisms have been utilized in setups by different research groups[28, 29]. A prism like this is used in order to cope with the small penetration depths which is common when mid-IR light is used. By making a prism of a high-index material, the light will be totally reflected at the surface of the prism which will be in contact with a sample. The totally reflected light will induce an evanescent field outside the prism which will interact with the sample and be absorbed by it. A sketch of the main principle of an ATR prism is shown in Figure 2.13. More theory related to ATR spectroscopy was presented in Section 2.2.

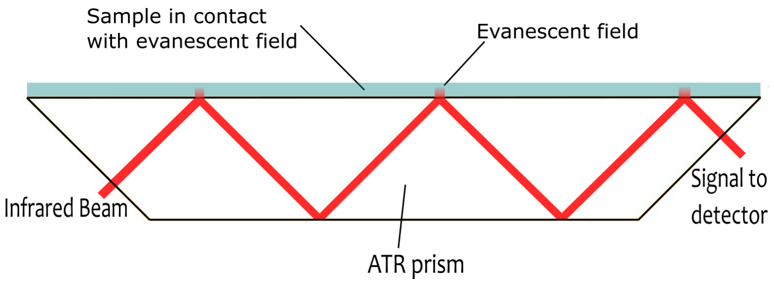


Figure 2.13: Setup utilizing an ATR prism. The infrared beam is sent into the prism and will experience total internal reflection which will induce an evanescent field on the top of the prism. This evanescent field will be able to interact with analytes on the contact surface.

Different designs have been investigated, e.g. a trapezoidal prism to get nine reflections was tested by Kino et. al. [28]. They based their design on the fact that a multi-reflection prism can contribute to achieving higher sensitivity and that a flat and wide contact surface will give higher measurement reproducibility. The group tested their device in vivo on the inner lip mucosa in people. The measurement site was chosen because oral mucosa has no keratinized layer which makes it possible for IR light to penetrate deeper tissues. In addition to this, the blood vessels are found to be closer to the surface at this point [28]. The same group later tried to use an ATR prism at the end of a probe to get a more flexible measurement device [29]. They place a rooftop-shaped prism at the end of their fibre probe and 73 % of the incident power was returned. This was significantly higher than when they tried using a cone-shaped ATR, both on the outside and the inside of the hollow fibre, these only got 26% and 34 % of the incident power returned, respectively. They used the rooftop-shaped prism to measure on skin with and without sunscreen as well as at the inner wall of an extracted small intestine of a pig.

Fibre Probe

Other groups have based their research on using a fibre probe for detection, these groups collect the backscattered light somewhat differently [7, 30]. Liakat et. al [7] collected the backscattered light through a regular hollow-core fibre. They used hollow-core fibres both to deliver the light and to collect the backscattered light from a human palm. Their measurements were done through the skin in the palm at a specific area that has blood vessels close to the skin surface.

Another paper by the same group [30], improved the design in order to collect more backscattered light by replacing the hollow-core fibre contact with an integrating sphere which was incorporated into the sensor. By using a sphere, the signal stability was seen to be significantly enhanced. They utilized an integrating sphere as it is known to be the best optical component to collect scattered light from a wide angular range. To collect the backscattered light, also from a palm, they used a TE-cooled detector in order to reduce the form factor and to make the sensor mobile. The backscattered light reflected from the human skin is very complex [30] and due to this a sphere was used to collect the scattered light. However, such a sphere has a bulky shape and a large physical size which makes it difficult to handle. It is also difficult to control the measured area accurately due to the shape [31].

Invasive Measurements

Not all previous work has been concentrated on non-invasive measurements of biological analytes in the body. Some groups have looked more into using optical measurement methods in an invasive manner. Vrančić et. al. have investigated making a CGM device in vivo with transcutaneous fibre-based mid-IR spectroscopy [32]. Two optical fibres were placed below the skin with a gap which was filled with the tissue ISF. The light source was a pulsed, nontunable QCL operated at 9.7 μm. When tested in rats a significant amount of the light transmitted over the gap was detected. A figure describing the set up in the paper is shown in Figure 2.14.

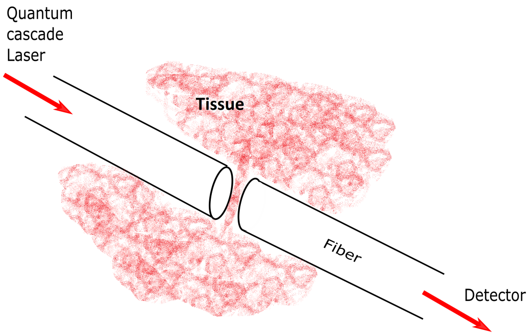


Figure 2.14: Fibre-based mid-IR spectroscopy where the light is sent through two fibres separated with a gap which can be filled with ISF.

Measurement of Different Analytes

Some of the previous papers have not only focused on the measurement of glucose, but also on the measurement of other analytes that can be found in the body. Groups that are working towards a CGM device usually concentrate on analytes in the body that can be related to diabetes and other disease monitoring. Analytes like this can be lactate, triglycerides, cholesterol and various proteins such as albumin. It has been shown in a paper by Brandstetter et. al. [27] that it is possible to measure and predict values for the analytes mentioned above. The prediction was not as precise as for glucose but it was shown to be sufficient for disease progression monitoring related to the different analytes. Interestingly it was also shown that the protein albumin, which is most distinct at 1625 cm^{-1} and the area around, could also be identified at lower wavenumbers as different QCLs have different ranges and not all include 1625 cm^{-1}. Albumin was shown to have relevant spectral features round 1148, 1165 and 1175 cm^{-1}. It may also be beneficial to analyze samples containing different analytes present in the blood which are not necessarily directly related to diabetes to investigate if they affect the measurement and detection of glucose.

Other Uses for Mid-IR Spectroscopy

Improving the mid-IR spectroscopy setup is not only beneficial for non-invasive or invasive measurements of glucose. Brandstetter et. al. [27] were able to make a portable sensory system that could be used for performing a direct mid-IR transmission measurement of samples with blood plasma. They did this with a broadly tunable EC-QCL that provided high spectral power density. Together with this, they had a $165\mu m$ flow cell which did not clog during the measurements of 300 samples. Since a broadly tunable laser (8.13 - 9.7 µm) is used, the device could be utilized to detect many biological analytes, not just glucose. This results in the device acting as a general blood analysis device. The advantages of such a device over the standard way of analyzing blood used in hospital laboratories are; faster acquisition time of the results and smaller sample sizes. This device has the ability to compete with the standard way of analyzing blood in laboratories, but the device is large and bulky and therefore not something that a person would be able to wear on their body. This would be a large drawback in relation to using the device as a CGM device.

Despite the fact that there are many different groups working on improving the measurement of analytes in the human body using mid-IR, there is still potential for improvement. The ultimate goal for developing this measurement method is, in general, to be able to use it in a CGM device. Such a device has the possibility of greatly improving the lives of people with diabetes type 1.

2.4.2 Other Research Directly Related to This Master Thesis

The following text includes previous work that is directly related to the setups and methods used in this master thesis. The main parts of the work will be described below and the most important results obtained in the articles will be presented here in order to have comparable values to the results that will be presented in this thesis.

Haas et.al. used a broadly tunable QCL in an ATR spectroscopy setup where the internal reflecting element (IRE) was a polycrystalline diamond [33]. They measured the glucose concentration in saliva of healthy people between the age of 20 and 30. In order to increase the saliva sugar levels, the volunteers had to drink 500 ml of Coca Cola one hour after the first measurement. The group was able to achieve a limit of detection of 0.02 mg/ml, which is well within the level that is acceptable for the detection of medical samples. The group only used a single band in the mid-IR range and only focused on the CO vibration of the glucose molecule. The main goal of this article was to see how a polycrystalline diamond would behave as the IRE, compared to single crystalline diamond. Through the simulation in this paper, it was possible to see that the mid-IR beam would diverge within the prism.

Another group, Haase, et. al. [34] used a fibre-based sensor in a fluidic chamber. They used the same laser as in this master thesis and measured pure glucose solutions and samples with a mixture of glucose and other saccharides. The glucose sensitivity in the pure glucose solutions was found to be 3 mg/dl. When glucose was detected with other analytes, the glucose selectivity was 8 mg/dl when only different monosaccharides were included and 14 mg/dl when other mono- and disaccharides were included. The glucose concentrations they used in the pure glucose solutions were; 0 mg/dl, 50 mg/dl, 100 mg/dl, 200 mg/dl, 250 mg/dl, 400 mg/dl, 500 mg/dl and 1000 mg/dl. When mixed with other analytes the maximum glucose concentration was 600 mg/dl. The setup in this article had a second detector which measured the lasers' output power before interfering with the sample in order to correct for power variations during the measurements. If the setup would be miniaturized the second detector will increase the size, which is not ideal for a personal CGM device which will be used on the body. Optimal design for the concentrations in the measured samples was also not used which means that not all relevant concentration combinations were measured.

The same group also used a transflectance setup with a specially designed optofluidic interface where the mid-IR beam was reflected at a gold surface below the liquid layer to be measured [35]. The solutions that were measured were; pure glucose solutions (average concentration = 214 mg/dl), glucose in the presence of other saccharides (0-600 mg/dl) and glucose in the presence of albumin (0-500 mg/dl). They found that the glucose sensitivity in the pure glucose concentration was 3.2 mg/dl. The glucose selectivity was 12 mg/dl when glucose was mixed with other monosaccharides, and below 10 mg/dl when glucose was in the presence of albumin. These results were achieved when a specific amount of wavenumbers were included and cross-validation was performed with ”leave-one-cycle-out”, i.e. using one of 5 measurement series obtained over 3 months as the validation data. A reference arm was also used in this article in order to monitor the lasers' output power. In addition to this, the setup in this paper included a second laser which covered the wavelength range 1500 - 1750 cm^{-1}. The second laser in this setup would increase the size of a miniaturized version of the setup as well as increase the price.

2.5 Data Analysis and Signal Processing

2.5.1 Regression Analysis

After data is obtained through spectroscopy, regression analysis has to be performed in order to estimate the amount of the different analytes in the sample. This type of analysis has to be used because of the weak detected signal and overlap between absorptions bands. It is not possible to simply use the intensity of individual peaks, instead we need to look at the entire spectrum. There are different regression models that can be used; linear, multilevel, nonlinear and principal components, to mention some. In the chemistry environment and in relation to spectroscopy the principal component analysis (PCA) is often used [7]. Partial least squares regression (PLSR) has been used as the analysis method here but PCA will be introduced first as it in many ways is a simplified version. Most of the following text is based on the book *Multivariate Analysis of Quality - An introduction* by Harald Martens and Magni Martens [36].

Principal Component Analysis

Principal components (PCs), also called latent variables, are variables that cannot be measured directly. In regard to PCA and PLSR these variables are the directions describing the largest variance in the data set. PCA is the analysis of one single set of input variables, **X** and the multivariate model for PCA can be seen in Equation 2.9.

$$\mathbf{X} = \mathbf{TP'} + \mathbf{P}. \tag{2.9}$$

T are referred to as scores and represent the original data in a rotated coordinate system. **P** are loadings and represent the weight that the original variable has to be multiplied with in order to get the component score, e.g. the rotated coordinates. The basic idea of PCA is to reduce the complexity of the data with matrix transformations, and use the N first PCs in the data analysis. The PCs can then be used in a regression model to predict, e.g. concentrations, and this is called principal component regression (PCR). It is thought that a small number of PCs are sufficient for explaining most of the variability in the data. Fitting a least square model to the PCs will usually yield better results than fitting a least square model to the input data. By reverse engineering the model for PCA, the model will be able to describe **X**.

Partial Least Square Regression

PLSR is a method that relates two data matrices, X and Y, by using a linear multivariate model. Other ways of performing regression also do this, but what is special with PLSR is that it also models the structure of X and Y. Another advantage of this model is that it has the possibility of analyzing data that is strongly correlated and experiences a lot of noise. The model is capable of analyzing data with numerous X-variables and then simultaneously model several response variables, $\mathbf{Y}$. In relation to absorption measurements which are done in this thesis, $\mathbf{X}$ = intensity at different wavenumbers and $\mathbf{Y}$ = concentrations. By increasing the number of relevant variables and observations included in the model the precision of the model parameters will also increase. PLSR reminds in many ways of PCR. The main difference between the two is that PLSR analyses the data from two data tables X and Y, whereas PCA only analyses the data from one data table X. PLSR also uses the response Y in order to identify the PCs, meaning that the PCs are found in a *supervised* way. This is not the case in PCA which means that there is no actual guarantee that the PCs that are found will be the best for predicting the response, even though they might be the best for explaining the predictors.

PLSR is, in similar ways to PCA, a dimension reduction method that is based on finding a new set of variables that summarizes the systematic pattern of variation between the samples. The aim is to collect all of the relevant and reliable information from the input variables into a reduced sequence that only consists of latent variables that are highly informative, the PCs. The input variables are usually referred to as $\mathbf{X} = [\vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots]$, the PCs are given as $\mathbf{T} = [\vec{t_1}, \vec{t_2}, \vec{t_3}, \ldots]$ and can be found by using the mathematical function $\mathbf{T} = w(\mathbf{X})$. In PLSR this function is defined in such a way that the first few PCs, in addition to maximizing the covariance within $\mathbf{X}$, are as Y-relevant as possible. The variables ($\mathbf{X}$ and $\mathbf{Y}$) can be modeled by the score ($\mathbf{T}$), their loadings ($\mathbf{P}$ and $\mathbf{Q}$) transposed (') and their residuals ($\mathbf{E}$ and $\mathbf{F}$) [36], see Equation 2.10 and 2.11. The residuals indicate how much of the data that has not yet been modeled. It is ideal to get the residuals as close to zero as possible. The X-residuals ($\mathbf{E}$) will be small if the principal components model the structure of the X-data more or less identically. The Y-residual ($\mathbf{F}$) will appear small if the modeled relationship between $\mathbf{X}$ and $\mathbf{Y}$ closely resembles the relationship between the original data. By increasing the numbers of PCs the residuals can be reduced.

$$\mathbf{X} = \mathbf{1}\bar{\mathrm{x}} + \mathbf{T}\mathbf{P}' + \mathbf{E} \tag{2.10}$$

$$\mathbf{Y} = \mathbf{1}\bar{\mathrm{y}} + \mathbf{T}\mathbf{Q}' + \mathbf{F} \tag{2.11}$$

In the model, Y has the ability to be expressed directly from $\mathbf{X}$ by using the regression coefficient $\mathbf{B}$. Y is therefore directly related to $\mathbf{X}$ in Equation 2.12. $\bar{\mathbf{b_0}}$ is here the offset of the model, meaning the expected value of Y if $\mathbf{X} = 0$.

$$\mathbf{Y} = \mathbf{1}\bar{\mathbf{b_0}} + \mathbf{X}\mathbf{B} + \mathbf{F} \tag{2.12}$$

It is important that the number of PCs are optimal. By using too few PCs there will be an underfitting of the model and the description of $\mathbf{X}$ and $\mathbf{Y}$ will not be good enough. The opposite happens when there are too many PCs and the model is overfitted. In this case, the model will be well fitted, but with little or no predictive power. It is common that a model experiences overfitting due to the fact that people like to get a good fit between model and data. Due to this, it is not always optimal to use the number of PCs that results in the smallest residual, as this may result in overfitting. The model consists of two phases; the training or calibration phase, and the prediction phase. The purpose of the calibration phase is to make a calibration model that has the ability to predict parameters $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, .., \mathbf{y}_n]$ from data $\mathbf{X}$. The purpose of the prediction phase will then be to use the calibration model found in the previous phase to predict the parameters, e.g. the concentration of solutions.

The root mean square error of prediction (RMSEP) and root mean square error of cross-validation (RMSECV) can be used for verifying a models' ability to predict and to investigate how stable the model is. RMSEP is a measure that calculates the estimate of how large the error is between the predicted and measured values from an entire data set. RMSECV predicts the error between the predicted and measured values when cross-validation is used. Cross-validation takes out a portion of the samples and uses the rest of the samples in the data set to train the model for predicting the removed samples. Leave-5-out cross-validation will take out 5 samples and use these as the validation data. The error is desired to be small between the predicted and measured values and consequently small RMSEP and RMSECV values are desired.

The coefficient of determination (R^2) is also important in regression analysis. R^2 is the proportion of the variance in the predicted values, that are predictable from the measured values. A high R^2 indicated that a lot of the predicted values fall within the regression line. If R^2 is 0 the validation values cannot be predicted from the measured values and none of the predicted values falls within the regression line. If R^2 is 1 the validation values can be predicted without any error from the measured values and all of the predicted values fall within the regression line. Consequently, is it desired that the $R^2$is high since a low or negative R^2 indicated that the model is a poor fit for the data.

2.5.2 Preprocessing

Before PLSR can be performed the data might need to be preprocessed. Depending on the data this might include normalization, standardization, aligning and/or filtering. Some filtering methods, a normalization method, a standardization method, and alignment will be described here since these can be used in the data analysis program in this master thesis.

Standardisation and Normalization

In order to achieve a common scale for all of the data sets used in the regression analysis, the data might need to be normalized or standardized if the amplitude of the data varies between the data sets.

Normalization will rescale the data to be fixed between 0 and 1. Normalization can be done different ways and a simple way is multiplying the data with the factor given in Equation 2.13. This factor will thus be multiplied with each data point from a given dataset and rescale the data to be fixed between 0 and 1.

$$\frac{1.0}{(\sum(\text{data}^2))^2} \tag{2.13}$$

Standardization is another method of rescaling the data. By using standardization the data is rescaled into having a mean (μ) of 0 and standard deviation (σ) of 1 (unit variance). In order to standardize the data(x) Equation 2.14 can be used. The equation will result in a standardized data set (z) that represents how much the standard deviation is above or below the mean for each observation.

$$z = \frac{x - \mu}{\sigma} \tag{2.14}$$

Alignment

If the starting point of the scans within one measurement differs with a few wavenumbers due to mode-hop they should be aligned. In order to align the data to each other a chi-square test can be used. This type of statistical test results in a value that reflects the difference between an observed and expected value [37]. By minimizing this value it is possible to find the offset between two spectra that results in a small chi-square value, i.e. a small difference between the two spectra. The statistical formula for the chi-square test is presented in Equation 2.15.

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} \tag{2.15}$$

Filtering

If the data contains a lot of noise it should be filtered to remove unwanted high frequencies. This is usually done by using a low pass filter, which removes the frequencies above a set frequency. There are different filter types that can be used for this and those that can be used in the preprocessing in this master thesis are described below.

Moving average (MA) filters are very simple low-pass filters. The impulse response of an n-order moving average has n elements with the value 1/n. This way of implementing the moving average filter will average each datapoint with the same value. It can also be implemented by averaging n points of the original data, see Equation 2.16. Here y_i is the i^{th} point of the filtered data, x is the original data and n is the order. These type of filters have a fairly good performance but will affect the amplitude if the order of the filter is high [38].

$$y_t = \sum_{i=t}^{t+n} \frac{x_i}{n} \tag{2.16}$$

A Butterworth filter is designed to have a frequency response as flat as possible in the passband [39]. If the order of the filter is high, the band becomes more confined and steep, more like a regular square response. To design such a filter the order has to be defined, together with where the stopband will be (usually normalized from 0 to 1). The Butterworth filter is an infinite impulse response (IIR) filter. The benefit of using such a filter type is that it is considered to be fast. The transfer function for a Butterworth filter with a specific order and stopband has to be found. The stopband defines the length of the filter, everything outside of the stopband will be set to zero. To actually filter the signal, the transfer function has to be multiplied with the original signal. The transfer function for filters in general is given as $H(S) = \frac{B(S)}{A(S)}$ and the filtering process is given as $Y(S) = X(S)H(S)$ [39]. Since the filter will be zero above a given frequency, the multiplication with the data will also set the frequencies in the data that are above the given frequency to zero.

It is also possible to filter a signal using a window function, which is a mathematical function that sets all the values outside some chosen interval to be zero. There is a family of windows referred to as the Cosine-sum windows. Hanning, Hamming and Blackman windows can be found in this window family [40]. The Hanning and Hamming windows are very similar and the Hamming window can be thought of as the modified Hanning window. The Hamming window is achieved by setting α = 25/46 = 0.54 in Equation 2.17. The window is defined by Equation 2.18 where n $\in$ [-N/2,N/2] and N is a positive integer [41]. This window has the α coefficient which achieves minimum sidelobe levels. Since the sidelobe cancellation is better here than for the Hanning Window, the initial sidelobe lever is lower [41].

$$w(n) = \alpha + (1-\alpha)cos\left[\frac{2\pi}{N}n\right] \tag{2.17}$$

$$w(n) = \begin{cases} 0.54 + 0.46cos\left[\frac{2\pi}{N}n\right], & n = -\frac{N}{2}, .., -1, 0, 1, ..., \frac{N}{2} \\ 0.54 - 0.46cos\left[\frac{2\pi}{N}n\right], & n = 0, 1, 2, ..., N-1 \end{cases} \tag{2.18}$$

To make sure that the signal isn't shifted when it is filtered, the filtering function should be zero-phase. This type of filtering will apply the given filter to the signal twice, once forward and once backward. If this type of filtering is not used, the signal can be shifted to the right which will move the peaks away from their respective wavenumber, which is not ideal for spectroscopy data.

Another type of low-pass filter, called Savitzky-Golay is well-adapted for data smoothing. This type of filter can be derived directly from a particular formulation of the data smoothing problem. The filter was initially used to render visible and relative widths and height of spectral lines in spectrometric data that were noisy [42]. In order to understand the Savitzky-Golay filter it is a good place to start by first looking at the simplest averaging procedure, Equation 2.19 [42].

$$g_i = \sum_{n=-n_L}^{n_R} c_n f_{i+n} \tag{2.19}$$

Here there is a fixed window, n_L to n_R, which is used to compute a variable g_i as the average of the datapoints within this window, f_{i-n_L} to f_{i+n_R} with a constant $c_n = 1/(n_L + n_R + 1)$. This method does not preserve the area under the spectral line and for a narrow spectral line the height will be reduced and the width increased. In spectral data the height and width of the spectral lines will be of physical interest and therefore this method of filtering is not ideal for this type of data. The idea of the Savitzky-Golay filter is instead to find filter coefficients c_n that can preserve higher moments. This is done by approximating the window with a polynomial of a higher order instead of a constant. For each data point f_i a polynomial is least-squared fitted to all $(n_L + n_R + 1)$ points in the moving window. g_i will then be set to be the value of the polynomial at position i. Since least-squares fitting is used, the coefficients will be linear in the values of the data. By doing this the precision of the data can be increased without distorting the signal tendency [42].

Differentiation

After the data has been filtered it can also be differentiated in order to enhance the differences in the spectrum and to remove additive or multiplicative effects. The method has been used for a long time and the main purposes of differentiating a spectrum are spectral discrimination and spectral resolution enhancement. When the data is differentiated, small structural differences can be enhanced and made easier to detect. It is also possible to enhance the resolution using differentiation since the overlapping bands will be easier to distinguish from each other so that the number of bands and the coherent wavenumber can be determined [43, 44]. However, it also has to be considered that the differentiation of the data might lead to increased noise.

2.6 Biological Analytes Related to Diabetes

The biological analytes in the body can be divided into two groups in relation to measurements by a CGM device. The first group consists of substances in the body that are directly related to diabetes, such as e.g. glucose or ketones. The second group of analytes are those that are not necessarily directly related to diabetes but have the ability to interfere with mid-IR glucose measurements. Substances that might have the ability to interfere are those that absorb at the same wavelengths as glucose. Biological analytes that have previously been measured are e.g. lactate, triglycerides, cholesterol, and various proteins, by e.g. Brandstetter et. al. [27].

In this master thesis the analytes that will be measured are presented in Table 2.1. Pure glucose and BHA solutions were measured as well as a more complex solution where glucose, urea, lactate and albumin were dissolved in a phosphate-buffered saline (PBS) solution.

Table 2.1: The relevant concentration ranges of the analytes chosen to be measured in this master thesis in blood/serum (mg/dl) [45, 46].

Blood Parameters	Normal Range [mg/dl]
Glucose	40 - 400
Lactate	10 - 100
Urea	15 - 200
Albumin	2000 - 7000
BHA	0 - 300

Diabetics are not able to regulate their own blood glucose level (BGL) like other people. To prevent the glucose level from increasing to an abnormally high level, a diabetic has to take injections of insulin to lower the glucose level. The relevant concentration range of glucose in blood/serum is 40 - 400 mg/dl. Glucose has some distinct absorption bands at 1045 cm^{-1}, 1080 cm^{-1} and 1110 cm^{-1}, which all lie inside the wavelength range of the laser used in this thesis.

The relevant concentration range for lactate in humans is 10 - 100 mg/dl. The motivation behind measuring lactate is that this substance absorbs at some of the same wavelengths as glucose in mid-IR and due to this, might interfere with the glucose measurement. Lactate levels also increase during sepsis/septic shock, which can be fatal, and lactate can therefore also be used as an indicator for detecting this [47].

Albumin is the protein with highest concentration in the plasma of healthy humans, as much as 50% of the proteins in the plasma are albumin [48]. The relevant concentration range of albumin found in the blood of healthy individuals is 2000 - 7000 mg/dl. In laboratory experiments bovine serum albumin is usually used, which is a serum albumin protein obtained from cows. Albumin can be an important analyte in order to diagnose kidney and liver diseases, as well as being a good indicator of acute and chronic inflammation. Brandstetter et. al. found that the relevant spectral features for albumin can be found around 1148, 1165 and 1176 cm^{-1} [27].

The relevant concentration range of urea related to the amount normally found in an adult is 15 - 200 mg/dl. The amount of urea in blood says something about renal failure, but the main motivation behind including this component in measurements is to see if it interferes with the measurement of glucose. In the mid-IR region urea has shown to absorb in the region from 1400 - 1800 cm^{-1} [49]. This spectral region is not in the wavelength region for the laser used here so samples with urea and water will have to be measured using FTIR spectroscopy in order to see if the absorption area for urea is included in the wavelength range of the laser.

One of the more serious complications of diabetes is diabetic ketoacidosis. The condition occurs if the body isn't able to produce enough insulin so that the body starts to burn fatty acids and fat. The burning process produces ketones and if the level of ketones in the blood increases significantly the blood can become acidic. This will leads to diabetic ketoacidosis [8]. When fatty acids are burned acetoacetate, acetone and β-hydroxybutyric acid (BHA) are formed [50]. These have all been measured before but in recent years ketones have only been measured in urine to the authors knowledge and acetoacetate is the only ketone that can be detected clearly in the urine [51]. Therefore it will be interesting to investigate if one of these can be detected using mid-IR, e.g. BHA. The normal range of BHA is between 4.16 and 5.21 mg/dl and the cut-off value for diabetic ketoacidosis ranges from 13.5 - 36.5 mg/dl [46]. The relevant concentration range for BHA is set to be 0 - 300 mg/dl. According to an article by Gamze Hoşafçı et. al. [52] BHA was detected in urine samples in the spectral region between 1800 and 800 cm^{-1}. FTIR spectroscopy should be used to investigate if the exact absorption area for the analyte is covered by the wavelength range of the laser used in this thesis.

3

Method

3.1 Simulation of ATR Prisms

Zemax OpticStudio 18.1 was used as the simulation tool in order to simulate how the light from the laser would act within ATR prisms of different sizes. The non-sequential mode was utilized in the program since the reflections within the prism follow multiple optical paths. Non-sequential ray tracing assumes that there are no pre-defined paths for the rays and that any ray will reflect, refract, diffract, scatter, etc. when hitting an object in its path [53]. The different steps in how the simulation was done will be gone through below and all of the steps are also summarized in a flow diagram, see Figure 3.1.

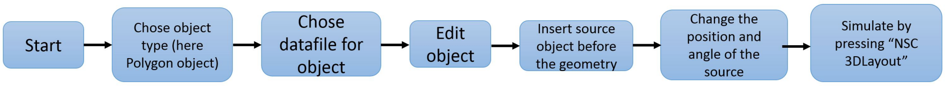


Figure 3.1: Flow diagram of the steps for how to simulate an ATR prism in Zemax OpticStudio.

The first step in simulating an ATR prism is to define which object the prism should be. This was done by modifying the example file for a dove prism which can be found in standard polygon objects in Zemax. The object is defined by an ASCII text file that defines the position of the corners and which corners that make up a rectangle. Each corner in the prism is defined by vectors, V_N = [x y z], N = 1,2,..,8 and the corners represented by the vectors V_1, V_2, V_3, etc. are shown in Figure 3.2. The geometry description for one of the ATR prisms can be seen in Figure 3.3. The vectors below **front face vertices** define the four corners to the left and the vectors below **back face vertices** define the four corners to the right. The x-component in the vector defines the width of the prism, the y-component defines the slanted sides and the z-component defines the top and bottom lengths. It is also defined in the object file that the top and bottom of the prism are reflective.

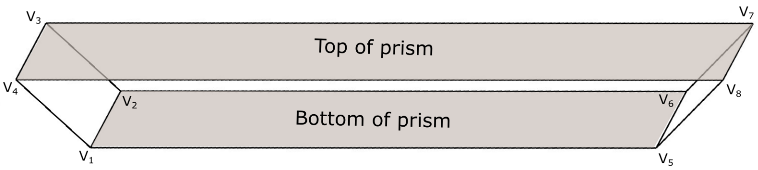
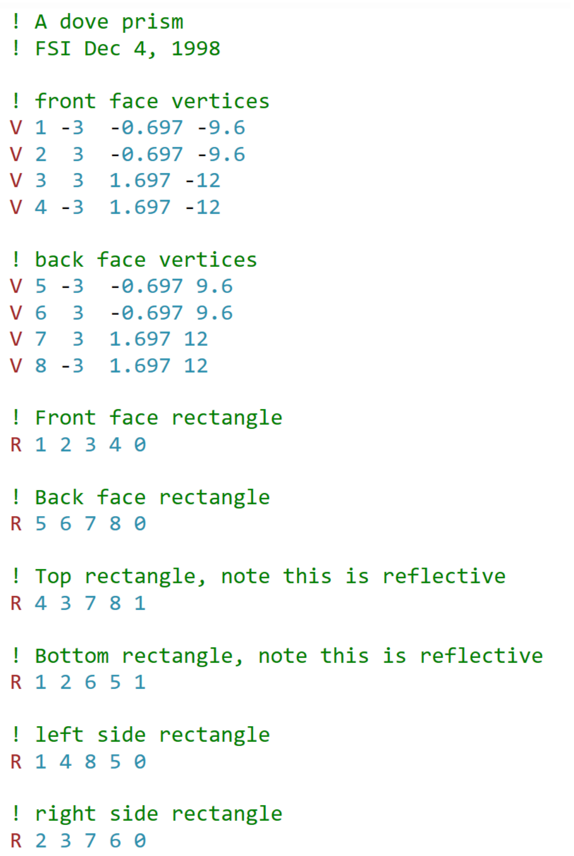


Figure 3.2: Geometry of a prism showing which vector represents the different corners in the prism. The vector coordinates are described in Figure 3.3.



```
! A dove prism
! FSI Dec 4, 1998

! front face vertices
V 1 -3  -0.697 -9.6
V 2  3  -0.697 -9.6
V 3  3  1.697 -12
V 4 -3  1.697 -12

! back face vertices
V 5 -3  -0.697 9.6
V 6  3  -0.697 9.6
V 7  3  1.697 12
V 8 -3  1.697 12

! Front face rectangle
R 1 2 3 4 0

! Back face rectangle
R 5 6 7 8 0

! Top rectangle, note this is reflective
R 4 3 7 8 1

! Bottom rectangle, note this is reflective
R 1 2 6 5 1

! left side rectangle
R 1 4 8 5 0

! right side rectangle
R 2 3 7 6 0
```

Figure 3.3: ASCII text file of the object to be simulated which is a dove prism in this thesis. Each vector represents one of the corners in the prism, see Figure 3.2. The top and bottom rectangles in the prism are defined to be reflective here.

The next step is to define a source of light that will propagate through the object. In the simulations done here a source ray is used in order to clearly see the amount of reflections. The source ray is placed on the left side of the prism, as close as possible to the slanted side by adjusting the z-position. It is then tilted in the x-direction with -45° so that the ray is sent into the prism with a 45° angle. To actually get a ray in the simulation it is important to set the number of layout rays to be higher than zero, here set to be 1. All of the simulations were performed with wavelength 9.347µm, which is in the middle of the operating wavelength range of the laser. This was chosen since the results of simulations using different wavelengths within the wavelength range of the laser were identical. Both the source ray and the object are defined in the non-sequential mode.

Five different geometrical designs were simulated using two different materials, ZnS and ZnSe. It is important to note that ZnSe is toxic and not ideal for any device that will interact with humans, such as a CGM device will. However, the different materials did not affect the simulations in any way and from the simulations there are no obvious benefits of using ZnSe over ZnS. The dimensions that were simulated can be found in Table 3.1 and a figure presenting the dimensions can be seen in Figure 3.4. By simulating the different dimensions together with a source ray with a 45° angle it is possible to investigate how many reflections each geometry will result in.

Table 3.1: Table of the five different geometries that have been simulated in this master thesis. Explanation of where the different lengths are placed on the prism can be seen in Figure 3.4.

Prism name	Length (top)	Length (bottom)	Height	Length X	Width
Prism1	24 mm	19.2 mm	2.40 mm	2.396 mm	6 mm
Prism2	14.4 mm	11.2 mm	2.406 mm	2.404 mm	4 mm
Prism3	14.4 mm	11.2 mm	1.608 mm	1.593 mm	4 mm
Prism4	22.24 mm	19.2 mm	1.523 mm	1.521 mm	6 mm
Prism5	42.2 mm	22.2 mm	10.016 mm	10.01 mm	10 mm

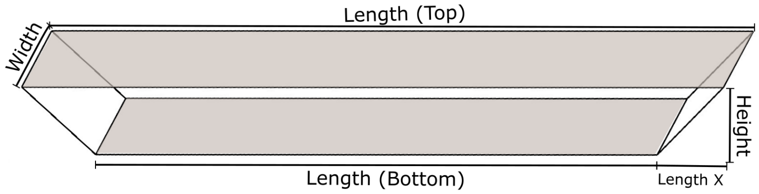


Figure 3.4: Sketch of a dove prism where the lengths from Table 3.1 are defined.

3.2 Fibre Polishing

Both of the setups used optical fibres to guide the light to and from the sample area and the purchased speciality fibres are received in one length. In order to get an interaction area between the light and the sample, the fibre has to be cut in two. The two fibre parts can then e.g. be placed in the transmission setup with a small gap between them, allowing the sample to fill the gap and interact with the light guided through the fibre.

The fibres are AgCl:AgBr polycrystalline fibre and the core of such a fibre is very soft and flexible. Due to this it is not possible to cut the fibre with a fibre cutting machine, which would result in a perfect cutting edge. Instead the fibre has to be cut by hand which results in the exposed end being rough and uneven. This will lead to a noisy and distorted signal in the setups which makes it difficult to predict the concentrations of the analytes. In order to reduce the noise of the signal as much as possible the two ends have to be even and smooth and to achieve this they have to be polished.

The polishing is done manually with the possibility of using five different polishing papers from ThorLabs (LF5P, LF3P, LF1P, LF03P, LFCF). The grit size of the sheets are; 5 µm grit using silicon carbide, 1 and 3 µm grits using aluminum oxide, 0.3 µm grit using calcined alumina and a finishing paper which has tiny diamond particles glued to a substrate of polyester with resin with grit size of 0.02 µm. The polishing papers that have grit size smaller than 1 µm were used together with drops of water. Each area of the sheet is only used once and to get a stable surface they were used together with a glass polishing plate and rubber mats from ThorLabs (NRS913A). Since the polishing papers are initially made for normal silica fibres, different combinations of the five papers were tested in order to find the combination that resulted in the least distorted signal for the fibres used here. All of the ten combinations that were tested can be seen in Table 3.2.

Table 3.2: Different polishing combinations using five different polishing papers from ThorLabs with different grit sizes.

Combination	5 µm	3 µm	1 µm	0.3 µm	Finishing paper (0.02 µm)
a)		X			
b)		X	X		
c)		X	X	X	
d)		X	X	X	X
e)			X	X	
f)	X	X	X		
g)	X	X	X		X
h)	X	X	X	X	
i)	X	X	X	X	X
j)	X	X			

To perform the polishing the fibre was placed in a polishing disk (D50-FC, ThorLabs) and then lightly pressed down onto the polishing paper in a figure-eight motion, see Figure 3.5. The figure-eight motion was repeated about 7 times for each paper. There is a source of error in the manual polishing as there might be an angle between the angle of the fibre end and the smoothness of the surface. Due to this each fibre had to be checked with a microscope (Olympus BX51M Metallurgical Microscope) after polishing with each paper to investigate if the edge was flat and perpendicular as required. Figure 3.6 shows a picture of the fibre right after cutting and a picture of the same fibre after polishing.

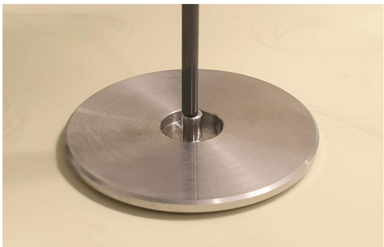


Figure 3.5: Picture showing the fibre placed in the polishing disk on one of the polishing papers.

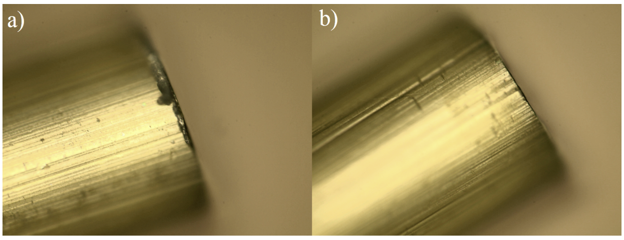


Figure 3.6: Picture obtained by microscope of the fibre end a) before polishing and b) after polishing. The fibre end was checked with the microscope after each polishing paper in order to see that the end was flat and perpendicular.

Before the fibre can be polished the two insulating layers have to be removed so that the fibre itself is exposed. This is done with a fibre stripper which is set to cut the two outer layers of the specific fibre type. When the inner layer is removed there is a possibility of the stripper touching the inner part of the fibre, which might lead to notches and scratches. If the notch is just in the cladding of the fibre and not the core, the signal will not be affected and distorted. However, if the core is damaged the signal might be noisy no matter how even and smooth the edge of the fibre is. This was thus also checked in the microscope after each time the fibre was stripped.

Another alternative to polishing the fibres by hand was to purchase speciality fibres that had already been cut in two and polished. Fibres like this was purchased for this thesis to investigate if the factory polished fibres would result in less noise than the hand polished fibres.

3.3 Experimental Setup

The QCL used in both setups is a Hedgehog-UT laser from Daylight Solutions, the data sheet is included in Appendix I. The wavelength range of the laser is 1200 - 905 cm^{-1} (8.33 - 11.04 µm) and the maximum average power is 22 mW. The solutions are only measured between 1200 - 925 cm^{-1} since the signal becomes very weak due to water absorption below 925 cm^{-1}. The QCL has triggering modes for both single wavelength emission and for scans. For pulsed operations the pulse length was 500 ns with a 5 % duty cycle and 100 kHz repetition rate. For each measurement the laser was scanned from 1200 - 925 cm^{-1} with the speed at 275 cm^{-1}/s, which results in each scan being 1 sec. The laser was thermoelectrically cooled and set to a temperature of 19°C for all measurements.

The photodetector used is from the PCI-4TE series by VIGO systems, the data sheet is included in Appendix J. This is an IR photoconductive MCT (mercury cadmium telluride) detector on a four-stage thermoelectrical cooler. The detector has high performance in the 2 - 14 µm spectral range, which covers the spectral range of the laser. The detectors in the series have different optimal wavelengths and the optimal wavelength for the one used here (PCI-4TE-10.6) is 10.6 µm. This is the detector in the series which provides high detectivity over the entire laser range. The detectivity at the peak wavelength for this detector is $\geq 8 \cdot 10^9$ $\mathrm{cm \cdot Hz^{1/2} \cdot w^{-1}}$.

The signal detected has to be converted into a digital signal and this can be done through an analog-to-digital converter. The converter used here is the M2P.5946-X4 card from Spectrum Instrumentation. It allows for recording up to eight channels with sampling rate up to 80 MS/s (mega samples per second). The card also has four digital inputs which can be connected with the trigger output from the laser. The recording program can thus wait for a trigger pulse sent out when a scan starts, before the recording is started. Software has to be used together with the card in order to control it, to start and stop the recording, as well as save the data. The software used here is a modified C++ driver initially developed by Spectrum Instrumentation for the specific card, Appendix B.

A digital microscope camera (RND 550-00025) from RND LAB was used in order to find the angles and distances in the two setups. The camera has 8 white LEDs, the adjustable magnification is 5x $\sim$ 200x and the photo resolution is 1.3 MPixels.

The fibres used in the setups are factory polished AgCl:AgBr polycrystaline fibres from Art Photonics, the data sheet is included in Appendix K. They are built up with a core, cladding and protective layer made of Polyether ether ketone (PEEK). The fibre that will further be referred to as the 'small-core fibre' has a core diameter of 410±15 µm and a cladding diameter of 500+0/-15 µm. The other fibre which will further be referred to as the 'large-core fibre' has a core diameter of 600±20 µm and a cladding diameter of 700+0/-15 µm.

3.3.1 Reflectance Setup

The most important component in the reflection setup utilizing an ATR prism is the prism itself. To get good measurements that are reproducible the prism needs to be stable and not move between measurements. The prism used in the ATR setup is made of ZnS (ZnS Large dove prism, *sin*optix, custom design) and the dimensions of the prism is presented in Figure 3.7.

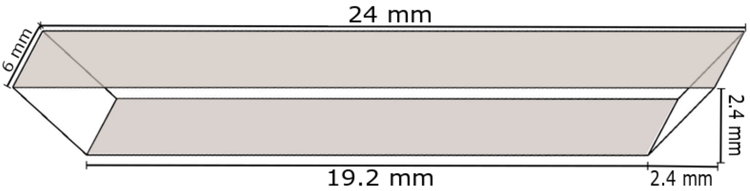


Figure 3.7: Dimensions of the ZnS prism used in the ATR spectroscopy setup.

In order to include the prism in the setup, a holder specially designed for the prism was fabricated in-house and tested. The holder was a plate made in aluminium with a hole in the middle which had almost the same dimensions as the prism, see Figure 3.8. Some measurements were done using this holder but it was not as stable as required. Therefore the prism was glued to a metal plate to ensure stability between the measurements. It was important that the prism was glued planar onto the metal plate, to have precise control over the interaction between the light and the prism. The result of gluing the prism onto the metal plate was a stable setup which made the measurements reproducible. A sketch of the ATR setup can be seen in Figure 3.9.

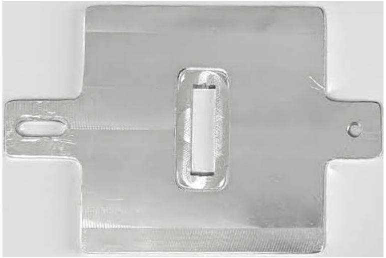


Figure 3.8: Metal holder which was designed for the specific prism. A few measurements were done using this but the result was that it was not stable and the measurements were not reproducible.

The light is transmitted into the prism through fibres. One fibre guides the light to the prism and one guides the light from the prism. The two fibres were placed in holders from ThorLabs (MSWCA/M) which have a rotating clamp that allows them to be directed towards the prism at an angle. This was ideal in order to place the fibres towards the prism at an approximately 45° angle. The first fibre which was the 'small-core fibre' was placed in contact with the prism and the angle was found to be 49.98°, using the microscope camera.

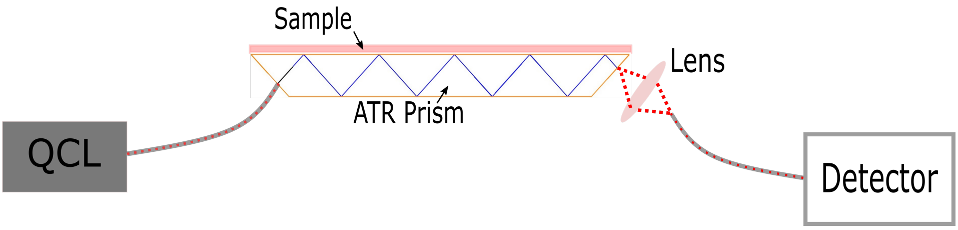


Figure 3.9: Sketch of ATR setup: The two fibres are bent so that the light is sent and received at approximately 45° angle to the prism. The sample is placed on top of the prism and the light leaving the prism is collected utilizing a lens.

The beam exiting the prism was investigated and found to diverge, which resulted in only a small portion of the signal being coupled into and transmitted through the fibre to the detector. This resulted in a low and distorted signal. To cope with this, a bi-convex ZnSe lens from ThorLabs (LB7639-E3, ThorLabs) was used to collect and focus the light into the fibre. The focal point of the lens was f = 20.0 mm and the lens offered high transmission from 0.6 - 16 µm. The lens was covered with a broadband anti-reflective coating optimized for 7 - 12 µm, which covers the spectral range of the laser. The distance between the lens and fibre was 22.37 mm, the angle of the lens was 30.27° and the angle of the fibre transmitting light to the detector was 50.8°. The output fibre was positioned in order to maximize the signal. Ideally the lens should be as close as possible to the prism to avoid the light diverging too much, but the metal plate which the prism was glued onto was partially blocking it. Pictures of the different dimensions and angles from the setup can be found in Appendix C. The 'large-core fibre' was used between the lens and the detector in order to collect more of the light to the detector.

Each sample was placed on top of the prism covering approximately the entire area each time. Paper was used to absorb the sample and the prism was then cleaned with water and a dry paper to make sure that all of the sample had been removed. Reference measurements with demineralized water were acquired before each sample measurement. The water was also placed on top of the prism, covering approximately the same area as the samples.

3.3.2 Transmission Setup

A sketch of the transmission setup can be seen in Figure 3.10. The QCL is coupled into the 'small-core fibre' which transmits the light to the sample which is placed between a small gap between the two fibres. The 'large-core fibre' was used to transmit the light from the sample to the detector and this was used to collect as much of the signal as possible in this setup. It is desirable with a long pathlength so that the light is transmitted through more of the sample, but without the signal becoming too weak. In this setup the pathlength was found to be 164 µm by using the microscope camera. The setup and pathlength were aligned so that the signal measured through water was 2V in order to get a signal at the longer wavelengths were the water absorbes more of the light.

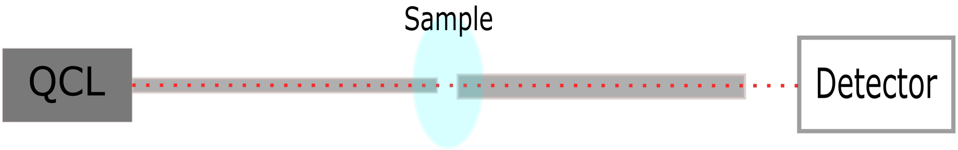


Figure 3.10: Sketch of transmission setup: Two fibres are placed close to each other with a small gap between them where the sample will be placed in order to interact with the light.

A reference measurement of water was performed before each sample measurement, in order to calculate the absorbance spectra through signal processing. Even though the setup was stable and the water spectra had the same shape between each sample measurement, the signal was found to drift so a water measurement for each sample measurement was necessary to achieve good results. Between each measurement the sample or water was removed using paper to absorb the water and a thin piece of paper was placed between the fibres to ensure that all of the sample was removed in the pathlength. The factory polished fibres were compact and it was therefore easy to remove the sample without displacing the fibres. Between some of the measurements the signal was checked in order to see that all of the sample had been removed.

3.4 Signal Processing

The raw data from each measurement has to be processed before it is possible to extract quantitative information from the spectra. One measurement is defined as an average of 10 scans of the lasing wavenumbers, from 1200 cm^{-1} to 925 cm^{-1}. The data was processed in two Python scripts which will be explained below. The measurement is recorded and saved into separate binary files for each scan. The files contain pulses for the wavenumber between 1200 cm^{-1} and 925 cm^{-1}. In the first script the pulses are located, averaged, aligned and signal averaged, and the Python script for this is included in Appendix D. All of the pulses in each scan are located by using the function **getpulseIntensity**. Here the periodicity in the data is found by finding the fast Fourier transform (FFT) frequency and dividing the FFT data by this frequency. After the periodicity is determined, the first pulse can be located and then the rest of the pulses can be localized by knowing the size/time of each pulse and the periodicity. This function is used on each scan to get all of the pulses in one measurement and then they are all plotted together, see Figure 3.11a). This plot shows how the pulses change in intensity with wavenumber. Each scan is then averaged over 255 pulses into a vector where each data point will correspond to the intensity below the pulse. A function called **CheckLength** is used on each scan to check the length of the averaged pulses. If the length is not equal to the most common length, here 100004, an appropriate amount of zeroes are added to the end. Figure 3.11b) shows the spectrum obtained from only one scan averaged over 255 pulses.

In some cases the spectra obtained from each of the 10 scans have an offset compared to the other scans due to mode-hops. It is not the case for every scan but since it occurs in some cases they are all aligned. The alignment is done by using a function called **chisqr_align** which minimizes the chi-square value. The Chi-square test is used in order to find the difference between an observed value and an expected value in statistics, the theory behind this was explained in Section 2.5. It is used here in order to find the difference between a target signal (observed) and a reference signal (expected). The Python code for this function is included in Appendix E. The function inputs a reference array that won't be shifted, a target array that will be shifted, the region of interest, an initial guess of the offset between the two data sets and a bound that constrains the shift search around the initial guess. The first scan is used as the reference signal for all of the other scans to find the difference between these and the first scan. The alignment is performed with a segment width of 580 datapoints. Each spectrum is then shifted with the offset between the target and reference signal so all of the spectra starts at the same wavelength. After alignment the spectra can be signal averaged by adding all 10 scans together and dividing by the number of scans, see Figure 3.11c). The result of this Python script is one spectrum for each measurement that is aligned and signal-averaged over the 10 scans obtained for the measurement.

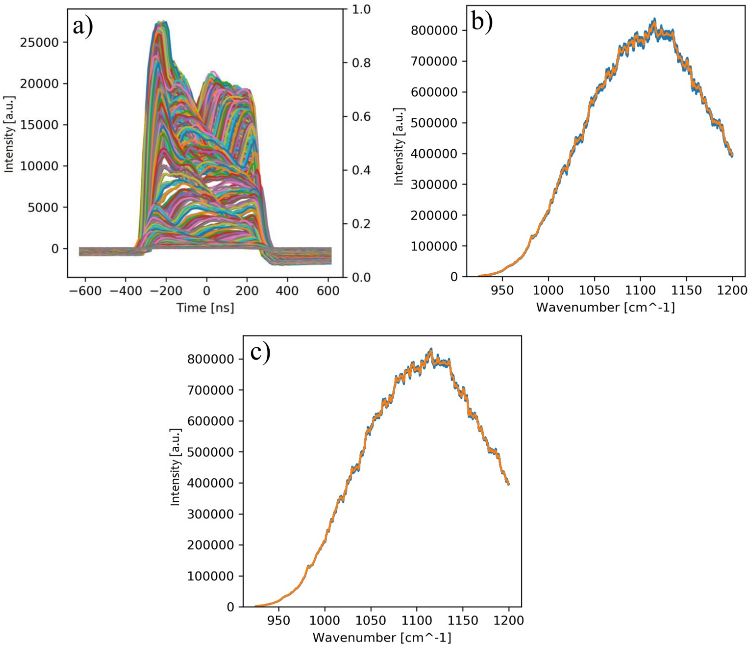


Figure 3.11: Plots of measurements after processing in python script. **a)** The intensity of every 500 pulse for 10 scans plotted together. **b)** Integrated pulses for only one scan where the blue plot is for one pulse, and the orange is the average over 255 pulses. **c)** Signal where the scans are aligned to the first scan and signal-averaged over the 10 scans.

The final absorbance spectrum for each measurement is obtained using Equation 3.1, where I is the spectrum of the sample and I_0 is the reference spectrum which was a water spectrum here.

$$A = -log\frac{I}{I_0} \tag{3.1}$$

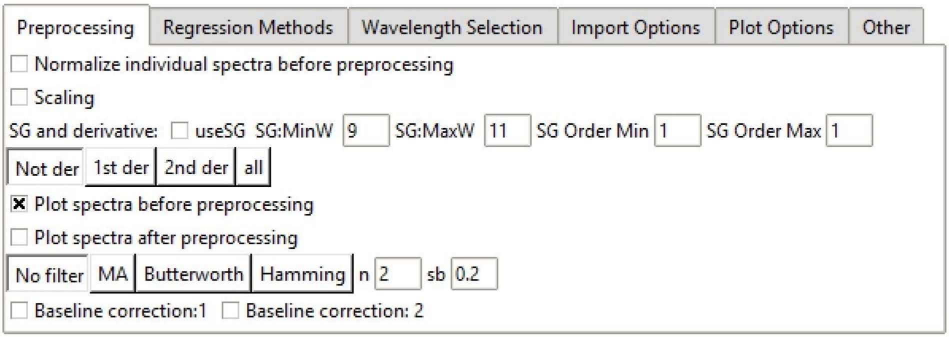
The Python plot that calculates the final absorbance spectra can be found in Appendix F. The script inputs all of the spectra for one run and calculates the final absorbance spectrum for each. The absorbance spectra are plotted in the script together in order to define outliers. The script iterates through the files, opens them and extracts the spectra. The absorbance spectra are then calculated using Equation 3.1 and the new spectra are saved into a given output folder. This method of making the absorbance spectra is used in order to be consistent with other literature and thus making it easier to compare the results obtained here with others.

3.5 Data Analysis

The program used in this master thesis for performing the multivariate analysis on the measurements, was developed by co-supervisor Ine Larsen Jernelv. The program is written in Python and is based on the Python libraries; SciPy, for scientific computing and technical computing, and Scikit-learn, which contains various classification, regression and clustering algorithms. The program reads in files containing columns with the paths to the measurement files together with the concentrations of the analytes in the measured sample. The analysis program has the ability of performing multivariate analysis in various ways but most importantly for this master thesis is PCR and PLSR. The program can perform cross-validation which is used in PLSR to estimate how well the model fits within one measurement.

The program was modified in this thesis work to include different filters in order to improve the analysis. The filters that were included were a low pass butterworth filter, MA filter, and a filter using a Hamming window; these are all explained in Section 2.5. It is important to not exaggerate the filtering since this will smooth away the important features of the spectrum. It is desirable to filter the spectra if contains fringes and other noise which may be a result of different reflections within the setup or a low signal. Another way of reducing the fringes is by averaging the 10 scans from one measurement and this was mentioned in Section 3.4.

All of the filter types are based on calculating the numerator (**b**) and in some cases the denominator (**a**) polynomials of an IIR or FIR filter. The two polynomials make up the transfer function of the filter and are used to define how the data (**inp**) should be filtered using the filtering function **filtfilt(b,a,inp)**. This is a zero-phase IIR or FIR (based on the input) filtering function which avoids that the filtered data is shifted. This is one of the many signal processing functions that can be found in the SciPy library, **scipy.signal**. The filters are added to the program as radio buttons, which initially are set to not filter the data, see Figure 3.12. If the data should be filtered the button for the desired filter has to be pressed. Together with this a variable **n** can be changed directly in the GUI, which changes how strong the filtering is. This is included in order to see how strong the filtering can be without removing the analyte identifying peaks and to see if strong or weak filtering results in lower a RMSEP/RMSECV value. The code for the filters that inputs the **n** value from the GUI and the data can be seen in Listing 3.1 below.



☐ Normalize individual spectra before preprocessing
☐ Scaling
SG and derivative: ☐ useSG SG:MinW 9 SG:MaxW 11 SG Order Min 1 SG Order Max 1
☒ Plot spectra before preprocessing
☐ Plot spectra after preprocessing
n 2 sb 0.2
☐ Baseline correction:1 ☐ Baseline correction: 2

Figure 3.12: GUI from PLSR program written in Python by co-supervisor Ine Larsen Jernelv. The modified part where the filters have been added, is found under the **Preprocessing** tab, at the bottom.

There is no built-in implementation of a MA filter in the SciPy library such as for many other filters. However, it is easy to implement. The numerator of the transfer function can be given as **b** = [1.0 / **n**] * **n**]. If **n** = 5, **b** will be a vector with length 5 where all the values are [1.0 / 5] = 0.2 (**b** = [0.2, 0.2, 0.2, 0.2, 0.2]). The second value **a** tells the function if both of the numbers should be normalized by the given value of **a**. If **a** equals 1, none of the numbers are normalized. Here it is set to be 1 in order not to normalize the data. The two polynomials of the transfer function are thus set and can be used together with the **filtfilt** function for filtering.

Low-pass Butterworth filter is able to remove fringes resulting from noise without distorting the signal shape. The Butterworth function was imported from the Scipy library and it inputs two values, **n** = the order of the filter and **sb** = the stopband (which is normalized between 0 and 1). A low n will filter the data a lot and a high n will filter the data less. The stop band can also be changed directly in the GUI, but in general **sb** = 0.2 was found to give the best result. The function returns both of the polynomials of the transfer function which is used for filtering.

scipy.signal.firwin is a function which designs a FIR filter using the window method. There are different window types for this but here the Hamming window has been used. The function inputs the length of the filter (**n**), the cutoff frequency (sb) and the window type and returns the numerator polynomial of the FIR filter. The denominator polynomial of the transfer function is set to be 1. The **filtfilt** function can then also here be used for filtering the data.

```
def MA(n,inp):
    b = [1.0/n]*n    #n = the order of the MA filter, the larger n is, the smoother curve will be
    a = 1.0
    return scipy.signal.filtfilt(b,a,inp)

def butterworth(n, sb, inp):
    [b, a] = signal.butter(n, sb)    #n = the order of the   Butterworth filter, the smaller n is, the smoother curve will be
    return scipy.signal.filtfilt(b,a,inp)

def Hamming(n, sb, inp):
    b = signal.firwin(n, sb, window = ”hamming”)    #n = the length of the filter, the larger n is, the smoother curve will be
    a = 1.0
    return scipy.signal.filtfilt(b,a,inp)
```

Listing 3.1: Filters implemented in the data anlysis program. Each filter inputs n, which can be changed in the GUI. The butterworth filter also takes in the stopband which can also be changed directly in the GUI.

The other preprocessing methods explained in section 2.5 can also be seen in the GUI shown in Figure 3.12. To use the Savitzky–Golay filter the **SG and derivative** box has to be checked. The minimum and maximum window sizes, which define the length of the filter window, have to be defined together with the minimum and maximum orders, which refer to the order of the polynomial used to fit the samples. The savgol_filter is imported from the SciPy library and it inputs the data to be filtered, the length of the filter window, the order of the polynomial used to fit the samples and the order of the derivative to compute. The default of the differentiation is set to zero so the filter can be applied without differentiation if wanted and the function returns the filtered data.

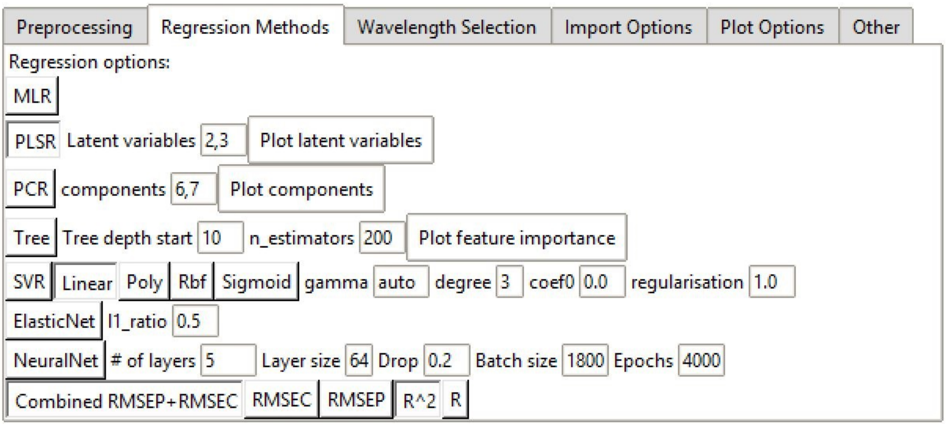
Differentiation can also be performed without the Savitzky–Golay filter and this is done in the program by using finite differences [54]. The function for differentiation inputs the wavenumbers and the data. The code for the function can be seen in Listing 3.2. A for-loop is then used to iterated from 0 to **n** through all of the wavenumbers and data, where **n** is the length of the wavenumber array. For each i the wavenumber at that position is added together with the next and divided on 2, $\frac{1}{2}(x[i] + x[i+1])$. For each i the derivative of the data is found by dividing the finite difference between two data points on the finite difference between two wavenumbers, $\frac{y[i+1]-y[i]}{x[i+1]-x[i]}$.

```
def Der(x,y):
  ”””Function for finding first derivative of spectral data. Uses finite differences.”””
  n=len(x)
  x2=np.zeros(n-1)
  y2=np.zeros(n-1)
  for i in range(n-1):
    x2[i]=0.5*(x[i]+x[i+1])
    y2[i]=(y[i+1]-y[i])/(x[i+1]-x[i])
  return(x2,y2)
```

Listing 3.2: lfilter in data analysis program

Baseline correction might also be used in the program before PLSR is performed on the data. Two different methods of doing this were tested in this master thesis. Baseline correction: 1 in the GUI was implemented so that it subtracts a linearly increasing baseline between the first and the last independent variable. Baseline Correction:2 in the GUI is very simple; the first independent variable of each spectrum is set to be zero and the difference between zero and the original value is then subtracted at each wavenumber in the spectra.

The GUI containing the parts of the program related to the regression methods can be seen in Figure 3.13. Here one can choose between different multivariate analysis methods but only PLSR is considered in this master thesis. The latent variables are set directly here. For example 2,3 will plot the concentrations plot with 2 and 3 latent variables and 3:6 will plot the concentration plot with 3, 4, 5 and 6 latent variables. At the bottom one can choose which value to be calculated; combined RMSEP+RMSEC, RMSEC or RMSEP.



Latent variables 2,3

components 6,7

Tree depth start 10 n_estimators 200

gamma auto degree 3 coef0 0.0 regularisation 1.0

l1_ratio 0.5

# of layers 5 Layer size 64 Drop 0.2 Batch size 1800 Epochs 4000

Figure 3.13: GUI from PLSR program written in Python by my co-supervisor. This part of the GUI shows everything related to the Regression method.

The output of the program is a concentration plot showing training and validation data together with the calculated RMSEP, RMSEC or RMSEP+RMSEC value, see Figure 3.14. The value associated with cross-validation is root-mean-square error of cross-validation (RMSECV). During cross-validation the sample size used as the validation data is defined together with how many iterations that should be performed i.e. how many times the cross-validation should be performed. The function for cross-validation in the program here shuffles the data for each iteration before defining which data should be included in the validation set. Due to this some of the data might be used as validation more than once. Each time the cross-validation is performed different measurements are used as the validation data and the concentration plot only presents the last cross-validation. However, the RMSECV value presented in the concentration plot is the root-mean-square value of the RMSECV values from each iteration. The concentration plot is shown in Figure 3.14 together with some explanation. In the figure the measurements predicted by the model are plotted as red stars and the measurements used for training the model are plotted as blue dots. The number of latent variables, the multivariate analysis model and information about how the spectra have been preprocessed are included in the concentration plot seen in the figure. Another important value presented in the concentration plot from the program is R^2.

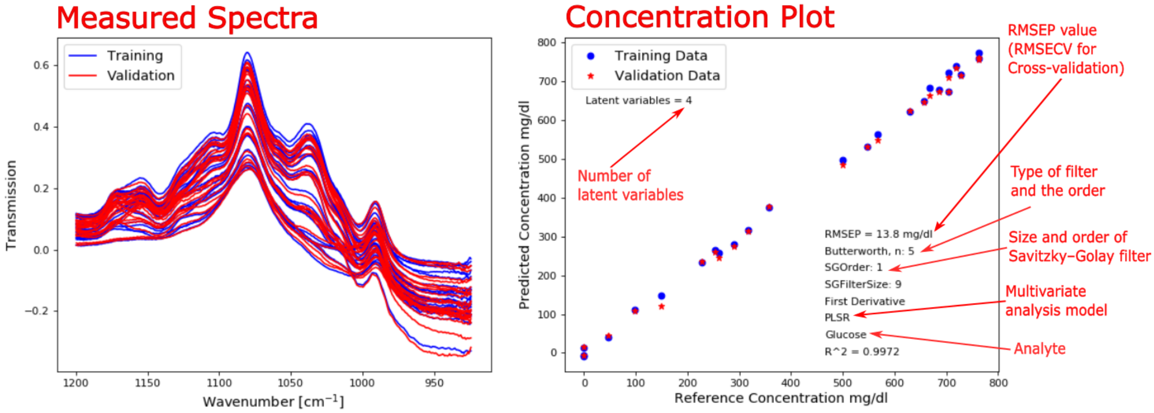


Figure 3.14: Example of a measurement, obtained through the regression analysis program, showing how the output will be together with an explanation.

3.6 Sample Test Matrix

Three sample sets were made and measured in this master thesis with the following analytes; glucose (G8270-100G, Sigma Aldrich), urea (U5378-100G, Sigma Aldrich), lactate (71718-10G, Sigma Aldrich), albumin (421501J, VWR Chemicals), BHA (H30426, Alfa Aesar) and PBS (404-200TABS, VWR Chemicals). The water used to dissolve the analytes is demineralized with a water purifying system (Millipore, Water purification systems, RiOs 50). 10 samples were prepared containing only glucose (50 - 1250 mg/dl) and water and 10 samples were prepared containing only BHA (0-338 mg/dl) and water, see Table 3.3 for the concentrations. 25 samples were prepared where the analytes were dissolved in phosphate-buffered saline (PBS). These samples contained glucose (0-800 mg/dl), lactate (0-90 mg/dl), urea (0-200 mg/dl) and albumin (0-6000 mg/dl), see Table 3.4 for the concentrations.

Before urea was added to the complex solution a pure urea solution was measured with FTIR spectroscopy in order to confirm that the absorption area for the analyte was included in the wavelength range of the laser. The pure BHA solution was also measured with FTIR spectroscopy in order to find the exact absorption area for the analyte before it was measured with the fibre-coupled transmission setup and the ATR setup.

Table 3.3: Concentrations (mg/dl) in the 10 samples only containing glucose and water and the concentrations in the 10 samples only containing BHA.

	—— Glucose —— mg/dl	—— BHA —— mg/dl
1	351.75	118.58
2	434.86	76.56
3	1250.60	91.48
4	212.47	9.30
5	501.98	295.86
6	809.19	39.67
7	676.14	184.90
8	288.78	172.80
9	50.09	263.68
10	138.34	231.18

For the complex solution set the combination of the different concentrations for the 4 analytes were found using Design-Expert 11. This is a statistical software from Stat-Ease Inc. which can design experiments. The number of mixture components were set to be 4 in the program and the minimum and maximum concentrations were defined; Glucose (0-800 mg/dl), Lactate (0-90 mg/dl), Urea (0-200 mg/dl) and Albumin (0-6000). A-optimal design was used, which minimizes the average variance of the polynomial coefficients. The program was then run until the concentrations of especially glucose was well spread over the entire glucose range, i.e. when no more than 3 samples were the minimum or maximum concentrations. The urea and lactate concentrations were concentrated around the center of the concentration range. When optimal design is used the sample sets will be robust without including too many samples. If this was not used a large amount of samples would have to be used in order to cover all of the concentration combinations for the analytes.

Table 3.4: Concentrations (mg/dl) in the 25 samples containing glucose, lactate, urea and albumin, solved in PBS.

	—— Glucose —— mg/dl	—— Urea —— mg/dl	—— Lactate —— mg/dl	—— Albumin —— mg/dl
1	704.90	0	88.81	6501.02
2	149.07	92.51	0	4318.75
3	666.80	0	44.70	1516.90
4	657.27	197.66	0	2640,08
5	762.06	0	46.88	3876.19
6	290.06	197.67	0	4517.14
7	317.68	128.88	45,68	3845.67
8	0	32.57	51.41	120.25
9	0	0	0	1170.48
10	719.19	4.02	0	3250.51
11	628.69	0	0	5081.78
12	253.38	58.82	88.81	2518.00
13	98.59	0	45.68	4135.62
14	228.14	197.67	88.81	3998.28
15	547.72	78.91	44.99	4547.66
16	704.90	0	88.81	4974.96
17	576.31	197.67	88.81	468.50
18	500.10	197.67	42.83	2029.66
19	728.72	109.11	41.25	4501.88
20	47.63	0	88.81	352.52
21	356.74	197.67	37.89	4578.18
22	762.06	156.55	88.81	3662.55
23	685.85	54.39	0	1141.49
24	259.58	0	0	2197.53
25	0	0	88.81	3204.73

4 | Results

4.1 ATR Spectroscopy

4.1.1 Penetration Depth

The penetration depth of the radiation into the sample is calculated using Equation 2.8 presented in Section 2.2. The angle of incident radiation is approximately 50°, the refractive index of the prism is n_2 = 2.2 [55] and the refractive index of the sample is set to be the same as for 10 $\%$ glucose solution in water n_1 = 1.34 [56]. This resulted in a minimum penetration depth $d_p(1200\text{cm}^{-1})$ = 3.950 µm and maximum penetration depth $d_p(905\text{cm}^{-1})$ = 5.403 µm .

4.1.2 Critical Angle

Based on the theory in Section 2.2, the critical angle is defined by the refractive index of the prism and the solution placed on top of the prism, see Equation 2.3 in Section 2.2. The refractive indices are as follows; n = 2.2 for a prism made of ZnS [55], n = 1.33 for water and n = 1.34 for 10$\%$ glucose solution in water [56]. The largest amount of glucose solution in the samples that was measured is 14 $\%$. By using these indices the critical angle is found to be θ_c = 37.2° when water is placed on top of the prism and θ_c = 37.52° if it is a 10 $\%$ glucose solution. However, some of the samples that will be measured have other analytes in them as well, such as albumin which has a large refractive index n = 1.9. The concentration of albumin in the samples are fairly high and the critical angle might be increased when these samples are measured. In order to ensure that TIR occurs for each measurement for all of the solutions, the critical angle is set to be a little higher θ_c = 39°.

4.1.3 Maximum Angle of the Incident Light

The angle of the incident light $\theta_{i,p}$ within the prism is dependent on the refracted angle in the prism, the angle of the light sent into the prism $\theta_{i,f}$ and the refractive indices of the material and the medium the light is transmitted through. It can be calculated using the theory and equations presented in Section 2.2. If there is a gap between the fibre and prism the medium will be air and when the angle of the incident light into the prism is $\theta_{i,f}$ = 50°, the refracted angle will be $\theta_{i,p}$ = 42.73°. If the entire beam is sent directly from the fibre, which has a refractive index n = 2.15, into the prism the incident angle has to be 45° and the refracted angle will also be $\theta_{i,p}$ = 45° which is above the critical angle. Different incident angles $\theta_{i,f}$ were tested and the results were as follows; $\theta_{i,p}(\theta_{i,f}\text{=}55^\circ)$ = 40.47°, $\theta_{i,p}(\theta_{i,f}\text{=}57^\circ)$ = 39.57°, $\theta_{i,p}(\theta_{i,f}\text{=}60^\circ)$ = 38.24°. Since the critical angle is 39° the maximum angle of the incident light into the prism $\theta_{i,f}$ is 57°.

In the setup the fibre is placed in contact with the prism and the entire light beam will not be sent through air. Since the angle of the fibre is 50° and the angle of the prism is 45° the fibre end will not be parallel to the slanted side of the prism and some of the fibre will not be in direct contact with the prism. It is difficult to say how much of the light is transmitted through air into the prism and how much is transmitted directly from the fibre into the prism. The incident angle within the prism $\theta_{i,p}$ was therefore calculated for different amounts of light being transmitted directly from the fibre into the prism. The approximated refractive index $n_1(a,b) = n_{fibre} \cdot a\% + n_{air} \cdot b\%$ was used for this, where a is the amount of the light transmitted directly from the fibre into the prism and b is the amount of the light transmitted through air. Different combinations were investigated, $n_1(50\%, 50\%)$ = 1.575, $n_1(70\%, 30\%)$ = 1.805 and $n_1(90\%, 10\%)$ = 2.035, in order to confirm that TIR still occur when $\theta_{i,f}$ = 50° no matter how much of the light is sent through air or fibre. $\theta_{i,p}(\theta_{i,f}{=}50°, n_1{=}1.575) = 41.42°$, $\theta_{i,p}(\theta_{i,f}{=}50°, n_1{=}1.805) = 40.90°$ and $\theta_{i,p}(\theta_{i,f}{=}50°, n_1{=}2.035) = 40.37°$. One can see that when a larger part of the light is transmitted directly from the fibre to the prism the incident angle within the prism $\theta_{i,p}$ is reduced and close to the critical angle. When this is calculated for larger angles the refracted angle becomes smaller than the critical angle when 50 $\%$ of the light is transmitted directly from the fibre into the prism. Based on these results the angle of the light sent into the prism $\theta_{i,f}$ should be no more than 50° when in contact with the prism to ensure TIR for all samples.

4.1.4 Fabry Pérot Interference

The angle of the lens used between the prism and the second fibre was initially 45° in the setup, but this lead to a distorted and noisy spectrum which was difficult to reproduce. A spectrum from a measurement when the lens was 45° can be seen in Figure 4.1 and it is seen that the fringes are periodic. The periodicity of the fringes is approximately 5.95 wavenumbers and since one scan is 1 s the distance between the fringes is approximately 21.62 mm. This is almost the same as the distance between the lens and the prism which indicates that they are a result of Fabry Pérot interference which occur due to multiple reflections between the lens and the prism.

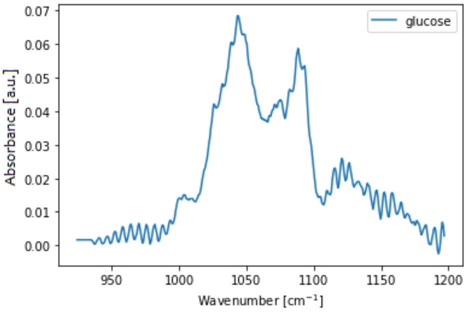


Figure 4.1: Measurement of a pure glucose solution with the ATR setup when the lens had a 45° angle and noise was induced most likely due to the FP interference.

To avoid FP interference the angle of the lens was reduced so that reflected light would not be reflected back to the prism but rather reflected out into the air, see Figure 4.2 and Figure 4.3. The angle of the lens that gave the best result, meaning the least distorted signal, was 30.27°. A spectrum from a measurement when the lens was 30.27° can be seen in Figure 4.4. It is clear that the periodic fringes have been reduced by changing the angle of the lens, but there is still some noise in the spectrum and in order to cope with this filtering was applied in the preprocessing. The shape of the spectrum in Figure 4.4 also resembles spectra that can be found in literature more than the spectrum in Figure 4.1 [34, 35].

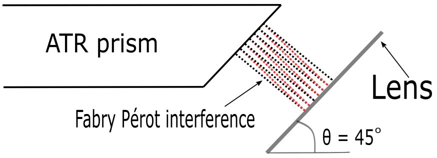


Figure 4.2: If the angle of the lens is 45° the lens and prism is parallel and FP interference can be induced between the lens and the prism. This interference might be the main reason for the noise that was seen when the angle of the lens was 45°.

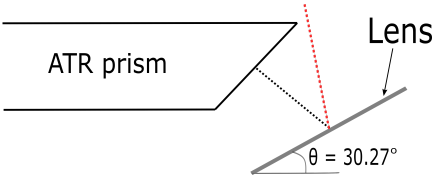


Figure 4.3: If the lens is not parallel to the prism all of the light will not be reflected between the lens and the prism. The light will then rather be sent out into the air.

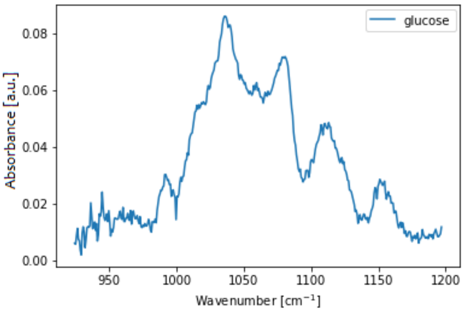


Figure 4.4: Measurement of a pure glucose solution with the ATR setup when the lens had a 30.27° angle, it can be seen that the noise due to FP interference is reduced and the shape resembles what can be found in the literature.

4.1.5 Pathlength and Absorption

The path length for the light within the prism can be calculated by using the definition of cosine. The incident angle within the prism was found to be 47.29 above and the height of the prism is 2.4 mm. The pathlength X will then be given by Equation 4.1 and is found to be 3.27 mm between the top and the bottom of the prism. The entire pathlength through the prism is approximately 29.403 mm for 9 reflections.

$$X = \frac{2.4}{\cos(47.29)} = 3.267 \tag{4.1}$$

Beer-Lambert's law can then be used to see how much of the light will be absorbed by the prism. The absorption coefficient for ZnS at 10 m wavelength is 0.096 cm^{-1} [57]. An approximate relation between the absorption coefficient (α) and the transmission (T) is given in Equation 4.2.

$$T = 10^{-\alpha l} \tag{4.2}$$

The equation originally uses the attenuated coefficient which is the sum of the absorption coefficient and the scattering [18]. However, since the prism is of high quality the scattering is assumed to be less than the absorption. The absorption can then be approximated by assuming that all of the light is either transmitted or absorbed, A$\%$ = 1- T$\%$. By using the pathlength found above the amount of light absorbed in the prism will be 47.8 $\%$. A thinner prism with height 1.5 mm that achieves 9 reflections will have a pathlength of 19.9 mm, which would reduce the absorption in the prism to 35.6 $\%$. This is just an approximation but it confirms that a thinner prism with shorter pathlength will reduce the absorbance.

4.2 Characterization of Analyte Absorbance

By measuring samples with urea and water using FTIR spectroscopy (Alpha II, Bruker) urea was found to absorb in the region from 1125-1175 cm^{-1}, see Figure 4.5 for Python plot of the absorbance spectrum. This spectral region is covered by the wavelength region of the laser and the analyte was therefore included in the complex solutions.

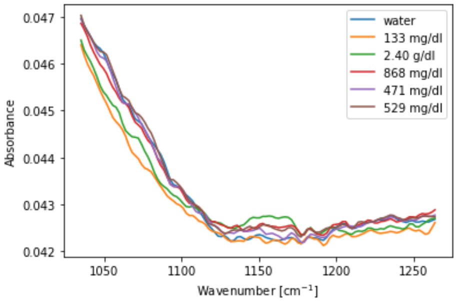


Figure 4.5: Measurements of different concentrations of urea in water using a FTIR spectrometer. This shows that urea have an absorption band in the region from approximately 1125-1175 cm^{-1}.

It was discovered that BHA have absorption bands between 1050 - 1090 cm^{-1} and 1110 - 1150 cm^{-1}. A figure of the absorption spectrum obtained from the FTIR spectrometer (Alpha II, Bruker) plotted in Python can be seen in Figure 4.16. Since the analyte was found to absorb in a spectral region covered by the laser the pure BHA solution was also measured with the fibre-coupled transmission setup and the ATR setup.

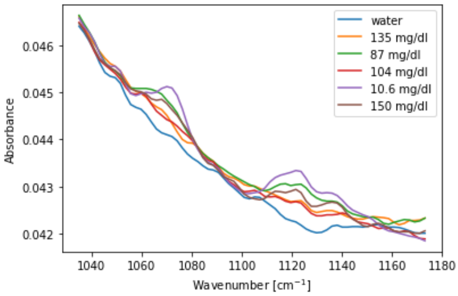


Figure 4.6: Measurements of different concentrations of BHA in water using a FTIR spectrometer. BHA has absorption bands in the region from approximately 1050-1090 cm^{-1} and 1110 - 1150 cm^{-1}.

4.3 Simulation of ATR Prisms

4.3.1 Reflections

Many reflections will result in a large interaction area with the analyte placed on top of the prism. From the simulations of the different dimensions presented in Table 3.1 in Section 3.1 together with a 45° source ray, the number of reflections was found, see Table 4.1. Here it is clear that the dimension that result in 13 reflections will result in the larges interaction area. However, a prism with these dimensions was not available in this master thesis. The three first dimensions listed in Table 3.1, Prism1, Prism2 and Prism3, are the ones that could be used. Of these, the largest prism result in 9 reflections and was used in the setup. The simulation of this prism with the angle of incident light used in the setup being θ_i = 50° can be seen Figure 4.7. The other simulations that were used in order to find the number of reflections for each prism dimension can be found in Appendix A.

Table 4.1: Simulation of the number reflections achieved with the five different geometries given in Table 3.1. The angle of the source ray in the simulations was 45° in order to decide which geometry would result in the highest numbers of reflections.

Prism name	**Number of reflections**	Associated Figure
Prism1	9	Figure 4.7
Prism2	5	Figure 6.1 (Appendix A.1)
Prism3	7	Figure 6.2 (Appendix A.2)
Prism4	13	Figure 6.3 (Appendix A.3)
Prism5	3	Figure 6.4 (Appendix A.4)

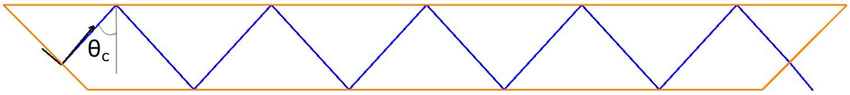


Figure 4.7: Simulation of reflected rays in the prism with dimensions of Prism1, found in Table 3.1. The simulation shows that a prism with these dimensions will achieve 9 reflections within the prism. The angle of the incident light within the prism is $\theta_{i,p}$ = 42.73° i.e. $\theta_i > \theta_c$ and TIR occurs. The black arrow indicates the angle of the light sent into the prism.

The critical angle is set to be θ_c = 39° based on the reflective indices of the samples and the prism. This results in that the angle of the incident light into the prism can be maximum 50°. The calculations of the critical angle and the maximum incident angle was presented in the previous section. $\theta_{i,f}$ larger than 50° will result in more reflections in the simulation, but TIR will not occur since the angle of the incident light within the prism $\theta_{i,p} < \theta_c$, the light will rather be refracted out of the prism. Therefore it is not possible to decide the angle based on the simulations alone and the calculated maximum incident angle has to be considered. See Figure 4.8 for definitions of the different angles. If the angle of the incident light $\theta_{i,f}$ is smaller than 36°, the number of reflections will be reduced. An example of this can be seen in Figure 4.9, here the angle of the incident light into the prism is 25° and only 7 reflections occur.

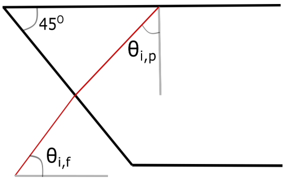


Figure 4.8: Schematics of the angle of the incident light into the prism and the angle of the incident light within the prism described and explained above.

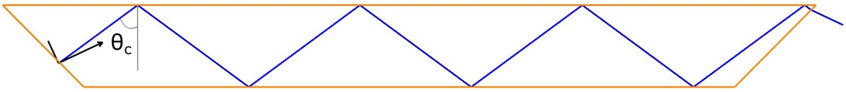


Figure 4.9: Simulation of the prism geometry which will be used in the setup. When the incident light is sent into the prism with a 25° angle instead of a 45° angle, one can clearly see that the amount of reflections are reduced from 9 to 7. The black arrow indicates the angle of the light sent into the prism.

4.3.2 Beam Divergence

The simulation tool was also used in order to see how the beam diverges and acts within the prism. This was done by using another source type called Source Two Angle and using a circular array. By doing this the radius of the circle can be set to the same as the fibre core at 200 µm and the angle of the light cone out of the fibre. The light cone out of the fibre will be the same as the acceptance cone into the fibre, which is defined as $NA = n \cdot \sin\theta$ [58]. The refractive index will be that of air n = 1 and NA = 0.30 is the numerical aperture for the fibre in the setup. Therefore the angle of acceptance cone will be $\theta = 17.5$. Since the light out of the fibre is circular, the light beam will diverge in both x and y directions and the source is therefore set to have an angle in both directions. In the simulation this means that the X and Y half angle in the non-sequential component editor is set to 17.5. A simulation where the beam only diverges in the x-direction was also performed to get a clear image of how large the beam exiting the prism will be, see Figure 4.10. The simulation of the actual beam can be seen in Figure 4.11. The rays within the beam change colour each time they are reflected. In order to get this result a coating had to be defined for the prism. The coating was set to reflect 70 $\%$ of the beam and absorb 30 $\%$, in the coating file. This can be done since the prism in itself will absorb some of the light in real life as well. If a coating is not used 100 $\%$ of the beam is reflected within the prism and is filled with rays, which is unphysical. This also makes it difficult to clearly see how the incoming light beam will diverge and act and therefore backscattered rays are omitted.

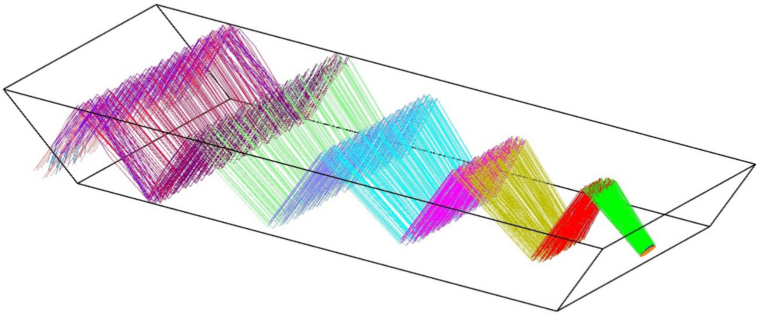


Figure 4.10: Simulation of the prism with a beam that diverges only in the x-direction with an angle of θ = 17.5°. This simulation does not fully represent real life, but it is included in order to clearly see how the beam diverges.

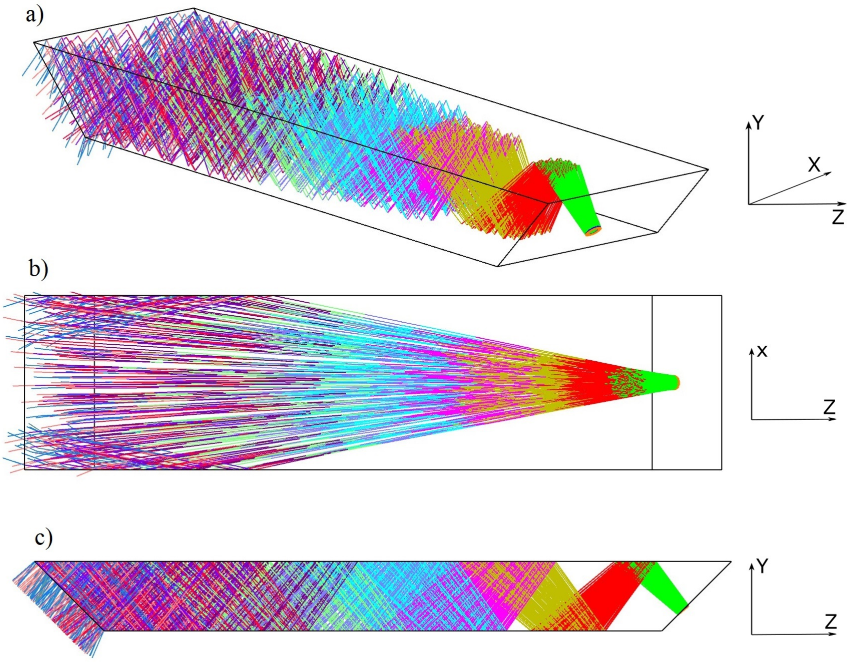


Figure 4.11: Simulation of the prism with a beam that diverges out with an acceptance angle of $\theta = 17.5^\circ$. It can be seen that the beam spreads out as it propagates through the prism. **a)** Simulation in 3D, **b)** simulation seen from above in X-Z plane, **c)** simulation seen from the side in Y-Z plane.

4.4 Fibres

4.4.1 Polishing

The fibres that were polished with 10 different polishing paper combinations were tested in the transmission setup described in Section 3.3.2. The fibres from each polishing paper combination was used to acquire a water spectra in order to see which combination would give the best fibres. The best fibres will obtain a water spectra with little noise and with a shape similar to that of a water spectra obtained from a free-space setup, see Figure 4.12. When the water spectra were obtained the pathlength was set so that the measured signal through water was approximately 2 V for all of the fibres.

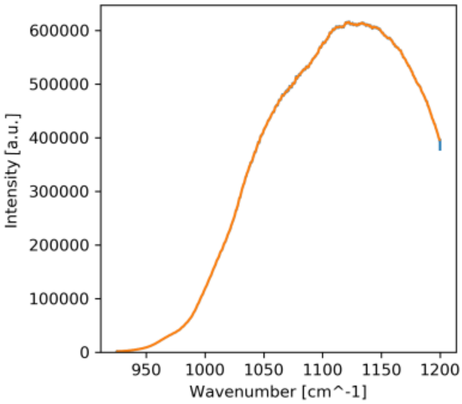
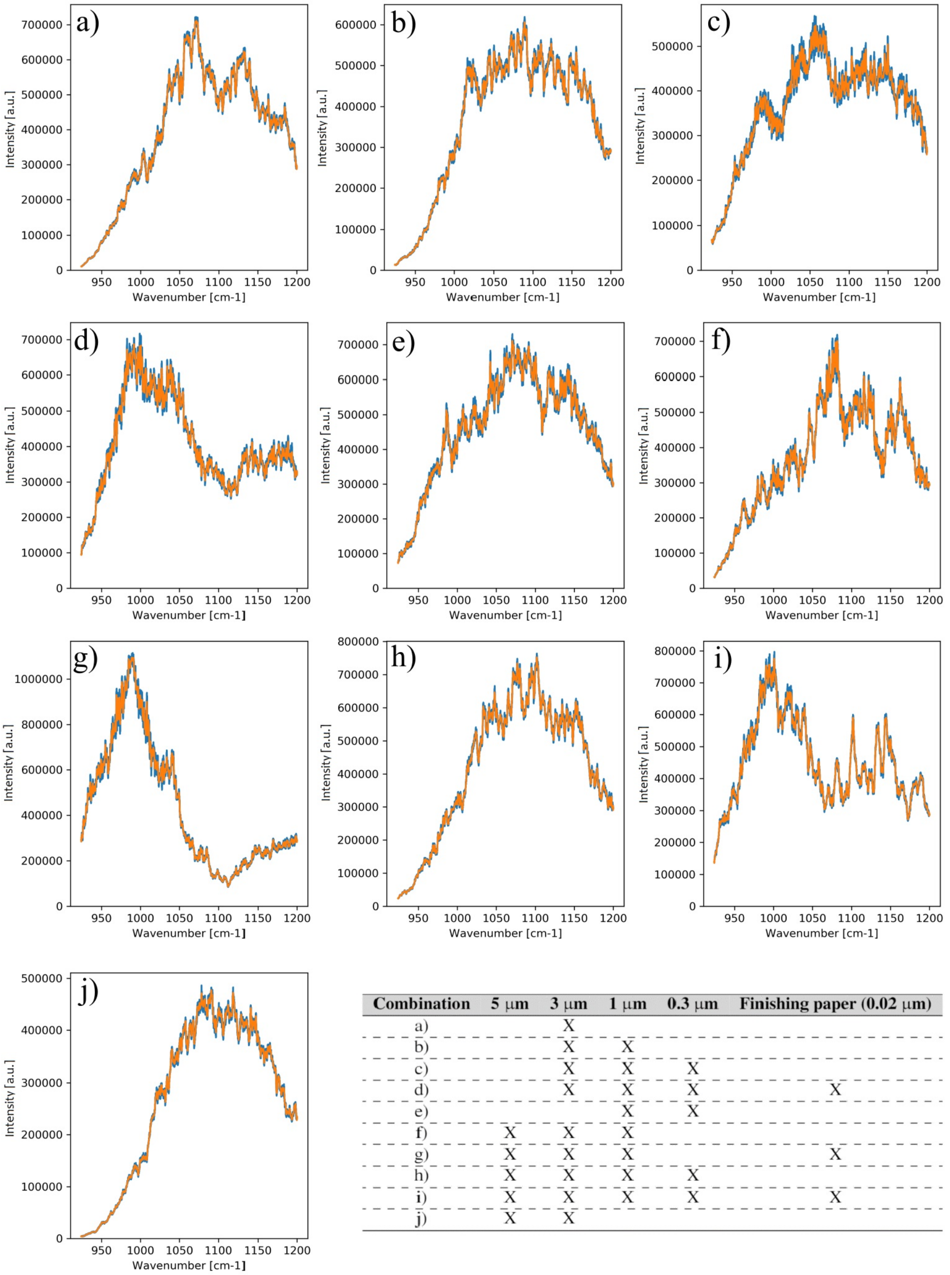


Figure 4.12: Water spectrum acquired with free-space spectroscopy using the same QCL as the setups used here. The spectrum was acquired for the project thesis for completeness.

Some of the different polishing combinations shown in Section 3.2 gave similar results, see Figure 4.13 for the diferent water spectra. By looking at all of the water spectra it is clear that the fibre polished with the 5 µm and then the 3 µm grit sized paper gave the best result. The spectra obtained by using these fibres have a shape which is very similar to the water spectra obtained with a free-space setup. It is also the spectra with the least amount of noise. When the finishing paper with the diamond particles was used to polish, the fiber ends often became rough and torn up. This is reflected in the water spectra obtained with these fibres, the shape of these spectra is very different from the others.



Combination	5 µm	3 µm	1 µm	0.3 µm	Finishing paper (0.02 µm)
a)		X			
b)		X	X		
c)		X	X	X	
d)		X	X	X	X
e)			X	X	
f)	X	X	X		
g)	X	X	X		X
h)	X	X	X	X	
i)	X	X	X	X	X
j)	X	X			

Figure 4.13: The water spectra obtained by using the fibres polished with the 10 different polishing paper combinations. The combinations is seen in the figure. The water spectra obtained with fibres that had been polished with the finishing paper have a very different shape than the others. From the spectra it is clear that polishing combination **j)** results in a spectrum with little noise and a shape very similar to the water spectrum acquired by a free-space setup seen in Figure 4.12.

4.4.2 Fibre Combination

Different optical fibre combinations were tested in the transmission setup with water in order to see which combination resulted in the least distorted signal. The best hand polished fibres were tested together and a measurement using these fibres can be found in Figure 4.14a). The spectrum has a shape that is similar to that of free-space spectroscopy, but the signal experiences some noise and could be improved. The measurement using hand-polished fibres were very similar to when factory polished fibres of the same size were used, see Figure 4.14b). Due to this the light transmitted out of the first fibre was investigated by using a laser viewing card for mid-IR. The beam was seen to diverge and was larger then the size of the core of the fibre. Due to this the beam was not collected by the next fiber after propagating through the sample and a lot of the intensity was lost. To correct for this a fibre with a larger core was used between the sample and detector to guide more of the diverging beam. The larger fibre was first tested together with the best self-polished fibre and the resulting measurements can be seen in Figure 4.14c). Here it is clear that the signal now experiences less fringes but have more defined dips in its shape. Therefore the larger fibre was also tested together with the smaller factory polished fibre between the sample and laser. The result of this measurement is found in Figure 4.14d). By using these fibres the SNR is seen to improve, the fringes have been greatly reduced and the shape of the spectrum corresponds very well with that acquired from free-space spectroscopy, see Figure 4.12. This fibre combination was therefore used in the setups.

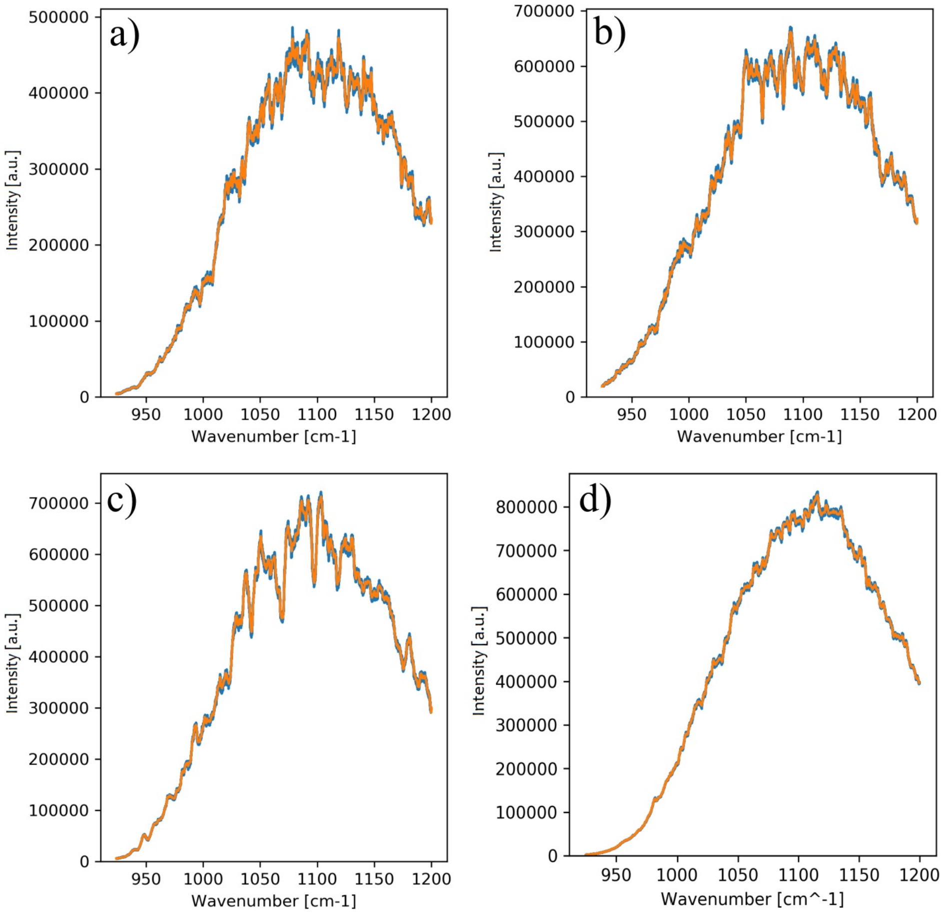


Figure 4.14: Absorption spectrum of water, acquired using; **a)** two fibers where both have been polished first using polishing paper with 5 μm grit size and then with 3 μm grit size.**b)** two fibres where both are factory polished. **c)** the fiber going to the detector is factory polished and has a larger diameter than the first fiber, which is the best of the hand polished optical fibres **d)** the fiber going to the detector is factory polished and the core has a diameter of 600 μm, the other fiber is also factory polished but has a smaller core diameter 410 μm.

4.5 Measurements

All of the samples were measured both in the ATR and the fibre-coupled transmission setup. The two measurements methods will further be referred to as the ATR and transmission setup. After all of the measurements were recorded and signal processed the absorption spectra were used in the data analysis program described in Section 3.5. Three sample sets were measured; 10 samples with glucose and water, 10 samples with BHA and water and 25 samples containing glucose, urea, lactate and albumin mixed in a PBS solution. The concentrations of the samples can be found in Section 3.6. Each of the sample sets were measured 4 times over 2 days with each setup. The samples were stored in a refrigerator between the measurements and to avoid temperature variations they were placed in the room with the setup at least one hour before the measurements were performed. For each setup 180 sample measurements were obtained and the same amount of reference measurements were acquired. The 4 measurements of the sample sets will be referred to as data sets, where each data set will contain the spectra for all of the samples in the sample set. The RMSEP value will be used in order to compare the two measurement methods when PLSR is used with a separate data set for validation and the three other data sets are used for prediction/training. The RMSECV value will also be used to see how well the concentrations can be predicted within one data set for each measurement method. The RMSECV value will be calculated by using the best data set for each sample set and method. Leave-one-out cross-validation will be used for the sample set containing 10 samples and leave-5-out will be used for the sample set with 25 samples. It is ideal that the RMSEP and RMSECV values are as low as possible since this indicates that the concentrations are predicted fairly well.

Some of the measurements were disregarded as outliers due to irregular shapes of the spectra. The amount of disregarded measurements can be seen in Table 4.2, the number of outliers from each measurement method is approximately the same. For the samples containing BHA the measurements that were removed as outliers are in general those that had the lowest concentrations.

Table 4.2: The number of outliers removed from the the analysis and the total number of measurements used in the regression analysis.

Sample Set	ATR: Outliers	ATR: Total number of measurements	Transmission: Outliers	Transmission: Total number of measurements
Glucose dissolved in water	1	39	1	39
BHA dissolved in water	10	30	11	29
Glucose, lactate, urea and albumin dissolved in PBS solution	2	98	2	98

4.5.1 Absorbance Spectra

The absorbance spectra for the three different sample sets measured with the two different setups can be seen below. The spectra are calculated using the absorbance equation, Equation 3.1 in Section 3.4 which makes the results comparable to results found in literature. The spectra have been filtered with a MA FIR filter to reduce noise and enhance the absorbance spectra. Figure 4.15 shows the absorbance spectra of the measured glucose solutions. Figure 4.16 shows the absorption spectra for the measured BHA solutions. The absorption spectra for the complex solution containing glucose, urea, lactate and albumin dissolved in PBS can be seen in Figure 4.17. A water measurement is included in all of the absorbance plots.

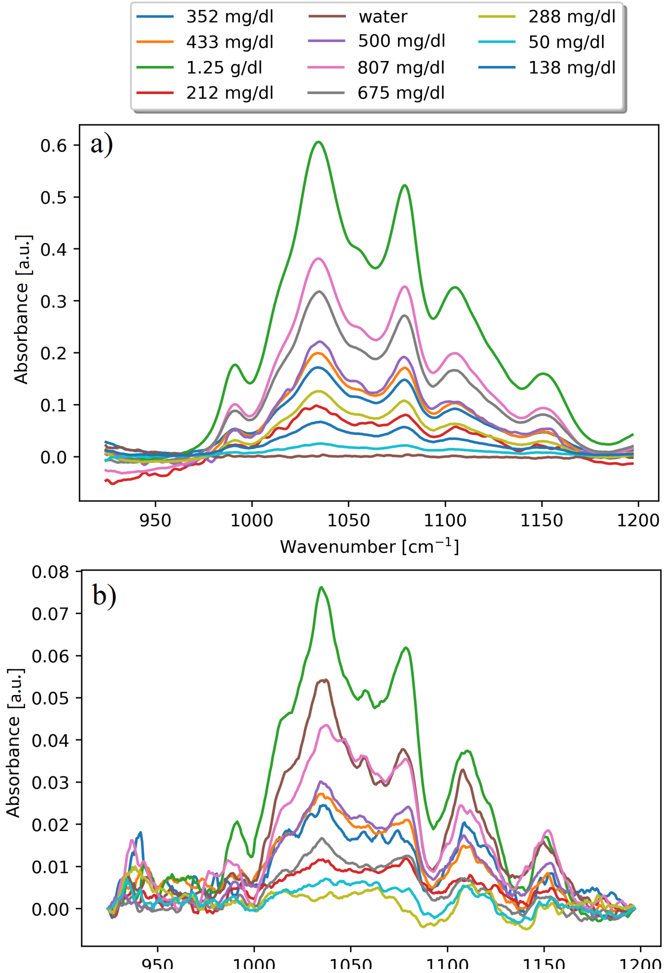


Figure 4.15: Absorbance spectra of glucose solutions with different glucose concentrations in a) the transmission setup and b) the ATR setup.

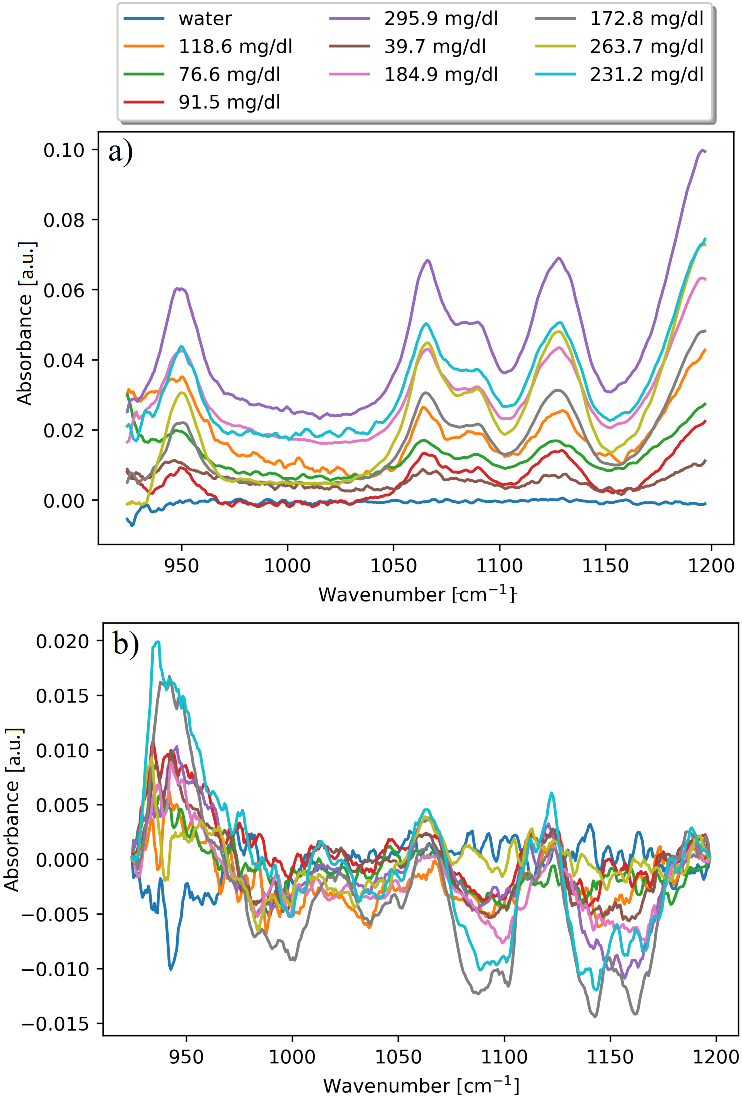


Figure 4.16: Absorbance spectra of BHA solutions with different concentrations in a) the transmission setup and b) the ATR setup.

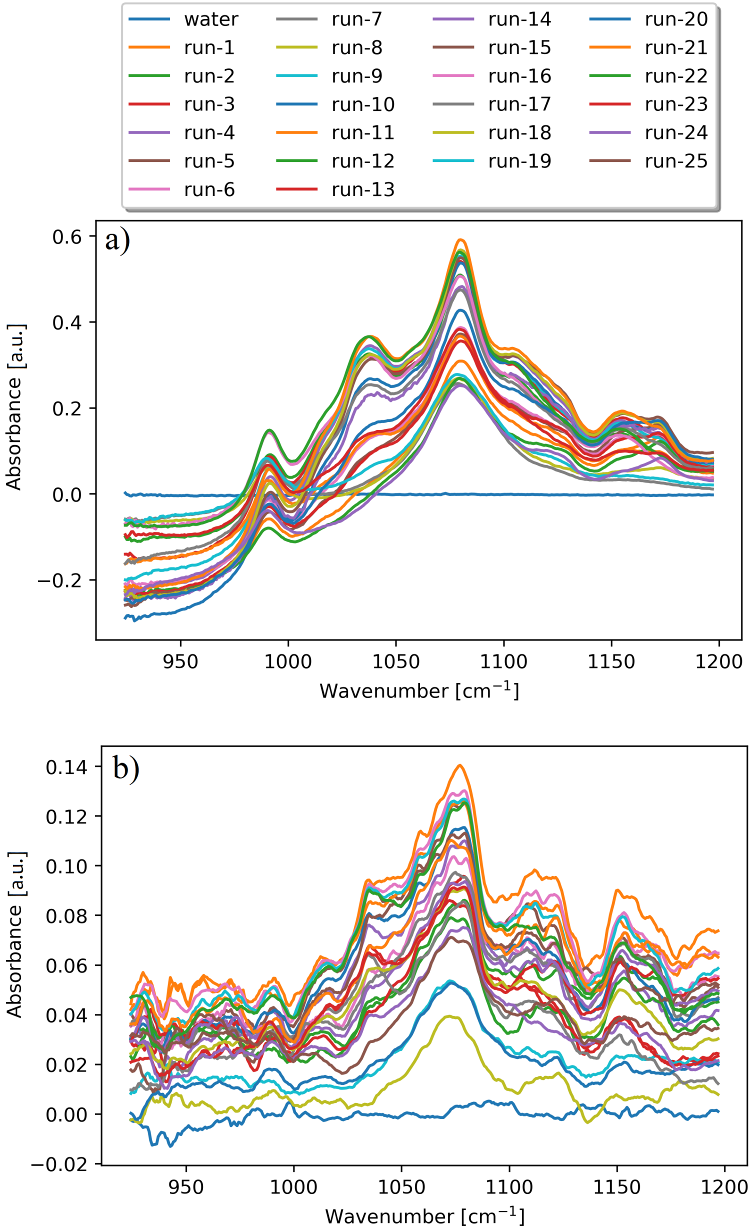


Figure 4.17: Absorbance spectra of the more complex solutions with different glucose, urea, lactate and albumin concentrations in a) the transmission setup and b) the ATR setup.

4.5.2 Concentration Plots

The concentration plots obtained from the data analysis program for both of the measurements methods are shown together in figures for easy comparison. PLSR is used with a separate data set for prediction and three data sets for validation. Cross validation is performed on the best data set from each measurement method and sample set. Leave-one-out cross-validation is performed with 7 iterations on the data obtained from the sample sets with 10 different samples. Leave-5-out cross-validation is performed with 10 iterations on the data obtained from the sample set with 25 different samples.

The RMSEP value is obtained when PLSR is performed using a separate data set for validation. The RMSECV value is obtained when cross-validation is performed. Both of these values are considered in order to compare the two measuring methods. Since the most important information in the results are the RMSEP and RMSECV values, these will be summarized in a table to clearly see the difference between the two measurement methods. See Table 4.3 for RMSEP/RMSECV values for the transmission setup and Table 4.4 for the RMSEP/RMSECV values for the ATR setup. If the coefficient of determination R^2 is very low or even negative the values do not correlate and the RMSEP and RMSECV values will not be representative. This will be represented with a line (-) in the tables presenting the RMSEP and RMSECV.

The sample set with pure glucose concentrations (sample set 1) is measured in order to evaluate the glucose sensitivity of the measurement setups. Small RMSEP/RMSECV values indicates a high glucose sensitivity. The BHA sensitivity of the setups is similarly investigated by measuring pure BHA with water (sample set 2). The more complex sample set (sample set 3) is measured in order to see how well the glucose selectivity of the setups is. If the glucose selectivity is high, the RMSEP/RMSECV values will be small. The concentration plots show the reference concentration plotted against the predicted concentration. The plots also includes the RMSEP/RMSECV values and other information about the preprocessing used in order to achieve the plots. The concentration plots for the glucose prediction in the complex sample set (sample set 3) can be seen in Figures 4.18 and 4.19. The concentration plots for the BHA prediction in sample set 2 can be seen in see Figures 4.20 and 4.21. The remaining concentration plots are included in Appendix G.

The reference measurements, i.e. the water measurements were performed before each sample measurement. For both setups it was tested how often the reference measurements should be taken in order to achieve a good result. Reference measurements obtained after every 5 measurement and after every second measurement were investigated but the best result was taking one reference measurement for each sample measurement. Even though the reference measurements from the transmission setup were found to be very similar to each other there was a small drift in the laser intensity and one reference measurements for each sample measurements gave a better result.

Summary of RMSEP and RMSECV obtained from concentration plots

Table 4.3: RMSEP and RMSECV values for the Transmission setup.

Sample Set	Analyte	RMSEP	Latent variables	RMSECV	Latent variables
1	Glucose	9.3	6	5.1	3
2	BHA	2.4	4	6.5	3
3	Glucose	10.8	4	10.5	4
	Urea	12.0	6	9.9	6
	Lactate	5.2	6	8.1	6
	Albumin	108.8	6	75.3	6

Table 4.4: RMSEP and RMSECV values for the ATR setup. * The RMSEP for BHA is found by only considering two measurement sets.

Sample Set	Analyte	RMSEP	Latent variables	RMSECV	Latent variables
1	Glucose	16.1	3	7.7	5
2	BHA	25.5*	4	28.0	4
3	Glucose	25.2	6	25.0	4
	Urea	-	-	-	-
	Lactate	-	-	-	-
	Albumin	362.7	3	184.7	4

Concentration plot predicting the glucose concentrations in sample set 3.

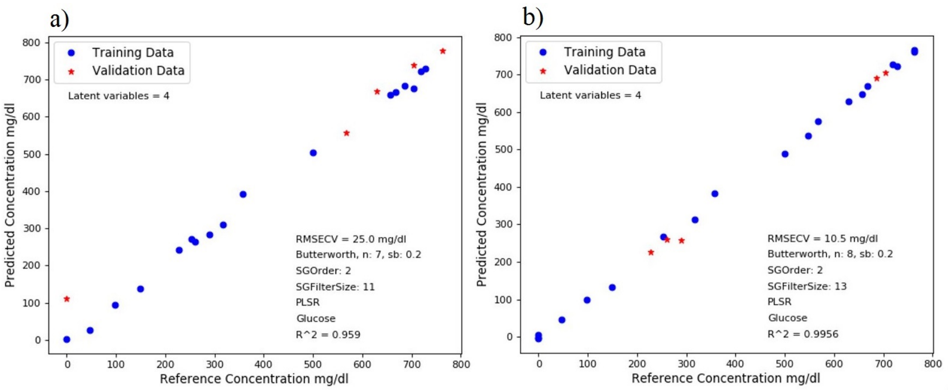


Figure 4.18: Glucose concentrations predicted with PLSR using cross-validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

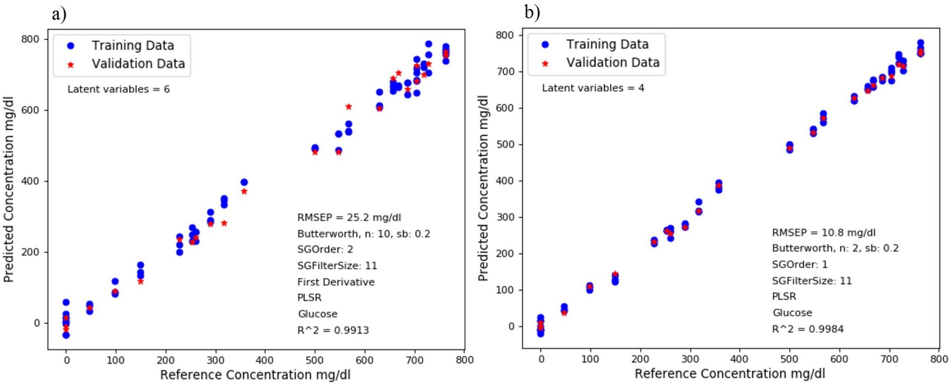


Figure 4.19: Glucose concentrations predicted with PLSR using a separate data set for prediction and 3 data sets for validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

Concentration plot predicting the BHA concentrations in sample set 2.

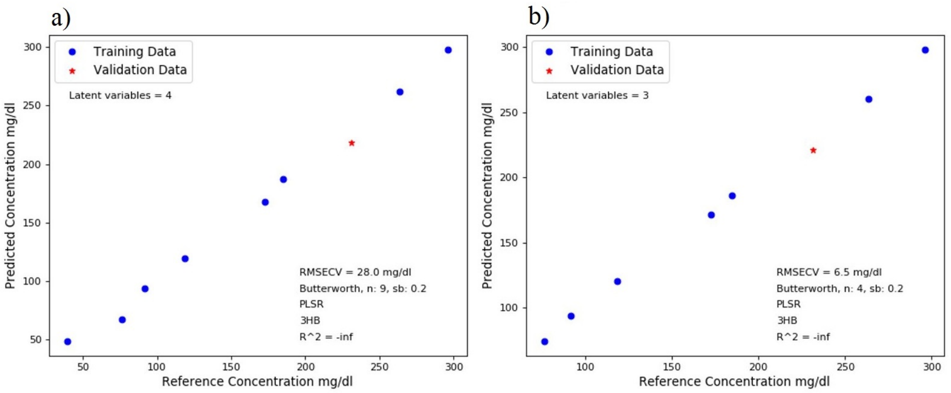


Figure 4.20: BHA concentrations predicted with PLSR using cross-validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

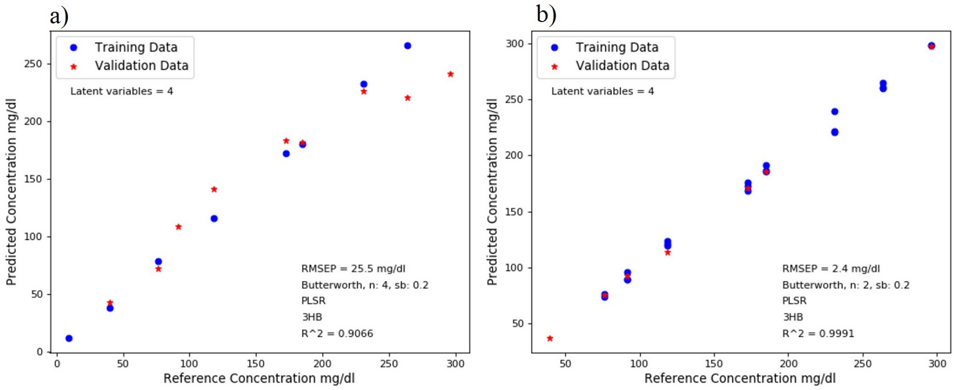


Figure 4.21: BHA concentrations predicted with PLSR using a separate data set for prediction and 3 data sets for validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

Regression analysis was also performed on absorbance spectra where baseline correction had been performed. When the absorbance spectra obtained from the transmission measurement were baseline corrected with the first method (Basline correction: 1) the predicted RMSEP value for glucose in the complex solution (sample set 3) was improved from 10.8 mg/dl to 10.3 mg/dl. However, the predicted RMSEP value for glucose in the complex solution was worse for the baseline-corrected absorbance spectra obtained from the ATR measurements, see Figure 4.22. The second baseline correction method was also tested but this did not give an improved result. In relation to a CGM device the most important value is the glucose selectivity RMSEP and RMSECV value in the complex solution, which is improved with baseline correction for the transmission setup. However, the results for the other analytes was that the RMSEP and RMSECV were either deteriorated or the same as without baseline correction. The RMSEP and RMSECV values were also worse for all of the ATR spectra when they were baseline corrected.

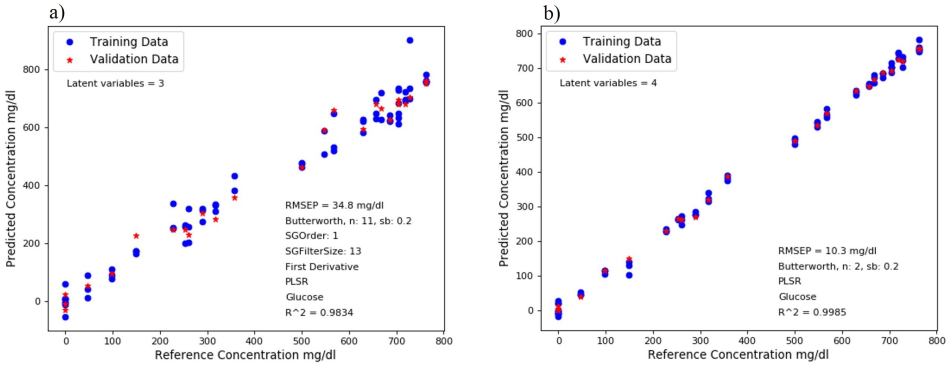


Figure 4.22: Glucose concentrations predicted in the complex solution (sample set 3) with PLSR using a separate data set for prediction, were the absorbance spectra have been baseline corrected, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

5 | Discussion

5.1 Glucose Meaurements

The measurements obtained with the fibre-coupled transmission setup were found to be both stable and with a high SNR. When glucose was predicted in the complex sample set the RMSEP was found to be 10.8 mg/dl and the RMSECV 10.5 mg/dl. The RMSEP was found to be 9.3 mg/dl when the glucose concentration were predicted in the pure glucose solution and the RMSECV 5.1 mg/dl. These values are well within the accepted values for prediction of glucose in the presence of interfering analytes in relation to a CGM device. Based on the RMSEP/RMSECV both the glucose sensitivity and selectivity is high for the measurements obtained by using the fibre-coupled transmission setup.

In the measurements obtained with the ATR setup the signal is very low which results in a lower SNR, as compared to the transmission setup. The measurements are still stable and some of the noise was removed with averaging and filtering. When glucose was predicted in the complex sample set the RMSEP was 25.2 mg/dl and the RMSECV 25.0 mg/dl. The glucose sensitivity RMSEP was found to be 16.1 mg/dl when the glucose concentrations were predicted in the pure glucose solution and the RMSECV was found to be 7.7 mg/dl.

Lower RMSEPs and RMSECVs have been achieved by other groups utilizing EC-QCL systems with both ATR and transmission setups. These values are though not directly comparable to the results achieved in this master thesis due to different setups, interfering analytes and concentrations. E.g. Haase et. al. [34] used a fibre-coupled transmission setup and reported on a glucose sensitivity RMSEP of 3 mg/dl in pure glucose solution. However, the group had 8 different concentrations and seven of these were 500 mg/dl or below. Their complex sample set did not contain the same analytes and they used a reference detector in order to correct for power variations during the measurements. Due to the second detector their setup is larger than in this master thesis, which will make it larger if miniaturized for a CGM device. They were also able to distinguish glucose from other analytes with a RMSEP of 14 mg/dl. The glucose selectivity is thus higher in the fibre-coupled transmission setup used in this master thesis. The same group also performed measurements utilizing a transflectance setup [35]. They reported on lower RMSECVs here as well, but the average concentration of glucose in the pure glucose samples was 214 mg/dl. Since the average concentration is low this indicates that most of the concentrations are below 214 mg/dl. The average glucose concentration in the pure glucose solution in this master thesis is 470.3 mg/dl. Since RMSEP and RMSECV values are not relative but scale accordingly to the predicted concentrations, the value will be larger for both of the measurement setups in this thesis due to more samples with larger concentrations.

Based on the glucose sensitivity RMSEP/RMSECV values and how good the setups are at distinguishing glucose from other analytes it is clear that the transmission setup, which is more stable and has a higher SNR, is preferable. In general both the RMSEP and RMSECV values for the glucose sensitivity and selectivity are similar to that acquired from previous studies such as Haase et. al. [34]. In this master thesis some of the measured concentrations are somewhat higher than one would find in the body, and as a result of this the RMSECV and RMSEP values are accordingly higher. One can see a linear response over the entire range and even at lower concentrations the samples are predicted very accurately.

5.2 Other Analytes

The prediction of BHA was found to be very different in the measurements obtained with the fibre-coupled transmission setup and those obtained from the ATR setup. The difference between the BHA sensitivity RMSEP in the two setups was found to be 23.1 mg/dl, due to a very low RMSEP for the transmission setup 2.4 mg/dl and a higher RMSEP for the ATR setup 25.5 mg/dl. From the preliminary characterisation of the BHA absorbance spectra obtained from both measurement setups it is clear that the ATR measurement is more affected by noise due to a lower signal. The RMSECV value was 6.5 mg/dl for the transmission setup and 28.0 mg/dl for the ATR setup.

Since it is also beneficial to measure the urea, lactate and albumin levels in the body these were also predicted with regression analysis. The selectivity of urea, lactate and albumin were high for the transmission setup. The RMSEP for urea was 12.0 mg/dl, 5.2 mg/dl for lactate and for albumin the RMSEP was 108.8 mg/dl. The RMSEP value for albumin is high due to the higher concentrations of this analyte in the samples. It was not possible to predict the concentrations for urea and lactate when the ATR measurement setup was used. When these were predicted from the measurements obtained from the ATR setup the coefficient of determination R^2 was very low or negative, which indicates that there is no correlation between the predicted concentrations and the reference concentrations. The reason for why this might occur is that the SNR is too low and the regression analysis is not able to predict any of the concentrations. It was possible to predict the concentration of albumin in the ATR setup and the RMSEP value was 362.7 mg/dl, which is very high compared to the transmission setup.

5.3 Noise

There is a lot of noise in the ATR measurements compared to the transmission measurements, most likely due to the relatively lower SNR. As mentioned in the results (Section 4.1) some noise was also induced due to the angle of the lens which resulted in FP interference between the lens and the prism. The angle of the lens was therefore changed so that the lens would not be parallel to the slanted side of the prism, in order to reduce the noise and distortion. However, based on the results and the measurements it appears as if the angle could have been more optimal in order to reduce the interference, and thus reduce the noise, even more. It is also possible that there was some reflections between the prism and the fibre end transmitting light into the prism. In order to avoid this the fibre was placed as close as possible to the prism. Since the angle of the fibre was 50° and the slanted sides of the prism 45°, some of the fibre was not in contact with the prism and there would have been some air between the fibre and prism that provided room for reflections. Aligning the setup was difficult and time consuming due to the many different components. The fibre between the prism and detector was adjusted with a 3D-stage which made it easy to control the fibre. A second 3D-stage was ordered to have more control of the first fibre to improve the alignment, but this was received too late so the first fibre was only positioned manually for the measurements done here.

It has been shown by some groups that there is some noise in the laser used in the setups here, EC-QCL. The noise that is observed in the laser is low-frequency flicker ($1/f$) noise. This type of noise is also typically found in semiconducting devices, but the origin of the noise is somewhat different in the two lasing types. It is difficult to identify the origin of the noise but the general thought is that fluctuations of electrons tunneling through the multi-barrier structure are the explanation of the observed flicker noise [59, 60]. Due to this, some of the noise that is observed in the spectra can be a result of the noise within the laser. The detector might also experience $1/f$ noise and the $1/f$ noise corner frequency is < 20 kHz. However, since the spectra obtained from the transmission setup experienced very little noise it is believed that the noise from both the laser and detector is small.

Based on the spectra obtained by using different fibre combinations presented in the results it is clear that it is difficult to achieve a spectrum with little noise when self-polished fibres are used. It is also clear that it is difficult to achieve a spectrum with little distortion when the size of the two fibres are the same. A larger fibre core was used since it was seen that the beam diverged out from the first fibre and some light was therefore lost. Based on these results the AgCl:AgBr polycrystalline fibre with a 410 ± 15 μm core was used between the laser and the sample, and a fibre with a 600 ± 20 μm core was used between the sample and the detector. The producers of the AgCl:AgBr polycrystalline fibre also offer a fibre with a 860 ± 20 μm core and since the spectra were improved by increasing the fibre core for the fibre transmitting the light to the detector it is possible that a fibre with an even larger core would improve the signal even more. However the minimum bending radius is increased for fibres with larger cores and it might not be suitable for both of the setups.

5.4 Preprocessing

The spectra were averaged over 255 pulses and signal-averaged in order to reduce the noise and fringes. By averaging, most of the noise was removed from the spectra, but filtering was also used in the data analysis to improve the results. Different filters were added to the regression analysis program in addition to the already implemented Savitzky-Golay filter. When the regression analysis was performed the different filters were tested with different orders to achieve the smallest possible RMSEP/RMSECV. In general the Butterworth filter reduced noise the most without over-filtering and was found to give the best result.

However, the improvement in the RMSEP and RMSECV was not as large as one would assume and other filters might improve the SNR even more. Due to time limitations the number of tested filter had to be limited. It was also seen that in general the absorbance spectra obtained from the transmission setup did not actually need filtering to get good results, but since the RMSEP and RMSECV were reduced some filtering was used on all of the measurements.

Based on an absorbance spectra of pure glucose solutions obtained with FTIR spectroscopy, included in Appendix H, there should not be any absorption outside the 980-1180 cm^{-1} range. This is not the case for the absorbance spectra presented in the results. The baseline shift might be due to a drift in the power from the laser or drift in the setup, e.g. movenements in the fibres. In order to get the expected absorbance spectra baseline correction could be used. The absorbance spectra of the complex solutions obtained by the transmission setup had negative values at the lower wavenumbers. The albumin in the complex solutions reduces the amount of water in the samples. When the absorbance spectra is calculated the water spectrum will be much larger than the sample spectrum resulting in negative values. This was not seen in the spectra obtained with the ATR setup most likely due to the low signal obtained with this setup.

When baseline correction was applied to the absorbance spectra, the glucose prediction RMSEP value in the complex solution was improved. However it was also seen that the baseline correction did not improve the prediction of the other analytes in this sample. The results when baseline correction was utilized was significantly worsened for the ATR measurements. To make the two results from the regression analysis directly comparable and since the improvement for the transmission setup was not that significant, the regression analysis for the baseline corrected spectra was not considered.

5.5 ATR Setup

Another factor in the ATR setup that might lead to noise is the dimension of the prism. Some of the light is absorbed in the ZnS prism due to the nature of the material and this results in the low signal obtained with this setup. In order to reduce the amount of absorbed light within the prism the height of the prism should be reduced. A thinner prism will reduce the pathlength of the light between the top and bottom of the prism and thus reduce the absorbance. If a prism with a 1.5 mm height was used instead the approximated absorbance would be reduced from 47.8 % to 35 %. It would also be beneficial if the prism was longer, since this leads to more reflections within the prism which increase the interaction area and this was seen in the paper by Haas et. al. [33]. However, there is a trade-off between the absorption and the number of reflections since a longer prism will increase the pathlength. If another material was used for the prism the absorbance could also have been reduced. If a prism made of diamond with the same pathlength was used only 37 % of the light would be absorbed due to a lower absorption coefficient (α = 0.7). However, a diamond prism is very expensive and not ideal for a personalized CGM device.

The simulations were used in order to find which dimensions would result in most reflections within the prism. It was also used in order to investigate how the beam would diverge within the prism and transmit out. The simulation was consistent with the observed radiation out of the prism by using a mid-IR laser viewing card. This indicates that the coating used in the simulations was an approximate representation of the real life absorption in the prism. From the simulation it was seen that the beam diverges within the prism and at the end it hits the sides of the prism and scatter. This will reduce the intensity of the beam and to avoid this a thinner prism could be used. A thinner prism with shorter pathlength will reduce the scattering as the light will not diverge as much.

5.6 Temperature

The setups were not in a temperature-controlled environment when the measurements were performed. The temperature was approximately 24°C when all of the transmission measurements were performed, and approximately 22°C when the ATR setup was used. According to the study on temperature variations' effects on spectroscopy setups by Haase et. al. [35] a large temperature variation is required in order to affect the signal. When the temperature is increased from 20°C to 40°C the the signal is increased with 9.85 mV. The setup in this thesis would only have experienced small temperature variations, which would not have had major effect on the signals.

5.7 Outliers

In both of the measurement setups there were some outliers for all of the sample sets. For the measurements of the pure BHA solutions the sample with the lowest concentration was removed for all of the measurements, from both setups. This indicates that the sample might have been mixed wrongly or might be polluted. The other outliers in the transmission measurements might be a result of the gap between the fibres not being completely filled or that the previous sample was not entirely removed. The outliers in the ATR measurements might also be a result of a sample not being entirely removed and thus interfering with the next sample measurement. The outliers might also occur due to the sample placed on top of the ATR prism not covering a large enough area of the prism. Noise in the laser and detector might also be the reason for some of the outliers.

6 | Conclusion

The main goal of this master thesis has been to improve spectroscopy measurement methods for detecting glucose with mid-IR light. Two spectroscopy methods were considered; fibre-coupled transmission spectroscopy and ATR spectroscopy. The glucose sensitivity was found to be higher for the transmission setup (RMSEP 9.3 mg/dl) compared to the ATR setup (RMSEP 18.1 mg/dl), but they were both found to be highly accurate over the entire concentration range (0 - 762 mg/dl). The predicted glucose selectivity was found by predicting the glucose concentrations in solutions containing analytes with interfering absorption bands such as lactate, urea and albumin. The selectivity was found to be very good especially for the transmission setup (RMSEP 10.8 mg/dl). The selectivity RMSEP value was almost the same as the sensitivity RMSEP found in the pure glucose solutions for the transmission setup. The selectivity for the ATR setup was somewhat lower (RMSEP 25.2 mg/dl) and based on the RMSEP/RMSECV values the transmission setup was found to be superior to the ATR setup.

Even though some of the concentrations in the samples measured in this thesis are relatively high compared to the physiological levels, the response was seen to be linear over the entire range for both cases. The error was thus seen to be low for both low concentrations ($<$ 100 mg/dl) as well as for higher concentrations.

The BHA sensitivity RMSEP/RMSECV was also found to be more accurate by measuring with the transmission setup (2.4/6.5 mg/dl). It is thus shown that fibre-coupled spectroscopy is a method that can be used for detecting diabetic ketoacidosis. The concentrations of lactate, urea and albumin could also be predicted in the transmission setup (RMSEP 5.2, 12.0 and 108.8 mg/dl respectively). In the ATR setup it was only possible to predict the concentration of albumin (RMSEP 362.7 mg/dl), due to a low SNR. The low SNR is a result of the light absorption in the prism, which could be reduced by changing the prism dimensions.

Due to the results, further work should be to optimize the ATR measurement setup in regards to the prism geometry. It is believed that a thinner and longer prism would provide better results since the absorption would be reduced and the interaction area increased. It should also be optimized further in regards to reduction of noise due to reflections within the setup. More work can also be done on the preprocessing, i.e. more filter types and other baseline functions could be tested in order to improve the results of the regression analysis. BHA should also be measured together with glucose in order to see if these interfere with each other or if both can be predicted with high accuracy in a solution containing both of these important analytes.

Bibliography

[1] I. K. Ilev and R. W. Waynant, "Mid-infrared Biomedical Applications," in *Mid-infrared Semiconductor Optoelectronicss* (A. Krier, ed.), pp. 615–631, London: Springer London, 2006.

[2] F. Toor, S. Jackson, X. Shang, S. Arafin, and H. Yang, "Mid-infrared Lasers for Medical Applications : introduction to the feature issue," *Biomedical Optics Express*, vol. 9, no. 12, pp. 6255–6257, 2018.

[3] L. Zhang, G. Tian, J. Li, and B. Yu, "Applications of Absorption Spectroscopy Using Quantum Cascade Lasers," *Applied Spectroscopy*, vol. 68, no. 10, pp. 1095–1107, 2014.

[4] J. Kottmann, J. M. Rey, J. Luginbühl, E. Reichmann, and M. W. Sigrist, "Glucose sensing in human epidermis using mid-infrared photoacoustic detection.," *Biomedical optics express*, vol. 3, pp. 667–80, 4 2012.

[5] WHO, "Diabetes." https://www.who.int/diabetes/global-report/WHD2016_Diabetes_Infographic_v2.pdf , [Accessed: 03.06.2019].

[6] A. P. F. Turner, "Biosensors: sense and sensibility.," *Chemical Society reviews*, vol. 42, no. 8, pp. 3184–96, 2013.

[7] S. Liakat, K. A. Bors, L. Xu, C. M. Woods, J. Doyle, and C. F. Gmachl, "Noninvasive in vivo glucose sensing on human subjects using mid-infrared light," *Biomedical Optics Express*, vol. 5, no. 7, pp. 2397 – 2404, 2014.

[8] D. Matthews, *Diabetes*. Oxford University Press, 1 ed., 2008.

[9] D. Bruen, C. Delaney, L. Florea, and D. Diamond, "Glucose Sensing for Diabetes Monitoring: Recent Developments.," *Sensors (Basel, Switzerland)*, vol. 17, 8 2017.

[10] M. J. Fowler, "Microvascular and Macrovascular Complications of Diabetes," *Clinical Diabetes*, vol. 26, pp. 77–82, 4 2008.

[11] J. Wang, "Electrochemical glucose biosensors," *Chemical Reviews*, vol. 108, no. 2, pp. 814–825, 2008.

[12] R. L. Weinstein, S. L. Schwartz, R. L. Brazg, J. R. Bugler, T. a. Peyser, and G. V. McGarraugh, "Accuracy of the 5-Day FreeStyle Navigator Continous Glucose Monitoring SYstem," *Diabetes Care*, vol. 30, no. 5, pp. 1125–30, 2007.

[13] T. S. Bailey, A. Chang, and M. Christiansen, "Clinical accuracy of a continuous glucose monitoring system with an advanced algorithm," *Journal of Diabetes Science and Technology*, vol. 9, no. 2, pp. 209–214, 2015.

[14] I. L. Jernelv, K. Milenko, S. S. Fuglerud, D. R. Hjelme, R. Ellingsen, and A. Aksnes, “A review of optical methods for continuous glucose monitoring,” *Applied Spectroscopy Reviews*, pp. 1–30, 2018.

[15] B. H. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*. John Wiley & Sons, Incorporated, 2004.

[16] J. I. Khan, T. J. Kennedy, and D. R. Christian, “Infrared Spectroscopy,” in *Basic Principles of Forensic Chemistry*, pp. 127–141, Totowa, NJ: Humana Press, 2012.

[17] N. Jaggi, “Fourier transform infrared spectroscopy,” in *Handbook of Applied Solid State Spectroscopy* (D. Vij, ed.), ch. 9, pp. 411–450, Springer, Boston, MA, 2006.

[18] D. F. Swinehart, “The Beer-Lambert Law,” *Journal of Chemical Education*, vol. 39, no. 7, pp. 333–335, 1962.

[19] B. E. A. Saleh, M. C. Teich, Teich, and M. C., “Polarization Optics,” in *Fundamentals of Photonics*, ch. 9, John Wiley & Sons, Incorporated, 2 ed., 2013.

[20] Z. M. Khoshhesab, “Reflectance IR Spectroscopy,” in *Infrared Spectroscopy – Materials Science, Engineering and Technology* (T. Theophile, ed.), ch. 12, pp. 233–244, InTech, 2012.

[21] A. Schwaighofer, M. Brandstetter, and B. Lendl, “Quantum cascade lasers (QCLs) in biomedical spectroscopy,” *Chemical Society Reviews*, vol. 46, no. 19, pp. 5903–5924, 2017.

[22] M. Pleitez, H. Von Lilienfeld-Toal, and W. Mäntele, “Infrared spectroscopic analysis of human interstitial fluid in vitro and in vivo using FT-IR spectroscopy and pulsed quantum cascade lasers (QCL): Establishing a new approach to non invasive glucose measurement,” *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 85, no. 1, pp. 61–65, 2012.

[23] F. Jerome, *Quantum Cascade Lasers*. Oxford University Press, 1 ed., 2013.

[24] F. Capasso, “High-performance midinfrared quantum cascade,” *Optical Engineering*, vol. 49, no. 11, pp. 1–9, 2010.

[25] F. Jerome, “Quantum devices,” in *Quantum Cascade Lasers*, ch. 1, pp. 1–8, Oxford University Press, 1 ed., 2013.

[26] A. Schönhals, H. Tholl, M. Glasmacher, N. Kröger-Lui, A. Pucci, and W. Petrich, “Optical properties of porcine dermis in the mid-infrared absorption band of glucose,” *Analyst*, vol. 142, no. 8, pp. 1235–1243, 2017.

[27] M. Brandstetter, T. Sumalowitsch, A. Genner, A. E. Posch, C. Herwig, A. Drolz, V. Fuhrmann, T. Perkmann, and B. Lendl, “Reagent-free monitoring of multiple clinically relevant parameters in human blood plasma using a mid-infrared quantum cascade laser based sensor system,” *Analyst*, vol. 138, no. 14, pp. 4022–4028, 2013.

[28] S. Kino, S. Omori, T. Katagiri, and Y. Matsuura, “Hollow optical-fiber based infrared spectroscopy for measurement of blood glucose level by using multi-reflection prism,” *Biomedical Optics Express*, vol. 7, no. 2, pp. 701–708, 2016.

[29] Y. Matsuura, S. Kino, and T. Katagiri, “Hollow-fiber-based flexible probe for remote measurement of infrared attenuated total reflection.,” *Applied optics*, vol. 48, no. 28, pp. 5396–400, 2009.

[30] A. Werth, S. Liakat, A. Dong, C. M. Woods, and C. F. Gmachl, “Implementation of an integrating sphere for the enhancement of noninvasive glucose detection using quantum cascade laser spectroscopy,” *Applied Physics B*, vol. 124, no. 5, p. 75, 2018.

[31] A. J. Welch and M. J. Van Gemert, eds., *Optical-Thermal Response of Laser-Irradiated Tissue*. Springer Netherlands, 2 ed., 2011.

[32] C. Vrançić, N. Kröger, N. Gretz, S. Neudecker, A. Pucci, and W. Petrich, “A quantitative look inside the body: Minimally invasive infrared analysis in vivo,” *Analytical Chemistry*, vol. 86, no. 21, pp. 10511–10514, 2014.

[33] J. Haas, E. V. Catalán, P. Piron, M. Karlsson, and B. Mizaikoff, “Infrared spectroscopy based on broadly tunable quantum cascade lasers and polycrystalline diamond waveguides,” *Analyst*, vol. 143, no. 21, pp. 5112–5119, 2018.

[34] K. Haase, N. Müller, and W. Petrich, “Towards a continuous glucose monitoring system using tunable quantum cascade lasers,” in *Biomedical Vibrational Spectroscopy 2018: Advances in Research and Industry*, vol. 10490, pp. 1–8, 2018.

[35] K. Isensee, N. Müller, A. Pucci, and W. Petrich, “Towards a quantum cascade laser-based implant for the continuous monitoring of glucose,” *Analyst*, vol. 143, no. 24, pp. 6025–6036, 2018.

[36] H. Martens and M. Martens, *Multivariate Analysis of Quality - An introduction*. John wiley & Sons, LTD, 2001.

[37] N. Pandis, “The chi-square test,” *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 150, no. 5, pp. 898–899, 2016.

[38] S. W. Smith, “Moving Average Filters,” in *The Scientist and Engineer’s Guide to Digital Signal Processing*, ch. 15, pp. 277–284, California Technical Publishing, 2 ed., 1999.

[39] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing: Pearson New International Edition*. Pearson Education Limited, 4 ed., 2013.

[40] “Generalized Cosine Windows.” https://www.mathworks.com/help/signal/ug/generalized-cosine-windows.html , [Accessed: 10.05.2019].

[41] F. J. Harris, “On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform,” *IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[42] W. H. Press and S. A. Teukolsky, “Savitzky-Golay Smoothing Filters,” *Computers in Physics*, vol. 4, no. 6, pp. 669–672, 1990.

[43] M. C. Gutiérrez, “Derivative Spectroscopy Applied To the Determination of Alpha- and Beta-Acids in Hops,” *Journal of the Institute of Brewing*, vol. 98, no. 4, pp. 277–281, 1992.

[44] M. A. Czarnecki, “Resolution enhancement in second-derivative spectra,” *Applied Spectroscopy*, vol. 69, no. 1, pp. 67–74, 2015.

[45] M. Brandstetter and B. Lendl, “Tunable mid-infrared lasers in physical chemosensors towards the detection of physiologically relevant parameters in biofluids,” *Sensors and Actuators, B: Chemical*, vol. 170, pp. 189–195, 2012.

[46] A. Perelas, “Beta-Hydroxybutyrate: Referance range.” https://emedicine.medscape.com/article/2087381-overview#a1 , [Accessed: 15.02.2019], 2015.

[47] S. M. Lee and W. S. An, “New clinical criteria for septic shock: serum lactate level as new emerging vital sign.,” *Journal of thoracic disease*, vol. 8, no. 7, pp. 1388–1390, 2016.

[48] P. Caraceni, M. Tufoni, and M. E. Bonavita, “Clinical use of albumin,” *Blood Transfus*, vol. 11, no. 4, pp. 18–25, 2013.

[49] R. A. Shaw, S. Kotowich, M. Leroux, and H. H. Mantsch, “Multianalyte Serum Analysis Using Mid-Infrared Spectroscopy,” *Annals of Clinical Biochemistry: An international journal of biochemistry and laboratory medicine*, vol. 35, no. 5, pp. 624–632, 1998.

[50] V. Stojanovic and S. Ihle, “Role of beta-hydroxybutyric acid in diabetic ketoacidosis: a review.,” *The Canadian veterinary journal = La revue veterinaire canadienne*, vol. 52, pp. 426–30, 4 2011.

[51] K. Dhatariya, “Blood Ketones: Measurement, Interpretation, Limitations, and Utility in the Management of Diabetic Ketoacidosis.,” *The review of diabetic studies*, vol. 13, no. 4, pp. 217–225, 2016.

[52] G. Hoşafçi, O. Klein, G. Oremek, and W. Mäntele, “Clinical chemistry without reagents? An infrared spectroscopic technique for determination of clinically relevant constituents of body fluids,” *Analytical and Bioanalytical Chemistry*, vol. 387, no. 5, pp. 1815–1822, 2007.

[53] Zemax, “Getting Started With OpticStudio 16,” 2016.

[54] A. Rinnan, F. v. d. Berg, and S. B. Engelsen, “Review of the most common preprocessing techniques for near-infrared spectra,” *Trends in Analytical Chemistry*, vol. 28, no. 10, pp. 1201–1222, 2009.

[55] M. Debenham, “Refractive indices of zinc sulfide in the 0405–13-μm wavelength range,” *Applied Optics*, vol. 23, no. 14, pp. 2238–2239, 1984.

[56] D. R. Lide, ed., *CRC Handbook of Chemistry and Physics*. CRC Press, 82 ed., 2001.

[57] "Zinc Sulfide (ZnS)." http://www.wavelength-oe.com/en/content.aspx?sid=1263 , [Accessed: 29.05.2019].

[58] B. E. A. Saleh, M. C. Teich, Teich, and M. C., "Guided Rays," in *Fundamentals of Photonics*, ch. 6, John Wiley & Sons, Incorporated, 2 ed., 2013.

[59] S. Borri, S. Bartalini, P. C. Pastor, I. Galli, G. Giusfredi, D. Mazzotti, M. Yamanishi, and P. De Natale, "Frequency-noise dynamics of mid-infrared quantum cascade lasers," *IEEE Journal of Quantum Electronics*, vol. 47, no. 7, pp. 984–988, 2011.

[60] S. Bartalini, M. S. Vitiello, and P. De Natale, "Quantum cascade lasers: A versatile source for precise measurements in the mid/far-infrared range," *Measurement Science and Technology*, vol. 25, no. 1, pp. 1–17, 2014.

Appendices

A Zemax Simulations: ATR Prisms

A.1 Simulation - Prism2

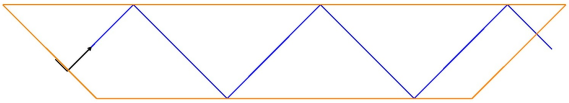


Figure 6.1: Simulation of reflecting rays in the prism with dimensions of Prism2, found in Table 3.1 in Section 3.1. The simulation shows that a prism with these dimensions will achieve 5 reflections within the prism. This prism was was available for the master thesis.

A.2 Simulation - Prism3

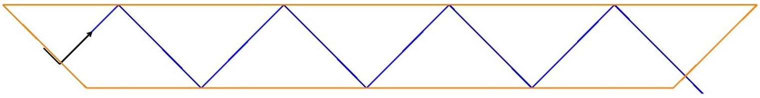


Figure 6.2: Simulation of reflecting rays in the prism with dimensions of Prism3, found in Table 3.1 in Section 3.1. The simulation shows that a prism with these dimensions will achieve 7 reflections within the prism. This prism was available for the master thesis.

A.3 Simulation - Prism4

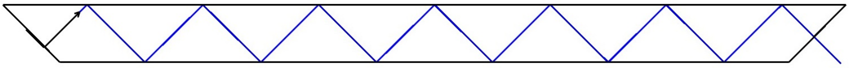


Figure 6.3: Simulation of reflecting rays in the prism with dimensions of Prism4, found in Table 3.1 in Section 3.1. The simulation shows that a prism with these dimensions will achieve 13 reflections within the prism. This prism was not available for the master thesis but the dimensions were simulated to see how the prism dimensions should be changed in order to get more reflections.

A.4 Simulation - Prism5

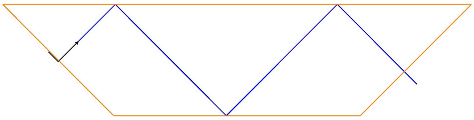


Figure 6.4: Simulation of reflecting rays in the prism with dimensions of Prism5, found in Table 3.1 in Section 3.1. The simulation shows that a prism with these dimensions will achieve 3 reflections within the prism. This prism was not available for the master thesis but a prism with these dimensions were possible to buy from ThorLabs.

B C++ Driver for ADC Card

The code for the original driver can be found under ”Downloads” on the Spectrum Instrumentation’s web page; `https://spectrum-instrumentation.com/en/m2p5946-x4`.

C Experimental Setup: Pictures

Pictures obtained with a digital microscope camera from RND labs of the experimental ATR and fibre-coupled transmission setup. Used in order to define the actual angle and length for the different components in the setup.

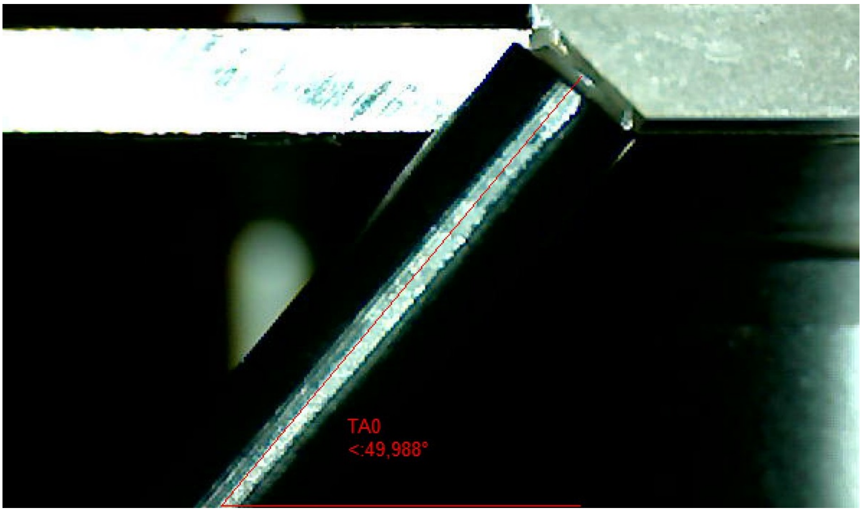


Figure 6.5: The angle of the fibre between the laser and the crystal in the ATR setup was found to be 49.988°.

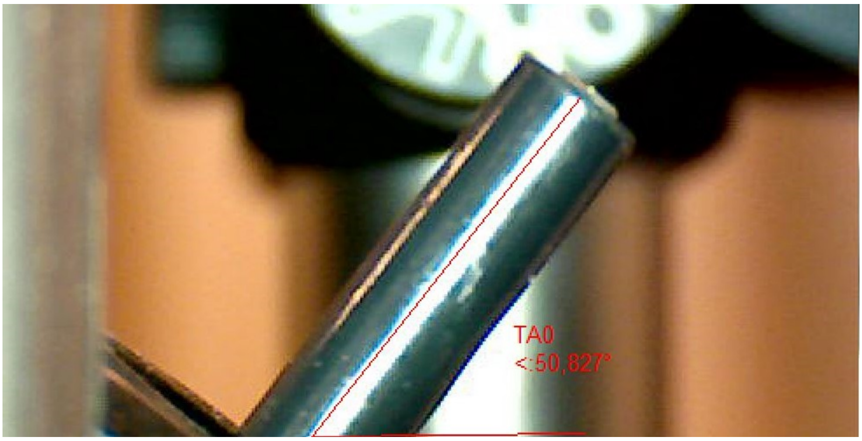


Figure 6.6: The angle of the fibre between the crystal and the detector in the ATR setup was found to be 50.827°.

Figure 6.7: The angle of the lens between the crystal and the fibre leading the light to the detector in the ATR setup was found to be 30.274°..

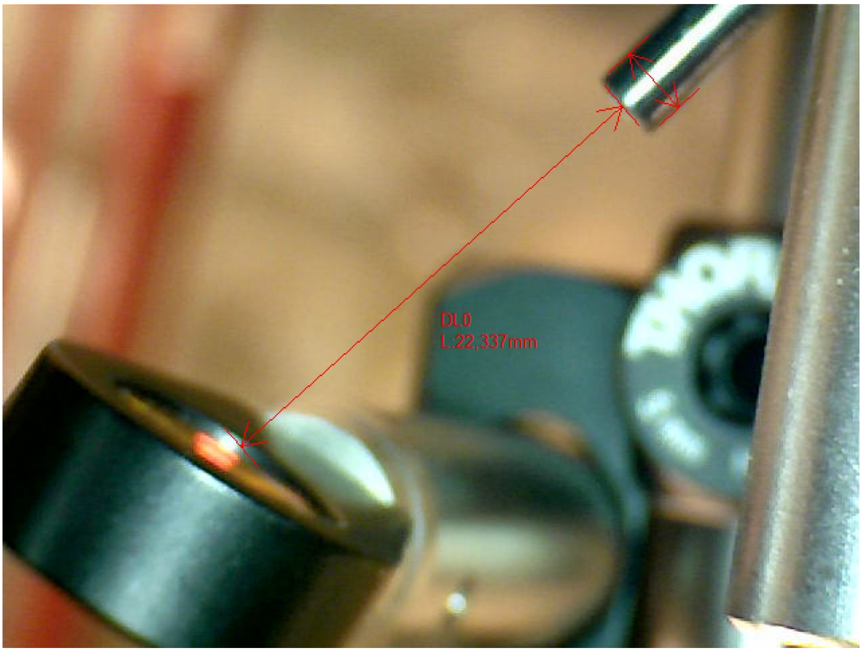


Figure 6.8: The distance between the lens and the fibre leading the light to the detector in the ATR setup was found to be 22.37 mm.

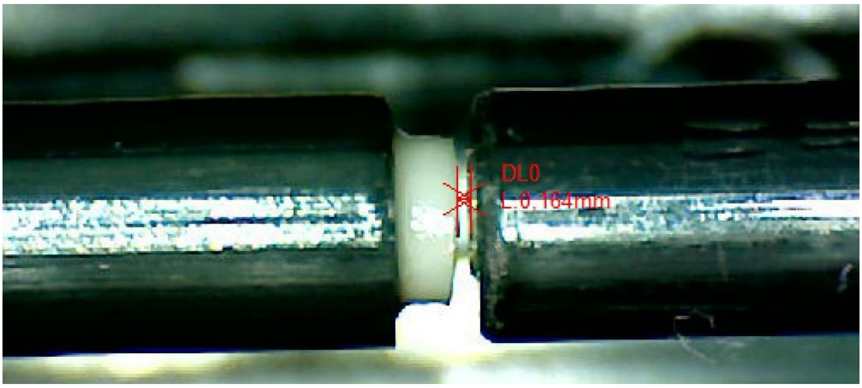


Figure 6.9: The distance between the two fibres in the transmission setup was found to be 164 μm.

D Python Code: Aligning and Averaging 10 Scans Into One Spectrum.

The following Python code average 10 scans over 255 pulses before it aligns the 10 scans to each other. The 10 scans are than averaged into one spectrum before the final spectrum is averaged over 255 pulses again in order to reduce the fringes. The code was described and explained in Section 3.4.

```python
import numpy as np
import matplotlib.pyplot as plt
import scipy.io.wavfile
from scipy.signal import lfilter
from scipy.ndimage import shift
from signal_alignment import phase_align, chisqr_align

################### function for normalization
def normalize(inp):
    inp=np.array(inp)
    factor=1.0/np.sqrt(sum(inp**2))
    return inp*factor

################# function for averaging over n pulses
def averageN(y,n=5):
  nh=int(n/2)
  y2=np.array(y.copy())
  l=len(y)
  y2[int(nh):l-1-nh]=0.0
  for i in range(-nh,nh+1,1):
    #print(int(nh)+i,l-1-nh+i,l)
    y2[int(nh):l-1-nh]+=y[int(nh)+i:l-1-nh+i]/float(n)
  return y2

#################### function for getting the pulses from each scan
def getpulseIntensity(fname):

    data=np.fromfile(fname,dtype=np.int16)
    rate=80000000 #only valid for 100kHz repetition rate

    ####################### get periodicity in data with FFT
    fftfreq=np.fft.fftfreq(24000)[1:]
    fft=np.fft.fft(data[0:24000])[1:]/fftfreq # divide by fftfreq
    period=abs(int(1/fftfreq[np.argmax(abs(fft))]))
    period=800

    ####################### Get first pulse
    somezeroes = np.zeros(100, dtype=np.int16)
    data = np.concatenate((somezeroes, data)) ##adding some zeroes to the
    beginning of data
    maxstart=max(data[0:2000])
    nextpulse=np.argmax(data[0:2000]-np.arange(0,maxstart,maxstart
    /(2000.0-0.5)))
    pulsewindow=50
    correctPlace=data[nextpulse-pulsewindow:nextpulse+pulsewindow]
    corvector=np.arange(-pulsewindow,pulsewindow,1)

    nextpulse+=int(np.sum(correctPlace*corvector)/np.sum(correctPlace))
```

```

  ####################### get max value of each pulse
  numpulses=int(len(data)/period)+100 #make array extra long, make sure it is long enough
  pulseIntensity=np.zeros(numpulses) #np.zeros(len(data)/period)

  i=0
  while nextpulse+20 < len(data) and nextpulse < 80000000:
    #print(nextpulse,i)
    if i%500==0:
      # every xxx pulses: plot pulse
      ax1.plot(corvector/rate*10**9,data[nextpulse-pulsewindow:nextpulse+pulsewindow])
    if i%500==0:
      # every 500 pulses: refine pulse position
      correctPlace=data[nextpulse-pulsewindow:nextpulse+pulsewindow]
      if np.sum(correctPlace) >1000:
        nextpulse+=int(np.sum(correctPlace*corvector)/np.sum(correctPlace))
    pulseIntensity[i]=np.sum(data[nextpulse-pulsewindow:nextpulse+pulsewindow]) # integrate the pulse
    nextpulse+=period
    i+=1
  ax1.set_xlabel(r'Time [ns]')
  ax1.set_ylabel('Intensity')

  ####################### cut off excess length of pulseIntensity
  i=-1
  while pulseIntensity[i]==0:
    i-=1
  numpulses=numpulses+i
  pulseIntensity=pulseIntensity[0:i+1]

  return pulseIntensity


###################### set up figures and subfigures
fig1 = plt.figure(num=None, figsize=(4, 4), dpi=300, facecolor='w', edgecolor='k')
ax1 = fig1.add_subplot(1, 1, 1)
ax1b=ax1.twinx()
fig2 = plt.figure(num=None, figsize=(4, 4), dpi=300, facecolor='w', edgecolor='k')
ax2 = fig2.add_subplot(1, 1, 1)

##################### Read in data from file

StartWL=1200
EndWL=925

filename = 'run2-4/run2-4';
#rate, data = scipy.io.wavfile.read(fname)




```

```python
#########Function for checking the length of each scan and adding zeroes to the end if too short
def checkLength (pulse):
    if len(pulse) == 100001:
        pulse = np.pad(pulse, (0, 1), 'constant')
    if len(pulse) == 100000:
        pulse = np.pad(pulse, (0, 2), 'constant')
    if len(pulse) == 99999:
        pulse = np.pad(pulse, (0, 3), 'constant')
    return pulse

NPTS = 100
SHIFTVAL = 4
NOISE = 1e-2 # can perturb offset retrieval from true

pulses = [None] * 10
s = [None]*9
pulseIntensity = [0]*100002

for i in range(0,10):
    # Get each pulse
    pulses[i] = getpulseIntensity(filename + '_' + str(i) +'.bin')
    # Average over n pulses
    pulses[i] = averageN(pulses[i],255)
    # Checks the length of each pulse
    pulses[i] = checkLength(pulses[i])

    if i > 0 :
        # Finds out how much each spectrum is shifter compared to the first spectrum
        s[i-1] = chisqr_align(pulses[0], pulses[i], [26010, 26590], init=-15, bound=50)
         #Shifts each spectrum with the distance to the first spectrum
        pulses[i] = shift(pulses[i],s[i-1],mode='nearest')

    #Adds all of the spectra together
    pulseIntensity = pulseIntensity + pulses[i]

pulseIntensity = pulseIntensity/10 #Average all of the spectra

############ Adds a zero to the end of the averaged spectrum in order to plot it
pulseIntensity = np.pad(pulseIntensity, (0,2), 'constant')
print('Lengde;',len(pulseIntensity))

x=np.arange(StartWL,EndWL,(EndWL-StartWL)/len(pulseIntensity -1))

####################### average data
averagenumber=255
lenaverage=int(len(pulseIntensity)/averagenumber)
averageint=np.mean(pulseIntensity[0:lenaverage*255].reshape(-1, averagenumber), axis=1) #averageN(pulseIntensity,averagenumber)
if len(pulseIntensity[lenaverage*255:])>0:
  np.concatenate((averageint,[np.mean(pulseIntensity[lenaverage*255:])]))



```

```python
####################### save to file
startwavelength=1200 #cm^-1
endwavelength=924 #cm^-1
wavelengthstep=(endwavelength-startwavelength)/len(averageint)
startwavelength=startwavelength+0.5*wavelengthstep
with open(filename.split('.bin')[0]+'.dpt','w') as f:
  for i, pIn in enumerate(averageint):
    string='{:.8f}'.format(startwavelength+i*wavelengthstep)+'\t'
    string+=str(int(pIn))+'\n'
    #string+=str(int(dichal1[i]))+'\n'
    f.write(string)
####################### plot data

x=np.arange(StartWL,EndWL,(EndWL-StartWL)/len(pulseIntensity -1))
ax2.plot(x,pulseIntensity)
wlstep=(EndWL-StartWL)/len(averageint)
xAverage=np.arange(StartWL+wlstep*0.5,EndWL,wlstep)
ax2.plot(xAverage,averageint)
ax2.set_xlabel(r'Wavenumber [cm^-1]')
ax2.set_ylabel('Intensity')

ylim=ax2.get_ylim()
ax2.set_ylim([0,ylim[1]])
#fig1.subplots_adjust(bottom=0.13,left=0.22, right=0.97, top=0.97)
#fig2.subplots_adjust(bottom=0.13,left=0.22, right=0.97, top=0.97)

fig1.savefig(filename.split('.bin')[0]+'-1.png')
fig2.savefig(filename.split('.bin')[0]+'-2.png')
```

E Python Code: Alignment by Using the Chi-Square Test

The following Python code uses the chi-square test in order to find the difference between two arrays of data. The chi-square test is then minimized in order to find how much the data array should be shifted in order to minimize the chi-square test. The code was described and explained in Section 3.4.

```python
def chisqr_align(reference, target, roi, order=1, init=0.1, bound=1):

    # convert to int to avoid indexing issues
    ROI = slice(int(roi[0]), int(roi[1]), 1)

    # normalize ref within ROI
    reference = reference/np.mean(reference[ROI])

    # define objective function: returns the array to be minimized
    def fcn2min(x):
        shifted = shift(target,x,order=order)
        shifted = shifted/np.mean(shifted[ROI])
        return np.sum( ((reference - shifted)**2 )[ROI] )

    # set up bounds for pos/neg shifts
    minb = min( [(init-bound),(init+bound)] )
    maxb = max( [(init-bound),(init+bound)] )

    # minimize chi-squared between the two signals
```

```
    result = minimize(fcn2min, init, method='L-BFGS-B', bounds=[ (minb, maxb)
    ])

    return result.x[0]
```

F Python Code: Absorbance Spectra

The following Python code takes in the raw spectra from all the measurements within a measurement set. The absorbance spectra are then calculated by using the definition of absorbance A = -log(I/I_0) where I is the sample measurement and I_0 is the reference measurement, which here is a water measurement. The code was described and explained in Section 3.4.

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.signal import lfilter
from scipy.ndimage import shift

laserlabels=[
'glucose 332 mg/dl',
'glucose 435 mg/dl',
'glucose 1.25 g/dl',
'glucose 212 mg/dl',
'glucose 502 mg/dl',
'glucose 802 mg/dl',
'glucose 676 mg/dl',
'glucose 288 mg/dl',
'glucose 50 mg/dl',
'glucose 138 mg/dl',
]
laserfiles=[
'Test2.dpt',
'Test4.dpt',
'Test6.dpt',
'Test8.dpt',
'Test10.dpt',
'Test12.dpt',
'Test14.dpt',
'Test16.dpt',
'Test18.dpt',
'Test20.dpt',
]
waterfiles=[
'Test1.dpt',
'Test3.dpt',
'Test5.dpt',
'Test7.dpt',
'Test9.dpt',
'Test11.dpt',
'Test13.dpt',
'Test15.dpt',
'Test17.dpt',
'Test19.dpt',
]
```

```python

inputfolder='run13-4-DPT/' #Where the spectra is found
outputfolder='run13-4-DPT_sub/' #Where the new spectra will be saved

for i,_ in enumerate(laserfiles):
#iterates through the files and extract the data
    wavenum=[]
    sampleIntensity=[]
    waterIntensity=[]
    with open(inputfolder+laserfiles[i]) as f1:
        for line in f1:
            if not '#' in line:
                wavenum.append(float(line.split()[0]))
                sampleIntensity.append(float(line.split()[1]))
    with open(inputfolder+waterfiles[i]) as f2:
        for line in f2:
            if not '#' in line:
                waterIntensity.append(float(line.split()[1]))
    wavenum=np.array(wavenum)

    absorbance= -np.log ( np.array(sampleIntensity)/np.array(waterIntensity)) #calculating the absorbance spectra

    with open(outputfolder+laserfiles[i].strip('.laser')+'SubtractedWater.laser','w') as f: #saving the absorbance spectra in new file
        f.write('#Wavelength[um]\tpulse intensity normalized\n')
        for j, pIn in enumerate(absorbance):
            string=str(wavenum[j])+'\t'
            string+=str(pIn)+'\n'
            f.write(string)
    plt.plot(wavenum[4:len(absorbance)], absorbance[4:len(absorbance)], label=laserlabels[i])

plt.legend()
plt.xlabel(r'Wavenumber [cm$^{-1}$]')
plt.ylabel('Absorbance')
plt.show()
```

G Concentration Measurements

The concentration measurements obtained for all of the analytes in all of the sample sets in this master thesis. The pure glucose solution is referred to as sample set 1, the pure BHA solution is referred to as sample set 2 and the complex solution is referred to as sample set 3.

Prediction of the glucose concentration in sample set 1.

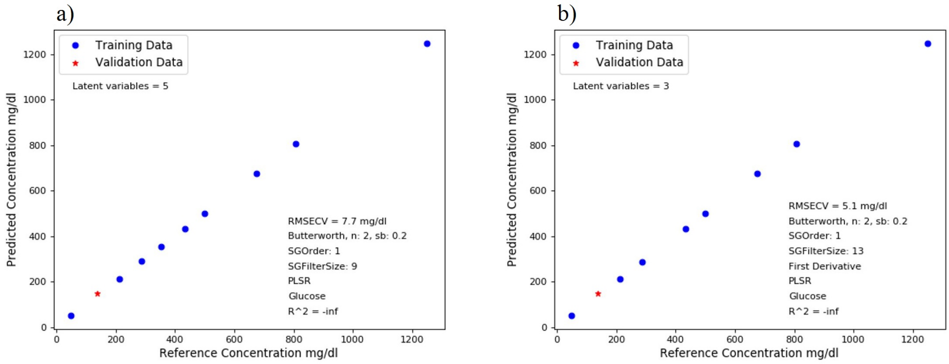


Figure 6.10: Glucose concentrations predicted with PLSR using cross-validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

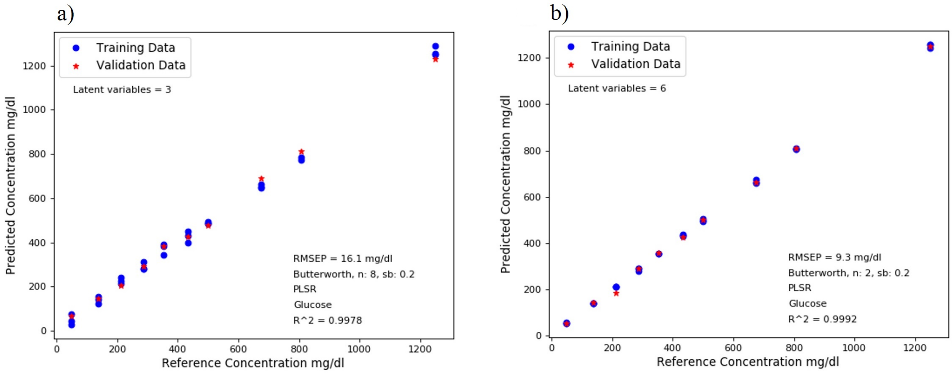


Figure 6.11: Glucose concentrations predicted with PLSR using a separate data set for prediction and 3 data sets for validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

Prediction of the urea concentration in sample set 3.

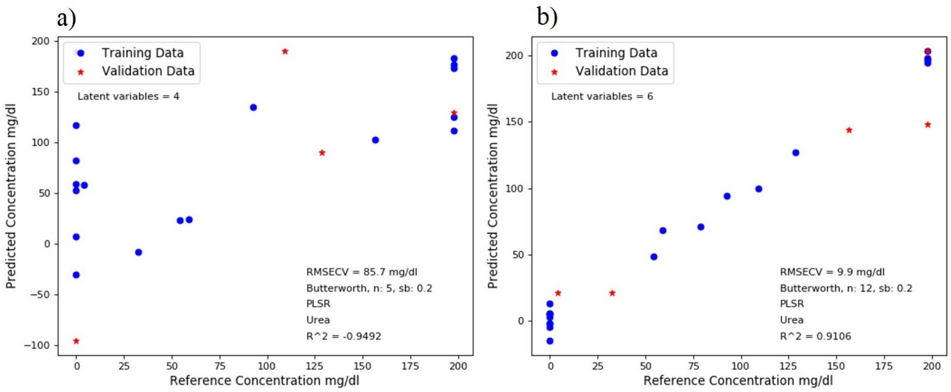


Figure 6.12: Urea concentrations predicted with PLSR using cross-validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

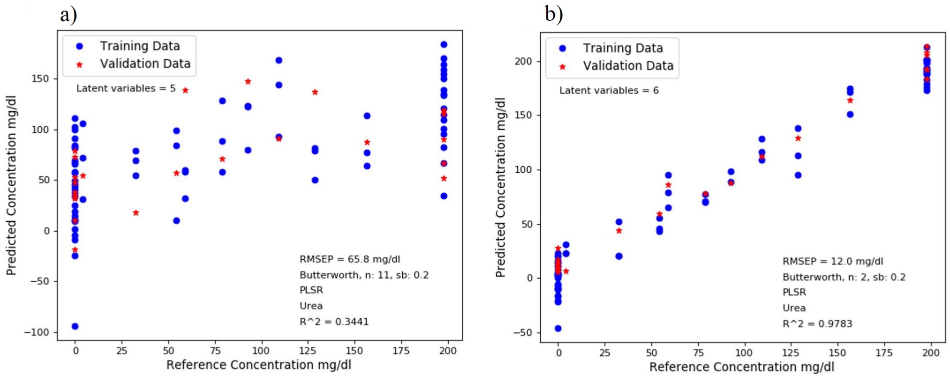


Figure 6.13: Urea concentrations predicted with PLSR using a separate data set for prediction and 3 data sets for validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

Prediction of the lactate concentration in sample set 3.

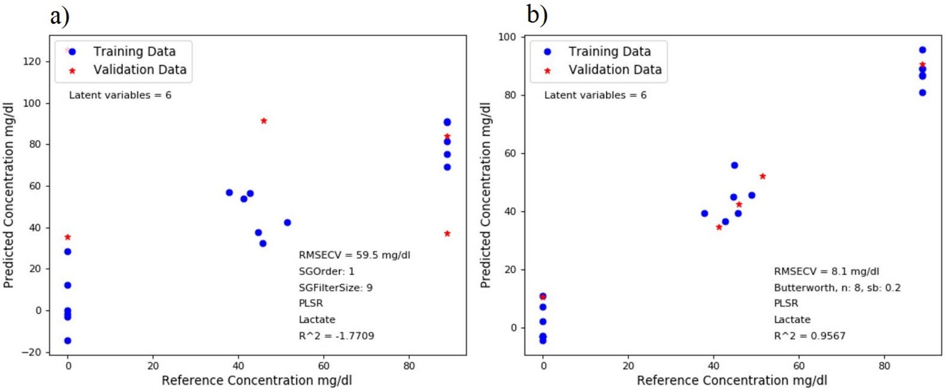


Figure 6.14: Lactate concentrations predicted with PLSR using cross-validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

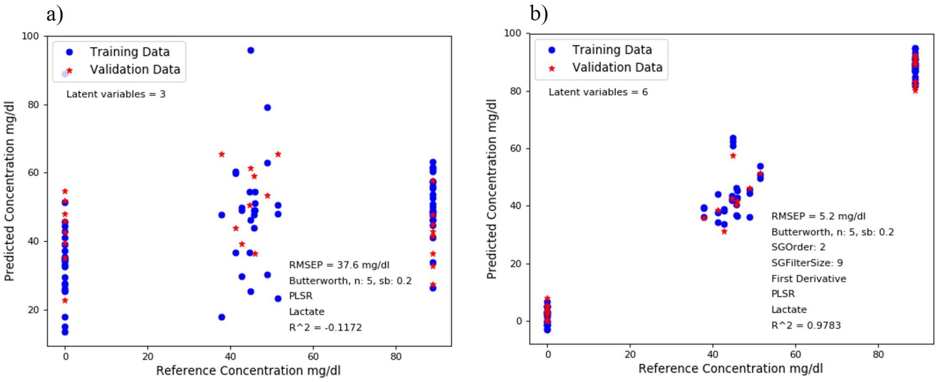


Figure 6.15: Lactate concentrations predicted with PLSR using a separate data set for prediction and 3 data sets for validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

Prediction of the albumin concentration in sample set 3.

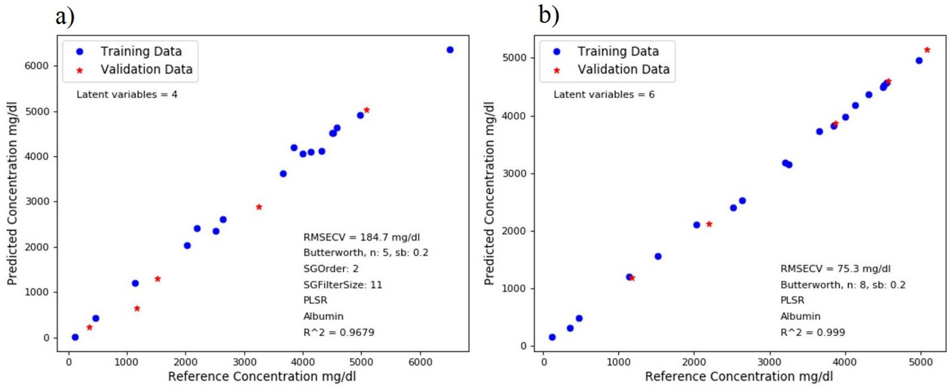


Figure 6.16: Albumin concentrations predicted with PLSR using cross-validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

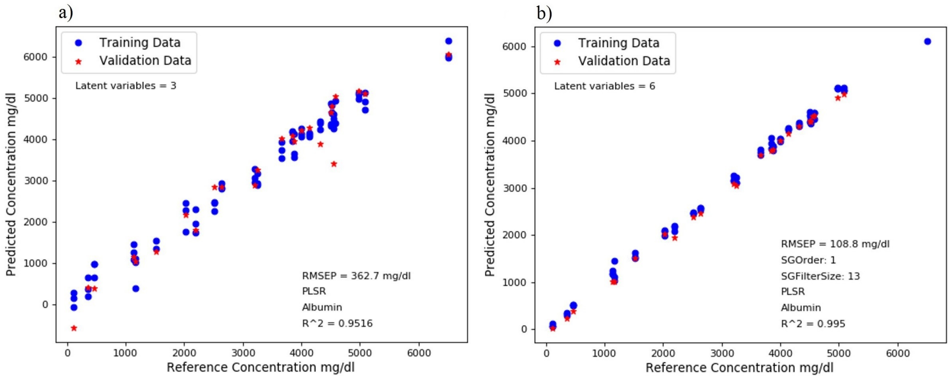


Figure 6.17: Albumin concentrations predicted with PLSR using a separate data set for prediction and 3 data sets for validation, for **a)** the ATR setup and **b)** the fibre-coupled transmission setup.

H FTIR Measurement of Pure Glucose Solutions

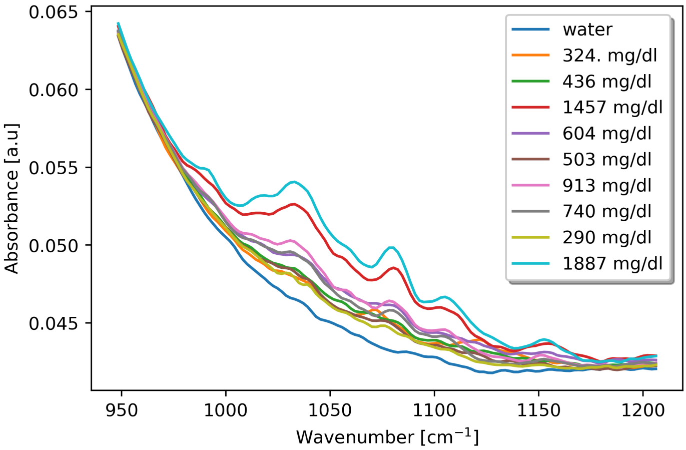


Figure 6.18: Absorbance spectra of pure glucose solutions obtained with an FTIR spectrometer (Alpha II, Bruker).

I Data Sheet: HEDGEHOG Laser

DRS DAYLIGHT SOLUTIONS

HEDGEHOG™

COMPACT, RAPID-SCAN, TUNABLE MID-IR LASER

Molecular spectroscopy applications benefit from rapid, high Signal-to-Noise Ratio data acquisition. This demands fast-scan mid-IR lasers delivering high-quality light. Until now, high tuning speed has come with compromises. The new Hedgehog from DRS Daylight Solutions changes this. For the first time, fast tuning and high-fidelity output is available from a compact, robust mid-IR laser. Hedgehog is built on Daylight's field-proven Quantum Cascade Laser (QCL) technology. Available center wavelengths span the mid-IR spectrum from < 4 µm to >13 µm, and Hedgehog can provide pulsed or CW output[2]. Users can select from three model types (HHG, HHG-UT, or HHG-LT) depending on their application power and tuning range requirements. All models include a GUI option for ultra-quiet CW operation, high wavelength repeatability, and multiple tuning modes.

Hedgehog's small size and rugged design make it ideally suited to either laboratory use or OEM integration. Each Hedgehog is shipped with a compact, easy-to-use SideKick™ multi-function QCL controller. All control functionality is via USB/Ethernet connectivity and an included GUI and SDK command set. Daylight's proprietary HFQD™ (High-Fidelity QCL Drive) circuitry also protects your QCL chip.

With Hedgehog, high-speed, high-quality mid-IR spectroscopic data acquisition is now a reality. Hedgehog brings new capabilities to a wide range of molecular sensing applications including process control, detection of pollutants, chemical and biological agents, time-resolved spectroscopy, and cellular imaging. Please contact us today to learn how Hedgehog, and our highly experienced team, can help your application

HIGHLIGHTS

- Tuning slew rates to > 30,000 cm^{-1}/s
- Ultra-low noise mode (CW RIN as low as -140 dBc/Hz)
- High wavelength accuracy, precision and repeatability
- Available center wavelengths: < 4 µm to >13 µm
- Compact head ideal for OEM integration or lab use
- New Hedgehog-LT: greater utility than DFBs

daylightsolutions.com

FOR SPECTROSCOPY AT SPEED, WITHOUT COMPROMISE

HEDGEHOG SPECIFICATIONS

PERFORMANCE SPECIFICATIONS[1]

MODEL	HHG	HHG-UT	HHG-LT
Tuning Range[2,3]	Up to 200 cm^{-1}	Up to 400 cm^{-1}	30 cm^{-1}
Average Power[2,3]	Up to 500 mW	Up to 500 mW	Up to 150 mW
Peak Power[3]	Up to 1 W	Up to 1 W	Up to 200 mW

PERFORMANCE

Center Wavelength Availability	< 4 µm to > 13 µm
Modes of Operation	Pulsed or CW[2]
Tuning Modes	Set λ, Step & Measure, Continuous Scans
Max. Tuning Speed (Step)	250 ms step-and-settle time to arbitrary λ
Max. Tuning Speed (Scan)	Slew rates to >5000 cm^{-1}/s
Wavelength Accuracy	≤ 1 cm^{-1}
Wavelength Repeatability	To ≤ 0.1 cm^{-1} [3]
Average Power Stability	< 2% (1 hr)
Spatial Mode	TEM_{00} (nominal)
Beam Divergence	< 4 mrad (full angle, $1/e^2$ intensity width)[4]
Beam Pointing Stability	< 1 mrad (beam centroid change)[3]
Spot Size	< 2.5 mm ($1/e^2$ intensity radius)[4]
Polarization	Linear, vertical, >100:1

CW PERFORMANCE[1,2]

Linewidth	≤ 100 MHz (FWHM, over 1s)[5]

PULSED PERFORMANCE

Energy Stability	< 3%, standard deviation
Linewidth	≤ 1 cm^{-1} (FWHM)
Pulse Width[6]	40 to 500 ns, 20-ns increments
Repetition Rate[6]	0.1 kHz to 1 MHz, 0.1-kHz increments
Maximum Duty Cycle[6]	10% (custom up to 30%)

OTHER SPECIFICATIONS

Triggering (Pulsed Operation)	Internal/External, External Pulse Input
Triggering (Scans)	External Wavelength Step, Scan Start
External Control Interfaces[7]	USB 2.0, Ethernet 10/100
Temperature Range (°C)	15 to 35 °C (operating)
Humidity	0—80% RH, non-condensing
Cooling	Passive Air (pulsed, up to 5% duty cycle) Water (CW, or >5% duty cycle pulsed)
Power Requirements	≤ 2A, 90 to 264 VAC, 47 to 63 Hz, single phase (or ≤ 3A, 24 VDC, OEM models)
Dimensions (L x W x H)	Head: 4.2 x 2.6 x 2.1 in. (11 x 6.5 x 5.2 cm)[8] Controller: 7.3 x 5.2 x 1.4 in. (19 x 13 x 4 cm)[9]

---

[1] All specifications are subject to change without notice and defined: at the tuning curve ceiling; after a 3-min warm-up; at the factory-recommended operating current.

[2] Requires CW-capable gain chip—please inquire.

[3] Depends on gain chip. Specifications to be agreed at time of order—please inquire.

[4] Measured at 4 µm; scales with wavelength—please inquire.

[5] With laser tuned for single logitudinal mode operation.

[6] Some chips can support pulses up to 10 µs, rep. rates to 3 MHz, and duty cycles up tp 30% —please inquire.

[7] GUI compatible with Windows© 7, 8.1, and 10. Please inquire for other OS.

[8] Head includes cooling plate for lab use. Head with plate: 5.7 x 2.6 x 2.8in. (14.5 x 6.5 x 7cm).

[9] Daylight Solutions' SideKick™ model SK-1000. Dimensions listed exclude connectors.

INVISIBLE LASER RADIATION
AVOID EXPOSURE TO THE BEAM
CLASS 3B LASER PRODUCT

CE

REV B-2018

LEONARDO DRS

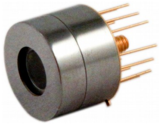
DRS Daylight Solutions
15378 Avenue of Science, Suite 200
San Diego, CA 92128
Tel. +1 858 432 7500

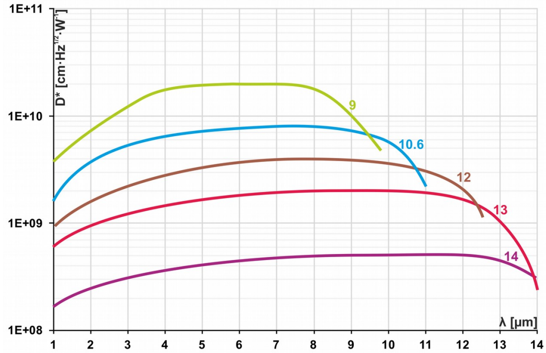
daylightsolutions.com

J Data Sheet: PCI-4TE-10.6 Photodetector

PCI-4TE Series

2 – 14 µm IR PHOTOCONDUCTORS
THERMOELECTRICALLY COOLED
OPTICALLY IMMERSED





Example of D vs Wavelength λ for PCI-4TE Series HgCdTe Detectors. Spectral Characteristics of individual detectors may vary from those shown on the chart.*

Features

- High performance in the 2 to 14 µm spectral range
- Fast response
- Convenient to use
- Wide dynamic range
- Compact, rugged and reliable
- Low cost
- Prompt delivery
- Custom design upon request

Description

The **PCI-4TE-λ_{opt}** photodetectors series (λ_{opt} - optimal wavelength in micrometers) feature IR photoconductive detector on four-stage thermoelectrical cooler, optically immersed to high refractive index GaAs hyperhemispherical (standard) or hemispherical or any intermediate lens (as option) for different acceptance angle and saturation level.

The devices are optimized for the maximum performance at λ_{opt}. Cut-on wavelength is limited by GaAs transmittance (~0.9 µm). Bias is needed to operate photocurrent. Performance at low frequencies (<20 kHz) is reduced due to 1/f noise. Highest performance and stability are achieved by application of variable gap **(HgCd)Te** semiconductor, optimized doping and sophisticated surface processing.

Custom devices with quadrant cells, multielement arrays, different windows, lenses and optical filters are available upon request.

Standard detectors are available in **TO8** packages with **wZnSeAR** windows. Other packages, windows and connectors are also available.

IR Detector Specification @20°C

Parameter	Symbol	Unit	PCI-4TE-9	PCI-4TE-10.6	PCI-4TE-12	PCI-4TE-13	PCI-4TE-14
Optimal Wavelength[*)]	λ_{opt}	µm	9	10.6	12	13	14
Detectivity[**)]: **@ λ_{peak}, 20 kHz** **@ λ_{opt}, 20 kHz**	D*	$\frac{cm\cdot\sqrt{Hz}}{W}$	$\geq 2.0\times10^{10}$ $\geq 1.0\times10^{10}$	$\geq 8.0\times10^{9}$ $\geq 4.0\times10^{9}$	$\geq 4.0\times10^{9}$ $\geq 2.0\times10^{9}$	$\geq 2.0\times10^{9}$ $\geq 1.0\times10^{9}$	$\geq 5.0\times10^{8}$ $\geq 3.0\times10^{8}$
Voltage Responsivity - Width Product @λ_{opt}, 1×1mm	$R_v\cdot w$	$\frac{V\cdot mm}{W}$	≥200	≥40	≥15	≥7	≥5
Time Constant	τ	ns	≤8	≤7		≤6	≤5
Corner Frequency	1/f	kHz	1 to 20				
Bias Current - Width Ratio	$\frac{I_b}{w}$	$\frac{mA}{mm}$	3 to 5				
Sheet Resistance	R_{sq}	Ω	80 to 250	60 to 200			50 to 150
Operating Temperature	T	K	~195				
Acceptance Angle, F/#	Φ, -	deg, -	36, 1.62				

[*)] Other Optimal Wavelengths available upon request.

[**)] Data Sheet states minimum guaranteed D values for each detector model. Higher performance detectors can be provided upon request.*

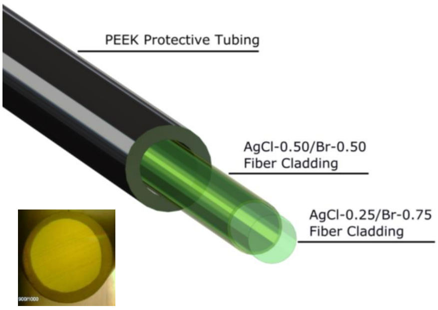
Type	Optical Area [mm×mm]									
	0.025×0.025	0.05×0.05	0.1×0.1	0.2×0.2	0.25×0.25	0.5×0.5	1×1	2×2	3×3	4×4
PCI-4TE-9					X	X	X	X		
PCI-4TE-10.6					X	X	X	X		
PCI-4TE-12					X	X	X	X		
PCI-4TE-13					X	X	X	X		
PCI-4TE-14					X	X	X	X		

X – standard detectors

K Data Sheet: AgCl:AgBr Polycrystaline Fibers

Polycrystalline InfraRed (PIR-) fiber

art photonics developed a volume production technology of the unique product – Polycrystalline InfraRed (PIR-) Core/Clad fibers – transparent over a broad spectral range 3 – 18 µm. The highest performance PIR Core/Clad fibers are extruded with the core diameters span from 240 to 860 µm. Continuously improved extrusion process provides PIR-fibers with a superior optical quality and high mechanical strength.
Low optical losses without absorption peaks over the mentioned spectral range ensure a successful use of PIR-fibers in a broad range of applications.



Applications:

- Mid IR spectroscopy
- Flexible IR pyrometry
- Flexible IR-Imaging systems
- Power delivery for Quantum Cascade Lasers
- Power delivery for CO- and CO_2-Lasers

Features:

- High transmittance in the range 3 – 18 µm*
- Low optical losses of about 0.2 – 0.3 dB/m in the range 9 – 13 µm
- Core/Clad structure with core diameters span from 240 to 860 µm
- Minimal aging effect
- Non-hydroscopic and non-toxic

* See transmission data

Parameters of standard Polycrystalline fibers

Code	Type	Core, µm	Cladding, µm	Protective Jacket, µm	NA	Min. bending Radius, mm
PIR240/300	Step Index Multimode	240±15	300+0/-15	no	0.30±0.03	45
PIR400/500	Step Index Multimode	410±15	500+0/-15	no	0.30±0.03	75
PIR600/700	Step Index Multimode	600±20	700+0/-15	no	0.30±0.03	100
PIR900/1000	Step Index Multimode	860±20	1000+0/-25	no	0.30±0.03	150

art photonics GmbH
Rudower Chaussee 46
12489 Berlin, Germany

+49 (0) 30-6779 887-0
www.artphotonics.com
sales@artphotonics.com

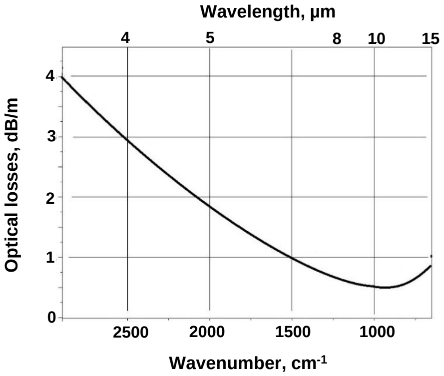
QAS Int.- certified
DIN EN ISO 9001:2015
Certificate No. A1887GER

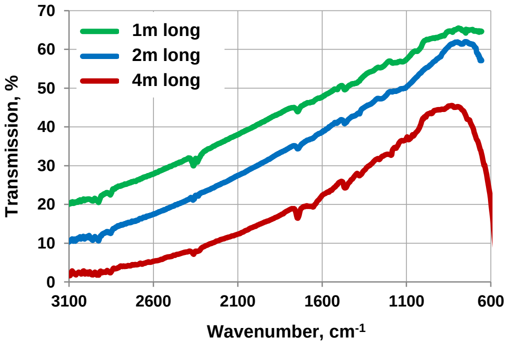
Polycrystalline InfraRed (PIR-) fiber

Specifications

Core/Clad composition	AgCl:AgBr
Spectral Range	3 – 18 µm
Core Refractive Index	2.15
Fresnel Reflection Losses	25%
Attenuation at 10.6 µm	0.2 – 0.4 dB/m
Effective Numerical Aperture NA	0.35 +/- 0.05
Melting Point	410 °C
Operating Temperature	–273°C to +140°C
Core/Clad Diameter (standard)	see table above
Laser Damage Threshold for CW CO_2-Laser	>12 kW/cm²
Tensile Strength	> 70 MPa
Minimum Bending Radius (fixed)	5x [Fiber Diameter]
Minimum Elastic Bending Radius	150x [Fiber Diameter]

Optical Losses and Transmission Spectra of PIR-fiber cables





NTNU
Norwegian University of
Science and Technology