Glucose Monitoring

Measuring Blood Glucose Using Vertical Cavity Surface Emitting Lasers (VCSELs)

by

Sahba Talebi Fard

B.Eng., Carleton University, 2006

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

The Faculty of Graduate Studies

(Biomedical Engineering in Electrical Engineering)

The University Of British Columbia

(Vancouver)

August, 2008

© Sahba Talebi Fard 2008
Abstract

Diabetes Mellitus is a common chronic disease that is an ever-increasing public health issue. Continuous glucose monitoring has been shown to help diabetes mellitus patients stabilize their glucose levels, leading to improved patient health. Hence, a glucose sensor, capable of continuous real-time monitoring, has been a topic of research for three decades. Current methods of glucose monitoring, however, require taking blood samples several times a day, hence patient compliance is an issue. Optical methods are one of the painless and promising methods that can be used for blood glucose predictions. However, having accuracies lower than what is acceptable clinically has been a major concern. To improve on the accuracy of the predictions, the signal-to-noise ratio in the spectrum can be increased, for which the use of thermally tunable vertical cavity surface emitting lasers (VCSELs) as the light source to obtain blood absorption spectra, along with a multivariate technique (Partial Least Square (PLS) techniques) for analysis, is proposed. VCSELs are semiconductor lasers with small dimensions and low power consumption, which makes them suitable for implants. VCSELs provide higher signal-to-noise ratio as they have high power spectral density and operate within a small spectrum. In the current research, experiments were run for the preliminary investigations to demonstrate the feasibility of the proposed technique for glucose monitoring.

This research involves preliminary investigations for developing a novel optical system for accurate measurement of glucose concentration. Experiments in aqueous glucose solutions were designed to demonstrate the feasibility of the proposed technique for glucose monitoring. In addition, multivariate techniques, such as PLS, were customized for various specific purposes of this project and its preliminary investigation. This research
lead to the development of a small, low power, implantable optical sensor for diabetes patients, which will be a major breakthrough in the area of treating diabetes patients, upon successful completion of this research and development of the device.
Table of Contents

Abstract .................................................. ii
Table of Contents ........................................... iv
List of Tables ................................................ viii
List of Figures ............................................... ix
Acknowledgements .......................................... xiv
Dedication .................................................... xv
Glossary ..................................................... xvi

1 Introduction and Background .............................. 1
  1.1 Motivation ............................................. 2
  1.2 History and Current State of Glucose Monitors ...... 3
    1.2.1 Chemical Sensors ................................. 4
    1.2.2 Optical Methods for Glucose Monitoring ........ 6
  1.3 Objective: mission statement of the project .......... 8
    1.3.1 Accomplishments Resulted in Publications ....... 9
  1.4 Aspects and Layout of the Project and Collaborations in the
      Project ................................................. 10
    1.4.1 Collaboration in Biocompatible Coatings ......... 11

2 Anatomy, Physiology, and Pathology ..................... 13
  2.1 Anatomy and Physiology related to Glucose Metabolism .. 14
## Table of Contents

2.2 Pathology related to Glucose metabolism and further complications .......................................................... 18  
2.3 Summary and Conclusion ......................................................................................................................... 22

3 Glucose Monitoring .................................................................................................................................. 23  
3.1 Chemical Based Methods for Glucose Monitoring .................................................................................. 23  
3.2 Optical Methods for Glucose Monitoring .............................................................................................. 23  
3.2.1 Spectroscopy Methods ...................................................................................................................... 24  
3.2.2 Absorption Spectra: Glucose Signature in Absorption Spectra .......................................................... 28  
3.2.3 Broad Source Spectroscopy for Glucose ......................................................................................... 31  
3.2.4 Laser Based Glucose Monitoring ................................................................................................. 32  
3.2.5 Suitability of VCSELs for Implantable Glucose Monitor ................................................................. 33

4 Data Analysis and Chemometrics ............................................................................................................. 35  
4.1 Introduction and Motivation .................................................................................................................... 35  
4.2 Basics of Chemometrics .......................................................................................................................... 37  
4.2.1 Principal Component Analysis (PCA) ............................................................................................ 37  
4.3 Preprocessing ......................................................................................................................................... 39  
4.3.1 Mean-Centering and Centering ...................................................................................................... 41  
4.3.2 Variance Scaling and Autoscaling .................................................................................................. 44  
4.3.3 Multiplicative Scattering Correction (MSC) .................................................................................. 45  
4.3.4 Standard Normal Variate (SNV) .................................................................................................... 47  
4.3.5 Orthogonal Signal Correction (OSC) ............................................................................................ 48  
4.3.6 Square Root Mean Scale (Sqrt Mean Scale) .................................................................................... 50  
4.4 Multivariate Regression .......................................................................................................................... 51  
4.4.1 Classical Least Squares (CLS) ....................................................................................................... 51  
4.4.2 Multiple Linear Regression (MLR) .................................................................................................. 52  
4.4.3 Principal Component Regression (PCR) ....................................................................................... 54  
4.4.4 Partial Least Square Regression (PLSR) ....................................................................................... 55  
4.4.5 Model Quality Measures ............................................................................................................... 56  
4.4.6 Identifying Optimum Number of Factors ........................................................................................ 57
# Table of Contents

4.5 Discussion ........................................ 58
   4.5.1 Choosing Multivariate Regression Method .......... 58
   4.5.2 Finding the Proper Preprocessing Method .......... 60

5 Spectroscopy for Glucose Sensing Using White Light ...... 67
   5.1 Description of Apparatus .......................... 68
   5.2 Determining sub-window ............................ 70
   5.3 Absorption Spectroscopy ............................ 72
   5.4 Results ........................................ 73
   5.5 Discussion ..................................... 74
   5.6 Further Improvements on Data Analysis and PLS ...... 76
      5.6.1 Improved Analysis on the Samples Analyzed in Previous Sections .................................................. 78
      5.6.2 Analysis on the Glucose Samples with Lower Concentrations .................................................. 92
   5.7 Conclusion ...................................... 102

6 Spectroscopy for Glucose Sensing Using VCSELs .......... 104
   6.1 Vertical Cavity Surface Emitting Lasers (VCSELs) .... 105
   6.2 Apparatus ........................................ 109
       6.2.1 PbS Pre-Amplified Detector .................... 109
       6.2.2 Lock-in Amplifier ............................. 110
       6.2.3 2.3 \textmu m VCSEL .......................... 112
       6.2.4 Temperature controlled VCSEL holder .......... 113
       6.2.5 Cuvette ...................................... 113
   6.3 Experiment ...................................... 114
   6.4 Analysis and Results ................................ 115
   6.5 Further Improvements on Data Analysis and PLS ...... 116
       6.5.1 Analysis for Glucose Samples with all Concentrations .................................................. 118
       6.5.2 Analysis for Glucose Samples with lower concentrations .................................................. 124
   6.6 Conclusion ...................................... 130
# Table of Contents

7 Optimization and Performance Improvements  
  7.1 Stability Optimization  
  7.1.1 Stability of VCSEL Output and Power Drift  
  7.1.2 Stability of the Measured Signal  
  7.1.3 Change of LI curve in 2.3 µm VCSEL  
  7.2 Analysis on Sources of Noise and Losses  
  7.2.1 Kinds of Noise  
  7.2.2 Detector’s Figures of Merit  
  7.2.3 Experiments for Noise Investigation in Detector  
  7.2.4 Electrical Attenuator to Improve Signal-to-Noise Ratio  

8 Summary, Conclusion, and Future work  
  8.1 Significance of the Work:  
  8.2 Summary, Conclusion, and Discussion of Improvements  
  8.2.1 Chemometrics  
  8.2.2 White Light Based Spectroscopy  
  8.2.3 VCSEL-Based Spectroscopy  
  8.3 Future work  

Bibliography
List of Tables

3.1 Glucose absorption bands... 30
List of Figures

1.1 Conventional blood glucose measuring procedure. ........ 3
1.2 Medtronic MiniMed Inc. continuous glucose monitoring sen-
        sor with infusion pump. ......................... 5

2.1 Summary of the inputs, outputs, and control factors on the
        glucose concentration in the body as the system. ........ 15
2.2 Regulation of Glucose in the body system. .................. 17
2.3 diagram representation of some of the complications associ-
        ated with diabetes. .................................. 19

3.1 Absorption coefficient for water ................................ 29
3.2 Structure of the glucose molecule. ............................. 29
3.3 Absorption Spectrum of Glucose [25]. .......................... 30

4.1 Graphical representation of Principal Component Analysis. ... 39
4.2 Absorption Spectra using white light. .......................... 40
4.3 Absorption Spectra using white light for lower concentration
        of glucose. ......................................... 41
4.4 Mean Centering ......................................... 42
4.5 Effect of mean centering on absorption spectra for glucose. ... 43
4.6 Effect of mean centering on absorption spectra for glucose. ... 43
4.7 Effect of auto-scaling on absorption spectra for glucose. .... 44
4.8 Effect of auto-scaling on absorption spectra for glucose. .... 45
4.9 Application of MSC preprocessing on absorption spectra for
        glucose. ............................................ 46
List of Figures

4.10 Application of MSC preprocessing on absorption spectra for glucose. .............................................. 46
4.11 Application of SNV preprocessing on absorption spectra for glucose. .............................................. 47
4.12 Application of SNV preprocessing on absorption spectra for glucose. .............................................. 48
4.13 Application of OSC preprocessing on absorption spectra for glucose. .............................................. 49
4.14 Application of OSC preprocessing on absorption spectra for glucose. .............................................. 50
4.15 RMSEC and RMSECV vs. latent variables (LVs) ............................................................... 59
4.16 Absorption spectra excluding water peak. ......................................................................................... 61
4.17 RMSECV vs. LV for determination of the preprocessing method. .................................................... 62
4.18 RMSECV vs. LV for determination of the preprocessing method for lower glucose concentration. .......... 63
4.19 RMSECV vs. LVs for various preprocessing methods using absorption spectra with one VCSEL. .......... 64
4.20 RMSECV vs. LVs for various preprocessing methods using absorption spectra of two VCSELs. .......... 65
4.21 RMSECV vs. LVs for the comparison of preprocessing methods. ......................................................... 66

5.1 A typical single mode optical fiber. ............................................. 69
5.2 The HL-2000 Tungsten Halogen Lamp spectral output [From www.oceanoptics.com]. .................... 69
5.3 Normalized Black Body Radiation ................................................. 70
5.4 Experimental set up with the white light. ......................................... 71
5.5 Glucose Absorption Spectrum ................................................. 72
5.6 Power vs. Concentration at one wavelength. ......................................... 73
5.7 Predicted glucose vs. measured glucose ........................................ 74
5.8 Predicted glucose vs. measured glucose, using the 2nd data set. ......................................................... 75
5.9 RMSECV vs. LV for determination of the preprocessing method. ........................................ 79
5.10 RMSECV vs. LV to compare various preprocessing method. ........................................ 80
5.11 RMSECV for determination of the method. ......................................................... 82
5.12 RMSECV for determination of the number of intervals. ............................... 83
5.13 Glucose predictions using PLS with 3 LVs and mean centering. ........................................ 85
5.14 Glucose predictions using PLS with 5 LVs and mean centering. ........................................ 86
5.15 Glucose predictions using PLS with 2 LVs and OSC along with mean centering. ........................................ 87
5.16 The 4 intervals chosen for glucose prediction. ......................................................... 88
5.17 Glucose predictions using PLS with 4 LVs and OSC along with mean centering. ........................................ 89
5.18 The 7 intervals chosen for glucose prediction. ......................................................... 90
5.19 Glucose predictions using combination of 7 intervals. ........................................ 91
5.20 RMSECV vs. LV for determination of the preprocessing method. ........................................ 93
5.21 RMSECV for determination of the best method. ......................................................... 95
5.22 Glucose predictions using PLS with 5 LVs and mean centering. ........................................ 96
5.23 Glucose predictions using PLS with 3 LVs and OSC along with mean centering. ........................................ 97
5.24 The 4 intervals chosen from lower concentration samples for glucose prediction. ........................................ 98
5.25 Glucose predictions using PLS with 5 LVs and OSC along with mean centering. ........................................ 99
5.26 The 5 intervals chosen for glucose prediction. ......................................................... 100
5.27 Glucose predictions using PLS with 5 LVs and OSC along with mean centering on five intervals. ........................................ 101

6.1 Structure of the VCSEL. ......................................................... 105
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2</td>
<td>LI curve of the VCSEL</td>
</tr>
<tr>
<td>6.3</td>
<td>I-V curve of the VCSEL</td>
</tr>
<tr>
<td>6.4</td>
<td>Effect of changing bias current on the spectrum of the VCSEL</td>
</tr>
<tr>
<td>6.5</td>
<td>Lasing wavelength vs. bias current</td>
</tr>
<tr>
<td>6.6</td>
<td>LI curves of a dying VCSEL</td>
</tr>
<tr>
<td>6.7</td>
<td>Relative response of PbS pre-amplified detector</td>
</tr>
<tr>
<td>6.8</td>
<td>Current dependent optical output spectra of the 2.3 μm LW-VCSEL, and wavelength shift with respect to bias current, at 300K, with the tuning slope of 0.3 nm/mA. The thermal tunability is 0.1 nm/°C [2]</td>
</tr>
<tr>
<td>6.9</td>
<td>Experimental set-up</td>
</tr>
<tr>
<td>6.10</td>
<td>Absorption spectra for concentrations of glucose solutions</td>
</tr>
<tr>
<td>6.11</td>
<td>Glucose prediction using the PLS model, derived from the absorption spectra data</td>
</tr>
<tr>
<td>6.12</td>
<td>Glucose predictions with the mean centered data from a single VCSEL</td>
</tr>
<tr>
<td>6.13</td>
<td>Glucose predictions with the Sqrt mean scaled and mean centered data from a single VCSEL</td>
</tr>
<tr>
<td>6.14</td>
<td>Glucose predictions using the model calibrated with OSC as preprocessing on data from single VCSEL</td>
</tr>
<tr>
<td>6.15</td>
<td>Glucose predictions with the mean centered data from both VCSELs</td>
</tr>
<tr>
<td>6.16</td>
<td>Glucose predictions with the Sqrt mean scaled and mean centered data from both VCSELs</td>
</tr>
<tr>
<td>6.17</td>
<td>Glucose predictions using the model calibrated with OSC as preprocessing on data from single VCSEL</td>
</tr>
<tr>
<td>6.18</td>
<td>Glucose predictions with the mean centered data from a single VCSEL</td>
</tr>
<tr>
<td>6.19</td>
<td>Glucose predictions with the Sqrt mean scaled and mean centered data from a single VCSEL</td>
</tr>
<tr>
<td>6.20</td>
<td>Glucose predictions with the mean centered data from both VCSELs</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.21</td>
<td>Glucose predictions with the Sqrt mean scaled and mean centered data from both VCSELs.</td>
<td>129</td>
</tr>
<tr>
<td>7.1</td>
<td>Plot of output power of VCSEL and temperature of its base vs. time.</td>
<td>134</td>
</tr>
<tr>
<td>7.2</td>
<td>Output power and lasing wavelength of VCSEL vs. temperature.</td>
<td>135</td>
</tr>
<tr>
<td>7.3</td>
<td>Measured power and temperature over time.</td>
<td>136</td>
</tr>
<tr>
<td>7.4</td>
<td>LI curve of a 2.3 µm VCSEL for two consecutive nights.</td>
<td>137</td>
</tr>
<tr>
<td>7.5</td>
<td>LI curve of a 2.3 µm VCSEL for several days in the burn-in process.</td>
<td>138</td>
</tr>
<tr>
<td>7.6</td>
<td>Normalized Detectivity for various detectors.</td>
<td>141</td>
</tr>
<tr>
<td>7.7</td>
<td>Measured noise signal for roll-off of 6 dB/oct vs. various time constants.</td>
<td>142</td>
</tr>
<tr>
<td>8.1</td>
<td>Preliminary sensor design</td>
<td>151</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to acknowledge the support of my supervisors: Dr. Lukas Chrostowski, and Dr. Ezra Kwok.

In addition, I would like to acknowledge NSERC as well as the BC Innovation Council for supporting me in this research.

Furthermore, Werner and Dr. Amann, from Walter Schottky Institute (WSI) in Germany, are acknowledged for providing 2.3 µm VCSELs.

Also, teachers of Eigenvector Research Inc. are acknowledged for the training they provided and their support.
Dedication

I dedicate this work to my dear parents, Sirous and Nahid, my dear brothers, Peyman and Pouria, and my beautiful dear sister, Saghar, and would like to appreciate their great support, both spiritually and materially, in choosing this research project and making a progress. Furthermore, it is dedicated to my grandparents, Laila, Habib and Banoo, who have supported me with their prayers.

In addition, and above all, I would like to dedicate this effort of mine to my real motivators, who taught me that the purpose of my life is to work for the betterment of the world through my service to Humanity; to my dear Bahai family and friends who suffered and are still suffering in Iran while striving to establish unity among all nations and races and the world peace.

My desire in life and the purpose of my education have always been to serve mankind. I hope this project can be a start in development of a device that can serve diabetes and their families to have a higher quality of life.
Glossary

**adsorb**  verb [ trans. ] (of a solid) hold (molecules of a gas or liquid or solute) as a thin film on the outside surface or on internal surfaces within the material.

**Spectroscopy**  : The study of the interactions between electromagnetic energy (light) and matter.

**Spectrum**  : A graphical display to show the amount of the interaction of light with the sample as a function of wavelength, wavenumber, or frequency.
Chapter 1

Introduction and Background

As diabetes mellitus is becoming a more widespread serious disease, a more convenient and accurate way of controlling blood glucose, which improves the patients life quality, effectively minimizes the complications associated with this disease and adds savings for health care systems, is desirable and many thriving companies are attempting to develop such a sensor. Currently, accurate blood glucose monitoring requires pricking of fingers for blood sampling. Enhancing glucose measurement techniques to allow easy and continuous monitoring has received a lot of attention from both academic and industrial researchers over the past three decades. Optical methods are one of the painless and promising methods that can be used for blood glucose predictions. However, having accuracies lower than what is acceptable clinically has been a major concern.

To improve the accuracy of the predictions, several improvements to the current challenges have been considered [1, 2]. Specifically, to improve on the accuracy of the predictions, the signal-to-noise ratio in the spectrum can be increased, for which the use of thermally tunable vertical cavity surface emitting lasers (VCSELs) is proposed. In addition, multivariate analysis is used to enhance the prediction accuracy by the means of data preprocessing and prediction techniques. The long-term goal of this research, including the improvements, is to develop a small and low power implantable glucose sensor that can last for a long time, and can easily be used by current biomedical companies which develop implantable devices. Experiments have been designed, and analyses have been performed to investigate the feasibil-
ity of the proposed techniques for developing such a sensor. This document will present and discuss the results of these investigations.

1.1 Motivation

Diabetes Mellitus is a common chronic disease that is an ever-increasing public health issue. It is characterized by the inability of the body to control blood glucose concentration, resulting in plasma glucose concentration elevated beyond the normoglycemic range (defined as blood glucose 70-100 mg/dL or 4-6 mmol/L). Most of the long-term health problems associated with diabetes, such as nephropathy, retinopathy, neuropathy, cardiovascular diseases and cerebral vascular events, result from the sustained hyperglycemia (blood glucose > 120 mg/dL). Statistics Canada data shows that in 2005, 4.3% of adults was diagnosed with diabetes mellitus. According to the Public Health Agency, there are an estimated 60,000 new cases of diagnosed diabetes every year in Canada. The age of onset of this disease has become lower over the past decade. Diabetes is ranked as the 7th leading cause of death in Canada, and has been a major contributing factor to other leading causes of death such as cardiovascular diseases and cancers.

Diabetes Mellitus is classified into two types: Type I is a form of autoimmune disease because the patient’s pancreas is no longer functioning. In Type II diabetes, the pancreas does not produce enough insulin, or the body does not properly respond to insulin. Although Type I (insulin-dependent) accounts for a small portion of the new cases in Canada, many Type II (insulin-independent) diabetic patients eventually require a high dose of insulin to maintain their poorly controlled blood glucose. Clinically, the key to improved long-term outcome for patients is frequent monitoring of glucose levels for both Type I and II diabetic patients.

Currently, the most common medical treatment for diabetes recommends three to four daily finger-prick glucose measurements, and an equivalent number of subcutaneous insulin injections or continuous insulin infusion. However, monitoring blood glucose and injecting insulin can be troublesome and patient compliance is always an issue. As a result of the predominantly
open-loop nature of the current treatment and poor compliance, insulin delivery cannot realistically be performed frequently enough to avoid blood glucose excursions in time. On the other hand, it has been shown that continuous glucose monitoring (CGM) is beneficial in detecting glucose highs and lows, and can significantly reduce HbA1C levels, as compared to finger stick testing [3, 4]. Reducing chronic hyperglycemia and intensive insulin therapy with continuous glucose monitoring (CGM) has been shown to effectively delay onset and slow the progression of long-term complications such as diabetic retinopathy, nephropathy, and neuropathy (DCCT study). Continuous glucose monitoring has been shown to help diabetes mellitus patients stabilize their glucose levels, leading to improved patient health and quality of life, and tremendous health-care benefits and savings.

1.2 History and Current State of Glucose Monitors

As explained in the previous section, a glucose sensor, capable of continuous real-time monitoring, provides many advantages both to the patients and to the health-care. Hence, the development of such a sensor has been considered a “holy-grail” in clinical diagnosis and treatment of diabetes and has received a lot of attention over the past decades. Current methods of glucose monitoring require taking blood samples several times a day, as shown in Figure 1.1 hence patient compliance is an issue.

This figure has been removed due to copyright restrictions. The information removed contained a pictorial presentation of the conventional finger pricking method for glucose monitoring for diabetes. The figure can be found at the following link: http://www.pennhealth.com/health%5Finfo/diabetes1/000265.html

Figure 1.1: Conventional blood glucose measuring procedure. [5]

Many companies have focused on developing a glucose sensor, capable of continuous real-time monitoring, to be used for clinical diagnosis and treat-
ment of diabetes. Approaches have included chemical sensors, including implantable continuous enzyme-based chemical sensors, and optical methods, including implantable and non-invasive optical sensors.

1.2.1 Chemical Sensors

Continuous monitoring using subcutaneous fluid in the abdomen, in particular, has been shown to be very effective [4] and successfully commercialized by several companies including Medtronic Minimed Inc., and DexCom, Inc. [6, 7]. Implemented using enzyme-based sensors [8, 9], the biosensors are restricted by their limited life spans, and have problems with drift and stability. The disposable and user-insertable sensors have a lifetime of three days, a limitation partly due to risk of infections. Most importantly, the chemical sensors suffer from signal drift and need daily calibration, a limitation due to the chemical sensors interacting directly with the subcutaneous fluid, leading to bio-fouling.

Despite these major challenges associated with surface interaction and bio-fouling, several companies and research groups are pursuing long-term-implantable chemical-based sensors, either subcutaneous (e.g., under abdomen) [10, 11] or intravenous [12, 13]. Continuous monitoring using subcutaneous fluid, in particular, has been shown to be very effective [4] and successfully commercialized by several companies including Medtronic Minimed Inc., and DexCom, Inc. [6, 7]. The Medtronic Minimed intravenous sensor has been shown to be functional for over one year [12], with approximately 1000 patients using the product. Figure 1.2 shows a sample of these CGM sensors with an infusion pump commercialized by Medtronic.

Digital Angel and VeriChip are also pursuing an implantable chemical-based glucose sensor, using RFID-tag technology to wirelessly power and communicate with the implant.

In summary the limitations of the biosensors that are using enzyme-based sensors [8, 9] are their limited life spans, and their problems with drift and stability.
Chapter 1. Introduction and Background

This figure has been removed due to copyright restrictions. The information removed contained a pictorial presentation of the Minilink real-time monitor, transmitter and pump.


Figure 1.2: Part of the prescribed MiniMed Paradigm Brand of Insulin Pumps, the MiniMed Paradigm Insulin Pump is small enough so that it can be worn almost anywhere under the clothing in a leg pouch, thigh pouch, bra pouch, or on the belt like a cell phone. It delivers insulin through an infusion set that has a soft tube called a cannula that sits under the skin for up to three days after being inserted in one virtually painless step. The insulin pump can easily and quickly be disconnected from the patient for activities such as bathing, swimming, or changing clothes. Real-Time Continuous Glucose Monitoring (CGM) is made possible through a tiny glucose sensor, wore for up to three days at a time. Like the cannula, it is easily inserted using an automatic insertion device provided with the system. Glucose sensor data is sent continuously to a MiniLink Real-Time Transmitter, a small lightweight device that attaches to the glucose sensor. The transmitter sends the glucose data to the insulin pump through advanced radio frequency (RF) wireless technology. (The glucose sensor, transmitter, and adhesive patch are all waterproof.) [from http://www.minimed.com].
1.2.2 Optical Methods for Glucose Monitoring

The other highly pursued method for glucose sensing is by optical techniques. Optical sensors avoid chemical biosensor limitations, since the light interacts with a larger volume of liquid, rather than just the surface of the sensor. In addition, optical biosensors are less susceptible to biofouling from blood protein adsorption, and light remains largely unaffected when interacting with the blood.

As such, various spectroscopic measurement techniques have been proposed [14–23], either in transmission or reflection mode [23], with near-IR wavelengths spanning up to 10.5\( \mu m \) [16]. The optical spectrum measurements, resulted from any of the above methods, are analyzed using multivariate techniques to determine glucose concentration.

Optical biosensors can be developed for non-invasive applications or bioimplants.

Non-Invasive Technologies

Some new technologies to monitor blood glucose levels will not require access to blood to read the glucose level, rather they can monitor glucose concentrations non-invasively. Non-invasive technologies include near IR detection, ultrasound and dielectric spectroscopy. These will free the person with diabetes from finger sticks to supply the drop of blood for blood glucose analysis. Most of the non-invasive methods under development are continuous glucose monitoring methods and offer additional advantages.

The ideal solution would be a non-invasive sensor, which involves passing a beam of light through some body parts such as finger, tongue, ear, eye, inner lip or body fluid such as tear; this has generated tremendous interest, with many companies pursuing this approach.

Companies that are currently developing non-invasive glucose sensors are Canadian-based CME Telemetrix, Inlight Solutions, who are involved in developing NIR glucose sensor, Sensys Medical GTS, Sontra developing Sontra ultrasonic Symphony Diabetes management system, Solianis Monitoring AG, VivaScan Corporation, Connecticut-based Infratec Inc., and etc.
In Light Solutions Inc. uses near infrared (NIR) optical spectroscopy and multi-variate analysis to detect glucose concentration by shining light on the skin. Canadian-based CME Telemetrix has developed a near-infrared glucose monitor that shines a beam of light for 30 seconds at patient’s finger and calculates glucose concentration. Connecticut-based Infratec Inc. is developing a glucose monitor for the use on ear. The patient slides the monitor in the ear and inserts the device for 10 seconds. The device calculates the glucose levels by measuring the body heat that is given off by blood in the vessels of the eardrum. The final product may be wearable like a hearing aid [24]. The scientists at the University of Pittsburgh are developing a contact lens that changes color according to glucose concentration in the body. The lens changes color from green (normal) to blue (high) or red (low) as the glucose concentration increases or decreases. A compact mirror with accompanying color wheel comes with the final product so that the patient can see the lens color and match the gradation to the colors in the color wheel.

Non-invasive sensors have challenges, however, due to interference, poor signal strength, and calibration issues, and these approaches are still not accurate enough for clinical use. These main challenges in accurate measurements would be significantly reduced if the optical sensor had close access to either interstitial fluid or preferably the blood plasma, so that the light did not have to interact with the several layers of tissue.

**Optical Bio-implants**

The accuracy of optical sensors is significantly improved if the optical sensor has close access to either interstitial fluid or whole blood [16]. The implantable optical sensor approach has been pursued commercially (e.g. Crothall’s team at Animas Inc.[25]), in which their sensor measures the transmission spectrum through a 3-4 mm blood vessel.

Pennsylvania-based Animas Corporation develops a pacemaker-like implant that will send readings to a wristwatch or a beeper-type monitor using radio waves; alarms will warn when extreme highs and lows are reached.
1.3 Objective: mission statement of the project

As it was mentioned in the previous sections, diabetes mellitus is becoming a more widespread serious disease; and a more convenient and accurate way of controlling blood glucose improves the quality of life for diabetes and adds savings for health care systems. Optical methods are one of the painless and promising methods that can be used for blood glucose predictions. This method addresses the limitations of enzyme-based glucose sensors, such as short life span, which is due to bio-fouling and light is less susceptible to it. However, having accuracies lower than what is acceptable clinically has been a major concern. These techniques suffer from low signal-to-noise ratio (SNR), and the packaging for the implant is not straight forward. To improve on the accuracy of the predictions, the signal-to-noise ratio in the spectrum can be increased, for which the use of thermally tunable vertical cavity surface-emitting lasers (VCSEL) is proposed, along with multivariate techniques such as Partial Least Square (PLS) techniques. VCSELs are semiconductor lasers with small dimensions and low power consumption, which make them suitable for implants. In addition, VCSELs operate within a small spectrum, and the high power spectral density provides a higher signal-to-noise ratio. The goal of this research is to develop small and low power implantable glucose sensor. Many companies (~30) either are currently developing CGMs, or have tried and failed. This project proposes a new method based on VCSELs, and addresses the major challenges encountered by previous approaches.

**Long Term Objectives:** The ultimate goal is to develop a miniature and low power implantable glucose sensor, that is sensitive, selective, stable, durable, and capable of monitoring blood glucose in real-time. The optical sensors will continuously monitor blood glucose levels via interstitial fluid found in subcutaneous tissue or by direct contact with blood. This sensor can also be designed to be placed on another implantable device such as stent, which is being developed by *Evasc Medical Systems* in Canada, or it can replace the chemical-based sensor in the Continuous Glucose Monitor device developed by *Medtronic*; or it may be designed to be injected or
implanted in the body and be minimally invasive due to the small dimension of the sensor. The sensor can send the data to an external device (which can be worn like a watch). This external device can then analyze the data and communicate with intelligent electronics that will deliver insulin accordingly.

Although several groups have demonstrated predicting glucose concentrations optically, it is uncertain whether it is possible to get accurate enough results from a small number of lasers, allowing for a small sensor that is amenable for implantable devices. For the first time, small-enough semiconductor laser sources are available at the ideal wavelength for optical glucose sensing, and this research project may contribute to solving this problem. To ensure biocompatibility, biocompatible coatings will be developed, and applied to the device package.

The current debate in implantable CGM approaches is between: 1) whole blood-based sensor, which requires minimal surgery, or 2) subcutaneous fluid-based sensor, which suffers from a \( \sim 20 \) minute time lag, but patient insert-able. Although our chip could in principle be used for both types (being small enough to be potentially mounted inside a needle for subcutaneous implantation), in our work we will be focusing on the more accurate approach of a whole blood-based implant.

### 1.3.1 Accomplishments Resulted in Publications

To achieve the purpose and long-term objective of this implantable real-time sensor, some preliminary investigations were performed to prove the feasibility of the proposed techniques.

First, white light absorption spectroscopy was used to verify the possibility of using segmented portions of wavelength to predict glucose concentration, resulted in Conference Proceedings [1]. The wavelength range in near-IR was investigated to determine the minimum number of wavelength segments required to predict glucose concentration. The selection process of wavelength segments was according to the absorption characteristics of glucose. The width of these segments was 7 nm wide for the purpose of using VCSELs instead of white light to improve signal-to-noise ratio. An
optical spectrum analyzer (OSA) was used to generate absorption spectra from the white light after passing through a 1 mm path-length in solution. A customized software program, after reading the spectra from OSA, selects small windows of these spectra. Partial Least Square (PLS) analysis, capable of making predictions from large data sample, which was included in PLS Toolbox of Matlab, was customized to analyze these selected wavelength windows in order to predict the concentration of glucose.

Having proven the possibility of predicting glucose concentration using sub-windows, we proceeded on to use lasers as the light source for absorption spectroscopy [2]. Hence, VCSELs at specific wavelengths, around 2.3 µm were used to increase the signal-to-noise ratio. Other improvements to the model include reducing the effects of flicker noise (1/f noise) to the absorption spectra by modulating the signal using a mechanical chopper. Flicker noise occurs in almost all the electronic devices and has a 1/f spectrum, meaning that it is lower at higher frequencies. The light beam, after passing through the chopper and the solution with a constant path length of 2 mm, reaches the PbS amplified detector. Lock-in amplifier reads the data from the detector and demodulates it to DC value, which was read by PC. Having access to the VCSELs at the 2.3 µm range is a unique opportunity for our research group since they are provided through collaboration with a group in Germany, that is the first and only group making them at this time.

In summary, with in-vitro studies, we established a strong correlation between optical measurements and actual glucose concentrations. Having verified the feasibility of proposed methods the path for future work to develop an implantable device opens.

1.4 Aspects and Layout of the Project and Collaborations in the Project

Development of this implantable glucose sensor includes the following aspects and steps.
• Background knowledge and investigation required to partake in such a project includes the following: Anatomy related to diabetes, as the final product would more or less mimic the body system. In addition, the anatomy and pathology is required to define specifications of the sensor. Reasonable knowledge about optics, sensors, biophotonics, instrumentation and electronics as well as data processing basics were part of the essentials.

• Spectroscopy Measurements: the next step was to get familiar with the various spectroscopic methods. Designing experiments to provide proof to the considered techniques according to the spectroscopic method under investigation required thoughts, planning, reflection, and modifications.

• Data analysis and Chemometrics: due to the special nature of the spectroscopic data, glucose absorption spectra, and interfering components and signals, multivariate signal analysis is an essential requirement in analyzing the data and building the model to predict glucose concentration from absorption spectra. In addition, preprocessing of the data is an important step for the nature of these data sets.

• Design of implantable sensor: the next step after preliminary investigation for feasibility of proposed techniques would include design of the actual sensor

• Biocompatibility of implantable device: a blood compatible coating for the proposed glucose sensor is in the process of development by a group in blood research center at UBC with the supervision of Dr. J. Kizhkkedathu, parallel to the development of the actual sensor.

1.4.1 Collaboration in Biocompatible Coatings

Any external device implanted in the body requires special coating in order not to be rejected by body or to prevent infections. To ensure the proper operation of a medical device, the challenges of the interaction of an im-
plantable device (e.g., optical glucose monitor) with biological environment need to be considered.

The biocompatibility aspect of this project is fulfilled through a collaboration with blood research center at UBC. Dr. J. Kizhkkedathu’s research group has had considerable success in developing biocompatible coatings.

Organization of the Thesis  In this Master thesis project, background investigations were performed, proposed techniques were defined in more details, experiments were designed to provide proof to the feasibility of the proposed techniques, data analysis was performed, and specific preprocessing methods were investigated to investigate the best method for predicting glucose concentration from its absorption spectra.

This thesis will include following chapters: Anatomy, Physiology, and Pathology related to Diabetes; Glucose Monitoring; Data Analysis and Chemometrics; Spectroscopy for Glucose Sensing Using White Light; Spectroscopy for Glucose Sensing Using VCSELS; Optimization and Performance Improvements; and Conclusion.
Chapter 2

Anatomy, Physiology, and Pathology Related to Diabetes and Glucose Monitor

This chapter will consider review of anatomy, physiology, and pathology related to diabetes patients who would require a glucose-monitoring device. The anatomy, physiology, and pathology related to the glucose monitor is extensive. The metabolism and function of glucose in the body is of significant importance since glucose serves as the major fuel for the brain function in producing ATP without which brain death may occur. As a result, it is important to maintain a reasonable concentration of glucose in body. There are many various factors that may affect the concentration of glucose; this makes the control of its concentration, using outside forces such as diet and exercise, difficult. Hence, the proposed method for the maintenance of the glucose concentration in patients, who do not have the ability to do so such as diabetes, mimics the technique that the body in healthy human uses.

The focus of this chapter is on the anatomy, physiology, and pathology related to diabetes. Therefore the topics included in this chapter are about metabolism of glucose in the body and the variables involved that affect the concentration of glucose in the body as a system, and the pathology of the patient with an inability to control glucose levels in blood and complications resulting wherefrom.

In this chapter, first the anatomy and physiology related to glucose cir-
culation in the body will be reviewed and then some of the pathology due to abnormal glucose metabolism and their complications will be discussed.

2.1 Anatomy and Physiology related to Glucose Metabolism

Glucose serves as the main fuel for cells to generate energy. Glucose may be considered as the main fuel which some of the organs, such as brain, rely on completely for their metabolism and operation. As a result, the tight control of the glucose level in blood, which is the provider to all cells, is vital. Blood circulation serves as the best averaging mechanism since the circulation mechanism circulates the whole blood at least once per minute in the whole body. Hence, the concentration of glucose in blood can be a good representation of its concentration in the whole body. In healthy people, absorption, creation, uptake, and utilization are under control such that the level of glucose concentration in blood stays within a certain limit to ensure that cells have the amount of glucose required for their proper functionality [26]. There are various determinants for the level of blood glucose in body.

Considering glucose as one of the constituents in body pool, the diagram in Figure 2.1 shows the input and outputs from this system of body pool.

The inputs to the concentration of the glucose in the body pool are as follows. Glucose enters human body with the food that is consumed. Glucose is extracted from food by the breakdown of dietary carbohydrates in the gut. In addition, glucose can be released from the glycogen stored in liver and muscles. Furthermore, Glucose is also made from protein in the liver to be circulated in blood and used by cells.

There are various outputs from this system of glucose concentration (Figure 2.1) that affects glucose concentration in body plasma. Blood glucose levels are normally tightly regulated, and the entry of blood glucose into the blood is balanced by the uptake of glucose into peripheral tissues [27]. Hence, one major output for glucose from blood is that glucose is used as fuel for cells, since it is the source of cells energy. In addition, Glucose in blood
can be stored as glycogen in liver and muscles; this is one of the mechanisms that prevent hyperglycemia, which results from high glucose concentration in blood. Furthermore, other outputs for glucose are large structural and functional organic or biochemical molecules that are composed of carbohydrate and may contain glucose. Examples of structural molecules include cell membrane and glycoprotein, and functional molecules include hormones that are glycoprotein.

Blood serves as the circulating and averaging mechanism to keep a uniform level of glucose in the body plasma. Providing the most efficient circulation of nutrition in the body, blood has the task of delivery and transportation of glucose to the cells, where it is used to generate the energy that cell needs to live and function. Hence, it can be well noted that not having well-controlled glucose concentration in blood easily affects all the organs and body systems and can bring about long-term complications in several major ones. For example, glucose plays a major role as a fuel for brain function, and lack of glucose can cause coma or brain death. In addition,
any damage to blood vessels would cause inefficiency in delivery of nutrition including glucose to the cells as well as blood pressure and heart diseases.

Pancreas has the task of controlling the level of glucose concentration in blood, which is a good representative of the concentration in the body fluid with an efficient averaging mechanism. Endocrine pancreas is the primary regulatory organ for fuel metabolism, namely glucose metabolism. Endocrine pancreas uses two key hormones for this purpose: insulin, and glucagon. The task of insulin, in glucose regulation, is increasing glycogen synthesis by enforcing storage of glucose in liver and muscles cells in the form of glycogen [28]. As a result, insulin’s regulatory task is lowering glucose concentration in blood and protecting body against hyperglycemia. In addition, glucose is needed for cells to uptake and burn glucose. Glucagon has a counter-regulatory function to insulin; in other words, it opposes the action of insulin to protect body against hypoglycemia. There are no backup hormones for insulin, but there are four other hormones that show similar catabolic functions as glucagons, namely, epinephrine, norepinephrine, cortisol, and growth hormone [10]. Insulin, however, has other metabolic actions: increased fatty acid synthesis by forcing fat cells to take in blood lipids to convert to triglycerides, increased esterification of fatty acids, decreased proteolysis by reducing protein degradation, decreased lipolysis, and decreased gluconeogenesis by decreasing production of glucose from non-sugar substrate. Hence, glucose may be produced from assorted substrates in the liver due to lack of insulin [28]. Pancreas has both endocrine part and exocrine glands. Insulin is synthesized within the beta cells of the islets of Langerhans in the pancreas. Millions of islets of Langerhans(pancreatic islets) form the endocrine part of the pancreas, which constitutes only 2 % of the total mass of pancreas.

In the case of an acute elevation of blood glucose concentration, the secretion of insulin happens in a biphasic manner. The first phase lasts for about 10 minutes, where the insulin release in this phase is linked to insulin granules located near the beta cell membrane or rapidly releasable pool. Then gradually increasing second phase depends on increasing synthesis of insulin and mobilizing insulin granules from storage to rapidly releasable
Hypoglycemia can easily cause death or permanent brain injuries since plasma glucose concentration falls to levels too low to sustain function in Central Nervous System (CNS) [30]. There are two changes required to maintain plasma glucose concentration within normal range and prevent hypoglycemia. One is that liver will act as an organ for glucose production. Initially, liver produces glucose by breaking down glycogen, and then it continues the production through gluconeogenesis, a process in which amino acids, lactate, pyruvate, and glycerol are converted to glucose by linked series of enzymatic reaction. Second change required to maintain blood glucose concentration is conversion of tissue other than CNS to a lipid. One fourth or one third of fatty acids taken up by the liver are converted to ketone bodies, which can be oxidized by CNS, and can be considered as back-up substrates for glucose [30]. Figure 2.2 shows the regulation of glucose in the body systems.

This figure has been removed due to copyright restrictions. The information removed shows the flow of fuel during feeding.

The figure can be found at the following link.
http://www.isletmedical.com/pages/define_diabphys.htm

Figure 2.2: Regulation of Glucose in the body system: glucose extracted from the food taken up, and transported to liver through blood circulation and the excess is stored as glycogen. Then the storage of glucose is used as the source of energy for brain, muscle, and tissue. Glucose is also stored in muscle as glycogen after being transported there. Picture reference: [26]

As it can be concluded, the level of glucose concentration in blood has various determinants. This makes external control of glucose concentration a complicated process and almost impossible to achieve with just controlling inputs and outputs. As a result, the reasonable approach to a tight control of glucose levels in body, for patients who are not capable of achieving this naturally, is to mimic the way that body controls glucose levels. This conclusion leads the researchers to investigate methods to monitor the glucose
levels continuously and send the results for directions on appropriate insulin injection.

2.2 Pathology related to Glucose metabolism and further complications

This section will include discussion about the pathology of diabetes. Diabetes mellitus is a group of metabolic diseases that is characterized by inability of pancreas to control glucose concentration in blood. In diabetes patients, body system cannot maintain the level of glucose concentration in body and blood. The chronic hyperglycemia of diabetes can result in long-term complications, such as damage, dysfunction, and failure of various organs, especially eyes, kidneys, nerves, heart and blood vessels [31].

There are two main types of diabetes mellitus. Type I diabetes mellitus, which is characterized by destruction of pancreatic beta cells, which produce insulin. As a result, type I diabetes patients have insulin deficiency, and are prone to ketosis. Type II diabetes mellitus patients are characterized by variable degree of insulin resistance in target tissue and insulin deficiency; however, they are not ketosis prone [32].

Other types of diabetes are: genetic defects of beta cell function, genetic defects of beta cell action, secondary to disease of exocrine pancreas, secondary to other endocrine diseases, and secondary to drug or toxin. Another common type is gestational diabetes that includes any form of glucose intolerance, which is first diagnosed during pregnancy [32].

Diabetes is associated with many various complications, which may be divided into two groups: chronic and acute metabolic. The diagram in Figure 2.3 summarizes these complications.

Acute complications can be caused due to severely high or abnormally low blood sugar concentration [33]. Acute metabolic complications that are characterized by hyperglycemia are divided into two groups: with and without acidosis, with coma as their common presenting feature. The two main complications associated with hyperglycemia are Diabetic KetoAcido-
Chapter 2. Anatomy, Physiology, and Pathology...

Figure 2.3: diagram representation of some of the complications associated with diabetes.

sis (DKA), and hyperosmolar non-ketotic syndrome (HONK). DKA happens due to insulin deficiency, which associates with increase in circulating levels of counter-regulatory hormones in type I diabetes. HONK, which appears exclusively in type II diabetes, has many common features with DKA without ketosis and acidosis [27].

Chronic complications may be divided into Macrovascular and Microvascular diseases. Macrovascular diseases are the ones involving larger vessels such as heart and blood vessels. Examples of macrovascular diseases include ischaemic heart disease (IHD), coronary heart disease (CHD) or coronary Artery disease (CAD), cerebrovascular disease (CVD), and peripheral vascular disease (PVD). Microvascular disease is the disease of finer blood vessel in the body such as neuropathy that can lead to loss of sensation, nephropathy, and retinopathy.

Damages to blood vessels is one of the major long term health complications associated with diabetes; since high levels of glucose concentration in blood causes damage to blood vessels [34]. Although the pathophysiology of
diabetes complications is not fully understood, it is known that hypertension has an important role in the development of microvascular and macrovascular complications. Increased capillary pressure may be one reason for the damages on the endothelium that causes the extravasation of proteins. Molecules such as collagen, once glycated, undergo series of chemical modifications to produce advanced glycation end products (AGEs), which in turn causes cascade of events which damages the endothelium and aggravates the effects of increased capillary pressure. Hyperglycemia causes AGE formation, which in turn affect vital functions such as elasticity of collagen in blood vessels and vessels loose their autoregulation, which causes microvascular hypertension and injury [27, 33]. Diabetes accelerates hardening of arteries and larger blood vessels leading to coronary heart disease [33]. Coronary circulation may be affected by Microvascular damages of coronary artery narrowing, which then results in diabetic cardiomyopathy.

Another long-term complication associated with high glucose levels in diabetes is heart disease and stroke. Poorly controlled blood glucose level increases the probability of atherosclerosis, which is furring up and narrowing of the blood vessels. This may cause angina due to the poor blood supply to the heart. In addition, there is a chance that blood vessels in brain or heart may be blocked completely causing heart attack or stroke [35].

Retinopathy, damage to the retina at the back of the eye, is another common long-term complication. Retinal arteries are end-arteries; hence to protect the retinal microcirculation against damages due to blood pressure fluctuations, it is important to have autoregulation and elasticity in these arteries. Pericytes, lining the endothelium, control the diameter of capillaries and proliferation of the endothelium, and provide blood-retinal barrier. Pericytes are lost in diabetes and hyperglycemia results in hyperperfusion and endothelial damage [8]. Hence, these small blood vessels in the retina that are damaged, hardened, or blocked, secondary to diabetes, may cause leakage of protein and blood in retina. Furthermore, blood vessels may leak, grow haphazardly or new vessels may grow on the retina, and get in the way of the light passing through to the retina, which can damage the vision if it is not treated [29, 33, 35].
Chapter 2. Anatomy, Physiology, and Pathology...

Diabetic nephropathy has become the major cause of end-stage renal disease. Kidney disease and its inefficiency may result from the fact that diseased and damaged small blood vessels may get blocked or become leaky to the point that kidneys lose their ability to cleanse and filter blood properly. Abnormally high glucose concentration in blood will result in extraction of glucose from body through urine [29, 33, 35]. Type I diabetes who have developed nephropathy are in the risk of macrovascular diseases since the same factors that cause endothelial damage within the glomerulus may cause damage to the vascular endothelium in larger arteries. Diabetes nephropathy characterized by thickening of glomerular basement membrane resulting in increased pore size in glomerular basement membrane and increased glomerular capillary pressure. In addition, hypertension, aggravated by salt and water retention in patients, increases capillary pressure even more, resulting in accelerated glomerular damage [27].

Complications regarding Nervous System (NS) include both Central Nervous System (CNS) and Peripheral Nervous System (PNS). A series of cascade events due to high concentration of glucose cause alteration in membrane potential and slowing in the nerve conduction velocity, due to a rise in sensory perception threshold. The symptoms of this complication include loss of sensation, and pain perception [27]. In addition, low concentration of glucose in blood means that NS does not get the required fuel to produce energy (or ATP) and cannot maintain resting potential or even generate Action Potential; as a result, damage to NS or eventually brain death may occur. Furthermore, damages to nerves may result in further complications such as foot problems and impotence in men. Foot problems may appear as the result of damage to nerves of the foot, which means that small cuts may not be noticed and may lead to the development of a foot ulcer [35].

There are many other complications that have shown strong correlation with diabetes disease such as bone problems. However, the length and scope of this thesis does not allow explanation of complications in detail.

Tight control of blood glucose levels reduces the risk of developing these complications greatly. As a result, the quality of life for diabetes improves significantly with an adequate, continuous blood glucose monitor, which can
lead to a tight control of glucose concentration in body.

2.3 Summary and Conclusion

In summary, it can be realized that glucose circulation and metabolism in body relies on various body systems, organs and hormones. It was also discussed that the concentration of glucose in body plasma has many determinants and hence, it may be very complicated and essentially impossible to control the concentration of glucose by its input and output (for example, food and exercise) solely. Body normally uses regulatory organs such as pancreas to control and regulate glucose levels. As a result, for patients with inability to control the level of glucose, the best treatment would be artificially controlling the glucose concentration by mimicking the regulatory organs as necessary.

Diabetes mellitus is one of the most widespread diseases, which is characterized by inability of the body to control the level of glucose concentration in blood. This wild swing in the levels of glucose in blood brings about various long-term complications for diabetes patients. Tight control of glucose concentrations will reduce these complications significantly. As a result, it is evident that a continuous glucose sensor leading to a tight control of glucose levels with normal variations over time, would greatly improve the quality of life for diabetes patients.
Chapter 3

Glucose Monitoring

This Chapter includes an overview of the glucose monitoring methods such as chemical-based glucose monitoring, and optical glucose monitoring techniques including lasers.

3.1 Chemical Based Methods for Glucose Monitoring

Used as early as 1962 [8], enzymatic sensors are used clinically for glucose sensing, and the biochemistry is still a topic of research today [9]. Current reliable chemical-based biosensor devices are restricted by their limited life spans. There are several reasons that explain the limited life span of these biosensors such as the degradation of the glucose sensitive enzymes, chemical interference, as well as biofouling [36], which is the strong tendency for protein, collagen, fibrin and organisms to physically adsorb to synthetic surfaces. These limitations arise mainly due to the chemical sensors needing to interact with the blood directly. Despite these limitations, several companies and research groups are pursuing implantable chemical-based sensors, either subcutaneous (e.g., under abdomen) [10–12] or intravenous [12, 13], working towards an implantable artificial beta-cell. However, these commercially available sensors have a limit of up to 3 days use.

3.2 Optical Methods for Glucose Monitoring

Optical sensors avoid chemical biosensor limitations because such sensors do not have to be in direct contact with the glucose, and light remains largely
unaffected when interacting with the blood. In addition, optical sensors
avoid chemical biosensor limitations, since the light interacts with a larger
volume of liquid, rather than just the surface of the sensor. This section
will review various optical methods and will state the origin and nature of
glucose spectra.

3.2.1 Spectroscopy Methods

Spectroscopy is the study of the structure and dynamics of molecules via the
interaction of light with matter. The different parts of the light spectrum
have different effects on molecules and atoms. Radio and microwave radia-
tions excite molecular rotations; infrared light excites molecular vibrations,
causing us to feel heat; and visible/ultraviolet light excites electrons [37].
Light at one wavelength region might get absorbed, scattered, reflected, or
cause emission at other wavelengths depending on the matter. These inter-
actions of light with matter are used to obtain chemical information about
the samples that light is interacting with. This chemical information can
be qualitative to identify the constituents of a sample or quantitative to
determine the concentration of constituents in the sample.

There are several spectroscopic methods that are frequently used for bi-
ological applications. Some of these methods are based on the electronic
transition of electrons between two electronic states, when coupling with
vibrations, they are called vibronic transitions which involves simultane-
ous changes of electronic and vibrational state. Absorption, which is the
transition of electron from a lower (usually ground state) to an excited en-
ergy state in the absorber molecule, and emission, which is the transition
of electron from a higher to a lower state, are two possibilities of electronic
transitions. There are other methods based on vibrational transitions such
as IR-absorption and Raman scattering spectroscopy [38]. The bonds in a
molecule can be modeled with a spring having attached weights, where the
whole system can vibrate [39].
Electronic Absorption/Transmission Spectroscopy: Electronic absorption is usually used for quantitative analysis of the sample. The basic absorption process is based on linear absorption of light. The photon picture of the linear absorption process is when a molecule absorbs a single linear photon to excite an electron from a lower (ground) level to an excited level. To measure the concentration of an absorbing compound in a sample, Beer-Lambert Law needs to be applied. The Beer’s law or law of absorption states that the fraction of absorbed light in a particular material is directly proportional to the thickness of the sample as well as to the concentration of the constituent in that sample that is absorbing the radiation. This proportionality constant is called absorption coefficient, $\varepsilon(\lambda)$, or extinction coefficient. The intensity of light attenuates exponentially with concentration and the path length that it travels through the material at each specific wavelength (equation 3.1).

$$I = I_0 e^{-\alpha(\lambda)x}$$

where $\alpha(\lambda)$ is attenuation coefficient, which is equal to the product of absorption coefficient, $\varepsilon(\lambda)$, and concentration, $c$; i.e. $\alpha(\lambda) = \varepsilon(\lambda)c$. The beer’s law establishes an empirical relationship between the intensity of the transmitted light and concentration of the absorber, considering the path length and the absorption coefficient at the exposed wavelength (equation 3.2).

$$A(\lambda) = -\ln(I/I_0) = \alpha(\lambda)c = \varepsilon(\lambda)cx$$

- $A$ is the absorbance, AU (Absorbance Unit)
- $I_0$ is the intensity of the incoming radiation
- $I$ is the intensity of the light after passing through the sample
- $\varepsilon$ is the absorption coefficient, L mol$^{-1}$ cm$^{-1}$
- $c$ is the concentration, mol L$^{-1}$
- $\alpha(\lambda)$ is the concentration dependent absorption coefficient, cm$^{-1}$
• x is the path length, cm

Beer’s law assumes that the incident radiation is monochromatic, that the absorbing units in the sample are independent of one another, and that absorption happens in a volume of uniform cross section. In addition, it assumes that the sample is non-scattering.

**IR Absorption Spectroscopy**  Near-Infrared and Mid-Infrared (NIR and MIR) spectroscopy are based on absorption of an IR photon, which causes change in the vibrational levels, specifically causing a move from lower to higher vibrational state. The signals in NIR and MIR spectra of chemical compounds are, in general, consequences of molecular vibrations. These molecular phenomena, based on absorption of radiation, can be described as when the vibrating masses \( m_1 \) and \( m_2 \) cause changes to internuclear distance in the molecule [40]. In this particular situation, potential energy, \( v \), generated by vibrating masses, \( m_1 \) and \( m_2 \), can be calculated using Hooke’s Law (equation 3.3).

\[
V = \frac{1}{2} k (r - r_e)^2 = \frac{1}{2} k(q)^2 \tag{3.3}
\]

- \( k \): force constant of bond
- \( r \): internuclear distance during vibration
- \( r_e \): equilibrium internuclear distance
- \( q = r - r_e \): displacement coordinate

The vibrational frequency calculated with this model is:

\[
\nu_0 = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \tag{3.4}
\]

Where reduced mass, \( m \), is calculated as:

\[
m = \frac{m_1 m_2}{m_1 + m_2} \tag{3.5}
\]
Chapter 3. Glucose Monitoring

Equations 3.4 and 3.5 show that vibrational frequencies are sensitive to the structure of the compound. Hence, the spectrum generated from shining light on the sample depends on the energy-level structure of that atom or molecule. Therefore, absorption spectrum is useful for identifying compounds in the sample. According to Schrodinger’s equation in quantum mechanics, the vibrational energy has only certain discrete values (equation 3.6) [40].

\[ E_n = h\nu_0(n + \frac{1}{2}) \]  \hspace{1cm} (3.6)

- \( h \): the Planck’s constant
- \( \nu_0 \): vibrational frequency as defined above
- \( n \): vibrational quantum number, which gets integer values 0,1,2,3, ...

The infrared radiation interacts with vibrating molecule when the electric vector of the radiation oscillates with the same frequency as the molecular dipole moment. In harmonic oscillators, the vibrational spacing between adjacent levels are equal, and transitions are only allowed between neighboring energy levels. Therefore, all these transitions would be at the same IR frequency. Since, according to Boltzmann distribution, molecules at room temperature mostly occupy the ground level, \( n=0 \), the fundamental transition from \( n=0 \) to \( n=1 \) is the dominant factor in the absorption spectrum [38, 40].

Molecules do not act as Hooke’s spring in all situations, they may get dissociated when the vibrating bonds are extremely extended, or upon extreme compression, the repulsive force increases rapidly resulting in rapid increase of energy. Allowed energy levels for an anharmonic oscillator can be described with the following equation 3.7.

\[ E_n = h\nu_0(n + \frac{1}{2}) - \chi h\nu_0(n + \frac{1}{2})^2 \]  \hspace{1cm} (3.7)

where \( \chi \) is the anharmonic constant.

As it can be implied from equation 3.7, the energy levels in the anharmonic oscillator are not equally spaced and transitions happen over more
than one energy level. The second term in the equation causes the energy levels more closely spaced as \( n \) increases, modeling the dissociation effect for the strongly extended vibrational bonds [37, 38, 40].

The fundamental transitions (\( \Delta n = 1 \)) generate fundamental bands; transitions with \( \Delta n > 1 \) are called vibrational overtones. When combination of two vibrational modes are excited, this type of coupled transition gives rise to combination bands [38].

Although vibrational transitions are weaker than the electronic transitions, they have richer structure with a large number of vibrational modes and corresponding bands. Therefore, vibrational spectroscopy has been used for applications in the structural characterization of biological materials [38]. The NIR region contains overtone and combination vibrations, and the overlap of these overtone and combination bands decreases the specificity of NIR spectroscopy. However, the availability of chemometric techniques for quantitative determination analysis made the NIR technique a more favorable technique as the low band intensities can go through larger sample thickness [40].

**Reflectance Spectroscopy** is the study of reflected or scattered light from a solid, liquid, or gas, as a function of wavelength. When photons enter a substance, some of the light gets reflected from the surface, some passes through it and some get absorbed. The reflected light from refraction through the particles is called scattering. Scattering is a random and nonlinear process, which makes the recovery of the quantitative information difficult. In the transmission spectroscopy, the path length is defined and most of the light passes through the material; whereas in reflectance, the optical path of photons is random [39].

### 3.2.2 Absorption Spectra: Glucose Signature in Absorption Spectra

**Water** constitutes major part of biological tissue, and has a strong combination and overtone spectrum that extends into the near IR. The assignment
of these near IR absorption bands for water is demonstrated in [41]. Figure 3.1 shows absorption coefficients for pure water.

Figure 3.1: Absorption coefficient for water [From: www.lsbu.ac.uk/water/vibrant].

The intensity of these water absorption bands are sensitive to solute concentration and temperature [42]. As the solute concentration increases, the intensity of water absorption bands decreases since the molar ratio of water changes. In addition, small change in temperature will affect the near IR spectra of water significantly since temperature has strong effect in the H-bonding of the hydroxyl (-OH) group in water [43]. However, multivariate analysis is capable of accounting for perturbations due to temperature as long as the expected temperature variation is present in calibration set.

**Glucose** molecule consists of Carbon, Hydrogen, and Oxygen represented as $\text{C}_6\text{H}_{12}\text{O}_6$. Figure 3.2 shows the structure of glucose. Glucose has fun-
damental IR absorption bands between 2.5 microns and 10 microns, which arise from C-C, C-H, O-H stretching, and bending vibrations. The strongest bands capable of generating intense combinations and overtones are the broad OH stretch and C-H stretch vibrations at $2.8 \, \mu m$ and $3.3 \, \mu m$ respectively. There are other possible combination bands at 939, 1126, 1408, 1538, 1688, 2261, and 2326 nm. In addition, glucose has various absorption bands at 1.67, 2.13, 2.27, and 2.33 $\mu m$. Table 3.1 summarizes the causes of these absorption bands [43, 44].

Table 3.1: Glucose absorption bands summarized according to the stretch and vibrations of the bonds in the molecule. [43, 44]

Figure 3.3: Absorption Spectrum of Glucose [25].

As table 3.1 shows, glucose absorption spectra is generated by combination and overtone molecular vibration with C-H and O-H bonds of glucose, which results in broad spectrum. Figure 3.3 shows absorption spectrum of glucose. In addition, several species in blood have overlapping absorption peaks (water, hemoglobin, protein, cholesterol, etc) with near or higher concentrations. Glucose concentration correlates with other physiological conditions. Therefore, there is no specific one wavelength point that reflects glucose concentration and there is a non-linear relation between glucose concentration and the absorption spectra. This suggests that the Beer’s law
may not be applied directly and multivariate techniques that can consider other variables would be required. Hence, multivariate analysis is required to model the correlation between absorption spectra and glucose concentration.

Water is the main component of the body and blood, and within certain wavelength regions the absorption is too strong to transmit light. There are two wavelength regions in the near infrared that are commonly used for glucose monitoring: first-overtone (1560 – 1850 nm), and combination band (2080 – 2325 nm) region, where glucose has numerous absorption bands and water has relatively higher transmittance.

3.2.3 Broad Source Spectroscopy for Glucose

In most of the spectroscopy applications, a broad source such as white light is used to illuminate the sample. Spectroscopic methods, to determine glucose concentrations, have been applied to various wavelength regions including near-IR or mid-IR.

Near-IR spectroscopy has been a highly studied optical approach, and it has been determined that these two most promising wavelength regions for glucose monitoring are the first-overtone band (1560-1850 nm) and the combination band (2080-2325 nm), where glucose has numerous absorption bands and water has relatively lower absorption. The absorption is stronger in the 2 μm combination band, which leads to a 20X higher signal-to-noise ratio as compared with the overtone band; hence it is a desirable wavelength region for glucose measurements using absorption spectroscopy [45]. Using 2.0-2.5 μm absorption spectroscopy, a good correlation in blood glucose has been shown in canine experiments [25], as well as in rats [17].

**Optical Glucose Monitoring:** Optical glucose sensing experiments in the near-IR wavelengths (especially 2.3 μm) have been typically conducted using a broad-spectrum source such as a white light. The required sensitivity in optical glucose measurements is $2 \times 10^{-5} AU/mm^1 mM^{-1}$, with a resolution of 1 millimole required for a clinically acceptable accuracy [46];
Chapter 3. Glucose Monitoring

this is difficult to achieve with a white light. Another limitation that a sensor using white light source may have is its unsuitable size for implants.

The alternative is to use a laser as the light source, which can provide various advantages that are required for the application of glucose monitoring.

3.2.4 Laser Based Glucose Monitoring

Most optical glucose sensing experiments in the near-IR wavelengths have been conducted using a broad-spectrum source such as a white light. The alternative is to use a laser as the light source, which can provide advantages such as higher signal-to-noise ratio in the absorption spectrum [21, 47], and is considered to be critical in improving glucose measurement accuracy. We have performed an analysis and optimization in our experiments to improve stability and SNR, and conclude that a laser, such as a VCSEL, is needed to achieve the required SNR, and is critical in improving glucose measurement accuracy.

Lasers, being an optical source that emits coherent photons, have been used as one of the possible optical methods when there is a tunable laser available for the analyzed wavelength range [47]. The amplified coherent light emitted from the lasers will bring about advantages such as higher signal-to-noise ratio in the absorption spectrum. Increasing the signal-to-noise ratio of the spectrometer has been considered as an important required improvement.

Since their inception, semiconductor diode lasers have been unique light sources with excellent spectral and beam properties. Vertical Cavity Surface Emitting Lasers (VCSELs) are a type of semiconductor laser that have received tremendous attention because of their low-cost, small size, array operation, lower power consumption, and circular beam pattern; these properties make them attractive for optical communications and biomedical applications.

Spectroscopy applications require the use of multiple wavelengths. There are several approaches possible: using tunable lasers, using an array with
different wavelength lasers, or by thermal tuning individual lasers. The most successful technique for fabricating tunable VCSELs involves micromechanically tunable mirrors \([48, 49]\). Such tunable VCSELs offer a tuning range of up to 40 nm, with the drawback that they require a tuning voltage to be applied. Alternatively, a fixed multi-wavelength array can be fabricated \([50–52]\). Finally, the easiest method of tuning a laser is by changing its temperature. Typically, thermal tuning provides a tuning range of 7\text{nm}, and can be achieved simply by varying the laser drive current \([53]\). The advantage of this approach is that a single laser can be swept over a molecule’s vibrational resonance feature, for example, such as that of glucose.

In our proposed approach, a small array of single-mode VCSELs will be used for the emission source. The array will be used to acquire a multi-point (e.g., 5 bands) optical spectrum. The lasers will be operated one by one with one common detector. These lasers will be thermally tuned by varying the bias current. Experiments have been designed for the feasibility study to determine if it would be possible to use VCSEL-based spectroscopy for the glucose monitoring.

### 3.2.5 Suitability of VCSELs for Implantable Glucose Monitor

Optical sensors are less susceptible to bio-fouling as the sensors do not degrade over time and light can pass through the collagen fouling, but have challenges in realizing high sensitivity detection required for the weak absorption signals. A laser, with a low noise and a high spectral power density, can aid in increasing the signal-to-noise ratio and provide accurate glucose measurements. Hence, the proposed method was to use the IR spectroscopy using VCSELs. Our preliminary investigations proved the feasibility of the proposed technique for glucose monitoring, by demonstrating that glucose measurements can be obtained using a VCSEL.

Most of the semiconductor lasers developed until now have been focused on wavelengths ranging from the visible to the near-IR (< 1.6 \mu m). It is only recently, however, that a semiconductor laser source has been developed
Chapter 3. Glucose Monitoring

for a 2.3 µm emission. These lasers are not only the first realized electrically pumped VCSELs at a 2.3 µm, but as well are the longest wavelength achieved so far for any InP-based interband laser [54]. The size of these lasers (Figure 6.1) and lower power consumption (<10 mW) makes them attractive for biomedical implants. Experiments have been conducted using these lasers, and the results have demonstrated that a 2.3 µm VCSEL can be used to demonstrate the feasibility of glucose sensing in an aqueous solution, using absorption spectroscopy and a Partial Least Squares (PLS) technique [2]. These experiments were done with a single VCSEL to demonstrate feasibility of the concept. Future work is aimed at extending the approach to multiple lasers to improve the accuracy, and working towards an implantable device.
Chapter 4

Data Analysis and Chemometrics

Chemometrics is the science that applies mathematical and statistical methods to extract information from multivariate chemical data and to relate measurements from chemical or biological systems to the state or interested properties of the system. It is also a science that designs optimal experiments for investigation of chemical systems [40, 55].

4.1 Introduction and Motivation

Glucose has various absorption bands including 1.12, 1.40, 1.54, 1.67, 2.13, 2.27, 2.33 $\mu m$, which arise from combination and overtone molecular vibration with C-H and O-H bonds of glucose molecule. These features result in a very broad absorption spectrum spanning the entire overtone (1560 - 1850 nm) and combination (2.0 - 2.3 $\mu m$) bands. Glucose characteristics in absorption spectra are correlated with other physiological and time dependent factors such as temperature, humidity, and status of the instrument. To avoid the chance of correlation between these factors and glucose concentration, and eliminate their effects in glucose prediction, the model should consider these environmental and physiological factors [56]. These glucose absorption bands lie on top of the water absorption spectra and the position of the water absorption band shifts to shorter wavelengths as temperature increases. Hence the shape of the glucose band changes due to temperature. In addition, there are overlapping spectral signatures from other components in blood or interstitial fluid. The broad and highly overlapped characteristics of the near infrared absorption bands dictates the need for a spectrum
(or several bands of spectra) rather than performing measurements at one or two discrete wavelengths [57]. Furthermore, due to the correlation of glucose with other physiological conditions and components of body fluid as well as lack of a specific one wavelength point reflecting the glucose concentration, glucose spectra have a broad shape with a non-monotonic increase with glucose concentration as would be expected from normal line-shape absorption spectra.

This non-linear relation between absorption spectra and glucose concentration suggests the use of Chemometrics and multivariate analysis [47, 58, 58]. Researchers have shown that multivariate models can be designed to predict glucose concentration, eliminating the effects of potential correlation and confounding factors, including physiological variations [59]. Multivariate analysis, such as Partial Least Squares (PLS), is required to find a correlation between the spectra and glucose concentrations [60, 61]. This method has been used extensively to determine glucose concentration from absorption spectra [60–62]. These multivariate methods cannot deduce any spectra of individual component in the solution, rather they compute regression vectors directly from the mixture spectra and the known concentrations of a target component [63].

As a result, various preprocessing and multivariate methods will be discussed in this chapter. For demonstration, verification, and discussions about these methods, a set of absorption spectra for various glucose concentrations has been used in this chapter. This set of absorption spectra have been collected using white light as the light source for the wavelengths from 1100 nm to 1700 nm. However, for the investigation of the suitable preprocessing, the wavelength window from 1419 nm to 1519 nm, which contains a strong water absorption peak that dominates the effects of absorption of other constituents, is removed from the spectra so that the effect from the change in water concentration does not override the effect from absorption of glucose. Two sets of data with these qualities but different ranges of concentrations are used. One is with concentrations of 500 to 20000 mg/dL (27.7 to 1111 mmol/L), which was used for preliminary investigations and the proof of concept, and the other with a lower range of 100 to 400 mg/dL.
Chapter 4. Data Analysis and Chemometrics

(5.56 to 22.2 mmol/L).

This Chapter consists of the following sections: Basics of Chemometrics including explanation of Principal Component Analysis (PCA), Preprocessing methods including description of various methods, Multivariate Regression discussing various methods leading to Partial Least Square method used for glucose predictions. In addition, the choice of multivariate method and especially the choice of preprocessing method(s) for the current application are discussed.

4.2 Basics of Chemometrics

When analyzing data using chemometrics, data are usually organized in matrix form, where the rows of the matrix correspond to the samples or observations and its columns correspond to variables. Variables in general can be absorbance at each wavelength, pressure, temperature, flow, etc. Samples can be obtained from different concentrations or data samples at different time points, etc [55].

The X block is a matrix of predictor variables, where the rows correspond to samples or observations and columns correspond to variables or wavelengths for spectroscopy applications. For the case of spectroscopy, the samples are absorption spectra for various concentrations of glucose in the solution. X has dimension of m × n, where m is the number of samples or observations or spectra in spectroscopy and n is the number of variables or wavelengths in spectroscopy applications. The Y block is a matrix or vector of predicted variables, which represents glucose concentrations for the current project.

This section will include explanation and discussion on the basics of chemometrics or more specifically Principal Component Analysis (PCA).

4.2.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a tool used for data compression and information extraction. Excluding irrelevant and redundant information
to enable extraction of essential information does information compression. Essential information is usually described in how variables change with respect to one another [55]. The American Society of Testing and Materials has defined PCA as following:

“Principal component analysis is a mathematical procedure for resolving sets of data into orthogonal components whose linear combinations approximate the original data to any desired degree of accuracy. As successive components are calculated, each component accounts for the maximum possible amount of residual variance in the set of data. In spectroscopy, the data are usually spectra, and the number of components is smaller than or equal to the number of variables or the number of spectra, whichever is less” [64].

If r is the rank of the data matrix $X$, PCA can decompose $X$ block data into the sum of $r$ $t_i p_i$ (equation 4.1) [55].

$$X = t_1 p_1^T + t_2 p_2^T + \ldots + t_k p_k^T + \ldots + t_r p_r^T$$  \hspace{1cm} (4.1)

where $r <= \min\{m,n\}$

Generally, the PCA model is truncated after k components. The remaining variance as a result of truncation is represented in a residual matrix $E$ (equation 4.2) [55].

$$X = t_1 p_1^T + t_2 p_2^T + \ldots + t_k p_k^T + E$$  \hspace{1cm} (4.2)

The order of $t_i p_i$ is according to the amount of variance they capture, and the first ones capture higher variance. $t_i$ vector, which forms an orthogonal set $T_k$ are called scores. They describe relationship between samples. $p_i$ vector, which forms an orthogonal set $P_k$ are called loadings. They describe direction of variation in data and relationship between variables. Figure 4.1 shows how PCs describe the direction of variation in data and construct a new coordinates. This Figure shows that the first PC is in the direction of
greatest variation in the data set, and the second PC describes the direction of the second greatest variation in the data set [55].

This figure has been removed due to copyright restrictions.
The information removed illustrates the concept of Principal Component Analysis (PCA).
It is figure 5-1 in the referenced document.

Figure 4.1: Graphical representation of Principal Component Analysis. The Euclidian distance of a data point from the plane described by the two PCs in this model is \( \sqrt{Q} \). Hotelling’s \( T^2 \) shows how the original variables deviate from the mean within the model. It is a measure of the distance between multivariate mean and projection of sample onto the plane of PCs [55].

Equation 4.2 can model a chemical and biological system. As an example, if \( X \) represents a number of absorbance spectra of a solution, which follow linear additivity of Beer’s Law, \( t_i \) can represent concentration and \( p_i \) can represent pure component spectra.

PCA assumes linear relationship between variables. Hence, non-linear data should be converted to a linear form if possible.

4.3 Preprocessing

Preprocessing is a method of preparing specific data type for its data analysis objective. The objective of preprocessing is to eliminate artifacts and non-linearities in the data to achieve simpler models. This includes linearizing relationships between variables; removing undesirable variances or clutters; and enhancing relevant variances.

Glucose absorption spectra, collected for the experiments explained in the subsequent chapters, will be used to demonstrate the effect of some of these methods on the specific data at hand. Two series of data were collected using Optical Spectrum Analyzer (OSA), where the light source is white light. Figure 4.2 shows the series of absorption spectra for the samples with higher concentrations (500 mg/dL to 20000 mg/dL), which were used for preliminary investigations and proof of concepts. Figure 4.3 shows
absorption spectra of some samples with lower concentrations (0:100:400 mg/dL). These two sets of data have been used in the next chapter, where the spectroscopy experiments for glucose sensing using white light is discussed.

Figure 4.2: Absorption spectra using white light as the source. The concentrations presented in the legend are in mg/dL. These concentrations are high to serve the purpose of preliminary investigations and proof of concepts. In addition, absorption spectra of high concentration samples unravel some characteristics such as high water absorption in the wavelength region 1419 to 1519 nm, which may bias the result of the analysis.
This section will describe several preprocessing techniques such as mean centering, variance scaling, Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV), Orthogonal Signal Correction (OSC), Square Root Mean Scale (sqrt Mean Scale).

4.3.1 Mean-Centering and Centering

Subtracting the mean spectrum of the data set from every spectrum in the data set is called mean centering. Most multivariate techniques are scale dependent, meaning that numerically larger variables appear more important in the data analysis or building the regression model. Mean centering is useful when variation of the data around the mean is of interest and not the absolute value. Mean centering is achieved by subtracting the mean of each column from all values of that column, and subsequently forming columns
with zero mean.

Mean centering results in a more simple and interpretable regression model. Figure 4.4 shows the effect of mean centering. In the Figure with no mean centering, the direction of the greatest variance is towards the mean of the spectra; whereas in the mean centered data set, the direction of the greatest variance is the same as the direction of the greatest variance within the data set [40].

This figure has been removed due to copyright restrictions.

The information removed illustrates the concept of mean centering.

It is figure 10.2 in the referenced document.

Figure 4.4: Demonstration of the effect of mean centering on the direction of the greatest variance. (a) non-mean centered data. (b) mean centered data. [40]

Figure 4.5 and 4.6 show the effect of mean centering on the absorption spectra presented in Figure 4.2 and 4.3 respectively.
Figure 4.5: Effect of mean centering on absorption spectra of glucose solutions presented in Figure 4.2. The concentrations presented in the legend are in mg/dL.

Figure 4.6: Effect of mean centering on absorption spectra of glucose solutions presented in Figure 4.3. The concentrations presented in the legend are in mg/dL.
4.3.2 Variance Scaling and Autoscaling

As multivariate analysis and regression techniques are scale dependent, variance is associated with importance. However, larger variance may not always mean higher importance. For example, when variables have different units, higher value of variance may not be an indication of higher importance as values are not comparable in different units and the regression model does not understand that. Therefore, scaling variables to have unit variance prevents the misinterpretation of the importance of various data points. To generate data with unit variance, each variable is divided by its standard deviation. Data set can also be auto-scaled by dividing each mean centered variable by its standard deviation.

The application of mean centering and variance scaling together is called auto-scaling. Figures 4.7 and 4.8 show the effect of auto-scaling on the absorption spectra presented in Figures 4.2 and 4.3 respectively.

![Figure 4.7: Effect of auto-scaling on absorption spectra of glucose solutions presented in Figure 4.2. The concentrations presented in the legend are in mg/dL.](image-url)
4.3.3 Multiplicative Scatter Correction (MSC)

The objective of this method is to reduce scattering artifacts in diffuse reflectance and transmission NIR spectra by accounting for scaling and offset (baseline) effects. MSC uses regression of new measured spectrum onto a reference spectrum, which is often the mean spectrum of a data set. MSC has also shown capability of removing varying background spectra that have not been originated from scattering effect. According to scattering theory, scattering has multiplicative effect on the transmittance and reflectance spectra, meaning that the spectra will have a broad changing background [40]. MSC models the scattering with an offset and slope. To correct the signal for the scattering effect, the offset is subtracted from the signal and the result is divided by the multiplicative factor [55].

Figures 4.9 and 4.10 show the effect of applying MSC on the absorption spectra presented in Figure 4.2 and 4.3 respectively.
Figure 4.9: Application of MSC preprocessing on absorption spectra of glucose solutions presented in Figure 4.2. The concentrations presented in the legend are in mg/dL.

Figure 4.10: Application of MSC preprocessing on absorption spectra of glucose solutions presented in Figure 4.3. The concentrations presented in the legend are in mg/dL.
4.3.4 Standard Normal Variate (SNV)

Standard Normal Variate (SNV) normalization is a weighted normalization method, where not all the points contribute equally to the normalization process. In SNV, the standard deviation of all the variables for a given sample (standard deviation of each row) is calculated. The entire sample is then mean-centered, using the mean of the row, and normalized by the standard deviation of the row. This normalization method aims to consider values that deviate more strongly from the individual sample mean. This method is most useful when overall signal is almost similar from sample to sample [55].

Figures 4.11 and 4.12 show the effect of applying SNV on the absorption spectra presented in Figures 4.2 and 4.3 respectively.

Figure 4.11: Application of SNV preprocessing on absorption spectra of glucose solutions presented in Figure 4.2. The concentrations presented in the legend are in mg/dL.
4.3.5 Orthogonal Signal Correction (OSC)

Orthogonal Signal Correction (OSC) removes variations in $X$ block that are irrelevant to the changes in predicted variable $Y$; in other words, it removes variations in $X$ block that are orthogonal to $Y$ block. This helps to construct a simpler regression model for predicting $Y$ block. OSC can be very helpful for the cases that, for example, Partial Least Square (PLS) model captures a large amount of $X$ variance in the first factor but gets very little variation or relevance of the predicted variable $Y$. OSC can also be used to remove variations between instruments that are not related to the predicted properties [55]. However, it might not necessarily improve the model performance, hence it has to be tested for the specific application in hand. As mentioned earlier, the main advantage of OSC is that it facilitates easier analysis and better interpretation of the corrected data rather than reducing prediction errors [65].

At the beginning of the OSC algorithm, the first principal component
Chapter 4. Data Analysis and Chemometrics

of the \( X \) is identified. Then, the loading is rotated and adjusted to have orthogonal scores to the \( Y \) block. Hence, the features of this loading are not influenced by the changes in the predicted variable (\( Y \) block). Afterwards, PLS model is built to predict these orthogonal scores from \( X \) block. The orthogonal loadings and predicted scores are then used to remove that orthogonal component from the calibration data before building the model and subsequently from all the data under analysis with this model [55].

Figures 4.13 and 4.14 show the effect of applying OSC on the absorption spectra presented in Figures 4.2 and 4.3 respectively.

![Absorption Spectra with OSC](image_url)

Figure 4.13: Application of OSC preprocessing on absorption spectra of glucose solutions presented in Figure 4.2. The concentrations presented in the legend are in mg/dL.

As it can be observed from Figures 4.13 and 4.14, OSC corrects for most of the variations that do not contribute to the predicted variable, glucose concentration. Applying mean centering along with OSC can be an ideal preprocessing method considering special application at hand.
Figure 4.14: Application of OSC preprocessing on absorption spectra of glucose solutions presented in Figure 4.3. The concentrations presented in the legend are in mg/dL.

4.3.6 Square Root Mean Scale (Sqrt Mean Scale)

The purpose of scaling is to adjust the magnitude of each variable such that the level of noise is almost the same for all variables.

Square root Mean Scale (Sqrt Mean Scale) is used for the cases where the noise at each variable is approximately equivalent to the square root of the signal at that variable. Hence, this method can effectively correct for the noises that are truly proportional to the square root of the signal. This method works by scaling each variable by the square root of the mean of that variable. This method, however, does not mean center the data; therefore, if data requires mean centering, it needs to be done as additional preprocessing.
4.4 Multivariate Regression

Regression analysis is used to identify the dependency between two blocks of data; specifically to relate a set of measured variables to a state of interest, and use these dependencies to make a regression model. Regression models are usually used to predict one block of data from the other.

Like Principal Component Analysis, Regression Analysis requires pre-processing and data should be linearized if possible. Data is mean centered most of the time and variance scaled when variables have different units or greatly different magnitudes. In addition, it is critical to eliminate outliers for regression model.

Notations that are used in this section are as follows. $\mathbf{X}$ is the matrix of predictor variables. $\mathbf{Y}$ is the matrix of predicted variables. Number of samples or observations is represented by $m$. Number of $\mathbf{X}$ and $\mathbf{Y}$ variables are represented by $nx$ and $ny$ respectively.

This section describes and discusses various methods of multivariate regression and some considerations. Specifically, this section will include descriptions of Classical Least Squares (CLS), Inverse Least Squares (ILS), Principal Component Regression (PCR), Partial Least Square Regression (PLSR), as well as Cross-Validation and some considerations with outliers.

4.4.1 Classical Least Squares (CLS)

Classical Least Squares (CLS) is one of the methods that can be used to develop calibration models. CLS develops a model to relate the spectral response to the analyte concentration [40]. The CLS model assumes that measurements can be represented as weighted sum of linearly independent signals. For example, in spectroscopy, measured component spectra are represented as the sum of two components; one is the pure component spectra weighted by the concentration of the analytes, and the other component is noise or the error matrix (equation 4.3) [55]. Spectroscopy observation of this equation follows the logic of Beer’s Law [40].

$$\mathbf{X} = \mathbf{CS}^T + \mathbf{E} \quad (4.3)$$
where

- $X$: The measured response $(m \times nx)$, where each row is a spectrum of a sample
- $S$: Matrix of pure component responses $(nx \times k)$ of each constituent in samples
- $C$: Matrix of weights $(m \times k)$, the extrinsic property of interest such as concentration.
- $E$: Noise or error matrix $(m \times nx)$

Given the measured response, $X$, when the pure component response, $S$, is known, the contribution degree of each component can be determined using equation 4.4 [55].

\[
C = XS^+ \tag{4.4}
\]

Where

\[
S^+ = S(S^TS)^{-1} \tag{4.5}
\]

For the term $(S^TS)^{-1}$ and consequently $S^+$ in equations 4.5 and 4.4 to exist, the pure component responses must be linearly independent. If the pure component responses are nearly collinear, $S^+$ can be calculated but will be unstable and very sensitive to small changes in the estimations of the pure responses due to noise.

The main disadvantage of CLS method is the requirement of knowing the pure responses, $S$, of all the components that are spectrally active in the compound or estimating them from the known concentrations of these components. Practically, this requirement may be hard to achieve [40].

### 4.4.2 Multiple Linear Regression (MLR)

Multiple Linear Regression (MLR) is one of the Inverse Least Square (ILS) methods, which assumes model of the form in equation 4.6. Note that this model associates the noise with the predicted property, instead of associating
with measured response as in CLS.

\[ Xb = y + e \]  \hspace{1cm} (4.6)

Where

- \( y \) is the property to be predicted (m \( \times \) 1), such as concentration vector
- \( X \) is the measured response (m \( \times \) nx)
- \( b \) is a vector of coefficients (nx \( \times \) 1)
- \( e \) is an error vector (m \( \times \) 1)

In linear calibration, \( b \) can be estimated from equation 4.7.

\[ b = X^+ y \]  \hspace{1cm} (4.7)

where \( X^+ \) is the pseudo-inverse of \( X \).

MLR is one of the methods that can be used to calculate the pseudo-inverse of \( X \) as shown in equation 4.8.

\[ X^+ = (X^T X)^{-1} X^T \]  \hspace{1cm} (4.8)

Although MLR is a straightforward technique, it has associated disadvantages. The main challenge is existence of collinearity in \( X \), which causes \((X^T X)^{-1}\) to be ill-conditioned. Collinearity occurs when columns of \( X \) can be expressed as linear combination of each other, or when there are more columns in the matrix than there are rows. Both of these causes exist in NIR spectra as absorbance in adjacent wavelengths are correlated and usually there are fewer samples available for calibration than number of wavelengths recorded in spectra. Hence, when collinearity exists in \( X \) (i.e. rank deficiency exists), the inverse will be highly unstable leading to unstable estimation of \( b \). In these cases, where regression vector is unstable, small random errors, such as noise, can result in large prediction errors [40, 55].

However, Inverse Least Square (ILS) models, which include Multiple Linear regression (MLR), Principal Component Regression (PCR), Partial
Chapter 4. Data Analysis and Chemometrics

Least Squares (PLS), and etc, have several advantages. The main advantage of ILS methods is that they do not require the prior knowledge of the concentrations of all analytes in the sample; whereas CLS required the knowledge of the concentrations of all the analytes including interferents. However, in calibrating ILS models, interferents must vary so that the ILS regression model is robust against them [55]. For this, the boundary of the model, regarding how much the interferents or other constituents in the sample change, should be known or estimated. Due to the useful advantage of ILS models for NIR spectroscopy applications, the following two sections will describe two more ILS models, and discuss their suitability for the current application.

4.4.3 Principal Component Regression (PCR)

Principal Component Regression (PCR) is another way of dealing with ill-conditioned problems. PCR is achieved by regressing the properties of interest, such as concentration, onto the principal component scores of the measured variables, such as spectral response variables. Since principal components are orthogonal by definition, the matrix of principal component scores, $T_k$, is well-conditioned and hence the pseudoinverse of the response data matrix, $X^+$, is stable [55]. Equation 4.9 shows the pseudoinverse of the response data matrix using PCR method.

$$X^+ = P_k(T_k^T T_k)^{-1}T_k^T$$  \hspace{1cm} (4.9)

where $P_k$ are loadings and $T_k$ are scores. The challenge in this method is finding the right number for $k$. Once the pseudoinverse of the response data matrix, $X^+$, is ready, equation 4.7 is used to estimate the vector of coefficients, $b$.

PCR method finds PCs that capture maximum variance in $X$, which may not be relevant for prediction. This shortcoming is due to not using the information about the property that needs to be predicted, $y$, when calculating PCs.
4.4.4 Partial Least Square Regression (PLSR)

Partial Least Square (PLS) is closely related to both MLR and PCR and can be considered as being in the midway between them. MLR method can capture factors that can best correlate predictor variables (\(X\)) with predicted variable (\(Y\)), and PCR can find the greatest amount of variance in the predictor variable (\(X\)). Therefore, PLS, being a combination of these two methods, finds factors that can both capture variance in predictor variables and achieve correlation between predictor and predicted variables. PLS finds latent variables that are not principal components; but they have a slightly different optimization criterion. PLS calculates and uses matrix of scores (\(T\)) and loadings (\(P\)), and an additional set of vectors called weights, \(W\), which is required to maintain orthogonal scores [40, 55].

In the PLS method, factors are calculated sequentially for each latent variable starting with the one that captures highest variance and correlation. PLS calculates and uses the pseudo-inverse of the response data matrix, \(X^+\), with the following equation 4.10. This pseudo-inverse of the response data matrix, \(X^+\), is then used in equation 4.7 to estimate the vector of coefficients, \(b\).

\[
X^+ = W(P^T W)^{-1}(T^T T)^{-1} T^T \tag{4.10}
\]

The matrix of scores and loadings are not calculated as those in PCR or PCA, but PLS has its own method of calculating these parameters. However, they can be thought of more as rotated versions of PCA scores and loadings, adjusted such that there are more relevant for predicting variable of interest, \(Y\).

Due to the unique role of weights in PLS model, and its relationship with loadings, scores for the new set of data is calculated differently from PCA and PCR model, where \(T_{new} = X_{new} P\) [55].

\[
T_{new} = X_{new} W(P^T W)^{-1} \tag{4.11}
\]

In summary, PLS attempts to find factors, which are called latent vari-
ables (LV), that captures the maximum amount of variation explained in \( X \) that is relevant for predicting \( Y \) [55].

### 4.4.5 Model Quality Measures

The purpose of quality measures is to estimate average deviation of the model estimates from the measured data. Root Mean Square Error (RMSE) Metrics, representing fit of the model to the data and assessing predictive ability of the model on new data, is used to measure the quality of the model. There are three RMSE measures: Root Mean Square Error of Calibration (RMSEC), Root Mean Square Error of Cross-Validation (RMSECV), and Root Mean Square Error of Prediction (RMSEP).

#### Root Mean Square Error of Calibration (RMSEC)

Root Mean Square Error of Calibration (RMSEC) is a measure of fit of the model to the calibration data. It is calculated using equation 4.12.

\[
RMSEC = \sqrt{\frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{m}}  \tag{4.12}
\]

where \( y_i \) is the measured variable, \( \hat{y}_i \) is the predicted variable evaluated when all the samples are included in calibration set, \( m \) is the number of samples used in calibration process, and \( i \)’s refer to all samples that are used to build the model.

#### Root Mean Square Error of Cross-Validation (RMSECV)

Root Mean Square Error of Cross-Validation (RMSECV) is an estimate of the predictive power on new data. It measures ability of the model to predict using samples that were not used to build the model.

RMSECV is based in cross-validation of the model. Cross-validation is achieved by dividing data into \( j \) subsets, building the model with the \( j-1 \) subsets, and calculating Predictive Residual Sum of Squares (PRESS) for the subset that is left out. This process is repeated \( j \) times until all the
subsets have been left out once [55].

\[ \text{PRESS} = e^2 = (y - Xb)^2 \]  \hspace{1cm} (4.13)

Value of RMSECV depends on the number of latent variable, LVs (or PCs for PCA method), and how the data set is divided for cross-validation.

\[ \text{RMSECV}_k = \sqrt{\frac{\text{PRESS}_k}{m}} \]  \hspace{1cm} (4.14)

where \( m \) is the number of samples used in the calibration set, and \( k \) is the number of factors (LVs, or PCs) used in the model.

**Root Mean Square Error of Prediction (RMSEP)**

Root Mean Square Error of Prediction (RMSEP) is used to validate a model with a true measure of the predictive power on new data.

For this purpose, independent set of sample measurements with known \( Y \) can be designated for prediction. When the model is applied to this new set of prediction samples, RMSEP can be calculated using equation 4.12, where \( i \)'s refer to samples not used to build the model. In this case \( \hat{y}_i \) is an estimate of samples in that new independent prediction set rather than samples involved in the calibration set.

**4.4.6 Identifying Optimum Number of Factors (LVs or PCs):**

Determining the right number of factors (LVs or PCs) in the model is important and not always simple. As the number of factors (LVs or PCs) involved in the model increases, the fit of the model improves; however, the validity of the model, with the new set of data, gradually declines. Therefore, there is an optimum number of LVs or PCs that can achieve the best validity of the multivariate regression model [55].

Although determining the right number of factors (LVs or PCs) to retain in the model is not always straightforward, there are several methods to
identify reasonably appropriate number of PCs. One method is looking for the “knee” in the plot of eigenvalues vs. LV or PC. There are other methods such as considering ratios of successive eigenvalues, and retaining LVs or PCs with percent variance of higher than noise level [55].

In addition, RMSEC and RMSECV can be used to determine the optimum number of factors (LVs, or PCs) for building the model. RMSEC, which is the measure of the fit of the model to the calibration set, decreases as the number of LVs or PCs increase. However, since the validity of the model with new set of data declines with the addition of new factors, RMSECV, which is an estimate of the predictive power of the model on a new set of data, will have a minimum point that is often an indication of the maximum number of LVs or PCs required to build the model. The optimum number of LVs can either be at the point from which the RMSECV does not improve significantly anymore (i.e. the ‘knee’ of the graph) or can be at the point of minimum RMSECV. There is a trade off between computational complexity and minimum error in choosing the number of factors from the ‘knee’ of the graph to its minimum.

Figure 4.15 is a plot of RMSEC and RMSECV vs. latent variables (LVs) for the absorption spectra collected with white light, which are collected and analyzed as explained in the next chapter.

4.5 Discussion

There are two main discussion points when applying chemometrics methods to a set of samples such as absorption spectra for various concentrations of glucose. One is the right choice of the multivariate regression method and the other is the choice of the preprocessing method.

4.5.1 Choosing Multivariate Regression Method

Multivariate methods are used when there are more components affecting the signal being analyzed. Multivariate regression models can be calibrated for prediction of more that one analyte. They also are best suited for predicting
Figure 4.15: RMSEC and RMSECV vs. latent variables (LVs). RMSEC is always decreasing as a function of LVs, whereas RMSECV decreases first and increases after certain number of LVs, which can be an optimum number of LVs to build the model with.

A variable of interest in the presence of various other interfering signals. Existence of various different constituents in blood, be the ones that are of interest or the interfering ones, suggests using one of the multivariate methods.

Among the aforementioned various multivariate methods, Partial Least Square Regression (PLSR) has the best capability in capturing variations in the predictor variables that are correlated to the variable of interest. Hence, using PLSR, choosing only factors that capture most variation in the direction that is correlated to the predicted variable, reduces the number of factors (LVs) required for reasonable or optimum prediction accuracy. Therefore, Partial Least Square Regression (PLSR) method, due to its capabilities, has been considered as one of the best optimized methods for applications of this kind and has been used by many researchers for similar chemometric application [60–62].
4.5.2 Finding the Proper Preprocessing Method

RMSECV is a method for investigating the optimum number of factors (LVs or PCs) for model calibration since RMSECV first decreases as the number of factors increases and after reaching a minimum, it will start increasing again. As mentioned earlier, the optimum number of factors (LVs) can be identified from the plot of RMSECV vs. Latent Variables. The number of LVs used for the calibration can either be the number of factors that give minimum error or the one after which RMSECV does not improve significantly (at the ‘knee’ of the plot). Choosing a number between these two depends on the accuracy required for the application as well as capacity of the system in handling the computational complexity generated by having more LVs. RMSECV also estimates the error using that model with the certain number of factors (LVs or PCs). Hence, it can be used to compare various preprocessing methods for a certain data sets with special characteristics.

Various choices of preprocessing methods have been investigated for the spectra collected from the samples of different glucose concentration using white light and VCSELs. RMSECV has been used to determine which preprocessing method will result in minimum cross-validation error with the least number of LVs.

Preprocessing for Glucose Absorption Spectra Collected Using White Light

Cross-validation was employed using PLS method on a set of absorption spectra collected using white light for different concentrations of glucose. Then RMSECV and RMSEC were recorded for each case and were plotted vs. number of Latent Variables (LVs). The wavelength window, in which water absorption is very strong and is the dominant factor that can bias the preprocessing and analysis, has been removed from the spectra before application of the analysis. Figure 4.16 demonstrates elimination of wavelength window corresponding to strong water absorption (1419 nm to 1519 nm) from absorption spectra in Figure 4.2.

As mentioned before, the absorption spectra using white light includes
Figure 4.16: Absorption spectra of glucose solutions with various concentrations, without wavelengths corresponding to the strong water absorption. This Figure demonstrates absorption spectra of Figure 4.2 without wavelengths corresponding to water absorption peaks. The concentrations presented in the legend are in mg/dL.

spectra from two sets of experiments. One is where higher concentrations of glucose samples (500 to 20000 mg/dL) are used and the other is for samples with lower concentrations (0:100:400 mg/dL).

These absorption spectra were preprocessed using the following methods, before being cross validated using PLS regression. In the presented investigations, the Y block has always been mean centered unless stated otherwise. Preprocessing methods applied to X are:

- Mean Centering (sometimes labeled as \textit{meanC} or \textit{meanCenter})
- AutoScale
- SNV : Standard Normal Variate
- SNV and Mean Centering (labeled as \textit{SNV+meanCenter} or \textit{SNVmeanC})
- MSC : Multiplicative Scatter Correction
- MSC and Mean Centering (labeled as \textit{MSC+meanCenter} or \textit{MSCmeanC})
- OSC : Orthogonal Signal Correction
• OSC and Mean Centering (labeled as OSC+meanCenter or OSCmeanC)

Figure 4.17 presents the RMSECV vectors resulted from the cross-validation of data with various preprocessing methods applied to the spectra in Figure 4.16.

![Figure 4.17: RMSECV vs. LV used to compare and determine a customized preprocessing method for the absorption spectra of glucose samples with high concentrations (500 to 20000 mg/dL), where water absorption band is eliminated from the spectra. RMSECV helps to determine the method that optimizes the number of LVs required with the relatively lower cross-validation error of prediction, according to the requirements of the application.](image)

Combination of applying OSC and mean centering together results in more relevant variations captured by the few first factors; meaning that the first few factors (LVs) capture more correlations resulting in lower RMSECV errors of cross-validation. The RMSECV graph for mean centering method has a knee at five LVs; at that point, preprocessing methods of OSC plus mean centering results in the same cross-validation error; Furthermore, OSC plus mean centering can even achieve similar RMSECV error with less number of factors (LVs).
Chapter 4. Data Analysis and Chemometrics

Figure 4.18 illustrates cross-validation error, RMSECV, resulted from applying various preprocessing methods to the absorption spectra excluding water absorption band for the lower concentrations of glucose. It is observed that OSC and mean centering results in a more optimized error in prediction.

![Figure 4.18: RMSECV vs. LV used to compare and determine a customized preprocessing method for the spectra of lower glucose concentrations (0:100:400 mg/dL) in this application. RMSECV helps to determine the method that results in lowest cross-validation error of prediction.](image)

Preprocessing for Glucose Absorption Spectra Collected Using VCSELs

The other set of data used for this project is from the experiment that will be explained in chapter 6, where VCSELs are used for glucose spectroscopy. In that experiment, two 2.3 μm VCSELs are used to measure absorption spectra. The data is analyzed using absorption spectra of only one VCSEL as well as using both VCSELs. Various preprocessing methods are applied to the absorption spectra using these VCSELs.

Figures 4.19 and 4.20 illustrates application of various preprocessing
methods to the absorption spectra using one VCSEL, and two VCSELs respectively. For this purpose, data were preprocessed and cross validated to estimate RMSEC and RMSECV.

Comparing the two Figures, 4.19 and 4.20 for the RMSECV errors from analyzing absorption spectra using one VCSEL and two VCSELs, shows that lower cross-validation error may be achieved using two VCSELs rather than one.

It is concluded from Figure 4.20 that the two preprocessing methods can be good candidates for model calibration: SNV and combination of OSC plus mean centering. Other available preprocessing methods from PLS toolbox have been examined for this kind of data when the sections of absorption spectra are generated by independent light sources (VCSELs). The preprocessing method on Y block, vector of concentrations, is mean centering as mentioned before. The additional preprocessing methods applied to X block, the absorption spectra of various samples, are as follows. The order of applying these preprocessing methods appears in the label.
Figure 4.20: RMSECV vs. LVs for the comparison of application of various preprocessing methods using absorption spectra of two VCSELs.

- Log10 and Sqrt Mean Scale (labeled as LogSq)
- Sqrt Mean Scale and SNV and Derivative (labeled as SqSNVder)
- Sqrt Mean Scale and Derivative and SNV (labeled as SqderSNV)
- Sqrt Mean Scale and Derivative (labeled as SQder)
- Derivative and SNV (labeled as derSNV)
- Mean Centering and Sqrt Mean Scale (labeled as meanCSq)
- Mean Centering and Sqrt Mean Scale and Derivative and SNV (labeled as SqmeanC)
- Sqrt Mean Scale (labeled as Sq)

Figure 4.21 shows RMSECV vector vs. LVs as a result of applying these additional preprocessing methods. The two promising methods from
Figure 4.21: RMSECV vs. LVs to identify the preprocessing method that results in the lowest RMSECV error and the optimum corresponding number of LVs. In the legend MeanC stands for mean centering, Sq stands for Sqrt Mean Scale, der stands for derivative (1st derivative).

previous graph (4.20), OSC plus mean centering and SNV, are also included in this Figure for comparison purposes.

Figure 4.21 identifies Sq plus mean centering with 11 LVs as one of the potential preprocessing methods. In choosing the right preprocessing, two main conditions are considered: lower number of LVs is desired for having less computational complexity, low enough RMSECV for the accuracy required for the given application. Considering these two points the choice may differ depending on the application and characteristics of data.
Chapter 5

Spectroscopy for Glucose Sensing Using White Light

Transmittance spectroscopy was used at this stage of the project for preliminary investigation of the proposed techniques. Two general sets of experiments were designed: white-light experiment, and VCSEL-based experiment. In the white-light experiment, a study was initiated to explore the possibilities of predicting glucose with only small segments of an optical spectrum, and to determine which laser wavelengths should be chosen for accurate glucose sensing. This experiment involved determining the wavelength segments in the 1.0 µm to 1.7 µm range capable of predicting glucose concentration, using multivariate data analysis. Ultimately, the optimal VCSEL wavelengths will be selected for the integration of a laser-based implantable chip. The study concluded that it is not necessary to use a wide spectral range to predict glucose; this positive outcome led us to begin laser-based experiments as the next set of experiments, explained and discussed in the next chapter. VCSEL-based experiments involved using actual VCSELs in the 2.3 µm range to investigate the possibility of using actual VCSELs as the light source for the experiment.

This chapter discusses the results of our investigation in the wavelength range of 1.0 µm to 1.7 µm. Our aim is to determine the minimum number of segments required to predict the glucose concentration within an acceptable clinical accuracy, and to analyze the accuracy of the partial least square technique, given the absorption spectra at specific wavelength segments. These experiments were performed using a white light (at wavelengths < 1700 nm), with the data analyzed for small portions of the spectrum, each
representing what would be one laser. Ultimately, the optimal VCSEL wavelengths will be selected for the integration of a laser-based implantable chip. Therefore, this chapter will present and discuss the results of applying Partial Least Square (PLS) techniques on small wavelength windows with the goal of determining the number of VCSELs required to predict glucose concentration, and verifying that it is possible to predict glucose concentration from a selected subset of absorption spectra using PLS.

This Chapter consists of the following sections: Description of Apparatus used for the experiment, Determining Sub-Window describes how the sub-windows of the absorption spectrum have been selected to identify glucose concentration, Absorption Spectroscopy describing the experiment, the Results and Further Improvements on Data Analysis and PLS.

5.1 Description of Apparatus

The set up for the experiment consists of a glass container for the solution, fiber collimators, a white light source, and an optical spectrum analyzer. The container is made of two slides of glass, with 1 mm thickness, separated by a 1 mm gap. Hence, the path-length of light in the solution is 1 mm.

**Optical fiber** is a waveguide for light. The structure of optical fibers includes core, cladding, coating or buffer, and jacket. Light propagates mainly along the core of the fiber, which is usually made of glass and surrounded by a layer of material called cladding. Coating or buffer is another layer of material used for extra fiber protection. Figure 5.1 shows the structure of a simple single mode optical fiber. The core of optical fiber guides the light using total internal reflection concept.

**Optical Spectrum analyzer (OSA)** ANDO AQ6315A was used for this experiment. Optical Spectrum Analyzer is capable of dividing a lightwave signal into its constituent wavelengths, which makes it possible to see the spectral profile of the signal over a certain wavelength range. The measurement wavelength range for this OSA is 350 to 1750 nm, with wavelength
White Light  

HL-2000 Tungsten Halogen Light source, which is a versatile lamp optimized for the visible and near Infrared (VIS-NIR) spectrum (360nm to 2000nm), was used for this experiment. These lamps emit through a fused silica lens and they usually operate at color temperature of 2960 K. Figure 5.2 shows the HL-2000 spectral output, and Figure 5.3 shows normalized black body radiation for color temperature of 2960 K. The combination of these two curves, along with the transmission curve of fused silica, defines the output of the Tungsten Halogen light source.

In this experiment, the light from the white light source goes through a fiber and after passing through the sample collected in another fiber, which transfers it to OSA. The OSA then scans through the wavelengths and produces absorption spectra for the sample. These spectra from OSA are
transferred to the computer using Matlab program with GPIB. Figure 5.4 shows the schematic of this experimental set up.

5.2 Determining sub-window

One of the objectives of this project is to use VCSELs to improve the accuracy of the glucose prediction to a more clinically relevant accuracy and increasing the signal-to-noise ratio. To achieve this objective, it is required to determine the number of absorption spectra segments that can reveal enough information to predict the glucose concentration. As a result, the number of VCSELs required in the wavelength range of 1.0\(\mu m\) to 1.7\(\mu m\), will be determined. The wavelength windows, which contain the glucose information, are selected to be centered on the glucose absorption peaks; namely these peaks are at 1126, 1408, 1536, and 1689 nm [44]. As a result, the wavelength windows are approximately 1126 ± 3.5 nm, 1408 ± 3.5 nm, 1536 ± 3.5 nm, and 1689 ± 3.5 nm. Figure 5.5 shows these selected 7 nm wide

![Normalized Black Body Radiation for color temperature of 2960K.](image.png)

Figure 5.3: Normalized Black Body Radiation for color temperature of 2960K.
wavelength segments. Later, in the subsequent sections, another method of choosing intervals will be discussed.

The partial least square (PLS) technique was applied to these selected spectral bands. Using all four windows, the PLS analysis is optimized with three principle components. It is noted that increasing the number of principle components further does not improve the model for glucose prediction, since noise gets incorporated into the model.

As it can be observed from Figure 5.5, the amplitude of absorption spectra does not change linearly with glucose concentration. For example, Figure 5.6 shows power vs. wavelength at one single wavelength, which illustrates this point further. This special characteristic in the correlation between the absorption spectra and glucose concentrations calls for a need in the multivariate regression analysis.
Chapter 5. Spectroscopy for Glucose Sensing Using White Light

Figure 5.5: Glucose Absorption Spectrum for the 7 nm wide wavelength segments: 1126±3.5nm, 1408±3.5nm, 1536±3.5nm, 1689±3.5nm. Inset: zoom-in for the 1536±3.5nm segment.

5.3 Absorption Spectroscopy

Solutions of glucose in distilled water for various concentrations have been used to construct absorption spectra for the wavelengths 1.0µm to 1.7µm. The absorption spectra are collected and subtracted from the absorption spectrum of the pure water; hence the resultant spectra contain mostly the absorption characteristics of glucose with added noise and less water. Figure 5.5 shows the average of these resultant spectra for six different concentrations in the chosen wavelength segments discussed in previous section. As can be observed from the figure (inset), there is no obvious linear relationship between the absorption spectra at these glucose absorption bands and glucose concentration; this issue has also been discussed in previous reports dealing with optical glucose measurements [47, 58]. As a result, a multivariate analysis such as partial least square (PLS) is required to find the relationship between the measured absorption spectrum and glucose concentration.

The concentrations of the glucose solutions are made high since this project is at the stage of preliminary investigation. In the future sec-
Chapter 5. Spectroscopy for Glucose Sensing Using White Light

Figure 5.6: Power vs. Concentration at one single wavelength in the absorption spectra. This shows that the amplitude of the spectra does not change linearly with concentration.

5.4 Results

Absorption spectra are collected in two separate experiments. In each experiment, 16 sets of spectra for each concentration were recorded. The first data set (i.e. 16 sets of absorption spectra) was used to build the PLS model for predicting glucose concentrations and for defining the 95% confidence interval. The second data set was used to validate the model and estimate the accuracy of the technique.

Figure 5.7 shows the results of building the PLS model from the average of the first data set. The same 16 sets of spectra are used to define the 95% confidence interval for predicted glucose concentration as shown by error bars. These error bars are an indication of the noise and provide an
estimation of the algorithm precision. In addition, on the same figure, the same 16 sets of spectra are plotted to illustrate the algorithm precision, when the input is the original data set. This figure (Figure 5.7) shows that a good agreement between measured and predicted values was obtained.

Next, the model is validated with the second data set, to test for accuracy. This is done by comparing the predicted concentrations from the second experiments with the 95% confidence interval defined by the first data set. The results, shown in Figure 5.8, show that 96% of the predicted concentrations from the second data set fall within the 95% confidence interval. Finally, averaging of multiple spectra can be employed to increase the precision of the glucose prediction.

5.5 Discussion

The results show that the chosen wavelength segments can be used to predict glucose concentration. Moreover, the results show that the model built from
Figure 5.8: Predicted glucose vs. measured glucose, using the 2nd data set. 96% of the predictions fell within the 95% confidence intervals, as defined by the 1st data set. The green dashed line indicates boundary for 10% variation in glucose concentration.
one set of data can successfully make predictions on different data sets with a good accuracy. As a result, the PLS technique applied on a small number of segments (4), with segments spanning 7 nm, can be used to predict glucose concentration. Additional improvements are needed to increase the accuracy of the prediction. Specifically, using lasers on the specific windows, which are shown to be useful for predicting glucose concentration, will improve the results by reducing the noise.

5.6 Further Improvements on Data Analysis and PLS

Several improvements and changes have been made to the analysis of data as explained in the preceding sections. In the new set of analysis, the absorption spectra of distilled water are included representing zero glucose concentration in the analysis rather than subtracting the water spectra from solutions of various glucose concentrations. All the absorption spectra are included representing a sample in either calibration set or validation process. These absorption spectra are not averaged for calibration of the model, rather they all fed into the model. This will allow the model to consider more background variations of the system when making the model; hence, the model will be more robust to those variations that exist naturally in the system.

The 7 nm wide wavelength windows from the absorption spectra, in the analysis of data in the preceding sections, were chosen according to the knowledge of existed glucose absorption peaks in 1000 nm to 1700 nm band. However, multivariate analysis considers many aspects when building a regression model such as having a reference baseline, or considering drifts and many more. As a result, the model may be more robust, if for example, the change in the slope of the absorption band is considered rather than the absolute amplitude change at the absorption peak.

The PLS toolbox has a build-in function, called Interval PLS (iPLS), that searches for best variable or combination of variables (Wavelength in-
Chapter 5. Spectroscopy for Glucose Sensing Using White Light

tervals) that can best predict the variable of interest. This function gets the desired number of intervals and the interval width, which can be a single variable or a window of adjacent variables. In the improved analysis presented in this section, \textit{iPLS} is used to find four (4) wavelength intervals of 7 nm wide. Furthermore, the number of intervals requested from \textit{iPLS} was increased to achieve an estimated prediction error close to the estimated prediction error when using all variables or complete spectrum, excluding the strong water absorption band.

The strong water absorption in the spectra influences the regression model as well as affecting the preprocessing results. This strong water absorption band can dominate the effect of the signature of other components in the solution or mislead the result of the analysis. For example, in the simple solution of glucose and distilled water, as concentration of glucose increases, there is less water in the fixed volume of solution, resulting in less absorption of light by water at the wavelengths, where water absorption is strong. Since the amount of water in fixed volume is strongly correlated with the glucose concentration, this wavelength region may result in unrealistic optimum prediction that is not taken from glucose signature. Hence, eliminating the wavelengths with strong water absorption from absorption spectra can result in a more reliable and stable analysis. Figure 4.16 demonstrates elimination of wavelength window corresponding to strong water absorption (1419 nm to 1519 nm). Considering the set of spectra excluding this wavelength window will assure that the result of analysis is not biased by the stronger effect of water absorption in favor of glucose prediction.

The new improved analysis is performed for two experiments, one with samples of higher glucose concentrations same as samples used for analysis in the preceding sections, and the other is using samples with lower concentrations of glucose closer to the clinical concentration range. For the data from each of these experiments, cross-validation was performed to discuss the choice of appropriate preprocessing method. The optimum choice of preprocessing was discussed in chapter 4, \textit{Data Analysis and Chemometrics}, under \textit{Finding the Appropriate Preprocessing Method}. 

77
5.6.1 Improved Analysis on the Samples Analyzed in Previous Sections

The samples for glucose concentrations from 500 to 20000 mg/dL (27 to 1111 mmol/L), which were analyzed in the preceding sections, are analyzed in this section again with the improvements considered in this section. As it was concluded from Chapter 4 (Figure 4.17), Orthogonal Signal Correction (OSC) along with mean centering is a promising customized preprocessing method resulting in a relatively lower RMSECV value with less number of required Latent Variables (LVs).

To find the intervals that can best predict the glucose concentration, iPLS is run. The iPLS is asked for four (4) intervals of 7 nm wide each to be comparable with the intervals chosen in the analysis of the previous sections corresponding to the glucose absorption peak.

To make sure that OSC along with Mean Centering is the best choice of preprocessing on the intervals as well, various preprocessing methods were applied to the intervals chosen with iPLS, which has chosen intervals from the original spectra without any preprocessing. Figure 5.9 shows RMSECV vs. LVs as a result of applying cross-validation with various preprocessing to the intervals chosen. The preprocessing methods, as before, are mean centering, autoscale, SNV, SNV along with mean centering, MSC, MSC along with mean centering, OSC, and OSC along with mean centering. It is confirmed from this figure that OSC along with mean centering is also a promising method when applied to the selected intervals for calibration and cross-validation. As a result, this combination of preprocessing (OSC and Mean Centering) will be applied to all the chosen combinations of intervals and spectra for calibration and cross-validation purposes.

To further confirm the suitability of this customized preprocessing method, Figure 5.10 compares various preprocessing methods on the intervals chosen corresponding to the glucose absorption peaks. Again, OSC along with mean centering is the best among the presented ones. This figure also shows that with mean centering as preprocessing method, having five (5) LVs results in the lowest estimate of the prediction error.
Figure 5.9: RMSECV vs. LV used to compare and determine customized preprocessing method for the intervals chosen, using iPLS from absorption spectra of glucose samples with high concentrations, where water absorption band is eliminated, without any preprocessing. RMSECV helps to determine the method that results in relatively lower cross-validation error of prediction with lower number of LVs. This figure confirms that OSC with mean centering is the most optimized preprocessing method for calibration even with the intervals chosen from spectra with no preprocessing.
Figure 5.10: RMSECV vs. LV used to compare and determine customized preprocessing method for the wavelength intervals corresponding to glucose absorption peaks in the spectra of glucose samples with high concentrations. RMSECV helps to determine the method that results in relatively lower cross-validation error of prediction with lower number of LVs. This figure confirms that OSC with mean centering is the most optimized preprocessing method for calibration even with the intervals chosen corresponding to glucose absorption bands.
Chapter 5. Spectroscopy for Glucose Sensing Using White Light

The set of absorption spectra can also be preprocessed before being supplied to iPLS. Hence, the set of absorption spectra have been preprocessed with the following methods before evoking iPLS, which will find the best four intervals. The \( Y \) block, vector of concentration, is mean centered in all cases unless stated otherwise.

- Mean Centering (labeled as \( \text{MeanC} \))
- AutoScale
- SNV : Standard Normal Variate
- OSC : Orthogonal Signal Correction
- No preprocessing for both \( X \) and \( Y \) block (labeled as \( \text{None} \))
- OSC and Mean Centering (labeled as \( \text{OSC+MeanC} \))
- OSC and autoscale (labeled as \( \text{OSC+auto} \))

The intervals chosen using iPLS as mentioned above, as well as the intervals chosen according to the glucose absorption peaks (the intervals that were analyzed in preceding sections) and the complete absorption spectra excluding water absorption band (labeled as \( \text{EXwater} \)), are preprocessed with OSC along with mean centering. Preprocessing methods of OSC along with mean centering is used for calibration as even the intervals chosen without any preprocessing results in lower RMSECV when PLS model is calibrated with the preprocessed data using OSC and mean centering method as observed in Figure 5.9. Cross-validation using PLS is then applied to these data and the resulted RMSECV vectors are plotted in Figure 5.11 to compare the suitability of the intervals chosen for glucose prediction. In other words, the RMSECV errors are compared to determine which set of intervals has the best correlation with glucose concentration.

The illustrated comparison in Figure 5.11 shows that choosing intervals from the data that has been preprocessed with OSC and mean centering results in the lowest error when calibrating with the same preprocessing method.
Figure 5.11: RMSECV used to compare and determine the method of choosing intervals that results in lowest RMSECV error. Cross-validation was performed, with OSC and mean centering as the preprocessing, on various selection of wavelengths. ‘EXwater’ refers to the full absorption spectra excluding wavelength window corresponding to strong water absorption band (1419 to 1519 nm). ‘int’ refers to the intervals chosen, in the previous section, corresponding to glucose absorption bands. The rest refers to the intervals chosen using iPLS from the preprocessed data that is indicated in the parenthesis. There are four wavelength intervals in all these cases.
Now that the method of choosing four (4) intervals as well as the pre-processing method for calibration is determined, the number of intervals chosen by iPLS can be increased to determine the number of intervals that are required to result in the lowest prediction error possible with this data. The complete absorption spectrum with excluded water absorption band, representing all of the available predictor variables, is used as an indication for the lowest achievable prediction error with this data set. Figure 5.12 compares RMSECV errors for predictions with four, five, six, and seven intervals with the RMSECV errors using all the predictor variables. It is observed that six and seven intervals result in estimated prediction errors close to the estimation of the best possible one.

![Figure 5.12: RMSECV used to determine the number of intervals required to have prediction with errors as low as prediction errors using complete spectrum excluding water peak. 'int' refers to the intervals chosen, in the previous section, considering glucose peaks in the spectra. The best prediction is achieved using whole spectrum without water peak. Number of intervals is increased to get closer to the errors when using whole spectrum without water peak. The number after iPLS in the legend shows the number of intervals chosen for the analysis.](image-url)
Glucose Prediction Using PLS Regression

To compare with the results presented in the preceding sections a gradual improvement of the analysis is presented here. First, the same set of data, choosing the same intervals corresponding to glucose absorption peaks, with the same preprocessing method (mean centering) and same number of Latent Variables (3 LVs) is analyzed considering the following improvements or changes in the analysis: 1) The absorption spectra for distilled water is included in the model as a sample with zero concentration rather than being subtracted from absorption spectra of the various glucose concentrations. 2) The absorption spectra are not averaged, but they are all used in either the calibration or validation process as separate samples.

For this set of analysis, there are two separate sets of absorption spectra. Each set contains 16 spectra for each glucose concentration (7 x 16 samples). One set is used for the calibration of the PLR regression model and the other set for validation. The result of this validation is presented in figure 5.13. The two dotted lines draw the boundary for the 10% variation in the concentration.

The next improvement is in determining the number of LVs that can result in lowest estimated error of prediction. According to Figure 5.10, when these selected intervals (according to the glucose absorption peaks) cross-validated using PLS with mean centering, the resulted RMSECV vector has a minimum at five (5) LVs. This means that the PLS regression model, which is built on these intervals preprocessed with mean centering, results in lowest RMSECV error when five (5) LVs are used for calibration of the model. Figure 5.14 shows the result of glucose prediction using the PLS regression model built with five (5) latent variable and mean centering as preprocessing. This figure can be compared with Figure 5.13 to observe the improvements that results from choosing an optimized number of latent variables for model calibration.

The next improvements is to use this same set of intervals, but preprocessed with OSC and mean centering, using the optimum number of LVs that give the lowest estimate of prediction error. The required number of
Figure 5.13: Glucose predictions using PLS with 3 LVs and mean centering as the preprocessing on the intervals corresponding to glucose absorption bands. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The green dashed line represents the boundary for the 10% variations in glucose concentrations. The RMSEP value for this prediction is 467.1316 mg/dL, equivalent to 25.95 mmol/L.
Figure 5.14: Glucose predictions using PLS with 5 LVs and mean centering as the preprocessing on the intervals corresponding to glucose absorption bands. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The green dashed line represents the boundary for the 10% variations in glucose concentrations. The RMSEP value for this prediction is 296.1313 mg/dL, equivalent to 16.45 mmol/L.
LVs can be obtained from Figure 5.10, where it shows that two (2) latent variables gives the minimum RMSECV error, for the case when this set of interval is cross-validated using PLS with OSC and mean centering as preprocessing methods. To demonstrate the effect of this improvement on glucose prediction, PLS regression model is built on the same set of data as before but with using the new preprocessing methods (OSC and Mean Centering) and two (2) latent variables. Figure 5.15 shows glucose prediction using this new model.

![Figure 5.15: Glucose predictions using PLS with 2 LVs and OSC along with mean centering as the preprocessing on the intervals corresponding to glucose absorption bands. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The green dashed line represents the boundary for the 10% variations in glucose concentrations. The RMSEP value for this prediction is 284.7602 mg/dL, equivalent to 15.82 mmol/L.](image)

Furthermore, the analysis is improved by using \textit{iPLS} function in PLS toolbox to find a set of intervals that shows the best correlation with glucose concentrations. As explained before, iPLS has been used to find four (4) in-
tervals from the spectra preprocessed with OSC and mean centering. Figure 5.16 shows the placement of these intervals in the absorption spectrum.

Figure 5.16: The 4 intervals chosen for glucose prediction from samples with higher concentrations (500 to 2000 mg/dL). The green solid vertical lines represent the glucose absorption peaks as given in the literature. The dashed lines show the chosen intervals, where each interval is 7 nm wide. The spectra have been preprocessed with OSC and mean centering before interval selection process. The wavelengths corresponding to these intervals are: 1319.2 to 1326.9 nm; 1411.6 to 1419.3 nm; 1519.4 to 1527.1 nm; and 1561.4 to 1569.1 nm.

It was also concluded that OSC and mean centering are the best combination of preprocessing when calibrating the PLS regression model. Cross-validation for this case shows that the model built using four (4) latent variables resulted in lowest RMSECV error. Figure 5.17 shows glucose predictions using the PLS regression model built using this new set of intervals, chosen by iPLS, preprocessed with OSC and mean centering using four latent variables. The RMSEP resulted from this model is 11.7588 mmol/L, which is improved from the previous model.
Figure 5.17: Glucose predictions using PLS with 4 LVs and OSC along with mean centering as the preprocessing on the 4 intervals chosen with iPLS. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The green dashed line represents the boundary for the 10% variations in glucose concentrations. The RMSEP value for this prediction is 211.6578 mg/dL, equivalent to 11.7588 mmol/L.
In addition, increasing the number of intervals improves the prediction error significantly. Figure 5.19 shows glucose predictions using the PLS regression model built on combination of seven intervals, preprocessed with OSC and mean centering using four latent variables. The RMSEP error is improved to 7.5785 mmol/L with this model. Figure 5.18 shows place of these intervals on the spectrum. The green solid vertical lines represent the glucose absorption peaks as given in the literature. The dashed lines show the chosen intervals, where each interval is 7 nm wide.

Figure 5.18: The 7 intervals chosen for glucose prediction from samples with higher concentrations (500 to 20000 mg/dL). The green solid vertical lines represent the glucose absorption peaks as given in the literature. The dashed lines show the chosen intervals, where each interval is 7 nm wide. The spectra have been preprocessed with OSC and mean centering before interval selection process. The wavelengths corresponding to these intervals are: 1058.8 to 1066.5 nm; 1168 to 1175.7 nm; 1268.8 to 1276.5 nm; 1319.2 to 1326.9 nm; 1411.6 to 1419.3 nm; 1519.4 to 1527.1 nm; and 1561.4 to 1569.1 nm.
Figure 5.19: Glucose predictions, on combination of 7 intervals, using PLS with 5 LVs and OSC along with mean centering as the preprocessing on the 7 intervals chosen with iPLS. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The green dashed line represents the boundary for the 10% variations in glucose concentrations. The RMSEP value for this prediction is 136.4129 mg/dL, equivalent to 7.5785 mmol/L.
5.6.2 Analysis on the Glucose Samples with Lower Concentrations

A new set of absorption spectra was collected with the same experimental set-up explained before but with lower concentrations of glucose. The concentrations of samples in this experiment were from 100 to 400 mg/dL and water (0:100:400 mg/dL) or from zero to 22.22 mmol/L.

To demonstrate and confirm that OSC along with Mean Centering is also the best choice for preprocessing of chosen discrete intervals for the spectra of low concentration samples, various preprocessing methods applied to the intervals corresponding to glucose absorption peaks as an example. Figure 5.20 shows plot of RMSECV errors resulted from applying cross-validation to these intervals preprocessed with various preprocessing methods. The preprocessing methods, as before, are mean centering, autoscale, SNV, SNV along with mean centering, MSC, MSC along with mean centering, OSC, and OSC along with mean centering. It is concluded from this figure that OSC along with mean centering is also a promising method when applied to the selected intervals, of absorption spectra of samples with lower concentrations, for calibration and cross-validation. As a result, when analyzing lower concentration data in this section, this combination of preprocessing (OSC and Mean Centering) will be applied to all the chosen combinations of intervals and spectra for calibration and cross-validation purposes.

Furthermore, after concluding on the choice of preprocessing methods for calibration of the model, this preprocessing method (OSC and Mean Centering) is applied to various combinations of intervals and spectra to compare and find the best set of intervals for glucose predictions. Similar to the case with higher concentration data, the built-in function iPLS in PLS toolbox is used to find intervals that show best correlation with glucose concentration. The methods used for preprocessing absorption spectra before applying iPLS are mean centering, SNV, OSC along with mean centering, while the concentration vector is mean centered in all the above cases, as well as having no preprocessing on set of absorption spectra and concentration vector. Figure 5.21 shows RMSECV resulted from PLS cross-validation.
Figure 5.20: RMSECV vs. LV used to compare and determine customized preprocessing method for the intervals corresponding to glucose absorption peaks in samples with lower glucose concentrations. RMSECV helps to determine the method that results in relatively lower cross-validation error of prediction with lower number of LVs.
applied to the selected intervals explained above as well as applied to the intervals corresponding to glucose absorption peaks and the complete spectra excluding water absorption band.

In addition, iPLS is evoked with increased number of intervals on the same preprocessed data to investigate the number of intervals required to result in the lowest achievable accuracy with this data set or to get to the accuracy required by the application. Figure 5.21 also presents improved RMSECV error of prediction when the number of interval chosen is increased to five (5) and six (6).

It can be concluded from Figure 5.21 that the intervals chosen from preprocessed data with OSC and mean centering is still a promising best option. Furthermore, applying cross-validation on six (6) intervals results in the same estimated prediction error as for the data using all the variables (complete absorption spectra excluding water absorption band), when using five (5) Latent Variables (LVs).

**Glucose Prediction Using PLS Regression**

PLS is used to calibrate a regression model for glucose prediction, using absorption spectra of the samples with lower concentrations (0:100:400 mg/dL), similar to the analysis performed for absorption spectra of samples with higher concentrations. Same as before, there are two sets of data for this set of samples. Each set contains 16 spectra for each glucose concentration. One set is used for calibration of the model and the other set is used for validation.

In this part, the PLS regression model will be built on two sets of intervals for comparison purposes. One is the set of intervals chosen according to the glucose absorption peaks. The other is the set of intervals chosen by PLS toolbox using iPLS function.

First, PLS is applied to build a regression model with the wavelength intervals corresponding to glucose absorption peaks. These windows of absorption spectra are preprocessed with mean centering to be used by PLS with 5 latent variables in building the regression model. Then, a new sep-
Figure 5.21: RMSECV used to determine the best method of choosing intervals as well as required number of intervals from the samples with lower concentrations. ‘EXwater’ refers to the full absorption spectra excluding wavelength window corresponding to strong water absorption band (1419 to 1519 nm). ‘int’ refers to the intervals chosen, in the previous section, considering glucose peaks in the spectra. The best prediction is achieved using whole spectrum without water peak. Number of intervals is increased to get as close to the errors when using whole spectrum without water peak. The preprocessing method indicated in the parenthesis for iPLS in legend refers to the preprocessing of the data that iPLS has been applied to. The number after iPLS in the legend shows the number of intervals chosen for the analysis.
parate set of data is used to validate the model. The prediction error (RMSEP) of this model is 1.4283 mmol/L, which is a bit more than acceptable error. Figure 5.22 presents the result of model validation. The two dashed red lines, parallel to the center line, shows the known concentration $\pm 1$ mmol/L, which is the required clinical accuracy for glucose prediction.

![Graph showing glucose predictions using PLS with 5 LVs and mean centering as the preprocessing on the set of intervals chosen from the glucose absorption peaks. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The dashed red line defines the boundary for the predictions to be clinically acceptable ($\pm 1$ mmol/L); and the dashed green line shows the 10% boundary. The RMSEP value for this prediction is 25.7093 mg/dL, equivalent to 1.4283 mmol/L.]

Next PLS regression model is built with the same set of intervals but preprocessed with OSC and mean centering. The PLS regression model is calibrated using three (3) latent variables, since three latent variables result in the minimum RMSECV error when cross-validated using PLS (illustrated in Figure 5.20). The prediction error (RMSEP) using this model improves to 0.8666 mmol/L, which is in the acceptable accuracy range of $\pm 1$ mmol/L.
Figure 5.23 shows the result of predicting glucose concentrations, from the other set of spectra, using this regression model.

Figure 5.24: Glucose predictions using PLS with 3 LVs and OSC along with mean centering as the preprocessing on the set of intervals chosen from the glucose absorption peaks. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The dashed red line defines the boundary for the predictions to be clinically acceptable (± 1 mmol/L); and the dashed green line shows the 10% boundary. The RMSEP value for this prediction is 15.5990 mg/dL, equivalent to 0.8666 mmol/L.

The other set of four intervals, which will be used in this section for glucose prediction, was found by iPLS function. The absorption spectra is preprocessed with OSC and mean centering before using iPLS to find the four intervals that best correlate with glucose concentrations. Figure 5.24 shows the placement of these intervals in the absorption spectrum.

These four intervals are then preprocessed using OSC and mean centering to be used by PLS in building a regression model with five (5) latent variables. The choice of five latent variables is taken from Figure 5.21, which
Figure 5.24: The 4 intervals chosen from lower concentration samples (0:100:400 mg/dL) for glucose prediction. The green solid vertical lines represent the glucose absorption peaks as given in the literature. The dashed lines show the chosen intervals, where each interval is 7 nm wide. The spectra have been preprocessed with OSC and mean centering before interval selection process. The wavelengths corresponding to these intervals are: 1025.2 to 1032.9 nm; 1084 to 1091.7 nm; 1218.4 to 1226.1 nm; and 1679 to 1686.7 nm.
shows that after five latent variables the estimated prediction error does not improve significantly. The separate set of absorption spectra, which was designated for validation purposes, was used to validate this regression model as presented in Figure 5.25. Predictions being in the range defined by the two dashed lines, shows that almost all the predictions are clinically accurate. The calculated prediction error (RMSEP) for this model being 0.3874 mmol/L also shows that the predictions on average have errors less than ±1 mmol/L.

Figure 5.25: Glucose predictions using PLS with 5 LVs and OSC along with mean centering as the preprocessing on the set of four (4) intervals found by iPLS function. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The dashed red line defines the boundary for the predictions to be clinically acceptable (± 1 mmol/L); and the dashed green line shows the 10% boundary. The RMSEP value for this prediction is 6.9730 mg/dL, equivalent to 0.3874 mmol/L.

In addition, iPLS was asked to find set of more than four intervals that could best predict glucose concentrations in the same way. This will help identify the number of intervals that are required to achieve the desired
accuracy. Specifically, it was used to find five intervals of 7 nm wide each from the absorption spectra preprocessed with OSC and mean centering. These chosen intervals are then used to build a PLS regression model. Figure 5.26 shows the placement of these intervals in the absorption spectrum.

![Graph showing absorption spectrum with intervals marked]

Figure 5.26: The 5 intervals chosen from lower concentration data (0:100:400 mg/dL) for glucose prediction. The green solid vertical lines represent the glucose absorption peaks as given in the literature. The dashed lines show the chosen intervals, where each interval is 7 nm wide. The spectra have been preprocessed with OSC and mean centering before interval selection process. The wavelengths corresponding to these intervals are: 1025.2 to 1032.9 nm; 1084 to 1091.7 nm; 1218.4 to 1226.1 nm; 1344.4 to 1352.1 nm; and 1679 to 1686.7 nm.

Therefore, a PLS regression model is calibrated using five latent variables on these intervals preprocessed with OSC and mean centering. The result of glucose predictions using this model is presented in Figure 5.27, which shows that all the predictions are within the acceptable clinical accuracy. The prediction error is improved to 0.2837 mmol/L with this model.

The function, iPLS, is applied to one set of data to find the intervals
Figure 5.27: Glucose predictions using PLS with 5 LVs and OSC along with mean centering as the preprocessing on the set of five (5) intervals found by iPLS function. The plot represents predicted glucose concentration using the model vs. known concentration of those samples. The dashed red line defines the boundary for the predictions to be clinically acceptable (± 1 mmol/L); and the dashed green line shows the 10% boundary. The RMSEP value for this prediction is 5.1071 mg/dL, equivalent to 0.2837 mmol/L.
that are best correlated with glucose concentration. One might wonder if that selection is valid for all the similar sets of spectra. In other words, if the intervals chosen using one set of spectra can actually be selected from another set of spectra to build the model with. This has been examined and the conclusion is that, once the intervals were identified, there were no noticeable change in predictions when choosing calibration intervals from either set, and validating with the other.

5.7 Conclusion

Four wavelength segments, each 7nm wide, have been identified in the wavelength range of 1.0µm to 1.7µm as promising windows that contain useful glucose information. As a result, four VCSELs in this range can be used to predict glucose concentration. The PLS signal processing technique has been successfully applied to predict glucose based on the optical spectra. In future work, we will implement a system using VCSELs to demonstrate VCSEL-based glucose measurements of whole blood. This work is published in Conference Proceedings [1].

However, a new set of analysis was performed on the same set of data used in the Conference Proceedings paper, as well as on data from samples with more clinically relevant concentrations. In this analysis, the most promising preprocessing method is identified for both experiments, with samples with higher and lower concentrations, to be OSC along with mean centering. The calibration of the model requires less latent variables to achieve same or lower prediction error with these preprocessing methods.

One of the improvements made to the analysis was feeding individual samples, as they are available, to the model for calibration. This made the model more robust and accurate when predicting glucose concentrations using a different set as indicated by the RMSEP error. Averaging samples eliminates some of the variations that normally exist in the samples, and calibrating model with averages cannot possibly consider those variations to be robust against them. Whereas, when the samples are fed individually, the model is aware of their existence and adapts itself to those conditions. In
addition, the spectra for distilled water are fed into the model as a solution with zero concentration instead of being subtracted from other solutions. Having all these individual spectra in the model, again, helps the model to estimate the variations that can generally exist in the spectra, even in the absence of any glucose, and to have a tolerance to these variations.

The best selection of intervals is not always at the absorption peak. Combination of many conditions can define the variations that correlate the most with the interested variable. For example, a point of reference that does not change with glucose variations is required for regression models, or sometimes considering the change in the slope of the absorption band may result in a more robust model than the absolute amplitude change at the absorption peak, especially where glucose has broader absorption bands.

The number of intervals required to achieve the best accuracy achievable with the given set of data are six and seven intervals for samples with lower and higher concentrations, respectively. The set of data from samples with higher concentrations consist of samples from 500 to 2000 mg/dL, which is higher than what normally exist in the human body. The concentration range of these samples is very wide, with fewer concentrations in between that can reduce the accuracy and predictive power of the model. Closer calibrating points in smaller range of concentrations can increase the accuracy.

The predictions after applying these improvements are either in or very close to the acceptable clinical range of ± 1 mmol. The RMSEP error, which is the most accurate (close to real) indication for prediction error, is down to 0.3874 mmol/L using four intervals, and to 0.2837 mmol/L using five intervals. This promising result opens the door for improvements towards experiments using blood serum, where there are more interfering particles in the samples.
Chapter 6

Spectroscopy for Glucose Sensing Using VCSELs

Using a laser as the light source instead of a white light can provide advantages such as higher signal-to-noise ratio in the absorption spectrum [58], leading to a more accurate estimation of glucose concentration. The two wavelength regions that are commonly used for glucose monitoring are: the first-overtone band (1560 – 1850 nm), and the combination band (2080 – 2325 nm), where glucose has numerous absorption bands and water has relatively lower absorption [60, 66, 67]. The absorption is stronger in the combination band, and is the most desirable wavelength region for glucose measurements using absorption spectroscopy.

There are two major wavelength bands that have shown promising correlation between their absorption spectra and glucose concentrations, one of which is 2.0-2.5 µm [60, 66, 67], and this band has been shown to be very effective for predicting glucose concentrations [25]. Recently, appropriate wavelength semiconductor laser sources have been developed. In this experiment, the lasers used are not only the first realized electrically pumped VCSELs at a wavelength of 2.3 µm, but also the longest wavelength achieved so far for any InP-based interband laser [54].

The feasibility of glucose prediction using 2.3 µm VCSELs and application of Partial Least Squares (PLS) technique are demonstrated in this chapter.

This chapter includes a brief description of Vertical Cavity Surface Emitting Lasers (VCSELs), an explanation of the apparatus of the experiment, followed by the experimental description, analysis and results of the exper-
iment, conclusion and further improvements on data analysis and the PLS regression models.

6.1 Vertical Cavity Surface Emitting Lasers (VCSELs)

Vertical Cavity Surface Emitting Lasers (VCSELs) are semiconductor lasers with dimensions of about $50 \mu m \times 50 \mu m$ and with low power consumption in the order of $10 \, mW$. VCSELs operating within a small spectrum with high power spectral density provide a higher signal-to-noise ratio. VCSELs have narrow circular beam suitable for direct fiber coupling, and provide single mode operation with vertical cavity. They can be thermally tuned by about $7 \, nm$ by varying their bias current. In addition, they have low cost and small packaging capability that makes them suitable for an implantable device. Due to these advantages and various other advantages, VCSELs have attracted the attention of researchers in this area.

Figure 6.1 shows the structure of a vertical cavity surface emitting lasers (VCSELs).

![Figure 6.1: Structure of the VCSEL. [68]](image)

Optical output characteristics of VCSELs are explained by their temperature dependent characteristics. These thermal effects on the optical characteristics of lasers can be viewed in two basic aspects: 1) the shift of the Fabry-Perot (FP) wavelength caused by thermally induced refractive index change and 2) the thermally induced change of the gain spectrum.
Chapter 6. Spectroscopy for Glucose Sensing Using VCSELs

However, in addition to these basic aspects, temperature affects threshold current in lasers.

VCSELs have a very small cavity and hence normally have only one Fabry-Perot (FP) wavelength within the gain spectrum, which is not necessary to match with the maximum gain. As the temperature increases, the gain spectrum shifts to longer wavelength or lower energy region, while its amplitude (or value of the gain coefficient $\gamma$) decreases and the spectrum gets narrower. For VCSELs, both the gain spectrum and the FP resonance shift to longer wavelength with increasing temperature. However, the temperature dependence of the FP resonance is determined by the temperature dependence of the refractive indices, and consequently temperature dependence of effective cavity length or optical path; whereas the gain spectrum has the temperature dependence of the energy band gap. Since the band gap shifts at a rate five times faster than the refractive index, a structure perfectly matched at one temperature will become mismatched as the temperature is varied [69]. This can make the output power of a VCSEL decrease as the temperature increased. Figure 6.2 shows output of the VCSEL, with serial number FAT 720, manufactured by the Bandwidth 9 Inc.

The source that powers the VCSEL is set to be the current source. As the bias (or drive) current is increased, the voltage drop across the source changes as illustrated in Figure 6.3, which is the I-V curve of the VCSEL FAT 720. The power supplied to the VCSEL at each current point is $P=I.V$.

An Optical Spectrum Analyzer (OSA) can be used to observe the spectrum of the VCSEL, namely output power vs. wavelength, including its lasing wavelength. Figure 6.4 shows the optical spectrum of the VCSEL. As the bias current increases, the lasing wavelength shifts towards longer wavelengths. The direction of the arrow indicates the increasing bias current.

From Figure 6.4, the maximum power for each bias current is extracted and is plotted vs. bias current, the result is demonstrated in Figure 6.5.

Like any other electronic device, VCSELs have a lifetime, which can be shortened if biased past its peak point. For example, the VCSEL that was used for investigation of the property of VCSELs, FAT 720, was biased passed its peak point as shown in Figure 6.2. The process of gradual
Figure 6.2: LI curve of the VCSEL. The decrease in output power is the result of combination of effects of gain spectrum shifting faster than FP lasing wavelength and gain coefficient decreasing in magnitude. Note that the coupling efficiency is included in this measurement; and this is one reason for low power measurements in this graph.

Figure 6.3: I-V curve of the VCSEL.
Figure 6.4: Effect of changing bias current on the optical spectrum of the VCSEL. The direction of arrow shows increasing bias current.
Figure 6.5: Effect of changing bias current on the lasing wavelength of the VCSEL. This is the plot of the wavelengths, corresponding to maximum power in the optical spectrum of the VCSEL, vs. bias current.

degradation of this VCSEL is illustrated in Figure 6.6.

As the bias current increases, the temperature of the laser increases, which in turn shifts the lasing wavelength to longer wavelengths. Hence, the tuning of these VCSELs are achieved using temperature effects of the increased bias current resulting in thermally tunable VCSELs.

6.2 Apparatus

This section includes description of apparatus of the experiment.

6.2.1 PbS Pre-Amplified Detector

PbS pre-amplified detector (ThorLabs PDA30G) is a voltage amplified detector that converts an optical signal to an electrical signal and is designed for the detection of modulated light signals ranging from 0.2 to 1 kHz. This detector has the wavelength range of 1.0 to 2.9 $\mu m$ with the peak wavelength
Figure 6.6: Effect of biasing VCSEL past its peak power. The LI curves gradually decrease and VCSEL death occurs.

at 2.2 µm. It supports the bandwidth of 0.2 to 1kHz, with NEP (noise equivalent power) of $1.5 \times 10^{-11}$ W/√Hz; detectivity of $(D^*)$ is $2 \times 10^{10}$ cm√Hz/W; and a typical peak sensitivity of $7.5 \times 10^5$ V/W for high impedance load. Its rise time is 250 µs. The light to voltage conversion, with 22 °C ambient temperature, can be estimated as shown in equation 6.1.

$$Output(V/W) = sensitivity(V/W) \times relativeresponsivity(\%) \quad (6.1)$$

Figure 6.7 shows the relative responsivity of this detector, and Figure 7.6 shows the detectivity of PbS detectors and how close it is to the ideal case.

### 6.2.2 Lock-in Amplifier

A Lock-in Amplifier (SR810, made by Stanford Research Systems) is used to measure small AC signals in the order of as low as nanovolts. It provides a DC output, in RMS volts, which is proportional to the AC signal under investigation. In the lock-in amplifier, the component of a signal at a specific
reference frequency and phase is singled out, whereas noise and interfering
signals at frequencies other than the reference frequency are filtered out,
and hence they do not affect the measurement reading.

The lock-in amplifier multiplies the input signal by a pure sine wave at
the reference frequency. The result of this multiplication is a DC output
signal representing the component of the input signal that has a frequency
equal to reference frequency (locked to the reference frequency). Then the
signal is passed through a low pass filter, which removes all the components
at other frequencies or AC fluctuations.

Noise is a varying signal at all frequencies. Hence, lock-in amplifier
removes noise at all frequencies other than the reference frequency, by low
pass filter. The width of the filter can be adjusted by the time constant
variable. Increasing the time constant results in a more steady output;
however, RC filters require five time-constants to stabilize to their final value.
Therefore, increasing the time constant would slow down the measurement
process. The signal-to-noise ratio is improved proportional to the square
root of the time constant of the filter.

The lock-in amplifier displays a DC signal, proportional to the AC signal

Figure 6.7: Relative response of PbS pre-amplified detector [From
www.Thorlabs.com].
read from the detector, in volts.

### 6.2.3 2.3 µm VCSEL

These lasers, fabricated by our collaborator Prof. Amann in Germany, are not only the first realized electrically pumped VCSELs at a 2.3 µm wavelength, but also are the longest wavelength achieved so far for any InP-based interband laser [54]. The small size of these lasers (less than 10 µm thick, less than 100 µm wide) and low power consumption (<10 mW) makes them attractive for biomedical implants.

The active region, which is used in the structure of these VCSELs, is optimized for emission near 2.3 µm. Fabricated devices can be operated in continuous wave mode up to 45°C with threshold voltages of about 1 V and output powers of 1 mW at 10°C. As presented in Figure 6.8, an emission wavelength of 2315 nm has been achieved and it can be tuned over 5.6 nm by increasing the driving current from 10 to 28 mA [2]. Figure 6.8.b shows how the wavelength shifts with changing bias current. The slope of the line denotes the dλ/dI, which is about 0.3 nm/mA.

![Figure 6.8: Current dependent optical output spectra of the 2.3 µm LW-VCSEL, and wavelength shift with respect to bias current, at 300K, with the tuning slope of 0.3 nm/mA. The thermal tunability is 0.1 nm/°C [2].](image-url)


6.2.4 Temperature controlled VCSEL holder

The Newport Temperature Controlled Mount 700C was used for this experiment. The operating temperature range of this mount is -20°C to +80°C. In this experiment, the temperature of the VCSEL holder, which defines the base temperature of the VCSEL, is set to 25°C. Matlab is used to read its temperature constantly to make sure the stability of the background temperature of the VCSEL.

The temperature controller, Newport 3040, was used to improve the stability of the VCSELs optical output. The power output of the lasers and its lasing wavelength vary with the temperature change, and such variations need to be considered during signal processing. In these bench-top experiments, we tried to maintain a constant temperature to reduce the variations in VCSELs output power due to the temperature fluctuation, and only tuned the VCSEL by changing the drive current. In this way, the temperature tuning (with drive current) happens on a constant base temperature to ensure consistency and repeatability. In the future, we will measure the background temperature and will include it in the PLS model along with the bias current.

The measurements were taken at 25 °C for the background temperature. Once we add the temperature coefficient in our model, we will test it for variability of body temperature. We are foreseeing several effects and the possible adjustments, and the model would accommodate the changes. Specifically, at a human body temperature (37 °C), we expect a decrease in the output power of the VCSEL, and a shift in the wavelength. However, the spectra due to glucose are broad and the adjustments may include changing the VCSEL to a one with more appropriate lasing wavelength.

6.2.5 Cuvette

The Eppendorf UVette is a plastic disposable cuvette that is used for glucose solutions in this experiment. It is made up of UV-transparent plastic and has its best transmission for wavelength range of 220 to 1600 nm. It can enable measurements on volumes of sample as low as 50 µL to 2000 µL. This cuvette provides choice of two optical path: 2mm and 10 mm, having four
optical surfaces.

### 6.3 Experiment

For the absorption spectroscopy experiments, a bench-top apparatus was designed to investigate the feasibility of the proposed technique, as shown in Figure 6.9. In this experiment, low-cost plastic cuvettes with a 2 mm optical path length are used to hold the solutions of distilled water and glucose. The 2.3 μm VCSEL is current-biased using a precision current source (Keithley 2602). A chopper modulates the beam at a frequency of 250 Hz before passing through the solution. This reduces the effect of flicker noise (1/f noise) as flicker noise decreases as a function of frequency. The beam is detected with a PbS voltage amplified detector (ThorLabs PDA30G). A lock-in amplifier (SR810) is used to demodulate the signal from the detector to a DC value. Matlab software is used to communicate with the devices, configure them, and collect and analyze the data.

![Figure 6.9: Experimental set-up.](image)

Solutions of distilled water containing concentrations of glucose ranging from 100-1000 mg/dL have been prepared and stored in sealed cuvettes. The VCSEL, biased with drive currents from 8 mA to 22 mA, was used to
generate absorption spectra for each solution. The measurements include the absorption spectral characteristics of the water, glucose, cuvette, with the addition of noise. To subtract the characteristics of the water and the cuvette from the analysis, absorption spectra of distilled water were subtracted from the absorption spectra of the glucose solutions. The data in Figure 6.10 show the resultant differential absorption spectra for different concentrations of glucose.

![Figure 6.10: Absorption spectra for concentrations of glucose, in distilled water, subtracted by the absorption spectra of distilled water.](image)

**6.4 Analysis and Results**

As discussed before and in chapter 4, Chemometrics and Data Analysis, PLS regression is one of the promising methods for the glucose predictions.

In this experiment, PLS is used to estimate the regression coefficients from a sample of data, and build a model that can be used to make predictions from new observations, but it cannot deduce any spectra of individual component in the solution [63].

In this experiment, nine (9) absorption spectra were collected for each concentration of glucose. From these 9 absorption spectra, 5 random spectra were used to compute regression vectors, or to train and build the PLS
model for predicting glucose concentration, using 4 principle components (PC). This PLS glucose prediction model was tested using the other 4 spectra to investigate the accuracy and precision of the model in predicting glucose concentration in a given solution. Figure 6.11 shows the result of this analysis. The estimated concentrations from the PLS modeling over the clinical range of interest are on average within 30% around the actual values. The results show that with only one VCSEL, glucose concentrations can be estimated from the small spectral window of 5-6 nm at a center wavelength of \(2.3 \mu m\). This result demonstrates the feasibility of glucose prediction with VCSELs operating within this small wavelength region.

![Figure 6.11: Glucose prediction using the PLS model, derived from the absorption spectra data.](image)

6.5 Further Improvements on Data Analysis and PLS

There are several methods of improving the regression model to predict the lower concentrations of glucose with clinically acceptable accuracies. One is using multiple VCSELs at different wavelengths, in order to capture a
larger portion of the glucose absorption features. In this section, the PLS regression model is improved with the following changes. One is that all the absorption spectra (or the LI curves) in the calibration set are fed into the calibration model individually rather than being averaged before being supplied to the calibration model. The other change is that the spectra (or the LI curve) for distilled water are supplied to the model as samples with zero glucose concentration, rather than being subtracted from the samples with absorption spectra of glucose solution. These above changes results in a more robust model for the glucose prediction as more system variations are being considered and are known to the calibration model.

Furthermore, the result of predictions using two VCSELs is compared with predictions using one VCSEL to illustrate the significant improvements of predictions.

In addition, the choice of preprocessing improves the predictions. As it was concluded from Figure 4.21, Sqrt Mean Scaling plus mean centering with 11 LVs is one of the potential preprocessing methods. This high number of LVs required is due to several reasons. One of the reasons is that the wavelengths at which these VCSELs are lasing do not contain strong signals due to the glucose absorption. Another reason is that two VCSELs provide very narrow wavelength band, which is not enough to achieve the best possible prediction error. As it was concluded in the final analysis for glucose spectroscopy using white light, there is a minimum number of intervals needed to achieve the accuracy achievable with that set of data. For that wavelength band of 1000 nm to 1700 nm about 6 or 7 intervals was required to achieve the same prediction accuracy using the whole spectra without water absorption peaks.

The data available for analysis in this section is the same as the ones used in preceding sections except that there are available data from two VCSELs rather than only one. Hence, there is an opportunity to compare and observe improvements in predictions using data from two VCSELs compared with using data from a single VCSEL. Both of these VCSELs have lasing wavelengths close to 2.3 $\mu$m. There are nine absorption spectra (or LI curve) for each glucose concentration from 0 to 1000 mg/dL (0:100:1000 mg/dL).
Five of these spectra are used in the calibration set and the other four are used in the validation set. The cross-validations, performed in Chapter 4 (Data Analysis and Chemometrics) to find the proper preprocessing, used all nine spectra. However, the optimum number of required latent variables sometimes differs when a smaller number of samples are available in the calibration. Hence, the analysis performed here is according to the optimum required number of latent variable for that specific calibration set.

6.5.1 Analysis for Glucose Samples with all Concentrations

The concentrations of glucose in the samples used for the analysis in this section are from zero to 1000 mg/dL (0:100:1000 mg/dL). Various PLS regression models were built with two different preprocessing selections on the data from one single VCSEL and data from two VCSELs. It will be observed that the predictions using spectra from two VCSELs are significantly improved compared to using data from one single VCSEL. However, the choice of preprocessing, as investigated to date, does not significantly improve the predictions, but the effect of preprocessing can still be observed. This can be due to the small number of wavelength intervals (i.e., small number of VCSELs) available. One limitation in these analyses is the small number of available samples relative to the available concentrations.

PLS Regression Model Using One VCSEL:

PLS regression model is built for the data from one single VCSEL, preprocessed with mean centering using five latent variables. The glucose predictions using this model are presented in Figure 6.12. The prediction error (RMSEP) with this model that is calculated in the verification step of the process is 3.4095 mmol/L.
Figure 6.12: Glucose predictions with the mean centered data from a single VCSEL. PLS is used to calibrate regression model using 5 latent variables. The dashed red line shows the 10% boundary. The RMSEP error is 61.3713 mg/dL, equivalent to 3.4095 mmol/L.
PLS regression model is built for the data from one single VCSEL, pre-processed with Sqrt Mean Scaling along with mean centering using 5 latent variables. The glucose predictions using this model are presented in Figure 6.13. The verification set is used to find the prediction error (RMSEP) of this model, which is 3.3568 mmol/L.

![Graph](image)

**Figure 6.13:** Glucose predictions with the Sqrt mean scaled and mean centered data from a single VCSEL. PLS is used to calibrate regression model using 5 latent variables. The dashed red line shows the 10% boundary. The RMSEP error is 60.4216 mg/dL, equivalent to 3.3568 mmol/L.

Using Sqrt Mean Scale with OSC and mean centering as preprocessing also shows a small improvement in prediction using data from one VCSEL. For calibration of this model three latent variables are used, which results an RMSEP error of 3.2999 mmol/L in validation. Figure 6.14 presents validation of the samples using this model.
Figure 6.14: Glucose predictions using the model calibrated with Sqrt Mean Scale along with OSC and mean centering as preprocessing on data from single VCSEL. PLS is used to calibrate regression model using 3 latent variables. The dashed red line shows the 10% boundary. The RMSEP error is 59.3980 mg/dL, equivalent to 3.2999 mmol/L.

PLS Regression Model Using Two VCSELs:

In this part, it will be illustrated that the predictions using spectra from two VCSELs improve significantly compared to using data from one VCSEL.

PLS regression model is built for the data from both VCSELs, preprocessed with mean centering using 11 latent variables. The glucose predictions using this model are presented in Figure 6.15. The prediction error, RMSEP, resulted from the validation process for this model is 1.8634 mmol/L.
Figure 6.15: Glucose predictions with the mean centered data from both VCSELs. PLS is used to calibrate regression model using 11 latent variables. The dashed red line shows the 10% boundary. The RMSEP error in the validation of this model is determined to be 33.5419 mg/dL, equivalent to 1.8634 mmol/L.
Chapter 6. Spectroscopy for Glucose Sensing Using VCSELs

PLS regression model is built for the data from both VCSELs, preprocessed with Sqrt Mean Scaling along with mean centering using 11 latent variables. The glucose predictions using this model are presented in Figure 6.16. The predictions using this model are not necessarily improved compared to the previous preprocessing used, especially considering that 11 latent variables are required, and that RMSEP value went up to 2.1848 mmol/L. However, it shows a change in the way that the predictions are spread.

![Figure 6.16: Glucose predictions with the Sqrt mean scaled and mean centered data from both VCSELs. PLS is used to calibrate regression model using 11 latent variables. The dashed red line shows the 10% boundary. The RMSEP error in the validation of this model is determined to be 39.3257 mg/dL, equivalent to 2.1848 mmol/L.](image)

Using OSC as preprocessing, however, shows a small improvement in prediction using data from two VCSELs. For calibration of this model nine latent variables are used, which results an RMSEP error of 1.8450 mmol/L.
in validation. Figure 6.17 presents validation of samples using this model.

Figure 6.17: Glucose predictions using the model calibrated with OSC as preprocessing on data from single VCSEL. PLS is used to calibrate regression model using 9 latent variables. The dashed red line shows the 10% boundary. The RMSEP error in the validation of this model is determined to be 33.2093 mg/dL, equivalent to 1.8450 mmol/L.

Concentrations of samples analyzed in this part are still higher than what possibly can exist in the human blood. Therefore, the set of samples corresponding to lower concentrations will be analyzed next.

### 6.5.2 Analysis for Glucose Samples with lower concentrations

In the proof-of-concept experiments, discussed in the previous sections, the glucose concentrations were higher than the typical clinical range (50-300 mg/dL). In this section, samples with concentrations from zero to 400 mg/dL (0:100:400 mg/dL), which are more clinically relevant, have been analyzed.
PLS regression models have been calibrated using data collected with one single VCSEL and data collected with both VCSELs. Two different combinations of preprocessing are considered to compare and to note the improvements: mean centering and combination of sqrt mean scaling along with mean centering.

**PLS Regression Model Using One VCSEL:**

PLS regression model is built for the lower concentration data from one single VCSEL, preprocessed with mean centering using five (5) latent variables. The glucose predictions using this model are presented in Figure 6.18. The calculated prediction error, RMSEP, in the validation process for this model is 1.1890 mmol/L.
Figure 6.18: Glucose predictions with the mean centered data from a single VCSEL for samples with lower concentrations. PLS is used to calibrate regression model using 5 latent variables. The dashed green line defines the boundary for the predictions to be clinically acceptable (± 1 mmol/L); and the dashed red line shows the 10% boundary. The RMSEP error in the validation of this model is determined to be 21.4015 mg/dL, equivalent to 1.1890 mmol/L.
Chapter 6. Spectroscopy for Glucose Sensing Using VCSELs

PLS regression model is built for the lower concentration data from one single VCSEL, preprocessed with Sqrt Mean Scaling along with mean centering using 4 latent variables. The glucose predictions using this model are presented in Figure 6.19. There is no significant improvement in the predictions observed when using combination of sqrt mean scaling and mean centering as preprocessing instead of mean centering alone. The RMSEP error for this model is 1.2646 mmol/L. This can be due to the different nature of noise existed in the combination of the data from two VCSELs, which may not exist in the data from only one VCSEL.

Figure 6.19: Glucose predictions with the Sqrt mean scaled and mean centered data from a single VCSEL for samples with lower concentration. PLS is used to calibrate regression model using 4 latent variables. The dashed green line defines the boundary for the predictions to be clinically acceptable (± 1 mmol/L); and the dashed red line shows the 10% boundary. The RMSEP error in the validation of this model determined to be 22.7624 mg/dL, equivalent to 1.2646 mmol/L.
PLS Regression Model Using Two VCSELs:

The results of predictions using spectra from two VCSELs show significant improvements in predictions using two VCSELs rather than one VCSEL. This will be illustrated in this part.

PLS regression model is built for the lower concentration data from both VCSELs, preprocessed with mean centering using 5 latent variables. The glucose predictions using this model are presented in Figure 6.20. The prediction error, RMSEP, calculated in the validation process for this model is 0.7955 mmol/L, which is within the desired range of ± 1 mmol/L.

![Figure 6.20: Glucose predictions with the mean centered data from both VCSELs for samples with lower concentrations. PLS is used to calibrate regression model using 5 latent variables. The dashed green line defines the boundary for the predictions to be clinically acceptable (± 1 mmol/L); and the dashed red line shows the 10% boundary. The RMSEP error in the validation of this model is determined to be 14.3187 mg/dL, equivalent to 0.7955 mmol/L.](image-url)

Figure 6.20 compared to Figure 6.18 shows that predictions are improved
significantly when data from two VCSELs are used rather than one.

PLS regression model is built for the lower concentration data from both VCSELs, preprocessed with Sqrt Mean Scaling along with mean centering using 6 latent variables. The glucose predictions using this model are presented in Figure 6.21. There is a small improvement in the predictions of this model due to the preprocessing used. The prediction error, RMSEP, calculated in the validation process for this model is 0.7721 mmol/L, which is within the desired range of ±1 mmol/L.

![Figure 6.21: Glucose predictions with the lower concentration data preprocessed using Sqrt mean scaling and mean centering, from both VCSELs. PLS is used to calibrate regression model using 6 latent variables. The dashed green line defines the boundary for the predictions to be clinically acceptable (±1 mmol/L); and the dashed red line shows the 10% boundary. The RMSEP error in the validation of this model is determined to be 13.8982 mg/dL, equivalent to 0.7721 mmol/L.](image)

Figure 6.21 compared to Figure 6.20 suggests a small improvements in predictions by using different preprocessing method. It also shows that most of the predictions with this model are within the clinically acceptable range...
Comparing the validation of the model represented in Figure 6.20 with the one in Figure 6.18; and comparing the model in Figure 6.21 with the one in Figure 6.19 suggests a significant improvement of predictions as the number of VCSELs used to measure spectra increases. This trend is expected to continue to a point, where there is no more significant improvement with the additional VCSEL or the error is low enough for the clinical use. Such a trend was observed in the white light studies of Chapter 5 (Spectroscopy for Glucose Sensing Using White Light).

6.6 Conclusion

In the first set of analysis on the data from single VCSEL, which was published in the PTL [2], a single 2.3 \( \mu \text{m} \) VCSEL was used to demonstrate the feasibility of glucose sensing in an aqueous solution, using absorption spectroscopy and a PLS algorithm. This result leads to further improvement in accurately predicting glucose in the body. It is worth mentioning that the set of experiments described in PTL presents the result of a preliminary exploratory research. The goal was to verify that lasers (with a very small tuning range) are feasible for this purpose and can provide enough signal power for the determination of glucose concentration. To prove this concept and test our hypothesis, it was best avoiding any interference to be able to focus the research on better understanding of the interference-free best case. These sets of experiments, which have no interference from plasma proteins, will serve as the control set. The next step will be using blood serum for the experiment. The results of this experiment will also serve as a based measurement so that the results from the experiments with blood serum will be correlated, analyzed, and compared.

Furthermore, the data analysis on the same set of data was improved and a set of data from another 2.3 \( \mu \text{m} \) VCSEL was added for the analysis as well. The predictions, especially on the samples with lower concentrations, were improved through improving calibration of the model such that the regression model is more robust to the background changes in the system.
In addition, preprocessing methods were investigated to find a combination of methods that can help the model eliminate some of the normally existed noise in the system. The effects of these improvements are clearer on the lower concentration samples, as samples are not spread over a large range.

The resulted prediction error, with the model for samples with clinically relevant concentrations, on the spectra from one single VCSEL was 1.19 mmol/L. This result was significantly improved using spectra from two VCSELs. The RMSEP error in validation of the model for data from two VCSEL was 0.77 mmol/L, which is within the required accuracy of ± 1 mmol/L.

Therefore, increasing number of VCSELs at different wavelengths, which means increasing the number of wavelength windows or absorption bands to capture a larger portion of the glucose absorption features, as well as increasing number of samples collected from glucose concentrations are potential steps towards improving glucose predictions using VCSELs. Furthermore, VCSELs can be chosen to better match the absorption bands, determined by iPLS technique described in Chapter 5 (Spectroscopy for Glucose Sensing Using White Light), that lead to a higher degree of correlation between the spectra and the glucose concentrations.

Although these predictions are very close to the clinically acceptable accuracies, they are not as good as the predictions stated in chapter 5, glucose spectroscopy using white light, for several reasons. One reason is that one or two VCSEL(s) provide too narrow range (or too small number of variables) in the spectra to be correlated to glucose concentrations. Hence, increasing number of VCSELs will improve the predictions significantly. In addition, the VCSELs may not be at the right wavelength to provide strong correlation with glucose concentration. As it was discussed in the previous chapter, the glucose absorption peaks are not always the only best choices for the wavelength intervals that can show the best correlation to glucose concentration.

The other reason behind these less accurate predictions is the unstable optical power output of the VCSELs, which will be discussed more in next the chapter as well as the drift in output. This problem is especially impor-
tant since spectra of all concentrations are collected over a long time. The time difference between collecting the spectra from the first concentration with the last one allows for the effect of these instabilities in the optical power output.

Furthermore, the number of available samples compared to the range of concentrations is small. This can be another reason for not having as good accuracies as there were in previous chapter, as there are less number of samples for calibration purposes (five vs. 16 for each concentration).
Chapter 7

Optimization and Performance Improvements

In the previous chapters, the feasibility of the proposed methods were illustrated and proven. The PLS techniques have shown strong promise in predicting glucose concentrations using small wavelength windows of the spectra, that will eventually be the result of using VCSELs as the light source. However, the performance should be improved to get more clinically accurate results.

Improving the quality and stability of the absorption spectra produced by lasers as well as increasing the signal-to-noise ratio of the system by reducing the noise to the reasonably lower possible level, are possible ways of improving the glucose prediction capability.

7.1 Stability Optimization

To have repeatable measurements using VCSELs, the measured power output needs to be reasonably stable over time. There are several optimizations that were identified to improve the stability. This section presents some of these performance optimizations that have either been implemented and established or identified for implementation in the future analysis.

7.1.1 Stability of VCSEL Output and Power Drift

Before the experiment for VCSEL-based spectroscopy was set up, it was observed that the output power of the VCSEL drifts and changes over time. After investigations, it was found that the change in output power is corre-
lated with the base temperature of the VCSEL (i.e. the temperature that its package experienced); the value of this correlation is 0.9184. Figure 7.1 presents one of the results of this investigation; it is the plot of output power over time and the base temperature of the VCSEL vs. time.

![Figure 7.1: Plot of output power of a 2.3 µm VCSEL and temperature of its base vs. time.](image)

The ‘resistance’ in legend refers to the resistance measured by a resistor placed at the base of the VCSEL, which is inversely proportional to the temperature. Depending on the type of the device used, the conversion formula between resistor and temperature varies; however, it is the correlation between changes that is of interest. The value of the correlation between output power and the measured resistance, representing the temperature, is 0.9184.

This strong correlation, between output power of the VCSEL and its base temperature, suggests using a Temperature Controlled VCSEL holder to improve the stability of the VCSEL output power. Using the temperature controlled VCSEL holder, the effect of base temperature of the VCSEL on the lasing wavelength and output power was investigated more systematically. In this investigational experiment, the temperature of the VCSEL was
set to numbers from 10 °C to 30 °C; the minimum of this range was decided based on the dew point temperature during the day of the experiment. Figure 7.2 illustrates how the lasing wavelength and the output power of the VCSEL changes with the temperature of the base of the VCSEL (FAT 796, with lasing wavelength of about 1300 nm).

Figure 7.2: Output power and lasing wavelength of VCSEL vs. temperature. The VCSEL, FAT 796, is biased at 3.6 mA.

From the data collected for this investigation, the variations in lasing wavelength (Δλ) corresponding to the change in the base temperature of the VCSEL (ΔT) can be derived (equation 7.1). This is close to the expected value of 1nm/10° C [70].

\[
\frac{\Delta \lambda}{\Delta T} = \frac{1.906\text{nm}}{20\text{°C}}
\]  

(7.1)

Although optimizations, such as controlling the temperature of the VCSEL, were done, still the output of VCSEL for the required accuracy in this application is not stable enough. There are other factors affecting the
measured signal as well.

### 7.1.2 Stability of the Measured Signal

After investigations, it was observed that the measured signal is also correlated to the temperature of the environment and possibly to the temperature of the solution if there is one in the path of the light, which will be discussed next. Figure 7.3 is a plot of the measured output power, both with and without passing through the solution, over time as well as the measured temperature of the environment and the solution over time.

![Figure 7.3: Measured output power, both with and without passing through the solution, over time as well as the measured temperature of the environment and the solution over time. This is a normalized plot illustrating how these variables change together and with respect to each other. The correlation between these variables are calculated to be: 0.8 between measured power of the VCSEL and temperature of the environment, 0.76 between measured power after passing through the solution and the temperature of the environment, and 0.84 between measured power after passing through solution and temperature of the solution.](image)

Figure 7.3: Measured output power, both with and without passing through the solution, over time as well as the measured temperature of the environment and the solution over time. This is a normalized plot illustrating how these variables change together and with respect to each other. The correlation between these variables are calculated to be: 0.8 between measured power of the VCSEL and temperature of the environment, 0.76 between measured power after passing through the solution and the temperature of the environment, and 0.84 between measured power after passing through solution and temperature of the solution.
Chapter 7. Optimization and Performance Improvements

This strong correlation can be due to several reasons, two of which can hypothesized to be due to 1) a change in the gain of the detector, 2) a small change in alignment due to temperature change in environment. However, independent of what the source of this correlation may be, one proposed solution can be to have the data analysis take care of this correlation by adding temperature as one of the variables in the model.

7.1.3 Change of LI curve in 2.3 µm VCSEL

The LI curve of the 2.3 µm VCSEL changed overnight drastically (Figure 7.4). The peak of the LI curve now happens at the bias current of 19 mA; and the maximum power read decreased from 35 mV to 20 mV.

![LI curve of a 2.3 µm VCSEL for two consecutive nights. The peak of the LI curve decreased in magnitude by about 15 mV (from 35 to 20 mV), and has shifted to the bias current of 19mA.](image)

The reason behind this change in output power of this VCSEL was that these VCSELs did not go through the burn-in process in their fabrication, as they are newly developed lasers. It was suggested to drive the VCSEL
for sometime at higher bias current and/or higher temperature to achieve a more stable output. This will serve as a pseudo burn-in process. As a result, the bias current of VCSELs was set to 22 mA, and every 10 s an LI curve was measured to observe the progress. Figure 7.5 shows how the LI curves changed over time with this pseudo burn-in process.

![LI curve of a 2.3 µm VCSEL for several days in the pseudo burn-in process. The peak of the LI curve decreased in magnitude, and the peak has shifted.](image)

As the research and development of these VCSELs progress, less of these sudden changes are expected.

### 7.2 Analysis on Sources of Noise and Losses

Every component and device in the system has its own limitations and as a result adds a noise level to the system. This section aims at reviewing some of these sources of noise in the system involved in the described experiments in previous chapter. In addition, some experiments were designed to analyze
Chapter 7. Optimization and Performance Improvements

and verify the noise limitations of the system will be discussed here as well.

7.2.1 Kinds of Noise

There are several types of noise that limit the detectivity of the detector; namely, Thermal (or Johnson) noise, shot noise, Flicker (or 1/f) noise, and generation-recombination noise.

**Thermal or Johnson Noise** is due to random motion of carriers in conductor. As a result, the detector’s internal resistance, or any resistance in series with detector’s terminals, fluctuates, and generated undesirable variations. Equation 7.2 shows the mathematical representation for Thermal noise. Thermal noise is a White Noise source.

\[
\text{Johnson noise} = \frac{dP_{\text{noise}}}{df} = kT \Delta f 
\]

where k is the Boltzmann’s constant in joules per Kelvin; t is the conductor temperature in Kelvin; and \( \Delta f \) is bandwidth in Hz.

**Shot Noise** is the result of the discrete nature of radiation, where photons arrive randomly in time. Therefore, the photoelectrons produced, from absorbed photons, at random intervals, causing a variation in current, which appears as noise. This noise may be generated by the actual desired signal or by background photons. Shot noise is a White Noise source.

**Dark Signal** of a detector is its electrical output in the absence of radiations from light source. It will include Thermal (Johnson) noise and shot noise.

**Flicker or 1/f Noise** is not well understood, and appears in detectors that require biasing current, such as photoconductors. The magnitude of flicker noise is proportional to \( 1/f^B \), where B is usually between 0.8 and 1.2.
7.2.2 Detector’s Figures of Merit

This section will describe some of the terms that are usually used to state detector’s specifications.

**Noise Equivalent Power (NEP)** is the r.m.s. value of the radiant power in watts that is required to give rise to an output equal to the r.m.s. of the detector’s dark noise. The detector’s response is assumed to be linear down to the noise level. The NEP values are usually stated at a specific wavelength, modulation frequency, detector area, temperature, and detector bandwidth. The unit for NEP is usually in WHz⁻¹/², and it is a commonly used version of Noise Equivalent Detector Input (P_N).

**Detectivity (D)** is defined as the reciprocal of the NEP as shown in equation 7.3.

\[
D = \frac{1}{NEP} = \frac{1}{P_n}
\]  
(7.3)

**Normalized Detectivity (D*)** is normalized for the area (cm²) and BW(Hz), as shown in equation 7.4. Detectivity is inversely proportional to the square root of the area of the detector.

\[
D^* = DA^{1/2}(\Delta f)^{1/2}
\]  
(7.4)

In addition, fluctuation in the overall temperature of the detector may cause temperature noise, which can be a problem for small thermal detectors with low thermal mass.

The unit for D* is cm Hz¹/² W⁻¹; Figure 7.6 shows normalized detectivity for various detectors with different materials.

Figure 7.6 indicates that for the wavelength of interest in this project (i.e. 2 to 2.3 μm), the best detector is a PbS detector, which is the detector used for the VCSEL based experiments of this project.
Chapter 7. Optimization and Performance Improvements

This figure has been removed due to copyright restrictions.
The information removed shows detectivity of various detectors fabricated using various materials.

Figure 7.6: Normalized Detectivity for various detectors. [From www.newport.com]

7.2.3 Experiments for Noise Investigation in Detector

An experiment was set up to investigate the noise from the detector and compare it with the specifications of the detector; i.e the Noise Equivalent Power (NEP) of the detector. This is to make sure that the lowest possible noise level is achieved with the VCSEL based experiments using this detector. The NEP is a measure of the dark noise, which is the electrical output of the detector in the absence of the light source. For this purpose, the VCSEL was turned off; then, the detected signal was measured with various filter settings of the lock-in amplifier. For example, for the filter roll-off of 6 dB/oct, the signal was measured 9 times with each time constant from 3 ms to 3s, each time waiting for 5×time-constant before measuring the signal. The standard deviation of each set of 9 measurements was calculated and recorded on the graph. Figure 7.7 shows these measured signals for various time constant settings. This continued similarly with other roll-off settings of the filter.

These results are compared with the Noise Equivalent Power (NEP) of the detector, given in specification to be 1.5e-11 Watts/$\sqrt{Hz}$. For this purpose, the NEP (in Watts/$\sqrt{Hz}$) is converted to V/$\sqrt{Hz}$, by multiplying with peak sensitivity of the detector (7.5e5 V/W) as shown in equation 7.5.

\[
NEP \text{ in volts} = 7.5 \times 10^5 \left( \frac{V}{W} \right) \times 1.5 \times 10^{-11} \left( \frac{W}{\sqrt{Hz}} \right) = 11.25 \times 10^{-6} \frac{V}{\sqrt{Hz}}
\]

(7.5)

Considering the time constant of 10 ms as an example for comparison, the corresponding bandwidth will be 100Hz, from which the noise level can
Figure 7.7: Measured noise signal for roll-off of 6 dB/oct vs. various time constants. 9 measurements at each time constant was made in dark. Before each measurement 5×time constant was elapsed. The number on the graph for each time constant is the standard deviation of the 9 measurements at that time constant.
be calculated for this specific bandwidth to be 1.25e-5 Volts. This can be compared with the standard deviation of the measurements with 10 ms time constant, 9.3e-5, which shows that the two results are reasonably close. Furthermore, with roll-off of 12 dB/oct, this noise goes down to 5.2e-5; and having electrical attenuator between detector and lock-in amplifier reduces this noise signal by a factor of less than but close to 10.

In comparing these results, it worth noting that NEP has been specified in certain situations such as 300K blackbody source, 600 Hz chopping, 635 Hz bandpass filter, and ambient temperature of 22 °C.

7.2.4 Electrical Attenuator to Improve Signal-to-Noise Ratio

In the experiments, the power output of the VCSELs was usually too high for the lock-in amplifier and would cause an overload in the reading. As a result, the power of the VCSEL was reduced by misaligning or by using an optical attenuator. However, the high level of noise competing with this small signal was causing a noisy output. The detector has the capability of detecting signals up to 10 V, whereas lock-in amplifier can only accept signals up to 1 Volts. To improve signal-to-noise ratio, instead of attenuating optical signal, it was decided to use an electrical attenuator between the detector and the lock-in amplifier. The attenuator was designed for the attenuation ratio of 10:1. This improvement increases the signal-to-noise ratio by about 10, which can be illustrated analytically. This improvement was implemented and will be used in future experiments.
Chapter 8

Summary, Conclusion, and Future work

The burden of care for patients suffering from diabetes mellitus has been increasing. Tight control of blood sugar is an important factor for effectively minimizing the complications associated with this disease. Since accurate blood glucose monitoring requires pricking of fingers for blood sampling, enhancing glucose measurement techniques to allow easy and continuous monitoring has received a lot of attention in the past decades. Nearly thirty companies either are currently developing Continuous Glucose Monitors (CGMs), or have tried and failed. This project proposes a novel and promising method based on VCSELs — the smallest lasers commercialized — and addresses the major challenges encountered by previous approaches, such as signal-to-noise ratio. VCSELs, operating within a small spectrum, and having high power spectral density, provide a higher signal-to-noise ratio. The proposed device is a small, low-power glucose sensor that may be injected to the interstitial fluid or mounted on the currently available biomedical technologies. This proposed sensor would be stable, cost-effective, low-maintenance and provide accurate real-time Continuous glucose monitoring (CGM), which will improve treatment of diabetes.

This chapter will include sections on the Significance of the Work, Summary and Conclusion of the current work, and Future work.

8.1 Significance of the Work:

Although many groups have demonstrated predicting glucose concentrations optically, none have realized a complete system that is both small enough
to be implantable, accurate enough for clinical relevance, and sufficiently biocompatible. The proposed methods for implantable glucose monitor in this research considered recent technological developments in semiconductor lasers and biocompatible materials, to address the challenges faced by current sensor technology. This research project aimed to demonstrate the suitability of VCSEL-based glucose sensors in an in-vitro study, by preliminary investigation of the proposed methods. The current project took an important step towards achieving our long-term goal of developing a durable long-term implantable sensor, integrated with electronics that would enable the realization of closed-loop glucose control using commercially available insulin pumps, i.e. an artificial pancreas. This proposed implantable glucose monitor is novel in using VCSELs (the first 2µm VCSEL-based glucose sensor), it is competitive, and has a strong potential for breakthrough in diabetes management. Upon successful design of the sensor and in-vitro tests, in-vivo animal model experiments will be pursued.

8.2 Summary, Conclusion, and Discussion of Improvements

To achieve the objectives of this project, several near-infrared spectroscopy experiments were designed to investigate the feasibility of the proposed methods. In addition, calibration of multivariate regression models was optimized for these experiments. The two major experiments that were performed are White light based and VCSEL based spectroscopy.

8.2.1 Chemometrics

Partial Least Square Regression, having the capability of capturing variations in the predictor matrix that best correlated with the variable of interest, is one of the best methods for prediction applications and has shown promising results. The calibration of PLS regression model was optimized for the available data for analysis. Specifically, the choice of preprocessing methods and the required number of latent variables for calibration of the
model were investigated for data from each experiments. The estimated prediction error (RMSECV) was used for these investigations. It was concluded that for the absorption spectra that resulted from white light experiments as well as for the intervals chosen from these spectra, Orthogonal Signal Correction (OSC) along with mean centering results in the most optimized predictions. With this preprocessing, the number of latent variables required for calibration is decreased while the prediction error is improved. For the spectra from VCSELs, applying the preprocessing methods of Sqrt Mean Scale, OSC, and mean centering have shown potential for improving predictions.

8.2.2 White Light Based Spectroscopy

The bench-top glucose measurement system was designed in the early stages of investigation to determine the optimal choice in VCSEL wavelengths, to evaluate the accuracy of VCSEL-based spectroscopy in solutions of glucose, to develop effective glucose prediction algorithms and quantify the performance, and to explore novel methods for optical signal measurements.

In this set of experiments, the possibility of using a number of narrow wavelength intervals, rather than using complete spectra, for glucose prediction was illustrated. The required number of wavelength segments and their wavelengths were identified through these experiments. The results of these experiments could achieve a prediction error (RMSEP) as low as 0.3874mmol/L with four intervals and to about 0.2837mmol/L with five intervals. Considering that the experimental set up for these set of experiments can be easily improved with the new technologies available in our lab, a more improved predictions are expected once the experiments are performed again.

The methodology of identifying proper setting, such as the optimum number of latent variables and preprocessing methods, for calibration of the model was developed. This identified method is now ready to be extended to absorption spectra of solutions with more blood constituents.
8.2.3 VCSEL-Based Spectroscopy

A set of VCSEL-based glucose spectroscopy experiments was designed to investigate the suitability and possibility of glucose predictions using VCSELs. It was demonstrated that, with the samples having clinically relevant concentrations, the prediction error (RMSEP) of about 1.2 mmol/L is achievable with spectra from one VCSEL, and the prediction error of about 0.77 mmol/L with spectra from two VCSELs.

In parts of the proof-of-concept experiments, the glucose concentrations were higher than the typical clinical range (50-300 mg/dL). There are several methods of improving the results to be able to predict lower concentrations of glucose. The first is increasing the number of VCSELs at different wavelengths, in order to capture a larger portion of the glucose absorption feature. Next, VCSELs can be chosen to better match the absorption bands, leading to a higher degree of correlation between the spectra and the glucose concentrations. Choosing cuvettes with a lower optical absorption, and optimizing the differential lock-in detection mechanism, can further increase the signal-to-noise ratio.

Once this VCSEL-based glucose measurement design has been refined for the clinical range, it will be evaluated for detecting glucose in blood serum or similar body fluids, where the system can be validated in the presence of interference. The ultimate goal is to package the design as an implantable device.

Other alternatives will be considered in the design of the implant. For example, instead of chopping the signal mechanically, the VCSEL may be modulated or other detectors may be used once the signal-to-noise ratio is sufficiently improved. In addition, CMOS electronics will be designed for the detection of the signal to replace the bench-top lock-in amplifier.

In addition, there is the potential of designing the detector with InGaAs, as their detectivity and response is comparable to the PbS detector used. This leads to the opportunity of having the detector from the same material to be implemented on one single chip, which makes the implant even more suitable for its purpose.
8.3 Future work

Having demonstrated in the current thesis that it is feasible to use lasers to measure the concentration of glucose in aqueous solution, the immediate step is to move to solutions of glucose in actual blood serum to evaluate the accuracy of VCSEL-based spectroscopy with scattered light in blood samples, and to study the effect of blood constituent interference on glucose prediction accuracy.

Furthermore, I am interested in pursuing implementation of the proposed optical sensor system, which will ultimately be part of an artificial pancreas. Hence, the objectives of this research are: developing an implantable optoelectronic device, having this device specialized as a glucose sensor, and testing this sensor in both in-vitro and in-vivo environment. To achieve these objectives, the following tasks and methods are being considered:

1. Although we have already demonstrated glucose detection with two lasers, we would like to improve the accuracy of the system by systematically studying the choice of VCSEL wavelength. For this, the bench-top system will be used with a white-light as the optical source, combined with a monochromator (1-2.3 μm), to obtain scattered light optical spectra from saline solution samples with varied glucose concentration, similar to the white light experiment in chapter 5 (Spectroscopy for Glucose Sensing Using White Light) but extended to 2.3 μm. Using a combination of global optimization (e.g. genetic algorithm, or simulated annealing), with the Partial Least Squares (PLS) technique we have employed, the spectra will be processed to identify the optimal choice of wavelengths for the vertical cavity lasers. This study is necessary, since each VCSEL can only be tuned over a narrow wavelength range (typically 7 nm when thermally tuned) as compared to the 300 nm-wide glucose band of interest. This study will answer the following questions: 1) what is the wavelength range necessary for a good correlation between actual glucose level and optical glucose prediction, 2) how many spectral segments (and hence number of lasers) are required to achieve a clinically relevant accuracy in the new
range. These answers will be the basis for an optimal solution for the VCSEL-based sensor experiments.

2. The bench-top system will then be used with Amann’s packaged 2.0 to 2.3 µm VCSELs [54, 71], and the scattered light will be analyzed using a commercial PbS detector, chopper, and lock-in amplifier. The VCSEL will be tuned by varying the bias current (typically in the range of 2 to 10 mA). Each VCSEL measurement will provide one spectral slice – by repeating the measurement using several lasers, we will obtain a sparse and segmented optical spectrum, which will be used in the PLS model for glucose prediction.

The first set of experiments will be conducted with saline-glucose solution samples, with the goal of optimizing the signal-to-noise ratio, validating the choice of optimal VCSEL wavelengths, and determining the best-case interference-free glucose measurement accuracy. This will allow us to test the hypothesis that a narrow optical linewidth (improved spectral resolution), and high optical power spectral density will improve the signal-to-noise ratio and will improve the glucose prediction accuracy, as compared to white-light measurements.

Next, the system will be used to verify that our approach of using the reflected or scattered light provides a high-enough glucose prediction accuracy, as compared to our previous optical transmission experiments. Another research group has compared reflectance versus transmission measurements, and shown that reflectance measurements provide a similar accuracy to 2 mm transmission experiments [23]. We will verify that this is the case for the VCSEL segmented optical spectrum approach we are using.

3. Having demonstrated the functionality of the bench-top system VCSEL sensor, we will undergo several algorithm design cycles, for in-vitro glucose measurements using human blood serum and human whole blood. We will establish a strong correlation between our optical measurements and actual blood glucose levels, using multiplicative
scatter correction (MSC) (for correcting for medium scattering) [23], spectral processing for baseline correction [72], independent component analysis (ICA) [73], multi-regression analysis signal processing tools, and will validate the measurements in the presence of interference (e.g. temperature, hemoglobin, water). Once the optical measurements have been correlated with glucose values in whole blood, statistical quality control methods will be employed to study signal consistency in the presence of simulated fluctuations of body temperature, osmolarity, viscosity and hemodynamics, as well as fluctuations in laser output power and wavelength.

4. Miniaturization of the optical table setup to a 1cm by 1cm device and design of the device components. The current set up on an optical table is very large; the size should be reduced in an implantable scale. Hence, we will investigate and evaluate the performance of the various approaches to choose the most suitable one. As an example, currently used mechanical chopper, used to modulate signal, is not feasible on implantable device. Hence, we will investigate the methods to modulate the signals electro-optically; for example, methods such as modulating the lasers (analysis will depend on the transient thermal responses), or modulating light at the detection site using frequency-selective structure. We will investigate the suitability of these models and evaluate the performance for comparison to have the optimized choice. In our bench-top setting, it was easy to set up experiments for transmission spectroscopy and align lasers and detectors efficiently; however, for the implantable case, it will be challenging to have lasers and detectors on the opposite sides of the sample. Hence, the use of reflectance spectroscopy [66] is proposed. We will investigate the suitability of this approach and compare it with transmission spectroscopy to optimize the device for the highest accuracy.

5. On-chip fabrication in collaboration with researchers in Germany and Canadian Microelectronics Corporation (CMC) + biofouling through collaboration with blood research group at UBC. The sensor consists of
an array of lasers and detectors, with electronics to transmit the data to a computer. Figure 8.1 shows a possible preliminary sensor device. Eventually, the sensor may be developed to a smaller dimension and the computer will be substituted by a commercially available insulin pump. The CMOS will be fabricated through CMC. The laser and detectors fabricated by researchers in Germany. The UBC clean-room facilities will be used to die-attach the photonic elements to the CMOS electronics.

![Figure 8.1: Preliminary sensor design](image)

6. Testing in-vitro and moving the measurement system towards real world application by conducting tests in in-vivo environment. The first step is to conduct experiments in the simulated in-vivo environment using blood serum. If successful, after ethics approval is received, the in-vivo test on diabetes mice will be performed through collaboration with the pharmacology and pathology department. We anticipate that in-vivo environment will reduce the accuracy of our system; for example, real-time interference from proteins, or variations of body temperature. We will a) systematically study the effects of physiological variables on accuracy; b) if necessary improve the accuracy to a more clinically acceptable level by: increasing the number of VCSELs, optimizing the PLS model, incorporating physiological variables in the PLS model, filtering wavelengths of interfering particles, and etc.

7. The optical biosensor chip developed and tested in the simulated in-vivo environment, will then be integrated with commercial biomedical implants for proof of concept demonstration.

This research is novel since it is the first time that 2.3 µm semiconduct-
tor lasers are available and applied to the biomedical area, leading to the innovation of the smallest optical implantable sensor.
Bibliography


[22] S. F. Malin, T. L. Ruchti, T. B. Blank, S. N. Thennadil, and S. L. Monfre, “Noninvasive prediction of glucose by near-infrared diffuse re-


Bibliography


157
Bibliography


