Development of Novel Mid-Infrared Spectrometers based on Quantum Cascade Lasers

Yin Wang

A Dissertation
Presented to the Faculty of Princeton University in Candidacy for the Degree of Doctor of Philosophy

Recommended for Acceptance by the Department of Electrical Engineering
Adviser: Gerard Wysocki

January 2014
© Copyright by Yin Wang, 2013.
All rights reserved.
Abstract

Sensitive detection of trace gas molecules has various important applications in environmental science, medical diagnostics and homeland security. The invention of quantum cascade lasers (QCLs) has triggered development of compact, efficient and highly sensitive mid-infrared (mid-IR) spectroscopic techniques. This dissertation is primarily focused on Faraday rotation spectroscopy (FRS) for detection of gas-phase radicals, and new methods to perform broadband, high-resolution mid-IR spectroscopy. The developed techniques allow the sensor to reach quantum limit in the real-world settings.

The noise in traditional FRS systems is typically far above the quantum shot-noise due to the strong laser noise at its spectral base-band. Here, a method employing heterodyne-enhanced FRS (H-FRS) is developed. Through optical heterodyning, the signal is shifted from the low frequency to radio frequencies (RF), where the noise is strongly suppressed, allowing significant improvement of the signal-to-noise ratio. An experimental demonstration of H-FRS was performed using a distributed feedback QCL and a mercury-cadmium-telluride photodetector. The cryogen-free system exhibited the total noise of 3.7 times higher than the quantum shot-noise.

The complex optical design of H-FRS limits its application only to laboratory conditions. To overcome this issue a dual modulation FRS method that requires much simpler set-up and is capable of even higher performance than H-FRS is proposed. A prototype was built as a robust transportable system and was delivered to Cleveland Clinic for the first, proof-of-principle isotopic studies of nitric oxide metabolism in human body. The total noise observed in this system is only two times higher than the quantum shot-noise.

A laser testing system for optimizing QCL chips is developed. The system allows for automatic optical alignment and characterization of the QCL chips in an external cavity QCL configuration. Thus it significantly improves the data quality and reduces
the manufacturing cost. These studies led to a better understanding of operation of Fabry-Perot (FP) QCLs, and allowed for development of a mid-IR spectroscopy based on multi-heterodyne of two FP-QCLs. Molecular absorption profile is down-converted into the RF spectrum by the heterodyne process. Both a multi-mode spectral retrieval and a high-resolution spectral scan capability based on the RF signal analysis are demonstrated.
Acknowledgements

First and foremost, I would like to express my deepest gratitude to my adviser, Professor Gerard Wysocki. Without his careful guidance and generous help this thesis would not have been possible. I am grateful to his persistent availability to discuss about the problems throughout these years. I am also thankful to his enthusiasm, his inspiration, and his efforts to guide me through every critical moment during development of this thesis. His way of doing research has been invaluable to my future career.

I would like to thank my dissertation readers for their efforts in polishing this thesis, Professor Mark A. Zondlo and Professor Hakan E. Tureci. I also want to thank my dissertation committee: Professor Claire F. Gmachl, and Professor Stephen A. Lyon.

I am very grateful for the chance to collaborate with Feng Xie, Abdou Diba, Catherine Caneau, Herve P. Leblanc, Christopher A. Page, Lawrence C. Hughes, and Dr. Chung-En Zah at Corning incorporated. Without their generous contribution my first intern project would not be successful. I also want to thank Frank Cikach, David Grove, Luma Dababneh, Jarrod Barnes, Liping Tian and Dr. Raed Dweik for the exciting collaborative opportunity of testing our breath analyzer prototype at the Cleveland Clinic. The project is nothing without your expertise on bio-medical area.

I would like to acknowledge Dr. Antoine Muller at Alpes Lasers for providing QCL chips for the H-FRS experiment and DM-FRS experiment. My dual-FP-QCL heterodyne spectrometer would not work without the QCLs from Professor Claire Gmachl, Professor Federico Capasso (Harvard University), Dr. Laurent Diehl and Dr. Mariano Troccoli. I must also acknowledge Peter Q. Liu, Xue Huang, Pierre Bouzi, Yu Yao, Anna Michel, Matthew Chang and Professor Paul Prucnal for helping me with the experimental setup. I must thank Dr. Yamac Dikmelik and Professor Jacob Khurgin (John Hopkins University) for fruitful discussion about the phase-
lock mechanism in FP-QCLs. I appreciate Mr. William H. Dix’s team at physics
department machine shop for their 100% excellent job. I must express my appreciation
to Lisa R. Lewis, Dorothy R. Coakley at the EE grant office, Linda M. Dreher, Lidia
H. Stokman at the department purchase office, and Sarah M. McGovern, Colleen C.
Conrad at the graduate studies office for their assistance and help.

I must also thank the members of my research group, Stephen So, Tracy Tsai,
Clinton Smith, Eve W. Wang, Michal Nikodem, Brian Brumfield, Andreas Hangauer,
Eric Zhang, Genny Plant and Alex Hallden-Abberton, who join together to become
an excellent team with cultural diversity. You are all great friends not only providing
mental and technical support, but also helping me to improve my English level. The
BBQ experiences, lab cleanings, ski trips, pizza lunches, as well as sharing Thorlabs
lab snacks with you will be the memories that last forever.

Five years is not short, I was fortunate to have Dr. Paul Nutkowitz as my English
tutor, and Yuyang Huang, Xiaolei Yang, Fei Liu, Bo Yang, Lin Mei, Chunye Wang,
Yi Rao, Jie Sun, Shu Jia, Jianhua He, Yi Ma, Wenjie Wan, Wenhao Jia, Andrew
Shu, Yuxuan Wang, Xue Huang, Yiyue Wu, Yi Shao, Yang Liu, He Wang, Liling
Wan, Ting Liu, Yang Yang, Yan Zhang, Lei Tao, Yue Tian, Yue-Kai Huang, Tzu-
Yu Chuang, Yao-Wen Yeh, Yen-Ting Chiu, Yu-Cheng Tsai, Amy Wu, Wayne Soon,
June Yeung, Chia-Chun Lin, An Dai, Chiao-Ti Huang, Chi-Hua Kuo, Jen-Tang Lv,
Chien-Hung Lu, Peiyuan Wu, Junwei Chuah and many other great friends.

Lastly, I am thankful to my parents who raised me with love and supported me in
all my pursuits. Most of all my loving, encouraging and patient wife Ting-Jung who
always stands with me is so appreciated. Thank you!
Contents

Abstract ................................................................. iii
Acknowledgements ...................................................... v
List of Tables ............................................................ xii
List of Figures ............................................................. xiii

1 Introduction .......................................................... 1
  1.1 Motivation for trace gas detection ............................... 1
    1.1.1 Environmental and atmospheric research .................. 1
    1.1.2 Combustion chemistry ...................................... 3
    1.1.3 Breath analysis ........................................... 4
  1.2 Techniques for trace gas sensing ................................ 8
    1.2.1 Absorption spectroscopy .................................... 8
    1.2.2 FRS ..................................................... 14
  1.3 Mid-IR laser sources ............................................ 20
    1.3.1 Why Mid-IR? .............................................. 20
    1.3.2 QCLs .................................................. 22
    1.3.3 QCL types .............................................. 24
  1.4 Organization of the thesis ...................................... 28

2 Faraday rotation spectroscopy theory .............................. 29
  2.1 Signal .......................................................... 29
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.1</td>
<td>Faraday rotation angle</td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>2.1.2</td>
<td>FRS signal generation</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>2.2</td>
<td>Noise in FRS</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Noise contribution</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>2.2.2</td>
<td>FRS system optimization</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>2.3</td>
<td>Shot noise limited condition and ultimate detection limit of FRS systems</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>3</td>
<td>Multi-pass cell enhanced Faraday rotation spectroscopy</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>3.1</td>
<td>Experimental demonstration</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>3.2</td>
<td>Chapter summary</td>
<td></td>
<td>51</td>
</tr>
<tr>
<td>4</td>
<td>Heterodyne enhanced Faraday rotation spectroscopy</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>4.1</td>
<td>Background and motivation</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>H-FRS operation principle</td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>4.3</td>
<td>Experimental</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td>4.3.1</td>
<td>System set-up</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Baseline distortion and interferometer stabilization</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>4.4</td>
<td>H-FRS results and discussion</td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Noise analysis</td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Sensitivity of the H-FRS system</td>
<td></td>
<td>69</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Comparison with other FRS systems</td>
<td></td>
<td>73</td>
</tr>
<tr>
<td>4.5</td>
<td>Chapter summary</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>Dual-modulation Faraday rotation spectroscopy</td>
<td></td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Background and motivation for isotopic NO detection</td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>5.2</td>
<td>DM-FRS principle</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td>5.3</td>
<td>DM-FRS instrument configuration</td>
<td></td>
<td>85</td>
</tr>
<tr>
<td>5.4</td>
<td>System characterization</td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>5.5</td>
<td>Applications</td>
<td></td>
<td>95</td>
</tr>
</tbody>
</table>
6 EC-QCL studies

6.1 Background and motivation

6.2 System design and configuration

6.3 QCL chip performance testing

   6.3.1 Tuning range vs. threshold current

   6.3.2 Selecting the best gain chips for EC laser

   6.3.3 Estimating the AR coating quality

6.4 Chapter summary

7 Multi-Heterodyne Spectroscopy based on Fabry-Perot Quantum Cascade Lasers

7.1 Background and motivation

7.2 Experimental set-up

7.3 Multi-heterodyne spectroscopy

7.4 Chapter summary

8 Conclusion

A A bit more about Zeeman effect

B H-FRS implementation details

   B.1 H-FRS customized photodetector

   B.2 H-FRS system photo

C DM-FRS implementation details

   C.1 DM-FRS signal spectrum calculation

   C.2 The Allan variance and the Allan deviation

   C.3 Signal spectrum comparison between DM-FRS and conventional AC-FRS
C.4 DM-FRS photos ......................................................... 148
C.5 Purge system for the liquid sample measurement .......... 149

D Implementation details of EC-QCL studies .................. 151
  D.1 EC-QCL bar tester software interface ...................... 151
  D.2 Error propagation for AR coating formula ............... 152

Bibliography .............................................................. 153
List of Tables

1.1 “Potential biomarkers” and their corresponding diseases . . . . . . . . 5

4.1 FRS system comparison . . . . . . . . . . . . . . . . . . . . . . . . 75

5.1 $^{15}\text{NO}/^{14}\text{NO}$ ratio measured in urine, blood and breath . . . . . . . 98

6.1 Threshold current of 5 $\mu$m width chips with different coating stages . 115

6.2 Threshold current of 4 $\mu$m width chips with different coating stages . 116
List of Figures

1.1 Nitric oxide synthesis and pathways in human body ............................................. 7
1.2 The basic principle of absorption spectroscopy ..................................................... 9
1.3 Beam spot pattern of a Herriott cell and an astigmatic mirror cell ......................... 11
1.4 The basic principle of WMS .................................................................................. 13
1.5 The basic principle of FRS ................................................................................... 15
1.6 Signal spectrum using AC-FRS and DC-FRS technique ........................................... 17
1.7 Line strengths of ro-vibrational transitions of common gas molecules ................... 21
1.8 Comparison of integrated line strengths of NO transitions in the mid-IR and near-IR region ............................................................................................................. 21
1.9 Principle of QCLs .................................................................................................. 22
1.10 External cavity laser configuration ...................................................................... 27
2.1 Principle of FRS .................................................................................................... 32
2.2 Faraday rotation angle for the entire $^{14}$NO fundamental band ......................... 34
2.3 Faraday rotation angle for the $^{14}$NO Q-branch ............................................... 34
2.4 Calculated Faraday angle peak signal vs. magnetic field strength ....................... 35
2.5 Calculated Faraday angle peak signal vs. sample pressure ................................... 37
2.6 Faraday rotation angle for the entire $^{14}$NO fundamental band at high B-field .... 37
2.7 Conceptual diagram for FRS signal and various noise dependence on analyzer offset angle ................................................................. 41
3.1 Multi-pass cell design .......................................................... 47
3.2 Calculated AC-FRS signal strength vs. multi-pass number ........ 48
3.3 Measured SNR vs. optical power on photodetector .................. 50
3.4 NO spectrum acquired at the optimum offset angle ............... 50
4.1 H-FRS configuration and principle ....................................... 56
4.2 Schematic diagram of the H-FRS setup .................................. 60
4.3 Laser source LIV curves ..................................................... 61
4.4 Laser wavenumber vs. current ............................................ 61
4.5 A simple 3-mirror configuration for 90° polarization rotation ... 62
4.6 H-FRS scan spectrum without interferometer stabilization ........ 64
4.7 Noise spectrum of the QCL .................................................. 66
4.8 The photodetected noise density at 1.08 kHz as a function of laser power .......................... 67
4.9 The total noise at ~30 MHz as a function of optical power .......... 68
4.10 Heterodyne efficiency characterization ................................. 70
4.11 An H-FRS spectrum of the NO R(17/2)f transition at 1906.73 cm⁻¹ .......................... 71
5.1 DM-FRS principle .............................................................. 82
5.2 Calculated DM-FRS 2nd harmonic signal amplitude at different modulation depth ........................................... 85
5.3 Schematic set-up for a DM-FRS system ................................ 86
5.4 DFB-QCL wavelength tuning vs. bias current at different temperature ........................................... 86
5.5 Spectrum showing Q branches of all three stable NO isotopes .... 87
5.6 Target wavelength region for two NO isotope detection ............ 88
5.7 Concept of monitoring multiple paramagnetic species using DM-FRS ........................................... 88
5.8 2nd harmonic DM-FRS signal output and 3rd harmonic signal from absorption cell output ........................................... 91
5.9 Allan deviation of measured total noise and calculated shot noise .. 92
5.10 Online breath $^{14}$NO (red) test. ................................. 94
5.11 Measurement of both $^{15}$NO (red) and $^{14}$NO (blue) concentration from a patient’s urine samples. ................................. 96
5.12 $^{15}$NO signal measured from a breath sample. .................. 98

6.1 Mechanical design of the EC-QCL bar tester. ......................... 105
6.2 System functional block diagram. .................................. 106
6.3 EC spectral tunability and threshold current spectrum for chip#28 on a bar operating at 3.5 $\mu$m. .................................. 109
6.4 Logarithm plot of the same spectrum as above. Side mode suppression ratio $\sim$20dB. .......................................................... 109
6.5 Schematic diagram of the gain and loss in a Littrow configured EC-QCL. 110
6.6 EC spectral tunability and threshold current spectrum for chip #22 on a wafer bar with an ultra-broadband gain region. .............. 111
6.7 Tuning spectrum of 6 3.5 $\mu$m chips ................................. 112
6.8 Threshold current vs. grating angle .................................. 113

7.1 Experimental configuration used for demonstration of the broadband, high-resolution, multi-heterodyne spectroscopy based on FP-QCLs. 120
7.2 A schematic operation principle of the multi-heterodyne spectroscopy with FP-QCLs. ...................................................... 122
7.3 Optical spectra of the two FP-QCLs and an RF spectrum of the beat notes they generate in the photodetector. ....................... 123
7.4 Instantaneous line-width and frequency correlation of the beat notes. 125
7.5 Measuring broadband NH$_3$ absorption features using all beat notes. 128
7.6 Measuring low pressure N$_2$O absorption line with high resolution. 130

A.1 Principle of Zeeman effect in hydrogen atom ....................... 139
B.1 A photo showing the H-FRS setup. ........................................... 143

C.1 Spectrum comparison between conventional AC-FRS (a) and DM-FRS (b). ......................................................... 147

C.2 A top view of the DM-FRS sensor optical set-up. .................. 148

C.3 A photo showing that the author is demonstrating how to use the DM-FRS based breath analyzer at Cleveland Clinic. ............... 149

C.4 Equipment for converting nitrite and nitrate in liquid samples to nitric oxide. .......................................................... 150

D.1 Software interface for EC-QCL testing platform. .................... 151
Chapter 1

Introduction

This thesis mainly focuses on the development of new Faraday rotation spectroscopy (FRS) techniques for sensitive detection of free radical species and studies of novel broadband high-resolution mid-infrared (mid-IR) spectroscopic systems. In this chapter I will introduce the motivation of sensitive detection of trace gas molecules, especially free radical molecules. A brief overview of the most popular laser absorption spectroscopic techniques for trace gas sensing as well as the technique primarily studied in this thesis, FRS, will be presented. In addition, this thesis also presents the development of an external cavity quantum cascade laser (EC-QCL) testing platform and a novel multi-heterodyne spectrometer based on Fabry-Perot (FP-)QCLs. Thus, this chapter will also discuss various types of QCLs, which are important mid-IR laser sources in broadband high-resolution spectroscopic applications.

1.1 Motivation for trace gas detection

1.1.1 Environmental and atmospheric research

Quantifying trace gas molecules in the Earth atmosphere with high sensitivity and high accuracy has various applications including climate modeling, enacting envi-
ronmental hazardous and air-pollution regulations, and monitoring industrial and agricultural processes [11] [12] [13]. For example, detecting various nitrogen compounds (e.g. ammonia (NH$_3$), nitrite (NO$_2^-$), nitrate (NO$_3^-$), nitrous oxide (N$_2$O), nitric oxide (NO) and nitric dioxide (NO$_2$), etc.) in the environment is particular interesting to environmental and ecological scientists for studying the global nitrogen cycle [4] [5]. Monitoring of these species is also very important to the anthropogenic atmospheric effects, such as formation of photochemical smog, production of acid precipitation, and depletion of stratospheric ozone. For public health safety, high NO and NO$_2$ concentrations can cause severe air quality deterioration which leads to wide public health concern [6] [7] [8]. To better understand and control the impacts of these nitrogen components on environment and eco-systems, there is an urgent need for instrumentation that can sensitively, accurately and continuously resolve and monitor the regional variances of nitrogen chemicals in our atmosphere.

Another example is the hydroxyl •OH radical – an important atmospheric chemistry intermediate species which helps to control the concentrations and distribution of greenhouse gases and pollutants in the Earth atmosphere. Reactions of SO$_2$, H$_2$S, and NO$_x$ (NO and NO$_2$) with •OH are also important atmospheric chemical processes that lead to photochemical smog and acid rain [9] [10]. Hydroxyl radicals have an ambient molecular density of $\sim 10^6$ molecules per cm$^3$ corresponding to the concentration of $\sim 0.1$ part-per-trillion by volume (pptv) [3]. The extremely low concentration and high reactivity rate of •OH radical make the detection and study of the species rather difficult. Therefore, a selective, sensitive and accurate •OH sensing technique to study the kinetics and mechanisms of the radical species reactions is needed.
1.1.2 Combustion chemistry

Combustion is a fast chemical reaction process in which a fuel combines with an oxidizer to generate combustion products and heat. It turns out that the steps of many combustion reactions are set by the reaction between H and O\textsubscript{2} molecules [11, 12]:

$$\text{H} + \text{O}_2 \rightarrow \bullet\text{OH} + \text{O} \quad (1.1)$$

This reaction generates \(\bullet\text{OH}\) radical, which plays an important role in many other elementary reactions. In fact, the chain reactions of almost all combustion processes are limited by the \(\bullet\text{OH}\) radicals generated from the elementary step 1.1 which serves as reactants of the subsequent steps. For example, the simplest gas phase hydrogen-oxygen reaction is supposed to consist of over 30 different elementary reactions. However the main rate-limiting process is the chain branching reaction that described by Eq. 1.1 [13]. Another example is the nitric oxide NO, which is mainly generated through the reaction of nitrogen and oxygen during combustion processes. The high temperature inside the internal combustion engines, power station boilers, incinerators or gas turbines splits the nitrogen from air nitrogen (N\textsubscript{2}) or from nitrogen bound in fuels, then the nitrogen react with oxygen to form NO and NO\textsubscript{2}. The principle of thermal NO reaction process can be described by Zeldovich mechanism [14]:

$$\text{N}_2 + \text{O} \rightarrow \text{NO} + \text{N} \quad (1.2)$$
$$\text{N} + \text{O}_2 \rightarrow \text{NO} + \text{O} \quad (1.3)$$
$$\text{N} + \bullet\text{OH} \rightarrow \text{NO} + \text{H} \quad (1.4)$$

Since the reactions of radicals (e.g. \(\bullet\text{OH}, \text{NO}_x, \text{HO}_2\), and \(\text{CH}_3\)\textsuperscript{•} etc.) control a large variety of combustion phenomena, actual experimental results recorded by
Effective gas sensors are necessary to compare and verify the development of the theory [15]. Moreover, the molecular kinetics and reactions in combustion processes are so fast and complex that only theoretical simulations assisted by computers are able to capture the chemical reactions in detail. On the other hand, in the areas with high motor vehicle traffic, the amount of combustion exhausts such as NO\textsubscript{X} emitted into the ground atmosphere can react to form smog and acid rain, causing significant air pollution issue. Therefore, monitoring techniques to measure concentrations of radical species with high temporal and spatial resolution in a wide pressure and temperature range have become extremely valuable tools to study many aspects of combustion, improve the efficiency of the engines, and regulate the emission of the combustion product [16, 17].

1.1.3 Breath analysis

Breath analysis is a promising medical diagnostic approach and has received serious clinical attention in recent years. It is a method to obtain information on the clinical state of an individual by monitoring chemical compounds present in the exhaled breath. Since there is a strong link between breath and blood, volatiles (metabolic or bio-markers) in the blood are transported into the lung via the alveoli, and then exhaled through the airway. Compared to other medical diagnostic methods such as blood/urine tests, biopsy, endoscopy, and imaging (e.g. CT, MRI) techniques, breath analysis offers a strong potential to significantly advance non-invasive, real-time disease monitoring, ease of repeating measurement. [18, 19, 20]

The earliest breath analysis was performed by Nobel prize winner Linus Pauling in 1971 who demonstrated that, despite the majority content of nitrogen, oxygen, carbon dioxide and water vapor, human breath contains various volatile organic compounds (VOCs) [21]. Latterly, more detailed studies discovered that normal human breath actually contains about four hundred major chemical compounds at the
concentration ranging from pptv levels to parts-per-million by volume (ppmv) levels [22, 23]. For example, Table 1.1 lists some of the identified compounds in the exhaled breath that have been established as “potential biomarkers” for particular diseases and metabolic disorders. Techniques for sensitive detection of these compounds without cross-interferences from other molecules in exhaled breath open a promising medical diagnostic field.

Table 1.1: Some established “potential biomarkers” and their physiological symptoms, table reproduced from [24]

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Metabolic status, disorders or diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon dioxide (CO₂)</td>
<td>Helicobacter pylori (H. pylori) infection, liver function, oxidative stress.</td>
</tr>
<tr>
<td>(¹³C-Isotopes)</td>
<td></td>
</tr>
<tr>
<td>Nitric oxide(NO)</td>
<td>Asthma, pulmonary hypertension, rhinitis, lung diseases</td>
</tr>
<tr>
<td>Acetone (OC(CH₃)₂)</td>
<td>Lung cancer, diabetes, dietary fat losses, congestive heart failure, brain seizure.</td>
</tr>
<tr>
<td>Acetaldehyde (CH₂CHO)</td>
<td>Alcoholism, liver related diseases, lung cancer.</td>
</tr>
<tr>
<td>Ammonia (NH₃)</td>
<td>Renal diseases, asthma.</td>
</tr>
<tr>
<td>Butane (C₄H₁₀)</td>
<td>Tumor marker in lung cancer.</td>
</tr>
<tr>
<td>Carbon monoxide (CO)</td>
<td>Oxidative stress, respiratory infection, anaemias.</td>
</tr>
<tr>
<td>Carbon disulphide (CS₂)</td>
<td>Schizophrenia, coronary, and artery diseases.</td>
</tr>
<tr>
<td>Carbonyl sulfide (OCS)</td>
<td>Liver related diseases.</td>
</tr>
<tr>
<td>Ethane (C₂H₆)</td>
<td>Vitamin E deficiency in children, lipid peroxidation, oxidative stress.</td>
</tr>
<tr>
<td>Ethylene (C₂H₄)</td>
<td>Lipid peroxidation, ultra violet radiation damage of skin.</td>
</tr>
<tr>
<td>Hydrogen peroxide (H₂O₂)</td>
<td>Asthma.</td>
</tr>
<tr>
<td>Hydrogen cyanide (HCN)</td>
<td>Pseudomonas aeruginosa in children affected with cystic fibrosis.</td>
</tr>
<tr>
<td>8-Isoprostane</td>
<td>Oxidative stress.</td>
</tr>
<tr>
<td>Isoprene</td>
<td>Blood cholesterol.</td>
</tr>
<tr>
<td>Methane (CH₄)</td>
<td>Intestinal problems, colonic fermentation.</td>
</tr>
<tr>
<td>Methanethiol (CH₃SH)</td>
<td>Halitosis.</td>
</tr>
<tr>
<td>Methanol (CH₃OH)</td>
<td>Nervous system disorder.</td>
</tr>
<tr>
<td>Methyl nitrate (CH₃NO₃)</td>
<td>Hyperglycemia in Type 1 diabetes</td>
</tr>
<tr>
<td>Nitrotyrosine (C₉H₁₀N₂O₅)</td>
<td>Asthma.</td>
</tr>
</tbody>
</table>

(Continued on next page)
Table 1.1: (continued)

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Metabolic status, disorders or diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pentane (C(<em>5)H(</em>{12}))</td>
<td>Peroxidation of lipids, liver diseases, schizophrenia, breast cancer, rheumatoid arthritis</td>
</tr>
<tr>
<td>Pyridine (C(_5)H(_5)N)</td>
<td>Periodontal disease</td>
</tr>
<tr>
<td>Sulfur compounds</td>
<td>Hepatic diseases and malordor, lung cancer</td>
</tr>
</tbody>
</table>

An interesting breath bio-marker is the free radical NO molecule. Detection of exhaled NO is of great importance in a number of bio-medical applications \[25, 26, 27\] and development of the sensing technique for NO molecule in the human exhaled breath is one of the major goals of this dissertation. NO synthesis and metabolism in human body is shown in Fig. 1.1. It is catalytically produced from biological systems by reduction from L-arginine with nitric oxide synthase (NOS) to form L-citrulline, which generates NO in the process. The metabolites mainly exist as nitrate (NO\(_3^-\)) in urine, nitrite and nitrate (NO\(_2^-\), NO\(_3^-\)) in blood, and NO in the exhaled breath. The presence of endogenous NO in exhaled breath of humans and animals was first reported in 1991 [28]. Since then, over 20 years of research has shown that exhaled NO can be used as an indicator of airway inflammation [19, 29, 30, 31, 32, 33], or transplant organ rejection [34]. It can play a critical role in cancer biology, treatment and prevention [35, 36]. Concentrations of exhaled NO originating from the lower airways in healthy patients typically vary from 5 to 20 parts-per-billion by volume (ppbv) [30]. Patients with pulmonary hypertension exhibit low levels of NO in their exhaled breath (< 5 ppbv) [38], while typical asthmatic patients have high NO levels range from 20 to 80 ppbv [39]. Currently, a major limitation to identify and treat asthma is that there are no compelling studies that are able to quantitatively reveal the correlation between NO metabolism and the related diseases. To do this, an NO analyzer with sensitivity of \(~ 1\) ppbv in a response time of \(~ 0.5\) second is needed. Ideally the instrument should also provide capability of selective quantification measurement of different NO isotopes for more precise studies of the kinetics of NO metabolism.
Figure 1.1: Nitric oxide synthesis and pathways in human body. (Figure is reproduced from website www.medscape.com [37].)

40 41 42. However, none of currently available NO analyzers can simultaneously provide all the above functionality required by the biomedical applications.

Since free radicals have unpaired valence electrons or open outer electron shells, they are highly reactive to other substances and play an important role in atmospheric chemistry, combustion, biochemistry and many other chemical processes. However, the highly reactive feature results in their very short lifetime and extremely low concentration in many chemical reaction processes, in which detecting radicals with high sensitivity, accuracy and selectivity is quite difficult. Development of novel detection techniques for quantitatively monitoring free radicals during the reaction processes will help scientists to better understand the functions of radicals in the en-
vironment and climate change, the kinetics and mechanism of radicals in combustion and energy technology, the pain and inflammation related to radicals in our body.

1.2 Techniques for trace gas sensing

A large number of analytical techniques have been developed for sensitive detection of trace gas molecules \cite{3,24,43,44,45,46}. There exist various technologies that can achieve high sensitivity, such as: (a) chemiluminescence \cite{47}, (b) electrochemical \cite{48,49}, (c) mass spectroscopy (MS) \cite{50}, (d) gas chromatography (GC) \cite{51,52} as well as purely laser based spectroscopic techniques including (e) absorption spectroscopy \cite{53,54}, (f) cavity enhanced spectroscopy \cite{55,56,57,58,59,60,61,62,63,64,65,66,67}, (g) laser induced fluorescence (LIF) \cite{66,67} and (h) FRS. Here I will mainly focus on the two most popular approaches: absorption spectroscopy and FRS. The remarkably high sensitivity and robustness of FRS technique that specifically targets radical species will be compared to the other methods and discussed in this section as well.

1.2.1 Absorption spectroscopy

Direct absorption spectroscopy

Absorption spectroscopy is generally realized by measuring the transmission of the light passing through an absorbing sample. The basis for assessment of the gas species concentration is the Beer-Lambert law which states that the transmission, $T$, of the light is related to the light pathlength $L$ and the concentration of the sample $N$:

$$T = \frac{I}{I_0} = e^{-\sigma(\nu) \cdot N \cdot L}, \quad (1.5)$$

where $I$ and $I_0$ are the intensity of the transmitted light and the incident light, respectively. $\nu$ is the frequency of the light. $\sigma(\nu)$ is the cross section of light absorption
by a single particle, which can be expressed in terms of the integrated molecular linestrength factor $S$ (cm$^2$cm$^{-1}$/molecule), and an (area-normalized) absorption lineshape function $\chi(\nu)$ (1/cm$^{-1}$):

$$\sigma(\nu) = S \cdot \chi(\nu).$$  \hspace{1cm} (1.6)

Fig. 1.2 shows a schematic layout of the basic principle of absorption spectroscopy. The laser source emits a monochromatic light whose frequency is closed to a transition of the target molecule sample. As light passes through the sample it is attenuated and the transmitted light is then focused on a detector that produces signal from which sample concentration can be retrieved. For low concentration trace gas detections, the absorbance is so small ($\alpha(\nu) \ll 1$) that Eq. 1.5 can be approximated by:

$$I = I_0(1 - \alpha(\nu)).$$  \hspace{1cm} (1.7)
Thus, the power absorbed is directly proportional to the sample concentration $N$:

$$
\Delta P \propto I_0 - I = I_0 \cdot \alpha(\nu) = I_0 \cdot S \cdot \chi(\nu) \cdot N \cdot L.
$$

(1.8)

The absorbed power $\Delta P$ is the desired signal that any absorption spectroscopic technique wants to resolve. It is a wavelength dependent parameter as the absorption cross section $\sigma(\nu)$ is proportional to the line-shape function $\chi(\nu)$ with maximum at the transition center. Ideally, with the laser wavelength coincident with the desired transition center, the system is able to continuously monitor the sample concentration by reading the photodetector output. However, in practical situation any free running lasers tend to drift in wavelength due to instability of current drivers or laser temperature controllers, which results in signal drift and reduces the sensor stability. Therefore, a common approach is to scan the laser wavelength over the absorption line and record the whole absorption profile. By fitting the measured line profile using established theoretical molecular line-shape functions (Lorentzian, Gaussian, or Voigt), the sample concentration can be retrieved.

A greater signal-to-noise ratio (SNR) or sensitivity is always desired for trace gas monitoring applications. Eq. 1.8 shows there are three ways to increase the signal strength: 1) to increase the incident laser intensity $I_0$, 2) to address transitions with larger line strengths $S$ and, 3) to increase the optical path-length $L$. However increasing the incident laser intensity does not lead to significantly higher sensitivity as the system background noise and interference fringes also increase with respect to the light intensity. The linestrength is a constant (e.g. obtained from HITRAN (High Resolution Transmission database) [68]) that does not change with external conditions. Among all the molecular transitions there is always a strongest absorp-

---

1At high pressure where collisional broadening dominates, line-shape function is a Lorentzian function and at low pressures (less than a few Torr), Doppler broadening dominates and the function is a Gaussian. At intermediate pressures a Voigt function is used which is a convolution of Lorentzian and Gaussian profiles.
tion transition which can be selected as the target transition. The dependence of line-strength to frequency and the selection of appropriate laser wavelength will be described in details in the Section 1.3.

**Multi-pass cell and cavity enhanced absorption spectroscopy**

Increasing the optical path-length $L$ is the simplest and the most feasible approach to enhance the system sensitivity in practical applications. Optical path-length can be enhanced either by using a multi-pass absorption cell or an optical cavity. White cells [69, 70] and Herriott cells [71] are the two most popular cells used in laser absorption spectroscopy. The curvature and the distance of the spherical mirrors are carefully designed so that the laser beam enters through a hole on one mirror and is reflected back and forth a number of times before exiting the other hole (or the same hole). Fig. 1.3 shows that the astigmatic mirror designs dramatically increase the effective optical path-length in a limited volume with minimal beam overlap (which might cause interference fringes), while preserving simple manufacturing and optical configurations [73, 75]. In a recent demonstration, Nelson et. al showed a NO sensor
based on a multi-pass astigmatic cell with effective path-length of 210 meters. The sensor achieved a minimum detection limit of 0.03 ppbv with 30 s integration time (~0.16 ppbv/Hz⁰.⁵ if normalized to detection bandwidth), which makes it useful for long-term and unattended monitoring of NO in the atmosphere [76].

Multi-pass cells usually create effective path-lengths up to several hundred meters. Effective path-lengths of up to a few kilometers can be achieved by using cavity enhanced method, such as cavity ring-down spectroscopy (CRDS) which uses two high-reflectivity mirrors (typical reflectivity > 99.9%) to bounce a short laser pulse back and forth on the order of ~10,000 times [55, 77]. The intensity of the trapped pulse will decrease exponentially due to both absorption and scattering by the gas sample within the cavity and mirror losses. Measurement of the pulse decay rate which depends on the loss mechanism gives the specific analyte concentration inside the cavity. However, absorption techniques based on multi-pass cell or cavity enhanced method require highly precise and stable optical alignment, which is quite challenging and increases the maintenance cost especially at harsh environmental conditions [78, 79]. In addition, the wall quenching issue in the relative large sampling cells causes measurement uncertainty, which has recently been blamed for the failure in detection of HO₂ radicals in a CRDS experiment [80, 81].

**Wavelength modulation spectroscopy**

Besides increasing the optical path-length, another approach to enhance the sensitivity in the absorption spectroscopy is to reduce the system noise by employing modulation technique and phase-sensitive lock-in detection. Fig. 1.4 shows the basic principle for the most common modulation technique, called wavelength modulation spectroscopy (WMS). WMS involves the modulation of the laser wavelength via current at a high frequency f with amplitude close to the width of the absorption line. The interaction between the wavelength modulation and the nonlinear absorption pro-
Figure 1.4: The basic principle of WMS. (Figure is reproduced from Ref. [82].)

file yields multiple harmonic components in the photodetector output signal, which can be demodulated with lock-in amplifiers. Generally the second harmonic (2f) is used because its profile is mostly similar to the direct absorption profile with signal peak locating at the transition center. The use of a lock-in amplifier is the key to enhance the system sensitivity because it serves as a band-pass filter to reject a large portion of noise that falls outside of the pass band, including laser intensity noise and photodetector electronic noise. Another noise reduction scheme is due to the fact that technical noise usually decreases with increasing frequency (often referred to as a 1/f noise [83]) and SNR improvement is achieved by encoding and detecting the absorption signal at a high frequency, where the noise level is low. In most situations, the WMS method is combined with multi-pass cell approach to utilize both noise reduction and signal enhancement benefits. WMS is also limited by the multiple re-
flections in the optical system which create optical interference fringes (an effect of periodic intensity modulation on spectrum due to constructive and destructive coherent interference) and result in the typical sensitivity in the range of $10^{-4}$ Hz$^{-1/2} \sim 10^{-5}$ Hz$^{-1/2}$ [46, 84, 85, 86]. Moreover, all absorption spectroscopic techniques would simply give the sum of all absorption contributions from the mixture, which limits the selective analysis of individual components in the mixture.

### 1.2.2 FRS

FRS is a technique that can provide ultra-sensitive and selective detection of small free radical species [87, 88, 89, 90, 91, 92, 93] such as NO, NO$_2$, O$_2$, ·OH, HO$_2^*$, and CH$_3^*$. As introduced in Section 1.1, there are urgent needs for monitoring free radicals in areas such as atmospheric research, medical diagnostics, and combustion chemistry, etc. Free radicals have unpaired valence electrons which creates net magnetic dipole moment, and therefore exhibit a paramagnetic property – an effect that the molecules (or atoms) are attracted by an externally applied magnetic field. The FRS technique utilizes the linear resonant Faraday effect which is also known as the Macaluso-Corbino effect [94]. Please note the difference between the resonant Faraday effect in paramagnetic species and the well-known Faraday effect in transparent magneto-optical materials such as terbium gallium garnet crystal (TGG) used in optical isolators. Taking the Verdet constant $V$ (a characterization parameter of the magnitude of magneto-optical rotation) as an example, typical crystals that are used in commercial Faraday polarization rotators and optical isolators have $V \approx 3 \times 10^{-5}$ rad G$^{-1}$ cm$^{-1}$. However $V$ can reach up to 100 rad G$^{-1}$ cm$^{-1}$ in resonant magneto-optical rotation observed in rubidium vapor [95]. Resonant Faraday effect was discovered by Macaluso and Corbino at the end of the 19th century, when the scientists were studying absorption spectra of the alkali atoms in the presence of magnetic fields. The connection between the Faraday and the Zeeman effects was first established by
Voigt, who also explained the observations of Macaluso and Corbino. Fig. 1.5 shows the basic principle for FRS. The presence of magnetic field results in the Zeeman split of the degenerated energy states of the molecules which causes magnetic circular birefringence, MCB (difference in refractive indices for left-handed, LHCP, and right-handed, RHCP circularly polarized components) and, magnetic circular dichroism, MCD (difference in absorption coefficients for LHCP and RHCP inducing ellipticity in the output light). Due to MCB, when linearly polarized light propagates through the sample it undergoes rotation of its polarization axis (the Faraday effect) that is proportional to sample concentration. The ability of indirect sample modulation via magnetic field allows achieving superior noise suppression, which enables extremely sensitive detection without the need for long optical paths. With short optical paths the MCD effects do not contribute significantly to the measured signals and are usually neglected. In a conventional FRS system a polarizer is placed before the sample cell to establish a well-defined polarization of light and a nearly crossed analyzer is placed after the sample to convert Faraday rotation to intensity changes that can be detected by a conventional photodetector.

Figure 1.5: The basic principle of FRS. (Figure is reproduced from Ref. [94].)
An early example of applying Faraday rotation to molecular spectroscopy was presented by Mann et. al. The authors used a zirconium arc source and a multi-pass cell with over 2000 Gauss magnetic field to demonstrate that the magneto-optical rotation spectrum of NO is easier to interpret than the absorption spectrum [96, 97]. Faraday rotation was firstly applied to analytical spectroscopy for trace mercury detection in 1974 by Church et. al., who used an incoherent lamp as the light source and showed that the new technique has sensitivity an order of magnitude higher than that could be obtained with absorption measurements [98]. The detection of trace gases by Faraday rotation was significantly advanced by employing laser sources with much higher frequency resolution and optical intensity. Using tunable color-center lasers, Litfin et. al. [88] demonstrated 50 times better sensitivity compared to absorption spectroscopy in detection of NO in the molecular overtone band of 2.7 μm with FRS. Another interesting result was obtained by Hinz et. al. [99] who worked on a CO laser targeting the fundamental band of NO molecules in the mid-infrared range.

To reduce noise, modulation techniques with phase-sensitive lock-in detection similar to those used in WMS is the basis of the two FRS methods which include an AC-FRS in which modulated magnetic field is used to modulate the magneto-optical properties of the sample [88, 90], and a DC-FRS in which a static magnetic field with a wavelength-modulated laser are used [93, 100]. Generally speaking there are mainly four reasons why FRS technique has high-sensitivity compared to absorption spectroscopy:

1). In principle, FRS measures the refractive index difference due to magnetic birefringence while absorption spectroscopy measures the intensity changes. The measurement of the change in refractive index (phase) can offer many potential advantages compared to the measurement of the change in intensity. The most significant factors include the linear relationship between the dispersion spectrum and the
sample concentration (unlike absorption that saturates with increasing concentration) which gives FRS large dynamic range on the measurement of the gas concentration, phase measurements that are immune to laser intensity fluctuations, and effects of the photodetector nonlinearity that are effectively eliminated in measurements that rely on detection of phase instead of intensity [101, 102].

2). In AC-FRS modulated magnetic field selectively modulates magneto-optical properties of the sample, and thus magnetically inactive background can be removed and spectroscopic signal can be effectively distinguished from parasitic etalon effects, which are the main limiting factors on the sensitivity improvement of absorption spectroscopy. In DC-FRS wavelength modulation allows all wavelength-dependent spectral features (e.g., spectral fringes, absorption features, etc.) to contribute to the modulation of laser power. Therefore the presence of parasitic etalon fringes and/or interfering spectral features can affect the performance of DC-FRS system [93]. Fig. 1.6 shows the typical AC-FRS spectra and the inset shows the DC-FRS

Figure 1.6: Signal spectrum using AC-FRS and DC-FRS technique (shown in inset) indicating strong etalon fringes on DC-FRS techniques.
spectra acquired with same magnetic field strength and calibrated NO sample using the same optical set-up. It was identified that the system was not optimized for DC-FRS, because its spectra contains too strong optical fringes. In AC-FRS the fringe noise was effectively removed, leading to a significant improvement in SNR.

3). Another key component giving FRS ultra-high sensitivity is the ability of effective suppression of the laser intensity noise. In the AC-FRS technique the analyzer is set to nearly crossed (so called “90-degree method”) with the input polarizer, therefore only light from the orthogonal polarization is detected and most of the laser power that containing the major portion of laser RIN is blocked. In the DC-FRS arrangement due to the high parasitic etalon fringes an analyzer that gives access to both polarization components (e.g. Wollaston prism) can be set at 45° and two photodetectors can be used to perform balanced detection of the FRS signals (so called “45-degree method”). In 45-degree method the laser RIN that occurs as common-mode on both detector elements can be suppressed through electronic subtraction of the photocurrents (balanced detection). The same operation retrieves the FRS signals, which have opposite phases and thus are constructively added [100].

4). AC-FRS offers high immunity to the spectral interference from other diamagnetic species (H$_2$O or CO$_2$) which do not respond to modulated magnetic field. Since H$_2$O and CO$_2$ are by products that largely contained in the exhaled breath or in hydrocarbon combustion exhaust, they generate significant detection interference in absorption based methods. FRS is sensitive only to paramagnetic species which is a significant advantage for the applications in biomedical and combustion areas. Recently, Brumfield et. al. successfully demonstrated the first direct in situ measurements of hydroperoxyl radical (HO$_2$) at atmospheric pressure from the exit of a laminar flow reactor using mid-IR FRS. The signals from diamagnetic molecules, such as H$_2$O, are removed, which significantly mitigates the problem of spectral interference [103].
In summary, compared to the commonly used absorption spectroscopy methods, FRS technique provides effective elimination of parasitic etalon effects, which affect the performance in most of the optical absorption based sensing systems. The system sensitivity can be largely improved compared to current absorption based spectroscopy techniques. As a result, to achieve the similar sensitivity as the absorption spectroscopy FRS does not require very long optical pathlengths, leading to much simpler and rugged opto-mechanical construction with significant system stabilization and miniaturization potential. All of these advantages make FRS a promising technique for the development of the next generation free radical analyzers.

Despite the variety of FRS signal retrieval methods, approaching fundamental limits of detection has been difficult so far. For example in DC-FRS based on 45-degree method despite efficient reduction of $1/f$ laser noise through high frequency wavelength modulation, the ultimate sensitivities are often limited by parasitic Fabry-Perot interference fringes [93] that are not sufficiently suppressed with balanced detection scheme [100, 104, 105, 106]. Moreover, in the mid-IR where molecular detection can be performed with the highest sensitivities, availability of commercial balanced photodetectors is very limited and custom devices can be costly. Therefore AC-FRS approach with a single photodetector element was the method of choice in mid-IR [88, 89, 87, 90]. Selective modulation of magneto-optical properties of the sample used in AC-FRS helps effectively distinguish the spectroscopic signal from parasitic etalon effects. However, high currents needed to produce required magnetic field generate electro-magnetic interference (EMI) that is difficult to control and usually deteriorates long-term stability of FRS systems (electronic pick-ups create fluctuating offsets in the measured signal or cause uncontrolled laser wavelength modulation). Moreover, when driving solenoids with high currents systems must be operated at relatively low frequencies (∼single-kHz), and $1/f$ noise significantly limits the sensitivity of AC-FRS spectrometers.
The thesis mainly focuses on new methodologies to improve the sensitivity for the FRS technique which gives us an opportunity to approach the fundamental shot-noise limit. NO will be the target molecule, however, the developed new methods can be applied to all FRS based sensors monitoring other small free radicals (i.e. NO₂, O₂, •OH, HO₂•, and CH₃• etc.).

1.3 Mid-IR laser sources

1.3.1 Why Mid-IR?

The mid-IR is the region of the electromagnetic spectrum with wavelengths from 3 to 30 µm. The spectroscopic sensing in mid-IR range has drawn enormous interest since many trace gas molecules have their fundamental rotational-vibrational absorption bands in this electromagnetic region. The transition spectrum at IR (also called molecular “finger-print” region) is very specific to a particular molecule, allowing highly selective trace gas detection if light sources with appropriate wavelengths are used. In addition, as discussed in Section 1.2.1 the transition line-strength are essential to achieve highly sensitive trace gas detection. The strongest vibrational bands occur within the mid-IR range, which are several orders of magnitude stronger than the overtone and combination bands in the near-infrared region (near-IR). As a result detection limits can be achieved down to ppbv or even pptv level. Fig. 1.7 shows the line-strengths and wavelengths of ro-bivrational transitions for several common target molecules. Fig. 1.8 illustrates that line strengths of NO in the mid-IR are two to three orders of magnitude stronger than its near-IR overtone transitions, meaning that an equivalent detection limit could be reached in the mid-IR with an optical path-length that is several hundreds times shorter than what could be achieved in the near-IR.
Figure 1.7: Line strengths of ro-vibrational transitions of common gas molecules taken from HITRAN 2008 database [68]. (Figure is reproduced from Ref. [107].)

Figure 1.8: Comparison of integrated linestrengths of NO transitions in the mid-IR and near-IR region. The linestrengths are roughly two and three orders of magnitude larger when comparing the fundamental rotational-vibrational band to the 1st and 2nd overtones, respectively. (Figure is reproduced from Ref. [108].)
1.3.2 QCLs

Development of room temperature, high power, continuous wave (CW)operation QCLs have provided convenient access to the fundamental ro-vibrational molecular transitions located in the mid-IR [109, 110]. As a result, QCL-based trace-gas sensing instrument reached the level of performance that was inaccessible before [111, 112, 113].

QCLs are unipolar injection lasers based on intersubband transitions in a repeated stack of semiconductor multiple quantum well heterostructures. Unlike laser diodes that rely on interband transitions, photons emitting from QCLs rely on the transitions of confined electronic states in the conduction band (Fig. 1.9 left), which can be achieved by means of band-structure engineering of the thickness of the quantum wells and barriers using molecular beam epitaxy (MBE) or metalorganic chemical vapor deposition (MOCVD) technique (Fig. 1.9 right). The first realization of a QCL
was reported in 1994 by Faist et. al. [109]. In the following years the optimization on the material systems, active region designs and QC structure growth using MOVCD technique present tremendous advantages of QCLs over the other mid-IR laser sources and paved the way for further applications of QCLs in laser spectroscopy world.

QCLs show excellent spectroscopic properties in terms of device dimension, continuous operation, optical power, temperature performance, spectral line-width and wavelength tunability. As a laser source the small foot-print of QCLs allows compact sensing systems to be built while gas lasers (e.g. CO$_2$ lasers), optical parametric oscillators or Fourier Transform Infrared Spectrometers (FTIRs) are typically large and bulky systems. Typical QCLs can operate at room temperature with either liquid or air cooling and are capable of generating hundreds of mW output optical powers. The emitting wavelength of a QCL can be arbitrarily designed (band-gap engineering) to target specific mid-IR wavelength region, while the best performing solid-state lasers (e.g. rare-earth ions in crystalline hosts) only emit in the 1-5 $\mu$m range due to fabrication challenges [114, 115, 116, 117, 118]; Current QCLs are able to operate in CW mode with thermoelectric cooling, but antimonide and lead-salt diode lasers with emitting wavelength longer than 3 $\mu$m still need cryogenic cooling [119, 120], which may ultimately degrade system performance. Additionally, QCL heterostructures can be designed with several closely spaced energy levels to create an ultra-broadband gain profile for large wavelength tunability. For example, quantum designs such as the “continuum-to-bound” and “bound-to-continuum” quantum structures or a heterogeneous cascade of multiple “bound-to-continuum” have demonstrated ultra-broadband gain spectra that are greater than 430 cm$^{-1}$ wide [121, 122]. High power QCLs are especially suited for applications such as stand-off detection [123, 124] and they help increasing the sensitivity of spectroscopic systems. Recently a QCL with master oscillator and power amplifier (MOPA) design emitting peak output power of 1.5 W at 300 K with single-mode emission was realized [125].
1.3.3 QCL types

FP-QCLs

This is the simplest structure of a QCL. An optical waveguide is first fabricated following the standard quantum cascade material growth procedure to form the gain medium. The crystalline semiconductor devices are then cleaved to form two parallel facets on either end of the waveguide, thus forming a Fabry-Perot resonator. The residual reflectivity on the cleaved facets from the semiconductor-to-air interface is sufficient to create a resonator with free spectral range (FSR) of \( \frac{c}{2nL} \), where \( c \) is the vacuum light speed, \( n \) is the effective waveguide refractive index, and \( L \) is the cavity length. As the injection current is large enough, all the waves satisfying Fabry-Perot constructive interference in the gain profile of the chip are available for lasing, leading to multiple longitudinal mode operation with mode spacing equal to the FSR. High power FP-QCLs can be used as IR emitter for target pointer and illumination. With multiple laser modes equally spaced in a wide wavelength region, an FP-QCL can be used as coherent IR comb source to replace the bulky and low efficient blackbody source typically used in an FTIR spectrometer, leading to much simpler and more compact design of the entire instrument. Recently, a compact, broadband, mid-IR frequency comb based on four wave mixing phase locking scheme using a semiconductor FP-QCL was demonstrated by Hugi et. al. [126]. Although it is still at a very early stage, it opens the way to utilize the simplest FP-QCL devices in realization of broadband, compact, all-solid-state mid-infrared spectrometers. This capability has been studied in this thesis.

DFB-QCLs

High resolution spectroscopic applications require single frequency (single mode) operation of a QCL source, which can be realized in several ways. The most popular
method is to fabricate a DFB grating structure on top of the laser active region. The grating acts as the wavelength selective element and provides the feedback to achieve lasing. The 1-D periodic grating is constructed so only a narrow band of wavelengths is reflected, and thus a single longitudinal lasing mode is produced with side-mode suppression ratio better than 25 dB. \[127\]. The Bragg wavelength, $\lambda_B$, is described by the following equation: $\lambda_B = 2n_{\text{eff}} \cdot \Lambda$ where $n_{\text{eff}}$ is the temperature-dependent waveguide effective refractive index and $\Lambda$ is the grating period. Frequency tuning of a DFB-QCL can be realized by means of thermo-optic effect, either directly by slowly varying the laser substrate temperature, or via a faster joule heating process of changing injection current. The frequency tuning coefficient by thermo-optical effect is about -0.1 cm$^{-1}$/K. The widest tuning range of a DFB-QCL by thermoelectric or Peltier cooler is on the order of 10 cm$^{-1}$. Therefore, DFB-QCLs are typically designed to target a single well resolved ro-vibrational absorption line of one specific molecule. \[128, 129, 130\]. While for WMS current tuning is the preferred method because it is simple and fast enough to generate laser wavelength modulation up to GHz frequency \[131\], with this the laser excess noise with $1/f$ dependence can be effectively avoided and the system sensitivity can be increased. The main drawback of DFB-QCLs is its narrow tuning range, which is quite small compared with the entire gain bandwidth of a QCL and limits DFB-QCLs to probe broadband absorption features such as glucose molecules or multiple gas molecules. To overcome this issue DFB-QC lasers with slightly different output frequencies has been fabricated out of the same epilayer with different grating periods \[132\]. An array of these DFB-QCLs covering a broad continuous wavelength range of 220 cm$^{-1}$ at 8.4 $\mu$m has been used to perform spectroscopy analysis of liquid samples \[133\]. Nevertheless, the drawbacks of this technique come from fabrication errors of the multiple laser chips, including errors on output wavelengths, different thresholds, slope efficiencies, output powers of the individual DFB lasers in the array. As a result, the lower yield of high quality
array devices compared to that of a single DFB laser increases manufacturing costs. Moreover, combining multiple laser beams into a single beam that can be focused onto a photodetector requires much effort. Therefore at the present time, it is easier and more economically feasible to produce FP chips with broadband gain profiles and to use an external grating configuration for wavelength tuning and control.

**EC-QCLs**

EC-QCLs overcome the limited tuneability of DFB-QCLs by using an external wavelength-selection element such as an optical grating, whose feedback wavelength depends on a complicated mechanical tuning approach. Although less compact, the external cavity configuration is capable of selecting the emitting wavelength anywhere within the available QCL gain spectra without changing the chip temperature, thus allowing more efficient utilization of the QCL spectral gain and significantly increasing the laser spectral coverage. The invention of the “bound-to-continuum” active region design [134] showed advantages of broadening the gain spectral bandwidth and enhancing the gain. EC-QCLs utilizing such a gain medium will therefore result in broad tuning range with large output powers. For instance, using substacks of 5 bound-to-continuum active regions with gain peaks at different wavelengths, Hugi et. al. demonstrated a room temperature ultra-broad-band EC-QCL tunable from 7.6 µm to 11.4 µm with a peak optical output power of 1 W and an average output power of 15 mW [122].

A typical EC-QCL consists of three basic components, a FP-QCL chip, a collimating lens, and a diffraction grating. The two most commonly used EC-QCL configurations are Littrow configuration and Littman-Metcalf configuration as shown in Fig. 1.10 [135]. Both configurations have been well studied and showed excellent performance in spectroscopy applications. In Littrow configuration, the laser output power is taken from the zero-order diffraction of the grating and the first order diffrac-
Figure 1.10: a)Littrow and b) Littman-Metcalf external cavity configuration. The large grey arrows indicate the directions of propagation of light and the small black arrows show how coarse tuning is achieved. \( \theta \) indicates the Littrow angle. (Figure is reproduced from [136].)

The feedback provides the laser feedback at a certain wavelength and a specific grating angle. Therefore wavelength tuning is accomplished by varying the diffraction grating angle (Littrow angle). In Littman-Metcalf configuration the grating orientation is fixed, and an additional mirror is used to reflect the first-order beam back to the laser. The wavelength can be tuned by rotating the mirror.

Despite the various configurations the FP-QCLs play an important role in EC-QCL setups as they are the gain medium inside the laser cavity. Compared to diode lasers used in telecommunications, manufacturing high quality broadband QCLs with high yield is more challenging due the complexity and accuracy required to control the superlattice growth of the QCL active region, especially for the ultra-broad-band QCL gain chips that contain hundreds of layers of quantum wells. Slow feedback from the chip test procedure also hinders improvement of the production. The traditional procedures to characterize EC-QCL chips are based on manual testing on chip at a time in an external cavity, which is highly time consuming and labor intensive. Techniques to improve the manufacturing efficiency and QCL testing efficiency are currently attractive to many QCL manufacturing companies.
1.4 Organization of the thesis

This thesis is organized as follows. Chapter 2 introduces the essential theory for FRS including signal strength simulation, SNR optimization and calculation of system ultimate sensitivity. Chapter 3 presents an AC-FRS system integrated with a multi-pass Herriott cell, which was proved to be an effective and straight way to enhance the FRS sensitivity. Further sensitivity enhancement can be realized by two new methods employing heterodyne-enhanced FRS detection (H-FRS) and dual modulation FRS (DM-FRS) schemes which are discussed in Chapter 4 and Chapter 5 respectively. New sources and techniques to perform high-resolution, broadband spectroscopy in the mid-IR are discussed next. An automatic QC laser bar testing system which has been proved to be a cost-efficient apparatus for characterizing EC-QCL chips in an efficient way is presented in Chapter 6. A novel broadband, high resolution spectroscopy technique based on multi-heterodyne of two FP-QCLs is introduced in Chapter 7. Chapter 8 concludes the thesis and discusses future research directions.
Chapter 2

Faraday rotation spectroscopy
theory

This chapter describes the theory of FRS. Based on the theory one can calculate the spectrum of Faraday rotation angle with given experimental setup and parameters. The best molecular transition line, sample pressure and magnetic field can be optimized based on the theoretical simulations. The noise contribution and the dependence of various noise sources on the analyzer offset angle are modeled and discussed. The chapter also addresses the possibility of achieving shot-noise-limited performance with an FRS system. The shot-noise-limited condition and ultimate detection limit are derived assuming the system noise is dominated by the fundamental quantum noise.

2.1 Signal

2.1.1 Faraday rotation angle

In mid-IR FRS, the presence of magnetic field results in the split of the rotation states of the paramagnetic molecules which causes magnetic circular birefringence, MCB.
(difference in refractive indices for left-handed, LHCP, and right-handed, RHCP circularly polarized components) and magnetic circular dichroism, MCD (difference in absorption coefficients for LHCP and RHCP). In most situations, it is applied in detection of trace gas with ultra low concentration, thus the absorption and MCD effect can be neglected. Due to MCB, when linearly polarized light propagates through the sample with paramagnetic species and under longitudinal magnetic field it undergoes rotation of its polarization axis that is proportional to sample concentration. The angle of the polarization rotation (so called Faraday rotation angle \( \Theta \)) can be described as \( \Theta = (n_R - n_L)\pi L/\lambda = \Delta n_{R-L}(\nu)\pi L/\lambda \), where \( L \) is the effective optical path length within the sample under magnetic field, \( \lambda \) is the wavelength of light and \( \Delta n_{R-L}(\nu) \) is a difference between refractive indices for RHCP and LHCP components at a given optical frequency \( \nu \) and magnetic field strength \( B \). \( \Delta n_{R-L}(\nu) \) can be calculated by summarizing all the differences between dispersion curves from the magnetic degenerated upper and lower rotation \( M \) states which obey quantum selection rule [89]:

\[
\Delta n_{R-L}(\nu) = \sum_p \sum_{M',M''} pD_{M'M''}(\nu),
\]

(2.1)

where \( M' \) and \( M'' \) are the magnetic quantum numbers for the upper and lower states, and \( p = M' - M'' \) gives the selection rule on \( M \). In FRS configuration since magnetic field is parallel with the optical axis, \( p = \pm 1 \). The \( D_{M'M''}(\nu) \) represents an individual \( M \) transition with shifted line center under magnetic field:

\[
D_{M'M''}(\nu) = X_{M'M''}\text{disp}(\nu, \nu_{M'M''}),
\]

(2.2)

where \( X_{M'M''} \) is the line strength factor:
\[
X_{M'pM''} = S_{J'pJ''} \cdot 3 \left( \begin{array}{cc} J' & J'' \\ -M' & M'' \end{array} \right)^2 e^{-\left( \frac{\hbar \nu_{J''}(T - T_0)}{kT T_0} \right)},
\]

(2.3)

where \( S_{J'pJ''} \) is the integrated line intensity of transitions under zero field strength and standard temperature and pressure, which can be acquired from HITRAN database \[68\]. The second matrix term is calculated by the corresponding Wigner 3-J symbol \[137\] which depends on the \( J \) and \( M \) quantum numbers of upper and lower level. The factor of 3 originates from the normalization requirement. The exponential term gives the temperature dependence of the Boltzmann population factor for the lower level assuming the upper rotational level is unpopulated. \( h\nu_{J''} \) represents the energy of the lower level state in the transition. \( k \) is Boltzmann’s constant, \( T \) is the sample temperature and \( T_0 = 296 \) K is the reference temperature.

In most situation contributions from pressure broadening and Doppler broadening are relevant, therefore the gas absorption line shape can be calculated by the convolution of a Lorentzian and a Gaussian line shape from the two broadening mechanisms (a.k.a. Voigt profile):

\[
\text{absp}(\nu, \nu_{M'pM''}) = \frac{1}{N_V} \int_{-\infty}^{\infty} \frac{1}{\left( \frac{\Gamma_D}{2} \right)^2 + (\nu - \tau - \nu_{M'pM''})^2} \times \exp \left( -\frac{\tau^2}{\left( \frac{\Gamma_D}{2\sqrt{\ln2}} \right)^2} \right) d\tau,
\]

(2.4)

where \( N_V \) is a scale factor, \( \Gamma_D \) and \( \Gamma_p \) is the Doppler and Lorentzian line width (FWHM) of the transitions, respectively. The \( \text{disp}(\nu, \nu_{M'pM''}) \) term in eq(2.2) is the dispersion line shape factor, which is the real part of the complex refractive index and can be calculated from Kramers-Kronig relations \[138\] using the Voigt absorption line calculated by Eq. 2.4.

For example, Fig. 2.1 shows the dispersion curves \( D_{M'pM''}(\nu) \) of all the 8 allowed transitions (as depicted in Fig. 2.1(c)) of the R(3/2) transitions for a paramagnetic
Figure 2.1: (a) Signal contributions of the individual transitions. (b) FRS signal shape. (c) Allowed transitions shown by vertical arrows. (reproduced from [87])
molecule. The 8 transitions include 4 transitions with \( p = 1 \) (thin solid curves in Fig. 2.1(a)) and 4 transitions with \( p = -1 \) (thin dot curves in Fig. 2.1(a)). The sum signal for \( p = 1 \) (thick solid curve) and \( p = -1 \) (thick dot curve) represent the dispersion lines for RHCP and LHCP light, respectively. Under non-zero magnetic field, the frequency shift for each of the individual degenerate dispersion curve is given by (see Appendix A)

\[
\Delta \nu_{M'\,M''}(B) = \frac{(M'g' - M''g'')\mu_B B}{h},
\]

where \( h \) is Planck’s constant, \( \mu_B \) is the Bohr magneton, \( g' \) and \( g'' \) are the g factors for the upper and lower states in the transition [139]. The sum of all these frequency shifted dispersion curves with the appropriate intensity factors finally creates the difference between refractive indices for RHCP and LHCP components as shown in Fig. 2.1(b).

Fig. 2.2 shows the calculated Faraday angle \( \Theta \) for the \(^{14}\)NO fundamental band around 5.3 \( \mu \)m. The simulation was implemented using software Matlab (MathWorks) with magnetic field \( B = 100 \) Gauss (G), sample pressure \( P = 35 \) Torr, optical path \( L = 15 \) cm, NO concentration of 2 ppmv, sample temperature \( T = 300 \) K. The other parameters such as zero field transition line center \( \nu_{J''} \), integrated line intensity \( S_{J',J''} \), line width pressure and temperature broadening coefficient were taken from the HITRAN database [68]. The \( g \) factors were taken from [139]. Simulation was performed for transitions of P(3/2)-P(29/2), Q(1/2)-Q(29/2), and R(1/2)-R(27/2). The wavenumber ranges from 1820 cm\(^{-1}\) to 1920 cm\(^{-1}\) with step size of 0.000125 cm\(^{-1}\).

Fig. 2.2 and Fig. 2.3 clearly indicate that Q(3/2) transition at 1875.8 cm\(^{-1}\) is the optimum transition that generates the maximum FRS signal for \(^{14}\)NO sensing. Although the Q(3/2) transition has lower integrate line intensity than those of the strongest P- and R-branch transitions, higher \( g \)-factors due to low \( J \) allow for optimum Zeeman split at relatively low magnetic field (100 G). Fig. 2.4 shows the dependence
Figure 2.2: Faraday rotation angle for the entire $^{14}$NO fundamental band. Simulation condition: Magnetic field $B = 100$ G, Pressure = 35 Torr, Optical path = 15 cm, NO concentration = 2 ppmv, Temperature = 300 K.

Figure 2.3: Faraday rotation angle for the $^{14}$NO Q-branch.
of Faraday angle on the strength of applied magnetic field with a constant sample pressure. The Faraday angle is the peak signal obtained from the calculated Q(3/2) spectrum. Calculation shows that the Faraday rotation signal shows maximum at an optimum strength of magnetic field. As shown by Fig. 2.1(a), the optimum Zeeman split for Q(3/2) transition occurs when the maximum of the RHCP dispersion curve ($\sum_{p=1} D_{M'M'^\nu}(\nu)$ in Fig. 2.1(a)) aligns with the minimum of the LHCP dispersion curve ($\sum_{p=-1} D_{M'M'^\nu}(\nu)$ in Fig. 2.1(a)). Higher magnetic field causes over-splitting that does not contribute stronger signal. On the other hand, when magnetic field strength is constant, the Faraday angle depends on the sample pressure which affects the width and intensity of the transition line. Fig. 2.5 shows the dependence of Faraday angle on the sample pressure with constant magnetic field. Calculation shows that there is an optimum sample pressure to achieve the greatest Faraday angle with 150 G magnetic field.

Fig. 2.6 shows the calculated Faraday angle spectrum for $^{14}$NO with 1000 G magnetic field. In this case, the optimum transition is no longer Q-branch but R-branch.
which is mainly because the strong field generates greater Zeeman split for R- and P-branch with greater integrate line-strength factor. Due to the system power limitation, achieving 1000 G is only feasible by using the rare-earth permanent magnets. Thus, a higher sensitivity can be potentially realized if DC-FRS method based on strong field permanent magnets is used.

### 2.1.2 FRS signal generation

Typically, a polarizer before the sample is used to establish a well-defined linear polarization state of light and a polarization analyzer after the sample is used to detect Faraday rotation. The function of the analyzer is providing effective suppression of the laser RIN \[88\] and transforming polarization modulated light into intensity modulation, which can then be detected by a photodetector. In order to further suppress the broadband noise from laser source and photodetector a phase sensitive lock-in amplifier incorporate with modulation of the signal is used to pick-up the signal and reject the noise effectively. As described by Section 1.2.2, there are two main approaches in conventional FRS: AC-FRS and DC-FRS. The principle of the DC-FRS is similar to wavelength modulation spectroscopy (WMS) and extensive studies \[140, 141, 142\] can be referred to investigate its signal generation and demodulation. Here I mainly focused on the theory of conventional AC-FRS as the basis for further studies.

The signal in AC-FRS is generated by altering the polarization of the linear polarized laser beam that passes through the paramagnetic species in the presence of the AC modulated magnetic field. A model for such detection can be described using Jones vectors and Jones matrices \[143\]. A linearly polarized electromagnetic wave before entering the sample can be represented by the Jones vector below:
Figure 2.5: Calculated Faraday angle peak signal vs. sample pressure (at B = 150 G).

Figure 2.6: Faraday rotation angle for the entire $^{14}\text{NO}$ fundamental band. Simulation condition: Magnetic field B = 1000 G, Pressure = 35 Torr, Optical path = 15 cm, NO concentration = 2 ppmv, Temperature = 300 K.
When the wave passes through the sample, its polarization axis is modulated around the initial polarization state by a Faraday rotation angle \( \Theta \) at modulation frequency \( \omega_m \). Calculation of the angle \( \Theta \) has been shown in Section 2.1.1. The Jones vector for such a wave with polarization modulation can be expressed as:

\[
J_1 = E_0 \begin{bmatrix} \sin(\Theta \sin(\omega_m t)) \\ \cos(\Theta \sin(\omega_m t)) \end{bmatrix} \approx E_0 \begin{bmatrix} \Theta \sin(\omega_m t) \\ 1 \end{bmatrix}, \tag{2.7}
\]

where one can consider \( \cos(\Theta \sin(\omega_m t)) \approx 1 \) because \( \Theta \) is typically very small as shown by Section 2.1.1. In conventional AC-FRS systems the analyzer is typically offset by a small angle \( \theta \) (up to several degrees) from the fully crossed position (90-degree method). In that case an analyzer can be represented by Jones matrix:

\[
M = \begin{bmatrix} \cos^2 \theta + \epsilon \sin^2 \theta & (1 - \epsilon) \sin \theta \cos \theta \\ (1 - \epsilon) \sin \theta \cos \theta & \sin^2 \theta + \epsilon \cos^2 \theta \end{bmatrix}, \tag{2.8}
\]

where \( \epsilon \) is the amplitude extinction ratio for the polarizer (please note that \( \epsilon^2 \) is the intensity or power extinction ratio). From Eq. 2.7 and 2.8 electric field after the analyzer is:

\[
E = M \cdot J_1 = E_0 \begin{bmatrix} (\cos^2 \theta + \epsilon \sin^2 \theta) \cdot \Theta \sin(\omega_m t) + (1 - \epsilon) \sin \theta \cos \theta \\ (1 - \epsilon) \sin \theta \cos \theta \cdot \Theta \sin(\omega_m t) + (\sin^2 \theta + \epsilon \cos^2 \theta) \end{bmatrix}. \tag{2.9}
\]

Eq. 2.9 represents the electrical field at the photodetector plane in a conventional FRS setup. Since photodetectors respond to optical intensity, a photocurrent generated by the optical radiation can be expressed by:
\[ I \propto E \cdot E^* = E_0^2 [(\cos^2 \theta + \epsilon^2 \sin^2 \Theta \sin(\omega_m t)) + (1 - \epsilon^2) \sin(2\theta) \cdot \Theta \sin(\omega_m t) \]
\[ + (\sin^2 \theta + \epsilon^2 \cos^2 \theta)] \]  
(2.10)

Since both \( \Theta \ll 1 \) and \( \epsilon^2 \ll 1 \), and because the signal demodulation is performed at first harmonic of \( \omega_m \), the term containing \( \Theta^2 \) can be neglected. This simplifies Eq. (2.10) to:

\[ I \propto P_0 [\sin 2\theta \cdot \Theta \sin(\omega_m t) + \sin^2 \theta + \epsilon^2 \cos^2 \theta], \]  
(2.11)

where \( P_0 \) represents the power of light entering the analyzer. The first term in Eq. (2.11) contains the Faraday angle modulation, while the other terms represent DC component of the photocurrent. The FRS signal can be demodulated using phase sensitive lock-in detection. Finally, we can get lock-in signal to be measured as:

\[ S = R_v \frac{\eta e}{h \nu} P_0 \sin(2\theta) \cdot \Theta, \]  
(2.12)

where \( R_v \) is the transimpedance of the photodetector, \( \eta \) is the quantum efficiency of the photodetector, \( e \) is the electron charge and \( h \nu \) is the photon energy.

### 2.2 Noise in FRS

#### 2.2.1 Noise contribution

Because in AC-FRS the spectral interference fringe can be effectively eliminated the total noise measured in an AC-FRS system is only composed of photodetector noise, laser intensity noise, and shot noise. Assuming laser intensity noise is linearly proportional to the laser intensity, which is transferred to the detector through the DC
terms in Eq. 2.11. If polarizers with high extinction are used so that $\epsilon \ll \theta$ the detected intensity noise can be expressed as:

$$N_{RIN} = \sqrt{\Delta f} \cdot \sigma(\omega_m) \cdot R_v \frac{\eta e}{h\nu} P_0 \sin^2 \theta,$$  \hspace{1cm} (2.13)

where $\sigma(\omega_m)$ is the bandwidth-normalized laser RIN at demodulation frequency $\omega_m$, $\Delta f$ is the demodulation bandwidth. The term $P_0 \epsilon^2 \cos^2 \theta \approx P_0 \epsilon^2$ is the actual laser power leakage through the non-ideal polarizer. For extinction ratio of $\epsilon^2 < 10^{-5}$, this term can be neglected as long as $\theta > 0.18^\circ$, which is a realistic assumption in a well-designed FRS system. Even the polarizer extinction ratio is poor, in small $\theta$ approximation, this part of noise can be treated as a contribution to the photodetector effective noise as a constant.

Photodetector noise can be expressed by:

$$N_{PD} = \sqrt{\Delta f} \cdot R_v \frac{\eta e}{h\nu} \cdot NEP(\omega_m),$$  \hspace{1cm} (2.14)

where $NEP(\omega_m)$ is the noise equivalent power at frequency $\omega_m$. Shot noise generated by the photocurrent $I$ is given by:

$$N_{shot} = \sqrt{\Delta f} \cdot R_v \cdot \sqrt{2eI} = \sqrt{\Delta f} \cdot R_v \cdot \sqrt{2e \frac{\eta e}{h\nu} P_0 \sin^2 \theta}.$$  \hspace{1cm} (2.15)

### 2.2.2 FRS system optimization

From Eq. 2.12 through 2.15 it can be concluded that offset angle $\theta$ determines which noise component limits SNR of the system. For small $\theta$ both the laser RIN and the shot noise can be reduced so that the total noise is dominated by the photodetector noise. Since the signal strength is proportional to $\sin(2\theta)$ and the NEP is independent of $\theta$ in the NEP-dominated regime the SNR can be improved by increasing $\theta$. This continues until $N_{shot}$ (proportional to $\sin(\theta)$) or $N_{RIN}$ (proportional to $\sin^2 \theta$) become
comparable to $N_{PD}$. In most conventional FRS systems operating at low modulation frequencies where the excess laser noise is high, the shot noise $N_{\text{shot}}$ is usually much lower than the laser noise $N_{RIN}$. Therefore the maximum SNR is commonly obtained by adjusting $\theta$ to equalize $N_{RIN}$ with the $N_{PD}$ (as shown in Fig. 2.7). At this optimum $\theta_{\text{opt}}$ the total noise becomes $N_{\text{total}} = \sqrt{N_{PD}^2 + N_{RIN}^2} = \sqrt{2} \cdot N_{PD}$ and the power received by the photodetector can be expressed as:

$$P_0 \cdot \sin^2 \theta_{\text{opt}} = \frac{NEP(\omega_m)}{\sigma(\omega_m)}.$$  \hfill (2.16)

Based on Eq. 2.16 if the laser and detector parameters are known one can conveniently calculate the optimum offset angle $\theta_{\text{opt}}$:

$$\sin \theta_{\text{opt}} = \sqrt{\frac{NEP(\omega_m)}{\sigma(\omega_m) \cdot P_0}}.$$  \hfill (2.17)

By setting $S = N_{\text{total}}$ (signal equals the noise) at $\theta_{\text{opt}}$ and with an approximation of $\sin \theta = \theta$ one can define noise-equivalent polarization rotation angle for conventional
AC-FRS system:

\[ \Theta_{NEA} = \sqrt{\Delta f \cdot \frac{NEP(\omega_m) \cdot \sigma(\omega_m)}{2P_0}}. \]  \hspace{1cm} (2.18)

Eq. 2.18 indicates that polarization rotation sensitivity is only related to the laser output power, laser RIN noise and photodetector noise. In fact it is a figure of merit that can be conveniently used to compare the results from different FRS systems regardless of their other features (e.g. wavelength, sample concentration, effective optical pathlength, magnetic field strength and analyzer offset angle etc.). It should be also noted that in conventional FRS \( \Theta_{NEA} \) is primarily limited by the large excess RIN of the laser source at low frequencies, where the modulation frequency of the magnetic solenoid \( \omega_m \) is located.

### 2.3 Shot noise limited condition and ultimate detection limit of FRS systems

Shot noise arises from the intrinsic quantum nature of the photogenerated detector current and is usually the ultimate sensitivity limitation of optical detection systems (except for some special techniques involving amplitude-squeezed light sources that allow measurements beyond the shot-noise limit) [144][145]. This fundamental limit represents the minimum detectable fluctuation of the field at the level of \(1/\sqrt{N} \), where \( N \) is a mean photoelectron number generated by the photodetector.

In order to assure shot noise limited photodetection in any optical system the following conditions should be met: a) the laser intensity noise should be smaller than the shot noise \( (N_{shot} > N_{RIN}) \), and b) the photodetector noise should be smaller than the shot noise \( (N_{shot} > N_{PD}) \). With an assumption of a general case of photodetection at wavelength \( \lambda \), with a signal frequency \( f \) and demodulation bandwidth \( \Delta f \) this two conditions can be expressed as (from Eq. 2.13 through Eq. 2.15):
$$\sqrt{\Delta f} \cdot R_v \cdot \sqrt{2e\frac{n_e}{h\nu}}P > \sqrt{\Delta f} \cdot R_v \cdot \sigma(\lambda, f)\frac{n_e}{h\nu}P, \quad (2.19)$$

$$\sqrt{\Delta f} \cdot R_v \cdot \sqrt{2e\frac{n_e}{h\nu}}P > \sqrt{\Delta f} \cdot R_v \cdot NEP(\lambda, f)\frac{n_e}{h\nu}. \quad (2.20)$$

If solved for optical power both equations can be simplified to:

$$\sqrt{P_{\text{min}}} = NEP(\lambda, f) \cdot \sqrt{\frac{\eta}{2h\nu}} < \sqrt{P} < \frac{1}{\sigma(\lambda, f)} \cdot \sqrt{\frac{2h\nu}{\eta}} = \sqrt{P_{\text{max}}}, \quad (2.21)$$

where $P$ is the optical power received by photodetector. Eq. (2.21) indicates that generally in a given optical system a shot noise dominated photodetection can be performed within a well-defined range of collected laser power $P$ between $P_{\text{min}}$ and $P_{\text{max}}$. If the power collected by the photodetector is smaller than $P_{\text{min}}$, photodetector noise dominates, while if the collected power is larger than $P_{\text{max}}$ the system becomes laser noise limited. A more general condition that can be obtained from Eq. (2.21) is:

$$NEP(\lambda, f) \cdot \sqrt{\frac{\eta}{2h\nu}} < \frac{1}{\sigma(\lambda, f)} \cdot \sqrt{\frac{2h\nu}{\eta}}, \quad (2.22)$$

which after simplification can be expressed as:

$$\sigma(\lambda, f) \cdot NEP(\lambda, f) < \frac{2h\nu}{\eta}. \quad (2.23)$$

Optical systems that do not satisfy this condition are not capable of shot noise dominated performance. Once the laser RIN, detector NEP and detector quantum efficiency are known, Eq. (2.23) can be used to determine if the shot noise limited performance is feasible with a given detection system. Please note that the optical power is not a factor that determines if the system is able to provide shot noise limited
performance. It also indicates that at shorter wavelengths (higher frequencies $\nu$) where diode lasers with lower RIN and photodetectors with lower NEP are available, shot noise limited performance of optical sensing is relatively easier to achieve \[146\]. However, as it will be shown later, although the shot noise limit represents the best performance for a given system, it does not guarantee the smallest noise-equivalent Faraday rotation angle among other FRS systems, because parameters such as optical power and photon energy used in the measurement must also be taken into account.

If laser RIN is negligible and there is sufficient optical power on the detector operating below its saturation regime, the general condition in Eq. 2.23 is met and the total noise becomes dominated by photocurrent shot noise. In such a regime the system SNR does not depend on analyzer offset angle $\theta$ as the shot noise is linearly proportional to $\sin \theta$ (see Eq. 2.15) and the signal is linearly proportional to $\sin \theta$ too (Eq. 2.12). There is no need to optimize the analyzer angle. From these two equations it can be calculated that the shot noise equivalent detectible polarization rotation angle for conventional AC-FRS based on 90-degree method:

$$\Theta_{SNEA_{AC-FRS}} = \sqrt{\Delta f} \cdot R_v \cdot \sqrt{\frac{2e}{h\nu} P_0 \sin^2 \theta} = \sqrt{\Delta f} \cdot \sqrt{\frac{2h\nu}{\eta P_0}} \cdot \frac{\sin \theta}{\sin 2\theta}. \quad (2.24)$$

For $\theta \ll 1$, which is true for most 90-degree FRS systems, the last factor in the Eq. 2.24 can be approximated as $\sin \theta / \sin(2\theta) \approx 1/2$. Therefore in shot noise limited regime, the sensitivity to the sample can only be enhanced by improving the optical power $P_0$ or by increasing the optical path within the sample to enhance the FRS rotation angle. One should note that by applying a multi-pass arrangement to increase the interaction path the optical power $P_0$ will be also affected due to the power loss from reflections on the metallic mirrors used, which must be taken into account in sensitivity estimations and will be discussed in Chapter 3. It should also be noted
that at shorter wavelengths/higher frequencies (given all other parameters are the same) the ultimate detection limit of FRS systems becomes gradually decreased.
Chapter 3

Multi-pass cell enhanced Faraday rotation spectroscopy

The sensitivity of an FRS system can be enhanced by utilizing multi-pass cell which can increase the effective optical path-length similar to the method used for absorption spectroscopy. In this chapter, a compact multi-pass Herriott cell with effective path-length of $\sim 3.5$ meters is integrated with a conventional AC-FRS system. Experimental optimization of the system signal and noise ratio based on the theory introduced in Chapter 2 is presented.

3.1 Experimental demonstration

The experimental configuration of the multi-pass cell based FRS spectrometer is similar to a conventional AC-FRS system. A collimated laser beam from a 5.2 $\mu$m DFB-QCL passes through first polarizer and is coupled into a multi-pass Herriot cell surrounded by a solenoid made of $\sim 800$ turns of 16 gauge copper wire. Second Rochon polarizer (analyzer) is placed after the cell to transform polarization rotation into intensity modulation that is detected by a MCT photodetector. The cell is formed using two 1 inch diameter gold coated spherical mirrors (with radius of
Figure 3.1: Multi-pass cell design.

Curvature of 150 mm) and provides a total path length of 3.5 m with overall dimension of only $20 \times 10 \times 10$ cm$^3$. Fig. 3.1 shows the design of the cell structure (cross section). Two grooves were machined on the inner surface of the cell tube. The 1 inch mirrors are secured on the pre-aligned grooves to provide fast optical alignment. The cell is made by non-metallic material ensuring no influence on the magnetic field. The transmission of the cell was measured to be 37.8%, corresponding to the mirror reflectivity of 96.9%. High reflectivity dielectric coating on the mirror could not be used due to deterioration of the polarization quality leading to lower performance of the FRS system. The average axial magnetic field strength versus driving current inside the solenoid was measured to be $\sim 63$ Gauss/A. During the experiment, the field strength, modulation frequency and sample gas pressure were set to 100 Gauss$_{rms}$, 2.3 kHz and 30 Torr, respectively. Within the available laser tuning range the NO
R(17/2)f transition at 1906.73 cm\(^{-1}\) has been selected as the target line. A calibrated gas mixture of 2 ppmv NO balanced by dry N\(_2\) was used for system characterization.

We followed the theory introduced in Chapter 2 to perform system optimization. The offset angle \(\theta\) between the crossed polarizers that guarantees maximal signal-to-noise ratio can be determined by characterizing the laser power, PD noise, laser noise and multi-pass cell throughput. The laser RIN and the photodetector NEP at \(\omega_m=2.3\ \text{kHz}\) were measured to be \(\sigma = 2.1 \times 10^{-6}\ \text{Hz}^{-1/2}\) and \(\text{NEP} = 1.03 \times 10^{-11}\ \text{W/Hz}^{1/2}\), respectively. The optimum system performance can be achieved when the laser noise equals the PD noise. This yields the optimized power on the PD of \(P = P_0 \cdot \sin^2 \theta = \text{NEP}/\sigma = 4.9\mu W\). With the output power after the multi-pass cell of \(P_0 = 6.8\text{mW}\) the optimal offset angle is found to be \(\theta = 1.51^\circ\). The FRS signal strength after the multi-pass cell is directly proportional to the number of passes.

Figure 3.2: Calculated AC-FRS signal strength vs. multi-pass number.
passes N whereas the output optical power $P_0$ decays exponentially with N due to the power loss on the mirrors. Therefore, the signal strength could be expressed as $S = P_i R^N \cdot \Theta N \cdot \sin(2\theta)$, where $P_i$ is the power before the multi-pass cell, $R$ is the power reflectivity on the mirror, $\Theta$ is the Faraday rotation angle per pass and $\sin(2\theta)$ indicates the first harmonic demodulation slope at offset angle $\theta$. The analyzer offset angle $\theta$ should always be optimized (to assure the laser noise to be equal the PD noise) and as a function of number of passes the optimized offset angle can be calculated as $\theta = \sin^{-1}(\sqrt{\text{NEP}/(\sigma \cdot P_i \cdot R^N)})$. By substituting it into the signal expression and assuming an arbitrary per-pass rotation angle of $\Theta=1 \mu$rad, the signal strength (that is proportional to SNR) as a function of number of passes N is obtained (Fig. 3.2).

The combination of linear enhancement of the FRS signal and exponential drop in output power yields a maximum signal for our system at N=63 passes. Our multi-pass cell designed for 31 passes is not optimal, however it can produce the SNR that is only 17% smaller than maximum achievable for our system (as indicated in Fig. 3.2).

Fig. 3.3 shows the measured SNR measured for 2 ppmv NO in N$_2$ as a function of optical power on PD (adjusted by varying the offset angle $\theta$). The maximum SNR of $625\sqrt{Hz}$ is achieved around $P_0 \sin^2\theta = 4.9\mu W$ and corresponding to the offset angle $\theta$ of 1.5°. This agrees well with the theoretical estimation of the optimum measurement conditions. Fig. 3.4 shows a typical FRS spectrum at optimized offset angle recorded by scanning the QCL wavelength over the target line with a lock-in time constant of 200 ms (corresponding to detection bandwidth of 0.83 Hz). The asymmetry in the measured spectrum is caused by the residual MCD as the path-length (absorption) increases. The SNR of $625\sqrt{Hz}$ has achieved for short integration times, and corresponds to a noise equivalent polarization rotation angle of $3.9 \times 10^{-8}$ rad/Hz$^{1/2}$ and NO minimum detection limit of 3.2 ppbv/Hz$^{1/2}$. The sensitivity can
Figure 3.3: Measured SNR vs. optical power on photodetector, achieved by adjusting the analyzer angle $\theta$.

Figure 3.4: NO spectrum acquired at the optimum offset angle.
be improved by one order of magnitude using the same optical arrangement and the optimal Q(3/2) line as a target transition.

3.2 Chapter summary

A DFB-QCL based FRS system integrated with a Herriott multi-pass cell was characterized by measuring the NO transition line R(17/2)f. Noise equivalent polarization rotation angle of $3.9 \times 10^{-8}$ rad/Hz$^{1/2}$ was achieved. We found that in multi-pass cell arrangement the signal strength reaches maximum for a certain number of passes. Our current system has shown NO sensitivity of 3.2 ppbv/Hz$^{1/2}$. However, one should note that the signal tends to drift in long-term performance, which is mainly caused by the optical misalignment of the multi-pass cell due to the solenoid thermal expansion. A multi-pass cell with higher thermal stability can be incorporated with techniques introduced in the following chapters to enable baseline-free, low drift, line-locked continuous NO detection with ultra-high sensitivity.
Chapter 4

Heterodyne enhanced Faraday rotation spectroscopy

In this chapter a new detection method for FRS is presented. Near shot-noise limited performance in the mid-IR is demonstrated using an H-FRS system without any cryogenic cooling. Theoretical analysis is performed to estimate the signal strength and ultimate sensitivity to polarization rotation for H-FRS. Sensing of NO has been performed with an H-FRS system using a thermoelectrically cooled 5.24 µm DFB-QCL and a MCT photodetector. The QCL’s RIN that dominates at low frequencies is largely avoided by performing the heterodyne detection in RF range. H-FRS exhibits a total noise level of only 3.7 times the fundamental shot noise. The achieved sensitivity to polarization rotation of $1.8 \times 10^{-8}$ rad/Hz$^{1/2}$ is only 5.6 times higher than the ultimate theoretical sensitivity limit estimated for this system. The path- and bandwidth-normalized NO detection limit of 3.1 ppbv-m/Hz$^{1/2}$ was achieved using the R(17/2)f transition of NO at 1906.73 cm$^{-1}$. 
4.1 Background and motivation

The most common practices reported to-date to improve FRS system performance were mostly focused on increase in the FRS signal strength, and on optimization of the analyzer offset angle $\theta$ (as performed in Chapter 3). Several strategies focused on the FRS signal increase have been reported including: selection of molecular transitions with higher intensity \cite{89, 87, 90, 147, 146}, providing optimal magnetic field strength \cite{87, 147, 92}, optimizing the sample pressures \cite{87, 92, 93}, using higher laser power or minimizing the losses in the gas cell \cite{147}, increasing the optical path with multi-pass cells \cite{148, 149}, as well as applying cavity enhanced techniques \cite{150}. Since majority of mid-IR FRS instruments reported to date \cite{87, 90, 147} showed significant domination of laser noise over the shot noise, one can conclude that in order to further enhance sensitivity of the mid-IR FRS systems, suppression of the laser RIN or decrease in the photodetector noise are the best strategies to approach the ultimate shot-noise limited performance. The photodetector noise can be reduced by selecting an appropriate detector element with a low-noise pre-amplifier. The performance of mid-IR detectors is often significantly improved through cryogenic cooling of the detector element, which helps reducing the thermal noise that dominates in this spectral range. However this approach represents a logistical challenge and cryogen-free detectors are preferred in sensor systems intended for autonomous long-term measurements especially in field settings. Therefore in this chapter we focus on reduction of the laser noise, which is challenging, but can significantly improve performance of FRS systems and provide more relaxed requirements for low-noise photodetectors. In general many spectroscopic systems are limited by laser noise, and the most popular techniques used to suppress the laser noise include: a) balanced detection incorporated with electronic noise canceller \cite{104, 105, 106, 151, 152}, b) active laser intensity stabilization \cite{153, 154, 155}, and c) heterodyne detection \cite{156, 157, 158}. All of these
techniques have been successful in avoiding or suppressing the excess laser noise in spectroscopic systems to the degree that the shot noise becomes dominant.

As recently demonstrated by Zhao et al., when very low optical power (at µW level) is used in combination with a sensitive TEC cooled MCT photodetector, a shot-noise-dominated FRS of OH radicals at $\sim 2.8 \, \mu m$ could be realized [146]. However up to now, no cryogen-free FRS system has been able to approach the fundamental limit in the mid-IR spectral range at wavelengths $> 3 \, \mu m$. For QCLs, that are currently very popular spectroscopic laser sources in the mid-IR, the main limitation is their intrinsic RIN that is relatively high. Due to their specific cascaded quantum-well structure and electron transport, the QCLs exhibit higher RIN than for example interband diode lasers [159, 160]. Additionally amplitude squeezing is difficult in QCLs, which makes sub-shot-noise limited detection significantly more challenging [160]. All three RIN suppression methods mentioned above have a chance to reduce the QCL noise and ultimately improve the FRS performance. For example, laser RIN can be effectively suppressed using FRS with a balanced detection scheme (referred to as “45-degree FRS method”, because the analyzer is set at 45° with respect to the incident polarization plane to give access to two polarization components with equal intensities used for balanced FRS measurement) [148]. Although very elegant, this approach is difficult to realize in mid-IR, because specialized dual element photodetectors with balanced differential preamplifiers are not readily available and costly if custom made.

To mitigate the need for specialized detectors, in this chapter we present a new H-FRS technique that provides significant noise reduction by employing heterodyne detection of the FRS signal. H-FRS effectively shifts the FRS signal detection to RF range where the QCL RIN is significantly lower. The developed system is cryogen-free and is based on a DFB-QCL and a single MCT photodetector, both thermoelectrically cooled. The sensitivity to the polarization rotation was improved by almost two orders of magnitude as compared to the previous conventional AC-FRS system for
NO detection reported in Ref. [90] and by one order of magnitude as compared to the work by Zhao et al. [146]. In the next Section 4.2, a theoretical analysis of H-FRS is presented. Section 4.3 describes experiment setup. Section 4.4 presents noise characterization, performance tests and discussion of the prototype H-FRS system.

4.2 H-FRS operation principle

In the H-FRS, by superimposing a local oscillator (LO) field with the FRS signal field emerging from the analyzer, a heterodyne beatnote can be created at frequency that is far above the noisy baseband region. The simplest way to create the LO is to recycle the extraordinary beam emerging from the analyzer that is polarized orthogonally with respect to the signal polarization. This beam contains most of the laser power that is usually lost in a conventional FRS system. In H-FRS a Rochon prism or any other high quality polarizer that gives access to both polarization components simultaneously (i.e. Wollaston prism) can be used as an analyzer as shown in Fig. 4.1. An optical frequency shifter (e.g. acousto-optical modulator, AOM) is used to create LO wave that is shifted by a well-defined RF frequency with respect to the signal wave. By recombining the signal beam and the LO beam on a beam splitter a heterodyne Mach-Zender interferometer is formed.

In AC-FRS, the photodetector collects the light transmitted through an analyzer, thus the intensity on the photodetector is proportional to \( \sin^2 \theta \) as shown in Fig. 4.1 (b). The FRS angle is sinusoidally modulated through alternating the magnetic field and the FRS signal is measured by demodulating the associated intensity changes at the first harmonic of the modulation signal. With a small signal approximation the measured FRS signal, \( S_{C-\text{FRS}} \), is proportional to the first derivative of the analyzer transmission curve in Fig. 4.1 which yields \( S_{C-\text{FRS}} = d(\sin^2 \theta)/d\theta = \sin(2\theta) \). Since the FRS signal vanishes at \( \theta = 0 \) the optimum SNR for conventional FRS system.
Figure 4.1: (a) A concept diagram for the H-FRS system configuration. The extraordinary beam exiting the Rochon prism is frequency-shifted by AOM, and serves as the LO wave. Since its polarization is orthogonal to the signal wave, a polarization rotator (PR) was used to transform its polarization axis by 90 and assure maximum heterodyne efficiency. A principle of signal generation is schematically shown for conventional FRS in (b) and for H-FRS in (c). (d) An electrical RF spectrum of the photocurrent in H-FRS is an equivalent of a carrier-suppressed amplitude modulated signal.

occurs at the analyzer angle that is slightly offset from the zero position. It has been shown that at optimal $\theta$ the laser noise equals the detector noise, and further uncrossing would deteriorate the SNR.

In contrast to AC-FRS the H-FRS employs detection of the wave amplitude (not intensity/power) that is proportional to $\sin \theta$ after the analyzer. If first harmonic demodulation is applied, the FRS signal can be approximated as $S_{H-FRS} \propto \cos \theta$, which at $\theta = 0$ exhibits maximum (Fig. 4.1(c)). Thus if the laser noise is sufficiently suppressed the $\theta = 0$ is an optimum analyzer position for H-FRS measurements. When FRS angle is sinusoidally modulated at $\omega_m$ the field in the ordinary beam transmitted through the analyzer at $\theta = 0$ can be described as:

$$E' = E_0 \Theta \sin(\omega_m t), \quad (4.1)$$
where \( E_0 \propto P_0^{1/2} \) is the optical field amplitude before the analyzer and \( \Theta \) is the Faraday rotation angle. Since \( \Theta \sin(\omega_m t) \) is usually very small an approximation \( \sin(\Theta \sin(\omega_m t)) \approx \Theta \sin(\omega_m t) \) has been applied in Eq. 4.1. For small \( \theta \) most of the laser power \( P_0 \) is transmitted into extraordinary beam. This beam is directed into an AOM, which is used to frequency-shift the extraordinary wave by \( \Omega \). This radiation is then passed through a 90 degrees polarization rotator and serves as the LO with amplitude \( E_{LO} \). The two beams (the signal and the LO beam) are then combined on a beam splitter and focused onto a photodetector. A square-law detector serves as a heterodyne mixer, which generates a photocurrent:

\[
I \propto (E' + E_{LO})^2 = (E')^2 + E_{LO}^2 + 2E' \cdot E_{LO} \cdot \cos(\Omega t + \Phi)
\]

\[
\propto P_{LO} + 2\sqrt{P_0 P_{LO}} \cdot \Theta \cdot \sin(\omega_m t) \cdot (\Omega t + \Phi),
\]

where \( \Phi \) is the phase difference between the LO wave and the signal wave. Since \( E'^2 \ll E_{LO}^2 \), the DC component of the photocurrent in Eq. 4.2 comes primarily from \( P_{LO} \). The second term is the H-FRS signal, which has a form of carrier-suppressed amplitude modulation (Fig. 4.1 (d)). The H-FRS signal \( S \) that is retrieved through amplitude demodulation of the RF beatnote can be expressed as:

\[
S = 2R_V \cdot \eta_{het} \cdot \frac{e}{h\nu} \sqrt{P_0 P_{LO}} \cdot \Phi,
\]

where \( R_V \) is the photodetector transimpedance, \( \frac{e}{h\nu} \) is the current responsivity of a photodetector with quantum efficiency \( \eta \) (\( e \) and \( h/\nu \) are the electron charge and a photon energy, respectively), and \( \eta_{het} \) is the heterodyne efficiency.

Similarly to AC-FRS (as described in Section 2.2), there are three main noise contributions: the photodetector noise, the laser-generated shot noise and the laser RIN noise. Given the LO accounts for most of the laser power collected by the photodetector all three individual noise contributions measured at the photodetector
voltage output can be expressed as bandwidth-normalized quantities:

\[ N_{PD} = \sqrt{\Delta f} \cdot R_V \cdot \frac{\eta e}{h\nu} \cdot NEP, \quad (4.4) \]

\[ N_{RIN} = \sqrt{\Delta f} \cdot R_V \cdot \sigma(\Omega \pm \omega_m) \cdot \frac{\eta e}{h\nu} \cdot P_{LO}, \quad (4.5) \]

\[ N_{\text{shot}} = \sqrt{\Delta f} \cdot R_V \cdot \sqrt{\frac{2e}{h\nu}} P_{LO}, \quad (4.6) \]

where \( \sigma(\Omega \pm \omega_m) \) in Eq. 4.5 denotes laser RIN measured at frequency \( \Omega \pm \omega_m \). As mentioned above, the AC component in Eq. 4.3 has a form of carrier-suppressed amplitude modulation. This is central to the concept of the H-FRS measurement, because with fully suppressed carrier the heterodyne conversion of the large noise in the baseband to high frequencies should not exist and the measured laser noise should only originate from the LO noise at frequency \( \Omega \pm \omega_m \) away from the baseband. Because in the RF frequency range the \( 1/f \) laser noise becomes extremely small, by setting \( \Omega \) sufficiently high the measured LO noise is strongly suppressed. Therefore with relatively high intensity of LO the shot noise cannot be neglected in the H-FRS, which is rarely the case in conventional FRS systems.

Similarly to Section 2.3, considering \( N_{\text{shot}} \) dominates in the system total noise, Eq. 4.6 and Eq. 4.3 can be used to calculate shot noise equivalent detectible polarization rotation angle for H-FRS:

\[ \Theta_{SNEA_{H-FRS}} = \sqrt{\Delta f} \cdot \frac{1}{\eta_{het}} \cdot \sqrt{\frac{h\nu}{2\eta P_0}} \quad (4.7) \]

The shot noise-equivalent Faraday rotation angle is inversely proportional to the square root of an average photon flux available before the analyzer adjusted by the quantum efficiency of the photodetector. The detection limit is also affected by the heterodyne efficiency which affects the signal and in practice is smaller than unity. In an ideal case with \( \eta_{het} = 1 \) this result is consistent with the phase sensitivity of
a shot noise limited Michelson interferometer $^{102}$. This confirms that FRS is actually a sensitive phase (dispersion) measurement and not an amplitude (absorption) measurement.

By looking at Eq. 2.24 and Eq. 4.7 one can conclude the two equations are actually identical. Therefore in theory, both H-FRS and AC-FRS can reach the same ultimate sensitivity when operated in shot noise limited regime.

4.3 Experimental

4.3.1 System set-up

A schematic diagram of the H-FRS experimental setup is shown in Fig. 4.2. The laser source is a 5.24 µm CW DFB-QCL (provided by Alpes Lasers SA), operating near room temperature. LIV characterization and wavenumber vs. current from the laser specs is reproduced in Fig. 4.3 and Fig. 4.4. The QCL temperature is stabilized using a TEC with a temperature controller (Arroyo Instruments 5305) and the bias current is delivered from a low noise current driver (Wavelength Electronics QCL500). The laser beam is collimated with a ZnSe lens (f = 1.9 mm) and directed through the first polarizer (RP1) that is used to improve the polarization quality and to precisely set the polarization axis of the light incident on the sample. RP1 is followed by a 15 cm gas cell surrounded by an electromagnetic solenoid driven with a power amplifier (RMX 850). Hard plastic material (e.g. Delrin) was used to construct the cell structure since conductive materials such as stainless steel and aluminum are not suitable to construct the Faraday cell since they create magnetic conductive or shielding effects that deteriorate the magnetic fields inside the gas sample cell. By supplying current of 3 A (root-mean-square, rms) to the solenoid, an axial magnetic field of 100 G$_{rms}$ (Gauss-rms) modulated at frequency $f_m$=1.08 kHz was obtained within the sample volume. The working magnetic solenoid is cooled by a
fan otherwise the overheating will cause thermal expansion of the optical component, leading to unwanted beam walk-off. The magnetic field strength in the FRS cell was carefully characterized with a Gauss-meter. The 100 G\textsubscript{rms} was an average value estimated for the active optical path inside the gas cell (the field is non-uniform along the optical axis with stronger field at the center and \(~50\%\) weaker at both ends of the cell).

When the laser optical frequency coincides with the target NO transition, the linearly polarized light undergoes Faraday rotation. The analyzer RP2 placed after the gas cell (MgF\textsubscript{2} Rochon prism with an extinction ratio of \(~1 \times 10^{-5}\)) transforms polarization rotation into the wave amplitude modulation as described by Eq. 4.1. The polarization axis of the ordinary beam emerging from Rochon analyzer is fully crossed (\(\theta = 0^\circ\)) with respect to the RP1 which guarantees maximum H-FRS signal. When there is no Faraday rotation, only a small fraction of the laser power determined by the effective extinction ratio leaks through the analyzer. The effective
Figure 4.3: Laser characterization curves: voltage and average power vs. current in CW operation (the solid squares indicate the maximum single mode emitted power)

Figure 4.4: Laser characterization: wavenumber vs. current in CW operation
extinction ratio of the system that consists of the RP1, the gas cell, and the RP2, was characterized to be $\sim 1 \times 1.86^{-4}$.

The measured extinction ratio is worse than manufacture’s specification of the Rochon prisms due to the deterioration of the polarization linearity by the cell windows. Therefore a third polarizer RP3 set to the same polarization as RP2 was placed within the ordinary beam path, which allowed for an additional factor of $\sim 5$ improvement in the effective polarization extinction ratio. The extraordinary beam emerging from Rochon analyzer undergoes a frequency shift by $\Omega = 30$ MHz with an AOM (IntraAction AGM-40) and then serves as the LO. The AOM is driven by a low noise function generator (Tektronix AFG 3102) and RF power amplifier (HP 230B) that provide more than 80 dB phase noise suppression ratio at $\sim 1$ kHz. Temperature stabilized water cooling (SolidState Cooling ThermoRack 650) is applied to the AOM in order to minimize refractive index variations caused by thermal drifts. Since the heterodyne efficiency is sensitive to the polarization of the two waves, the extraordinary beam is directed through a periscope assembly (Thorlabs RS99) constructed by three mirrors (shown by Fig. 4.5) to perform a simple 90° rotation of the polariza-
The two beams (signal and LO) are then re-combined using a calcium fluoride (CaF$_2$) beam splitter (BS) with 50/50 split ratio coating and then are focused by an aspheric ZnSe lens (Thorlabs AL72512-E) onto a TEC-cooled MCT photodetector (VIGO PV-2TE-5) customized for mid-IR heterodyne detection (For photodetector details please see Appendix B). A band pass filter (BPS) with 6 MHz bandwidth centered at 30 MHz is used for signal filtering and rejection of large noise at low frequencies. A double balanced frequency mixer (Mini-circuits ZX05-1L-S+) is used to down-convert the RF heterodyne signal at 30 MHz to the base band (IF output in Fig. 4.2) by mixing it with a reference signal from the function generator. The IF output of the mixer is demodulated by a lock-in amplifier (Signal Recovery 7265) at the magnetic field modulation frequency $f_m$ to retrieve the FRS signal. The FRS signal is then recorded by a PC equipped with a data acquisition card (NI USB-6251).

4.3.2 Baseline distortion and interferometer stabilization

Due to imperfections of the polarizers, a small portion of the light in the orthogonal polarization state leaks through the analyzer even when there is no Faraday rotation present. This residual light carries small amplitude modulation at $\omega_m$ due to electro-magnetic pick-up from the solenoid. Since this occurs as amplitude modulation sideband that is offset from the carrier optical frequency by $\omega_m$ it yields unwanted signal demodulated by the system after heterodyne detection. Because this is a coherent detection process this heterodyne demodulated pick-up is sensitive to the LO phase $\Phi$ in Eq. 4.2. This causes an unstable residual baseline if there is a drift in the opto-mechanical stability of the interferometer setup (as shown in Fig. 4.6). However if the phase of this interference is stabilized, this additional heterodyne signal is only a constant offset in the H-FRS signal that can be easily accounted for. This stabilization can be performed by equalizing the optical path lengths of the ordinary beam.
Figure 4.6: H-FRS scan spectrum without interferometer stabilization. Large baseline fluctuations are caused by the LO beam path length variation due to temperature fluctuations in the AOM.

and the extraordinary LO beam and by active compensation of the drifts occurring in the system (mostly due to temperature drift of the water-cooled AOM).

In order to keep those two paths equal in length, an actively controlled optical delay line is incorporated into the ordinary beam path. A total observed RF power of the heterodyne beatnote is used to derive an error signal and to perform active control of the optical delay line. This is possible because the LO beam carries a small portion of residual amplitude modulation (RAM) at frequency Ω introduced by the AOM, and its phase does not change with the interferometer drift. However the superposition of the phase-stable RAM and a phase-drifting heterodyne signal (both at 30 MHz frequency) results in a varying power of the RF carrier at 30 MHz which is then used as a feedback signal for interferometer stabilization. The active path
length control is implemented as shown in Fig. 4.2. A piezoelectric actuator with 10 nm resolution (Thorlabs PE4) and a PID controller (SRS SIM960) are used to control the optical delay line and to stabilize the path length. The error signal for the feedback loop is derived from portion of the photodetector output power measured by a high sensitivity Schottky diode (Herotek DHM020BB) compared to a constant set point. Despite introducing additional complication into the optical system this active feedback method has an advantage of being able to perform the H-FRS measurement with only one photodetector element. This complication could be avoided if a balanced heterodyne receiver is used to measure both outputs of the interferometer. However, similarly to conventional balanced photodetectors discussed earlier, there is no mid-IR balanced heterodyne photoreceivers available commercially and custom devices can be costly. Therefore the single photodetector approach implemented here is more viable solution at this time. A photo of the complete optical setup of the H-FRS system is shown in the Appendix B.

4.4 H-FRS results and discussion

4.4.1 Noise analysis

The system noise spectrum in the range between 10 kHz and 100 MHz was recorded directly at the output of the photodetector using an RF spectrum analyzer (Tektronix RSA6106A) and the results are shown in Fig. 4.7. The photodetector noise was measured with the laser light blocked and is shown as a black line in Fig. 4.7. The red trace shows the total noise measured with the QCL light incident on the detector. During this measurement, the laser driving current was kept at 160 mA, the laser temperature was 0°C, and the power on the detector was set to 120 µW. At such a low power level the shot noise is lower than photodetector noise. Therefore, the increase in the noise level observed between the black and red traces shown in Fig. 4.7
Figure 4.7: Noise spectrum recorded at the detector output without (black) and with the laser light shining at the detector (red). The contribution of the laser noise is clearly noticeable. The resolution bandwidth (RBW) of the RF spectrum analyzer is 100 Hz. QCL was operated at 0°C with 160 mA bias current, and the optical power on the photodetector was set to $\sim 120 \, \mu W$.

indicate primarily an influence of the laser noise. As expected in the low frequency range the laser noise generally follows $1/f$ trend, and at frequencies $> 1$MHz there are some distinct peaks that can be avoided by selecting an appropriate measurement frequency. We have not studied the origin of those distinct peaks in the noise spectrum between 1 MHz and 10 MHz, but we suspect the noise of the laser current source, power supplies, and TEC controller electromagnetic interference are the most possible causes of this specific spectral noise structure. Generally the noise floor at higher frequencies ($> 500$kHz), is much lower than the noise level at 10 kHz. We found that in the present system the frequency of 30 MHz is optimal for heterodyne detection (as shown in Fig. 4.7 inset). At 30 MHz, both the photodetector and the laser noise are
Figure 4.8: The photodetected noise density at 1.08 kHz (black dots) as a function of laser power. A slope of a linear fit (blue line) to the data in laser noise dominated regime is used to measure RIN. For comparison, shot noise calculated from Eq. 4.6 is shown by red dots. The QCL is operated at 160 mA, and at heatsink temperature of 0°C. Optical power was adjusted by the polarizer. Green dash line indicates the measured photodetector noise, which corresponds to NEP = 4.8 × 10^{-11} W/Hz^{1/2} at 1.08 kHz.

The detector NEP is strongly reduced while the detector is still fast enough to detect signals (it operates well below its 3dB cut-off frequency). An AOM designed for frequency range between 30 MHz and 50 MHz was used to produce the LO.

In order to compare how far from the fundamental shot noise limit both the AC-FRS and the H-FRS would operate in this setup, the noise contributions relevant in both configurations were carefully characterized. The noise for AC-FRS has been evaluated at the coil modulation frequency \( f_m \)=1.08 kHz and for H-FRS the noise at \( \Omega \) was taken into account. Because the available RF analyzer cannot measure signals below 9 kHz, the noise characterization at \( f_m \)=1.08 kHz was performed using
Figure 4.9: The total noise (black dots) at $\sim$30 MHz measured with an RF spectrum analyzer as a function of optical power. The same QCL operating conditions were used as in Fig. 4.8. The shot noise (red dots) and laser noise (blue dots) were calculated based on photodetector current responsivity of 2.0 A/W, and detector transimpedance of 8350 V/A. The green dashed line indicates the measured photodetector noise, which at $\sim$30 MHz corresponds to $\text{NEP}=1 \times 10^{-11} W/Hz^{1/2}$.

a lock-in amplifier. In both cases the noise measurement was performed as a function of optical power. A constant laser current of 160 mA was used in both measurements and the optical power on the photodetector was varied by rotating a polarizer placed between laser and detector. The AC-FRS noise measurement results are shown in Fig. 4.8. The RIN analysis is performed using noise data that are far above the photodetector noise and the shot noise. The slope of a linear fit to the noise data as a function of optical power level shown in Fig. 4.8 represents the laser RIN of $\sigma(\omega_m)=2.94 \times 10^{-6} Hz^{-1/2}$. For comparison, Fig. 4.9 shows the noise vs. optical power characteristics for the same laser operating conditions as in Fig. 4.8 but
measured at $\Omega = (30 \text{ MHz} + 1.08 \text{ kHz})$ using RF spectrum analyzer. It is clear that the shot noise calculated from Eq. (4.6) is not negligible at this frequency. The laser noise $N_{RIN} = P \cdot \sigma(\Omega)$ (shown as blue triangles in Fig. 4.9) was calculated as:

$$N_{RIN} = \sqrt{N_{total}^2 - N_{PD}^2 - N_{shot}^2}$$  \hspace{1cm} (4.8)

where the $N_{total}$ is the total noise (measured), $N_{PD}$ is the detector noise (measured), and the $N_{shot}$ is the shot noise (calculated). As shown in Fig. 4.9, the laser noise at $\sim 30 \text{ MHz}$ is almost at the same level as the shot noise. The calculated RIN of $\sigma(\Omega) = 7.98 \times 10^{-9} \text{Hz}^{-1/2}$ is $\sim 370$ times lower than the RIN measured at 1.08 kHz. The H-FRS enables performing shot noise dominated measurements in a system which at low frequencies is dominated by laser noise. With a photodetector NEP of $1.0 \times 10^{-11} \text{W/Hz}^{1/2}$ (measured at 30 MHz) Eq. 2.21 gives a range of LO power of $0.66 \text{ mW} < P_{LO} < 2.39 \text{ mW}$, within which shot noise limit can be reached. This is in excellent agreement with the data in Fig. 4.9, which clearly show that for optical powers above 0.66 mW the shot noise becomes higher than the photodetector noise. For optical power of $\sim 1 \text{ mW}$, the total noise is only 1.56 times higher than the shot noise. The high power limit could not be verified, because the photodetector preamplifier approaches saturation for optical powers above 1 mW.

### 4.4.2 Sensitivity of the H-FRS system

Heterodyne efficiency $\eta_{het}$ is one of the factors that affect the signal strength and thus the detection limit in H-FRS. In the theoretical analysis above we have assumed $\eta_{het} = 1$, which requires the interfering signal and LO waves to be perfect plane waves with exactly the same polarization state and ideally aligned wavefronts at the photodetector plane. However in practice there are multiple factors such as wavefront quality distortions, polarization-degradation, imperfect mode matching, or photode-
Figure 4.10: Heterodyne efficiency characterization.

tector nonlinearity that can reduce the heterodyne efficiency. In the H-FRS system discussed here the heterodyne efficiency was evaluated experimentally using fixed LO power at 60 $\mu$W and several values of signal beam power varied between 10 and 70 $\mu$W (as shown in Fig. 4.10). The measured heterodyne beat note is divided by calculated beat note strength assuming 100% heterodyne efficiency, which gives the average heterodyne efficiency in our system $\eta_{\text{het}} \sim 50\%$ (the photodetector quantum efficiency has been accounted for in the calculation). This is a reasonable value given many transmissive components that potentially can deteriorate wavefront and polarization quality in this H-FRS system prototype.

Fig. 4.11 shows a typical H-FRS spectrum acquired with the system. The QCL wavelength was scanned across the NO R(17/2)f transition by stepping the bias current of the QCL from 160 mA to 163 mA with 0.017 mA steps. The laser frequency
Figure 4.11: An H-FRS spectrum of the NO R(17/2)f transition at 1906.73 cm$^{-1}$. Experimental conditions: 2 ppmv NO in N$_2$ mixture, sample pressure of 30 Torr, sample temperature of 300 K, magnetic field of $\sim$100 G, active optical pathlength L = 15 cm, optical power before analyzer $P_0 = 14$ mW, LO power $P_{LO} = 1$ mW, photodetector current responsivity 2.0 A/W, transimpedance $R_V = 8350$ V/A, heterodyne efficiency $\eta_{het} = 0.5$, and a measurement bandwidth of 0.83 Hz. The system electrical gains have been factored in and the y-scale reflects voltage at the detector output.
tuning rate was approximately $\sim 0.033 \text{ cm}^{-1}/\text{mA}$. The LO power on the photodetector was set to the maximum value of $\sim 1 \text{ mW}$ permitted by the photodetector specifications. The lock-in amplifier time constant was set to 200 ms (with 12dB/octave filter slope) which gives effective detection bandwidth of $\Delta f=0.83 \text{ Hz}$. We have identified that the asymmetry in the measured spectrum was caused by the combination of MCB and MCD effects similarly to references [92, 163]. Contribution of the MCD signal can be suppressed when analyzer offset angle is set very precisely at $0^\circ$. In the present system a manual adjustment of the polarization axis did not provide sufficient precision to eliminate the MCD contribution below detectible level. In future work, a motorized rotary stage with resolution of $0.001^\circ$ should be used to increase the adjustment accuracy and to automatize the process of suppressing the MCD effects in the acquired spectra. However the peak of the H-FRS spectrum is not affected by the presence of the MCD contribution (which has derivative-like shape with a zero-crossing at the line center). This allows for reliable and reproducible estimation of the systems SNR by using the signal amplitude at the line center and the noise measured as a standard deviation of the spectral data points away from the transition. Based on data in Fig. 4.11, the signal strength at the peak is $S = 76.8 \mu \text{V}$ (measured with respect to the baseline offset) and the noise spectral density of $N = 0.792 \pm 0.066 \mu \text{V/Hz}^{1/2}$ have been measured. This corresponds to an SNR of $\sim 105$, which yields a bandwidth-normalized NO concentration detection limit of $20.6 \text{ ppbv/Hz}^{1/2}$.

The measured noise spectral density is approximately 2.4 times larger than the total noise observed for this system with RF spectrum analyzer in Fig. 4.9 and 3.7 times larger than the fundamental shot-noise level calculated for $P_{LO}=1\text{mW}$. This slight deterioration with respect to Fig. 4.9 was expected because the frequency mixer used in the H-FRS system contributes additional noise (the noise figure specified by the manufacturer is $\sim 5.6 \text{ dB}$), as well as the system was found to be sensitive to
mechanical vibrations that also contribute noise that cannot be corrected by the relatively slow active stabilization of the interferometer implemented in this system. Despite this small increase in the total system noise an excellent noise equivalent Faraday rotation angle of $\Theta_{NEA}/\sqrt{\Delta f} = 1.79 \times 10^{-8}\text{rad/Hz}^{1/2}$ has been obtained (the value is estimated by equating the measured noise $N$ to the signal $S$ in Eq. 4.3 reduced by factor of $\sqrt{2}$ to account for signal amplitude loss on the 50/50 beam splitter). When compared to the ultimate shot noise equivalent Faraday rotation angle of $3.2 \times 10^{-9}\text{rad/Hz}^{1/2}$ calculated for this system using Eq. 4.7 the experimental result is only about factor of 5.6 times worse (which agrees with the signal loss at the beam splitter by $\sqrt{2}$ times and with the noise that is 3.7 times higher than the calculated shot noise level).

### 4.4.3 Comparison with other FRS systems

For comparison with other FRS results we have surveyed experimental results obtained with cryogen-free FRS systems published recently. The three experimental setups published in Refs. [90, 147, 146] show different sources of performance limitations. In the NO sensing system reported by Lewicki et. al. [90], the noise equivalent Faraday rotation angle measured with a TEC cooled detector was $2.53 \times 10^{-7}\text{rad/Hz}^{1/2}$ (calculated as rms value for the experimental conditions used in this work). As shown in Section 2.1.1 The system uses the $^{14}$NO Q(3/2) transition that is optimum for FRS measurements of NO at low magnetic fields [89, 87, 90]. The minimum detectible polarization rotation in this system corresponds to NO detection limit of $3.8 \text{ppbv-m/Hz}^{1/2}$ (the detection limit is normalized to optical path and measurement bandwidth of $\Delta f=0.25\text{Hz}$ estimated for 1 s time constant and 6dB/octave filter slope). However the fundamental shot noise limit of $3.62 \times 10^{-9}\text{rad/Hz}^{1/2}$ calculated for the system parameters using Eq. 2.24 is $\sim 70$ times better, which indicates that system is strongly dominated by laser/detector noise. Therefore after application of
a liquid nitrogen cooled InSb detector in this work an improvement in the detection limit by factor of 11.3 was achieved. The overall performance of the system with cryogenically cooled detector corresponds to \( \sim 6.2 \) times the shot noise limited performance. Kluczynski et al. in Ref. [147] used a DFB QCL to target the same Q(3/2) transition of NO and a single spectral point NO detection limit of 1.03 ppbv-m/Hz\(^{1/2}\) (also path- and bandwidth-normalized) was obtained. A spectroscopic modeling performed for the reported detection limit achieved with 15 cm optical path, sample gas pressure of 67 mbar (\( \sim 50 \) Torr), and magnetic field of 156 G\(_{rms}\) yields the noise equivalent Faraday rotation angle of \( 7.89 \times 10^{-8} \text{rad/Hz}^{1/2} \), which is two orders (\( \sim 99 \)) of magnitude worse than the ultimate shot noise limited performance of \( 7.96 \times 10^{-10} \text{rad/Hz}^{1/2} \) estimated for the available laser power of 60 mW. This suggests that the laser noise was still the primary limiting factor in Ref. [147]. Theoretically, if the laser RIN was comparable in both experiments, a 20-fold increase in optical power and 0.72 times lower NEP in Ref. [147] should improve the by approximately 5.3 times (\( \sqrt{20/0.72} \)) with respect to Ref. [90]. Since there is only \( \sim 3 \times \) improvement in minimum detectible rotation angle, this indicates 3 times higher laser RIN in Ref. [147] as compared to the laser used in Ref. [90]. In conclusion both NO systems could tremendously benefit from a reduction of their laser noise below the fundamental shot noise limit with a prospect of up to two orders of magnitude improvement in sensitivity. A conventional FRS system with shot noise dominated performance was reported by Zhao et al. [146]. In this setup targeting detection of OH-radicals the measured minimum detectible Faraday rotation angle of \( 1.39 \times 10^{-7} \text{rad/Hz}^{1/2} \) was only 2.1 times higher than the theoretical shot noise limit of \( 6.62 \times 10^{-8} \text{rad/Hz}^{1/2} \). By analyzing the fundamental limit provided in Eq. 2.24 one can clearly see that the detection performance of the OH system was impeded by relatively low optical power (15\( \mu \)W) and by the higher energy of the photon at 2.8\( \mu \)m. This yields \( >10 \) times worse ultimate detection limit as compared to NO sensing systems operated at
5.3 μm. Therefore despite close to shot-noise limited performance the actual noise equivalent Faraday rotation angle achievable by this system was very similar to the results obtained with the other two conventional FRS systems discussed above. The performance of the H-FRS described in this paper can now be directly compared to the three conventional FRS systems mentioned above. Table 4.1 shows a summary of parameters for all FRS systems discussed (including a separate column for the system from Ref. [90] based on cryogenically cooled detector).

Table 4.1: Performance comparison of selected conventional FRS systems with H-FRS.

<table>
<thead>
<tr>
<th></th>
<th>AC-FRS w/ TEC MCT detector [90]</th>
<th>AC-FRS w/ LN₂ cooled InSb detector [90]</th>
<th>AC-FRS [147]</th>
<th>AC-FRS [146]</th>
<th>H-FRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured [rad/Hz₁/²]</td>
<td>2.53×10⁻⁷</td>
<td>2.24×10⁻⁸</td>
<td>7.89×10⁻⁸</td>
<td>1.39×10⁻⁷</td>
<td>1.79×10⁻⁸</td>
</tr>
<tr>
<td>Molecule</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Target transition</td>
<td>Q(3/2)</td>
<td>Q(3/2)</td>
<td>Q(3/2)</td>
<td>Q(3/2)e,f</td>
<td>R(17/2)f</td>
</tr>
<tr>
<td>Line intensity [cm/molecule]</td>
<td>6.63×10⁻²¹</td>
<td>6.63×10⁻²¹</td>
<td>6.63×10⁻²¹</td>
<td>9.45×10⁻²⁰</td>
<td>1.2×10⁻²⁰</td>
</tr>
<tr>
<td>Frequency [cm⁻¹]</td>
<td>1875.8</td>
<td>1875.8</td>
<td>1875.8</td>
<td>3568.5, 3568.4</td>
<td>1906.7</td>
</tr>
<tr>
<td>g-factor</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.93</td>
<td>0.0996</td>
</tr>
<tr>
<td>Detection limit [ppbv-m/Hz₁/²]</td>
<td>3.78</td>
<td>0.334</td>
<td>1.03</td>
<td>8.88</td>
<td>3.09 (~0.5)²</td>
</tr>
</tbody>
</table>

1^The shot noise ΘSNEA/√(Δf) is calculated with Eqs. 2.24 and 4.7 based on the actual experimental conditions: P₀ = 2.9 mW for Ref. [90], P₀ = 60 mW for Ref. [147], P₀ = 0.015 mW for Ref. [146], and P₀ = 14 mW was used for the H-FRS. The detector quantum efficiency in all the AC-FRS systems was assumed to be η = 0.5, which is a moderate number that should provide a reasonable approximation of the systems fundamental limit. A measured ηhet = 0.5 was used for H-FRS.

2^Detection limit estimated for Q(3/2) transition of NO at 1875.8 cm⁻¹.
In terms of the ultimate shot-noise limited sensitivity all NO sensing systems were very similar (within factor of 5 between the highest and the lowest limit) and the OH system shows the lowest performance limit due to the reasons described above. In terms of the actual noise equivalent Faraday rotation angle the H-FRS system outperforms all AC-FRS systems by: a) a factor of 14 with respect to Ref. [90], b) a factor of 1.3 with respect to system with cryo-cooling in Ref. [90], c) a factor of 4.4 with respect to Ref. [147] and d) a factor of 7.8 with respect to Ref. [146]. Most importantly the H-FRS system approaches the shot noise limited operation (with a noise floor of 3.7 times the shot noise) by utilizing the heterodyne process and a non-cryogenically cooled detector. This is obtained through recycling of the optical power that is normally lost in 90-degree FRS systems and using it as a LO in the H-FRS measurement. It is obvious that despite significant improvement in terms of the noise equivalent Faraday rotation angle the sensitivity to NO concentration for the H-FRS system is up to 3 times lower than for cryogen-free AC-FRS systems referenced above. The higher concentration sensitivity of the two AC-FRS sensors is determined primarily by the target NO transition that is optimal for FRS measurements at low magnetic fields. Although the Q(3/2) transition of NO at 1875.8 cm$^{-1}$ has nearly two times lower absorption line intensity than the R(17/2)f transition at 1906.73 cm$^{-1}$, higher g-factor in the Q-branch (see Table 4.1) allows for more optimal Zeeman split at relatively low magnetic field ($\sim 100$ G$_{rms}$). For our experimental conditions we have estimated $\sim 6$ times stronger FRS signal that can be achieved with the Q(3/2) transition in comparison to the R(8.5)f transition targeted in this work. After accounting for this signal enhancement the cryogen-free H-FRS system equipped with a laser targeting Q(3/2) transition is expected to show concentration sensitivity of 0.5 ppbv-m/Hz$^{1/2}$. Such a performance would be comparable to the AC-FRS system based on LN$_2$ cooled InSb detector in Ref. [90] but without the need for cryogenic cooling (see Table 4.1 for details).
4.5 Chapter summary

A heterodyne enhanced Faraday rotation spectroscopy has been introduced. When compared to a conventional FRS, the H-FRS offers several advantages that improve system performance and allow close-to fundamental noise limited operation. Especially when combined with mid-IR QCL sources, which show significant laser noise in the low frequency region (DC to \(\sim 100\) kHz), the shift of FRS detection to the RF frequencies enables near shot noise limited performance of the system. In our system the laser RIN at 30MHz used for H-FRS is \(\sim 370\) times lower than that observed at the \(f_m\) of 1.08 kHz for AC-FRS. Moreover, theoretically the H-FRS can suppress electromagnetic interference (EMI) caused by the high AC currents used to generate modulated magnetic fields. Those, if not appropriately shielded, affect the performance of the conventional FRS systems by creating uncontrolled pick-up at the modulation frequency \(f_m\) in various parts of the system electronics. Since the signal demodulation in AC-FRS is performed at the same frequency, this uncontrolled pick-up becomes a significant issue. In H-FRS system FRS signal is encoded in the heterodyne beatnote at RF frequency \(\Omega\), which is significantly higher than \(f_m\) and noise at low frequencies (including the pick-up) can be effectively avoided and suppressed by high-pass filters. Therefore, as shown in Fig. 4.11, H-FRS signal shows significantly lower parasitic offset compared to common 90-degree AC-FRS systems (the small offset in Fig. 4.11 was caused by an insufficient EMI shielding of the electrical components located after the frequency mixer or within the laser driver itself that creates small residual intensity modulation). An experimental demonstration of H-FRS was performed using a cryogen-free system based on 5.24 \(\mu\)m DFB-QCL targeting nitric oxide as the test molecule and a TEC-cooled MCT photodetector. The optical bench-top prototype H-FRS system exhibited the total noise of only 3.7 times (5.7 dB) higher than the quantum shot noise, which corresponds to the noise equivalent Faraday rotation angle of \(1.79 \times 10^{-8}\) rad/Hz\(^{1/2}\). When the R(17/2)f tran-
sition of NO at 1906.73 cm\(^{-1}\) is used as the target line with 15 cm optical path and magnetic field of 100 G\(_{\text{rms}}\), this noise equivalent Faraday rotation angle translates to the NO detection sensitivity (1\(\sigma\)) of 20.6 ppbv/Hz\(^{1/2}\). The performance of this system can be further improved by optimizing the optical design of the system (e.g. 95/5 beam splitter can be implemented to preserve optical power in the signal beam; a balanced heterodyne receiver can be applied etc.), or a six-fold improvement can be simply achieved by targeting the optimum Q(3/2) transition of NO at 5.33 \(\mu\)m.
Chapter 5

Dual-modulation Faraday rotation spectroscopy

In Chapter 4 a new H-FRS detection scheme was developed that achieves close to fundamental shot noise limited performance without application of cryogenically cooled detectors. The H-FRS work has demonstrated that by shifting the detection frequency from several kHz to RF >30 MHz region, allowing for effective elimination of laser noise and achieved sensitivity that was only 3.7 times the fundamental shot noise limit performance. However, the complex optical set-up as well as the interferometer instability due to the AOM frequency shifter limit its application only to laboratory use. In this chapter I discuss a new DM-FRS that utilizes much simpler laser wavelength modulation to suppress laser noise and other unwanted parasitic effects and drifts in an FRS system. Based on this measurement method a transportable, robust, cryogenic-free instrument for detection of NO isotopes has been constructed and tested in clinical conditions. The instrument operates at only \( \sim 2 \) times the fundamental quantum noise level. The following sections will discuss the design and principle of operation as well as the implementation of the instrument to
bio-medical sensing of two NO isotopes ($^{14}$NO and $^{15}$NO) in human breath, blood and urine.

5.1 Background and motivation for isotopic NO detection

Currently there are several commercially available technologies commonly used for NO detection and most popular include chemiluminescence analyzers [164] and electro-chemical sensors [165]. However these technologies are not capable of distinguishing between different NO isotopes, which prevents them from being used in advanced bio-medical applications that use isotope labeling techniques (e.g. in metabolic studies). NO isotope measurements are often performed with significantly more complex instrumentation such as mass spectrometry (MS) equipped with resonance enhanced multi-photon ionization (REMPI) [166]. A significant barrier with such a complex systems is a high ownership and maintenance cost as well as the requirement of trained personnel, which limits their use only to highly specialized laboratories. Optical methods based on laser spectroscopy are also capable of sensitive NO isotope detection and with increasing availability of new turn-key laser sources these methods show potential for instrumentation that can be used by non-experts. Since specimens studied in bio-medical applications are often quite complex mixtures of various (and often unknown) species, high chemical selectivity is an important feature that has to be provided by the measurement method of choice. Not all spectroscopic methods capable of NO-isotopic detection can fulfill the requirement of high sensitivity and selectivity simultaneously. For example LIF [167], although very sensitive to NO, exhibits dependence of the measured signal on gas sample composition (primarily due to non-radiative quenching effects). To assure satisfactory accuracy in NO-isotope detection LIF would require sophisticated calibration procedures performed
by a qualified operator, which prevents this technique from being broadly applied. A number of mid-IR laser absorption spectroscopy techniques have also been successfully employed to isotopic NO detection. Spectral region around 5.3 $\mu$m containing the strongest fundamental ro-vibrational band of NO is usually targeted to assure the highest sensitivity. In addition to that a variety of sensitivity enhancement schemes such as multi-pass cells \cite{76} or high finesse cavity enhanced detection \cite{168,45} have been applied to achieve NO detection at sub-ppbv/Hz$^{1/2}$ levels. All absorption based techniques require careful selection of the target transition to avoid or minimize spectral interference from other species, but in case of complex gas mixtures a possibility of unintended spectral interference still exist. One of the mid-IR laser spectroscopic techniques successfully used for NO detection that distinguishes itself by providing ultra-high sensitivity and significantly enhanced selectivity is FRS. FRS is sensitive exclusively to paramagnetic species (including NO) \cite{88,91,92,89,87,90,93}, which allows to achieve high selectivity even in the case of spectral interference from diamagnetic species (such as H$_2$O or CO$_2$) that could be present in the sample gas. In biomedical applications this represents a significant advantage over methods discussed above. The novel DM-FRS technique presented in this chapter allows performing measurements of NO with sensitivities approaching the fundamental quantum limits, while using relatively simple opto-mechanical design that allows for construction of a reliable instruments for applications in clinical settings.

5.2 DM-FRS principle

It is important to realize that the shortcomings of the conventional FRS approaches (both AC-FRS and DC-FRS) are strongly related to the process of signal demodulation that occurs exactly at the first or higher harmonics of the modulation frequency. This causes unwanted pickups of EMI generated background in AC-FRS or residual
Figure 5.1: a) modulation/demodulation process in DM-FRS and b) Frequency spectrum of the AC-, DC-, and DM-FRS signals
absorption signals/parasitic fringing in DC-FRS. Moreover low frequency modulation of the magnetic field in AC-FRS makes it difficult to achieve shot noise limited operation unless high-stability laser drivers and cryogenically cooled detectors are used to lower the technical noise at low frequencies. We have already demonstrated that at high frequencies (>10 kHz) FRS can be performed with sensitivities dominated by the fundamental quantum noise floor without the need for custom laser drivers or cryogenic cooling. However to achieve that goal a rather complex H-FRS has been implemented (see Chapter 4 and Ref. [169]), which would be difficult to use outside specialized optics laboratories.

In the new DM-FRS method developed here, an SNR enhancement comparable to H-FRS has been achieved with a significantly simpler and robust optical layout used for conventional AC-FRS systems. The noise and parasitic effects have been efficiently suppressed through application of dual modulation process: 1) magnetic field modulation occurs at $f_0$, while 2) the laser wavelength is modulated at the frequency $f_1$ that is much higher than $f_0$ (e.g. for $f_0$ in the single kHz range, $f_1$ is selected in the tens to hundreds of kHz range). Within an instant $\tau$ (assuming $1/f_1 < \tau \ll 1/f_0$) the FRS signal can be essentially analyzed using a static magnetic field approximation similar to DC-FRS [100], which can be demodulated at harmonics of $f_1$. Over longer time the magnetic field modulation at $f_0$ will cause the amplitude of the FRS signal at $f_1$ to oscillate. Thus the photodetector signal contains different harmonics of the carrier frequency $f_1$ with the amplitude modulation (AM) sidebands separated by $f_0$. Schematic generation of the DM-FRS spectrum as well as a comparison of the conventional AC- and DC-FRS signals with the DM-FRS signal in the frequency domain are schematically shown in Fig. 5.1a and b respectively.

The detection of the DM-FRS signal can be performed through straightforward AM-demodulation at the carrier frequency that corresponds to the desired harmonic of the wavelength modulated FRS signal. To calculate the DM-FRS spectrum at any
of the detected harmonics of $f_1$, DM-FRS can be analyzed as a wavelength modulated version of AC-FRS. A modeling of the Faraday rotation angle and the expected AC-FRS signal has been studied by many groups and is well understood [89, 87, 170]. Thus the DM-FRS signal can be conveniently modeled using the same method as described by G. Wilson [171] for wavelength modulation spectroscopy (WMS), and the Fourier expansion the $n_{th}$ harmonic spectrum $H_n(x)$ can be calculated as:

$$H_n(x) = \frac{2}{\pi} \int_0^\pi G(x + W \cos(\theta)) \cos(n\theta) d\theta \quad (n \neq 0), \quad (5.1)$$

where $G(x)$ is a known AC-FRS signal line shape function, $W$ is the modulation amplitude with the same unit as the wavelength $x$. Dependence of the DM-FRS 2$^{nd}$ harmonic signal amplitude on the wavelength modulation depth is shown in Fig. 5.2 (See Appendix C.1 for Matlab code). $R$ is the ratio of the modulation amplitude $W$ to the full width half maximum (FWHM) of the AC-FRS spectrum. The DM-FRS signal amplitude reaches maximum at $R=1.19$. It should also be noted that the DM-FRS maximum amplitude is a factor of 1.82 smaller than the maximum AC-FRS signal amplitude. In the actual DM-FRS system, the laser modulation depth was optimized to be around the line-width of the target transition, which provides the maximum DM-FRS signal. Despite the maximum signal amplitude of DM-FRS being one half of the AC-FRS signal, the system SNR can be still improved due to the significant reduction of the noise at high frequencies ($1/f$ noise). Technically the DM-FRS signal detection can be performed by down-conversion of the signal at any harmonic of $f_1$ to the baseband using a frequency mixer and by subsequent lock-in detection of the DM-FRS signal at $f_0$. For practical purposes demodulation at $2 \times f_1$ is performed in this work. The 2$^{nd}$ harmonic DM-FRS spectrum exhibits maximum at the center of the transition, and its peak value is proportional to the sample concentration. This
Figure 5.2: Calculated DM-FRS 2\textsuperscript{nd} harmonic signal amplitude at different modulation ratio R (R = W/(FWHM of AC-FRS)). The inset shows the spectrum of an AC-FRS system (gray) and the spectrum of a DM-FRS system (black) demodulated at the 2\textsuperscript{nd} harmonic with modulation ratio of R = 1.19.

approach allows performing continuous monitoring of sample concentration with the laser frequency locked to the transition peak.

5.3 DM-FRS instrument configuration

The configuration of a prototype DM-FRS spectrometer is shown in Fig. 5.3. The sensor was designed to be able to target two isotopes of NO (\textsuperscript{14}N\textsuperscript{16}O, \textsuperscript{15}N\textsuperscript{16}O, in the following text \textsuperscript{14}NO and \textsuperscript{15}NO will be used for simplicity) and achieve the highest sensitivity for the minor isotope \textsuperscript{15}NO which occurs at much lower concentration (natural abundance of 0.363\%). A room temperature DFB-QCL that gives access to the Q(3/2) transition at 1842.76 cm\textsuperscript{-1} in the \nu_3 fundamental ro-vibrational band of
Figure 5.3: Schematic set-up for a DM-FRS system. The green box indicates all functional blocks for signal generation and processing, which are accomplished by a single digital lock-in amplifier (Zurich Instrument HF2LI).

Figure 5.4: DFB-QCL wavelength tuning vs. bias current at different temperature.
Figure 5.5: All three stable NO isotopes’ (i.e. $^{14}\text{N}^{16}\text{O}$ ($^{14}\text{NO}$), $^{15}\text{N}^{16}\text{O}$ ($^{15}\text{NO}$) and $^{15}\text{N}^{18}\text{O}$) transitions in the wavenumber region of $1825\text{ cm}^{-1} \sim 1975\text{ cm}^{-1}$. The shaded area indicates that a minimum wavenumber range of $\sim50\text{cm}^{-1}$ is required to access the most optimum $Q(3/2)$ transitions for monitoring all three stable NO isotopes using FRS technique.

$^{15}\text{NO}$ has been used as a source. The measured QCL wavelength coverage at different temperatures is shown in Fig. 5.4. As discussed in Chapter 2, $Q(3/2)$ transition yields the strongest FRS signal at low magnetic fields, which guarantees the best SNR. The same laser source can be tuned (by laser current) to access the $P(19/2)e$ transition at $1842.96\text{ cm}^{-1}$ in the $\nu_3$ fundamental ro-vibrational band of the major isotope $^{14}\text{NO}$ as shown by Fig. 5.5 and Fig. 5.6. Such selection of the laser source proposed earlier by Sabana et. al. [172] enables detection of both isotopes with a single laser. In the future this system can be further improved to enable simultaneous concentration monitoring of multiple paramagnetic species (e.g. all NO isotopes or NO and NO$_2$ measured simultaneously) by targeting multiple optimum transitions (as indicated by
Figure 5.6: Zoomed in spectrum between the wavenumber 1840 cm$^{-1}$ and 1850 cm$^{-1}$ indicating $^{15}$NO and $^{14}$NO transitions. The target region using the DFB-QCL is shown by the green box.

Figure 5.7: Concept of monitoring multiple paramagnetic species using DM-FRS

Fig. 5.5 for all NO isotope detection) using multiple laser sources sharing a common optical path and processing unit. This unique technique can be realized by individual encoding/decoding of multiple molecular signals using different carrier frequencies applied to different lasers as depicted in the Fig. 5.7.

In this single laser configuration, the laser beam was collimated with a ZnSe lens with focal length of 1.9 mm and diameter of 4 mm. The laser was operated at 12.5 $^\circ$C with a bias current of $\sim$390 mA, which provided an output power of $\sim$8 mW.
Two MgF$_2$ Rochon prisms were used as the polarizer and the analyzer providing an effective extinction ratio of $1.86 \times 10^{-4}$. To increase the effective path-length without adding alignment complexity associated with multi-pass cells, a simple triple-pass configuration was formed in the 15-cm-long cell which resulted in an effective path-length of 45 cm. The cell was surrounded by an air core solenoid (796 turns of 16 AWG copper wire) and was driven with a power amplifier creating an axial AC magnetic field modulated at frequency of 3.26 kHz (this is frequency of the resonant circuit formed by adding a 0.34 $\mu$F capacitor in parallel with the solenoid). To ensure the maximum signal amplitude the pressure and the magnetic field strength inside the gas cell were optimized to 35 Torr and 158 G respectively [90, 147]. Gas pressure was stabilized using a pressure controller (MKS $\pi$PC PC99). The ordinary beam from the analyzer is then focused by an aspheric ZnSe lens onto a thermoelectrically cooled mercury cadmium telluride photodetector. After frequency down-conversion from 2$^{\text{nd}}$ harmonic of $f_1$ (2 $\times$ 50 kHz) using a mixer, the DM-FRS signal is measured using lock-in demodulation at $f_0=3.26$ kHz. In the line-locked mode that enables continuous concentration monitoring, the extraordinary beam emerging from the analyzer is recycled and directed through a 5 cm-length reference cell filled with $\sim$ 1% of NO in N$_2$ mixture (the NO content was composed of both $^{15}$NO and $^{14}$NO isotopes in amounts assuring equal absorption strengths). The 3$^{\text{rd}}$ harmonic of the absorption WMS signal demodulated by a lock-in amplifier in the reference branch is used as an error signal for a PID controller that performs feedback controlled locking of the laser wavelength to the center of the target transition.

The system was constructed as transportable instrument with the overall size of approximately 60$\times$45$\times$100 cm$^3$ (this was about half of the size of the conventional AC-FRS sensor system based on an EC-QCL light source developed earlier by Lewicki et. al. [90]). The signal generation and demodulation, lock-in detection, data acquisition, and PID control was realized using a single digital lock-in amplifier unit (Zurich
Instrument HF2LI performing functions marked with a green box in Fig. 5.3, which allowed for significant reduction of the instrument size and the number of table-top electronics units required by the system.

5.4 System characterization

In order to optimize the DM-FRS performance a full characterization of the noise sources has been performed. The laser RIN and the photodetector NEP at signal frequency were measured to be $\sigma(\omega) = 1.78 \times 10^{-7} Hz^{-1/2}$ and $\text{NEP}(\omega) = 1.39 \times 10^{-12} W/Hz^{1/2}$, respectively. As predicted by Section 2.2.2 the optimum performance (maximum SNR) of FRS spectrometer is achieved at analyzer offset angle that provides laser noise equal to the detector noise. For this DM-FRS system the $\theta_{opt}$ is set to 1.8°. From Eq. 2.18 and assuming a signal loss factor of $\sim 1.82$ due to the second harmonic detection, a noise-equivalent Faraday rotation angle $\Theta_{NEA} = 7.11 \times 10^{-9} rad/Hz^{1/2}$ has been determined for this system.

Another important assessment of the system performance can be performed by comparing the systems $\Theta_{NEA}$ to the theoretical limit set by the fundamental quantum noise floor determined by the shot noise of the measured photocurrent. With this assumption the shot-noise-equivalent Faraday rotation angle (see Eq. 2.24) of $\Theta_{shot} = 3.22 \times 10^{-9} rad/Hz^{1/2}$ has been calculated. In terms of figure-of-merit DM-FRS prototype provides sensitivity that is only 2.21 times worse than the theoretical quantum limit and outperforms by 2.5 times our previous cryogen-free FRS system based on a relatively complex heterodyne-enhanced FRS technique (see Chapter 4). The excellent performance is a result of the signal demodulation shifted from single kilohertz to $\sim 100$ kHz range where laser RIN and NEP allow for shot noise-limited operation.
Figure 5.8: 2\textsuperscript{nd} harmonic DM-FRS spectrum acquired in the sample channel (a) and 3rd harmonic WMS signal from absorption cell acquired in the reference channel (b). The laser wavelength was scanned across the spectral region containing the target transitions for \(^{15}\text{NO}\) and \(^{14}\text{NO}\) measurements.

Fig. 5.8a shows a typical DM-FRS spectrum acquired with the laser wavelength scanned across the target transitions of \(^{14}\text{NO}\) (P\((19/2)\)e doublet) and \(^{15}\text{NO}\) (Q\((3/2)\) line). A certified mixture of 2 ppmv NO in N\(_2\) was used in this measurement. Based on the natural isotopic abundance the mixture contains \(\sim 2\) ppmv of \(^{14}\text{NO}\) and 7.3 ppbv of minor isotope \(^{15}\text{NO}\) (assuming 0.363\% natural abundance). The minor isotope line is clearly visible in the scan. The conversion factors for the output signal to gas concentration of 25.6 ppbv/\(\mu\)V and 3.65 ppbv/\(\mu\)V were estimated based on the peak FRS signal values for the \(^{14}\text{NO}\) (P\((19/2)\)e) and \(^{15}\text{NO}\) (Q\((3/2)\)) respectively. Seven times higher sensitivity to heavy isotope clearly shows that the performance of the system has been optimized for this isotope. In a continuous concentration moni-
Figure 5.9: The Allan deviation plot (red) for the measured DM-FRS signal (raw data shown by upper panel) and for the system shot noise (green dot). The measurement was performed with the laser wavelength locked to the $^{15}$NO Q(3/2) transition and dry N$_2$ was flown through the gas cell.

In monitoring mode, only the peak DM-FRS signal is measured and converted into molecular concentration units after prior calibration. Fig. 5.8b shows the 3rd harmonic ($3f_1 = 150$ kHz) of the WMS signal acquired during the same scan in the reference arm and used as an error signal for an active line-locking.

The system long-term performance has been tested with the laser wavelength locked to the Q(3/2) line of the minor isotope. To assess the system minimum detection limit a zero gas (dry N$_2$) was continuously purged through the sample cell. Fig. 5.9 shows an Allan deviation plot (see Appendix C for details) of the DM-FRS signal acquired with 1 Hz bandwidth and data points sampled in 1 s intervals (raw data shown by upper panel). Similarly to the spectrum in Fig. 5.8a, the line-
locked DM-FRS signal is free from background offset that is usually observed in other FRS systems and typically originates from the EMI pickup (in AC-FRS) or parasitic etalons (in DC-FRS) (see Appendix C for the spectrum comparison). The DM-FRS measurement shows nearly ideal white-noise performance up to integration times $>3000$ s. Based on the measured conversion factors the noise spectral density of 145 nV/Hz$^{1/2}$ measured in the line-locked mode corresponds to a detection limit of 0.53 ppbv/Hz$^{1/2}$ and 3.72 ppbv/Hz$^{1/2}$ for $^{15}$NO and $^{14}$NO, respectively. A green dot line indicating the Allan plot for the shot-noise level of 71.2 nV/Hz$^{1/2}$ ($\sim2$ lower than the system noise) calculated for the measured photocurrent is also shown in Fig. 5.9. The factor between shot noise and system noise is 20% lower than that calculated based on figure-of-merit $\Theta_{NEA}$ and shot-noise-equivalent Faraday rotation angle $\Theta_{shot}$. The discrepancy can be attributed to the inaccurate on the measured system parameters (e.g. laser RIN, photodetector NEP etc.). Given the optimized performance for $^{15}$NO isotope, the DM-FRS detection limit corresponds to a bandwidth-normalized equivalent fractional absorption of $6.89 \times 10^{-8} Hz^{-1/2}$. This clearly shows that in comparison to the best mid-IR absorption based systems achieving a bandwidth-normalized fractional absorption down to $\sim 2 \times 10^{-6} Hz^{-1/2}$ levels [74], the DM-FRS system offers $\sim30$ times better sensitivity. This allows performing measurements of paramagnetic species such as NO, NO$_2$, O$_2$, •OH, HO$_2$ etc. at ppmv and low ppbv levels without the need for multi-pass absorption cells [74] or high finesse cavity enhancement [173]. Therefore the current setup employing an effective optical path of only 45 cm offers simple and robust optical alignment and shows potential for further miniaturization of the optical setup.
Figure 5.10: Online breath $^{14}$NO (red) test with normal breath. The DM-FRS sensor is integrated with a LOCCIONI single breath sampler which measures CO$_2$ (blue) and airway pressure (black) at the mean time.
5.5 Applications

Detection of NO is of great importance in a number of bio-medical applications such as non-invasive medical diagnostics based on exhaled human breath analysis [174, 175] or study of the regulation of biological and physiological processes in human and mammalian cells [26, 27, 25, 176]. One of the major limitations facing the study of NO metabolism in human health and disease is the inability to easily measure the total body NO production. While many researchers have measured gas phase NO in exhaled breath and NO metabolites (mainly nitrate and nitrite) in the blood and urine, there has not been a good method to determine the relative contribution of any or all of these to the total body production of NO. In one study total body NO production was assessed by giving an intravenous infusion of L-[15N]2-arginine (50 µmol/min for 30 min) and measuring isotopic urinary enrichment of 15N-nitrite and 15N-nitrate [177]. The main difficulty for the study was the inability to measure the different isotopes in exhaled breath which is a major portion of the total NO produced by the body. Clearly, studies of the kinetics of NO metabolism are in need for instrumentation that offers simple operation, robust design and high sensitivity, and which enables monitoring of isotope-labeled NO directly in human breath as well as in urine and blood samples.

To test the DM-FRS system performance as a bio-medical analyzer of the exhaled isotope-labeled NO we have performed preliminary performance test using a LOCCIONI single breath sampler integrated with the system. First measurements of the exhaled major isotope 14NO (red line) concentration together with the CO2 sensor data (blue line) and mouth pressure (black line) provided by the LOCCIONI breath sampler have been acquired (shown in Fig. 5.10). This provided characterization of the sampling procedure and time delays associated with the gas flow through the system. There is a 10 second lag time for the NO signal as compared to the CO2 and mouth pressure signal measured within the mouth piece of the LOCCIONI sam-
Figure 5.11: Measurement of both $^{15}$NO (red) and $^{14}$NO (blue) concentration from a patient’s urine samples obtained before and after 15N labeled arginine was introduced intravenously.

pler. This can be attributed to the relatively long tubing (∼1 m) between the breath sampler and the NO instrument. The measured NO concentration shows a signal rise-time of ∼3 seconds and a clean-up time of ∼5 seconds. This confirms that the flow of ∼200 mL/min used in this experiment provides the sample cell flush time of <5 s, which is sufficient to capture a clear NO concentration plateau for NO concentration (either isotope) during the collection in a single breath maneuver as suggested by the clinical guidelines [42].

In order to enable measurement of NO in urine and blood samples the DM-FRS was coupled with a commercial kit for chemical conversion of nitrate and nitrite to NO (purge system for Sievers NOA 280i analyzers, see Appendix C for details). The system test was performed using urine samples collected from unknown patients before and 10 hours after infusion with L-$^{15}$N-arginine serving as a precursor for the synthesis of biogenetic NO [178]. Because the sensor has been designed to measure one isotope at a time, each sample was divided into two equal portions and each
portion was tested for presence of different NO isotope. Fig. 5.11 shows the concentration peaks of both isotopes when the diluted urine samples were injected into the reducing agent. In the sample collected before \( ^{15} \text{N-arginine} \) injection the peak \( ^{14} \text{NO} \) concentration was \( \sim 1.2 \text{ ppmv} \) and the peak \( ^{15} \text{NO} \) concentration was \( \sim 4 \text{ ppbv} \). The area under the signal peak, which is proportional to the amount of nitrite and nitrate in the sample, is used for calculation of the heavy isotope content. The percentage of the heavy \( ^{15} \text{NO} \) isotope in the sample before arginine infusion is 0.36%, which is in excellent agreement with the natural abundance. For the sample collected 10 hours after the arginine injection, the peak concentration of \( ^{15} \text{NO} \) and \( ^{14} \text{NO} \) in the urine sample acquired from the same patient was found to be \( \sim 8 \text{ ppbv} \) and \( \sim 0.54 \text{ ppmv} \) respectively. The heavy isotope content increased to 1.41%, clearly indicating an arginine catabolism process in the body. In fact, measurement of \( ^{15} \text{NO} \) content in blood samples from patients with \( ^{15} \text{N-arginine} \) injection was also performed and all the results are consistent with those measured from urine samples.

\( ^{15} \text{NO} \) content in breath samples have also been measured. Due to the limitation on the instrument sensitivity only breath samples from patients with \( ^{15} \text{N-arginine} \) injection were measured. The breath samples were stored in small volume (\( \sim 10 \text{ mL} \) tubes), therefore the flow rate of the DM-FRS sensor was adjusted to \( \sim 50 \text{ mL/min} \) to allow the samples flow through the cell with enough acquisition time. Please note unlike the chemiluminescence NO analyzer DM-FRS signal strength does not depend on the sample flow rate. Fig. 5.12 shows an example of the signal measured from one of the breath samples. At this small levels of \( ^{15} \text{NO} \) a small background drift was observed and a baseline correction has been performed before calculating the area integral of the peak. The average percentage of \( ^{15} \text{NO} \) isotopes in breath samples is \( \sim 1.33\% \), which agrees quite well with that measured in blood and urine samples. Table 5.1 summaries the typical isotope ratio measured from the blood, urine and breath samples before and after \( ^{15} \text{N-arginine} \) dose.
Figure 5.12: $^{15}$NO signal measured from a breath sample.

Table 5.1: $^{15}$NO/$^{14}$NO ratio measured in urine, blood and breath collected before and after $^{15}$N-arginine injection.

<table>
<thead>
<tr>
<th></th>
<th>$^{15}$NO/$^{14}$NO ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
</tr>
<tr>
<td>Urine</td>
<td>0.36%</td>
</tr>
<tr>
<td>Blood</td>
<td>0.37%</td>
</tr>
<tr>
<td>Breath</td>
<td>N/A</td>
</tr>
</tbody>
</table>
In general, we have successfully observed $^{15}$NO isotope (although at quite low level of $\sim$0.4 ppbv) existing in breath samples from patients with $^{15}$N-arginine enrichment. The measured concentration of $^{15}$NO in blood, urine and breath can be a good reference for characterization of the whole body NO production.

5.6 Chapter summary

Overcoming drawbacks of conventional FRS systems and increasing performance of instrumentation based on this very elegant and sensitive physical principle has been a real challenge. In the previous chapter we have presented heterodyne-enhanced FRS technique that allowed significant improvement in sensitivity of FRS systems with the total noise of only 3.7 times above quantum shot-noise [169]. However, high complexity of the optical set-up prevented this technology from being used in the environments outside specialized optical laboratory. In this chapter we have presented a new DM-FRS system that is free from this limitation and outperforms H-FRS in terms of sensitivity. DM-FRS allows for truly zero-baseline, large dynamic range, in-situ, and flow rate independent measurements of NO isotopes with sub-ppbv sensitivity. The noise-equivalent polarization rotation angle of $7.11 \times 10^{-9} \text{rad/Hz}^{1/2}$ represents at least an order of magnitude more sensitive than those previously reported systems based on the Faraday effect [90, 147, 146]. This translates into the NO detection sensitivity ($1\sigma$) of 0.53 ppbv/Hz$^{1/2}$ and 3.72 ppbv/Hz$^{1/2}$ for $^{15}$NO and $^{14}$NO respectively (with 45cm optical path and magnetic field of 158 G). The total noise observed in this system is only 2 times higher than the quantum shot-noise limited values, which has not been achieved before with a non-cryogenic FRS system. The simplicity and robustness of the optical setup and integration of the data acquisition and signal processing electronics allowed a prototype system to become a reliable transportable

$^1$Due to the limitation of the sensitivity on $^{15}$NO, the DM-FRS sensor is not capable of detecting the natural content $^{15}$NO concentration ($\sim$0.1 ppbv level) in the breath samples before $^{15}$N-arginine injection.
instrument that could be used in a true clinical environment and enable studies of NO metabolism in human body. This makes DM-FRS a very promising approach for a next generation NO isotope spectrometers for applications in environmental and medical studies.
There is a strong demand for broadband, high resolution spectrometry in the mid-IR spectral region where many molecular species have their strongest rotational-vibrational absorption bands. QCLs have become indispensable compact mid-IR laser sources for the development of selective and sensitive molecular spectroscopic sensing techniques. Moreover, application of broadly tunable EC-QCL technology is particularly attractive since it allows for efficient multi-species detection using only a single laser source. The main goal of this dissertation was to develop a powerful FRS based bio-medical NO sensor with unique capability of simultaneous detection of all stable NO isotopes (i.e. $^{14}\text{N}^{16}\text{O}$, $^{15}\text{N}^{16}\text{O}$ and $^{15}\text{N}^{18}\text{O}$). An EC-QCL operating at $\sim$5.3 $\mu$m and with >50 cm$^{-1}$ tuning range is needed to access all three Q branches of NO isotopes which are far apart (see Fig. 5.5). To achieve this, appropriate selection of a FP-QCL as the gain medium, proper optimization of the EC configurations and parameters such as the collimating optics, diffraction grating, and anti-reflection (AR) coatings on the QCL’s facet need to be performed.

In this chapter, a new EC-QCL testing platform for reliable testing QCL chips in an EC configuration is presented. The system allows for automatically acquiring a set of data, including threshold current, output spectra and tuning ranges, from
QCL chips located on a wafer bar. The following sections describe the background of this work, the system design and functionality, the preliminary test results and the method for estimating the coating quality across the wafer using the measured threshold current data.

6.1 Background and motivation

Spectroscopic detection of large molecules (e.g. explosive materials) with a large number of dense unresolvable ro-vibrational lines, require broadband optical frequency coverage of over > 50 cm$^{-1}$ to measure these broad absorption features [179, 180]. On the other hand presence of other small molecular species in the gas mixture might require high spectral resolution to resolve narrow spectral lines and distinguish them from absorption features of the broadband absorbers [181, 182]. Fourier transform infrared spectroscopy (FTIR) is a popular technique for the broadband spectral analysis in the IR region. Commercial FTIR units can achieve resolution of 0.1 cm$^{-1}$ in relatively compact table top instruments down to ~0.001 cm$^{-1}$ (or ~30 MHz) for larger instruments (> 3 m$^2$ footprint, e.g. see Bruker IFS125 series). However, large physical dimensions, long data-processing times, along with moderate spectral resolution and low spectral brightness of the incandescent light sources are the main limitations of this spectroscopic technology, which restricts its use mainly to specialized laboratories [183]. The development of room temperature, high power, CW QCLs enabled development of high resolution mid-IR spectrometers for applications outside the well-controlled laboratory conditions [109]. Many mid-IR spectroscopic applications rely on QCL sources with single longitudinal mode emissions which are typically realized by fabricating distributed feedback (DFB) gratings into the laser waveguide [127, 184, 132] or by integrating an FP-QCL with an external grating-feedback cavity (EC) [181, 185, 121, 136]. DFB-QCLs are compact and simple to
operate, however, the wavelength tuning range of a DFB-QCL achievable by thermal changes to the waveguide properties is usually limited to 1-3 cm\(^{-1}\) if tuned by variation of the injection current and to \(\sim 10-20\) cm\(^{-1}\) obtainable through variations of laser operation temperature. Therefore fabrication tolerances for DFB structures targeting a specific absorption line of the target molecule are very tight (\(\pm 5\) cm\(^{-1}\) if thermal tuning can be utilized for frequency correction), which is difficult to achieve and increases the cost of device fabrication. On the other hand, QCLs are known to provide extremely broadband spectral gain profiles [121], which can be fully exploited by application of EC configurations with broadband high-resolution (\(\sim 0.001\) cm\(^{-1}\)) spectral scanning capability [111, 122].

To select the broadband QCL gain chips for EC-QCL applications, conventional method requires several steps: 1. Wafer preparation; 2. Cleaving QCL wafer based on a specific QCL chip length into bars; 3. Facet coating; 4. Further cleaving QCL bars into chips; 4. Chip die bonding; 5. Chip testing and optimum chip selection (one at a time). Since the bare facet can be contaminated in the die-bonding process, it makes any further coating process ineffective thus, lasers must be coated before bonding process. Once the AR/HR (high-reflection) coating is applied to laser facets, the chips must be die-bonded in order to be characterized one at a time in an external cavity. This is a lengthy process which is time consuming, labor intensive and requires a qualified operator. A system called laser bar prober has been proved to be a cost-efficient equipment to efficiently characterize the electrical parameters of DFB-laser chips [186]. Thanks to the highly uniform and equally distributed arrangement of laser chips on a laser bar, once the first and last laser on the wafer bar are aligned with measurement optics, all the other chips in the bar can be tested by simply translating the whole bar in the horizontal and in the direction perpendicular to the waveguide without adjusting the other components. This method is quite successful for testing laser chips that do not require external optical feedback components (e.g. DFB- and
DBR-lasers etc.). However, to test the laser chips in an EC configuration, in which a $\sim 1 \, \mu m$ optical alignment error will lead to significant change of the lasing threshold, much higher accuracy on the EC alignment is needed. To our best knowledge, no such a system has been reported in literature.

The new system developed here uses a more advanced hexapod motion stage, which allows for 6-degree-freedom optical alignment with 0.1 $\mu m$ accuracy. Instead of time consuming installation and collimation of each QCL chip when mounted separately, the system is able to perform the EC coupling, LI curve and lasing spectra characterization of the laser chips with much less human operation, leading to an improved chip testing and selecting efficiency.

### 6.2 System design and configuration

Fig. 6.1 shows the optical configuration and mechanical design of the system. It is basically a typical Littrow EC configuration [111]. The output beam is collimated by an one inch AR coated (3-12 $\mu m$) germanium aspheric lens. The external cavity uses the $0^{th}$ order reflection from the grating as an output beam and the $1^{st}$ order diffraction beam to provide the laser feedback at Littrow angle as shown in Fig. 6.1. The lasing wavelength $\lambda$ is determined by the Littrow angle as the following equation:

$$
\lambda = 2d \cdot \sin \theta_L,
$$

(6.1)

where $d$ is the distance between grooves of the grating, $\theta_L$ is the Littrow angle. A beam-steering mirror is mounted on the same rotary platform with the diffraction grating to maintain the same optical axis during the tuning process. The diffraction grating and the steering mirror are integrated into one pre-aligned structure mounted on a motorized rotary stage for automatic wavelength tuning by changing the Littrow angle. The gratings were installed in custom-made mounts that were designed to be
Figure 6.1: Mechanical design of the EC-QCL bar tester.
interchangeable parts in the system. Since different gratings are designed for a certain wavelength region, this modular design allows a convenient replacement without further alignment for testing chips at different wavelengths. The whole system functional block diagram is shown in Fig. 6.2. The wafer bar is secured in place by a customized vacuum chuck with machined slot so that it keeps its position every time when a new bar is reloaded. The copper chuck also acts as a thermal and electrical conductor for stabilizing the chip’s temperature with a TEC module as well as for providing an electrical contact. A motorized electrical probe manipulator (constructed by NAI MM-3 series miniature translation stages) is automatized to provide the top electrical contact for any of the QCL chips on the wafer bar. The electrical probe stage is co-located with the chuck on the same moving platform to assure when EC alignment is in motion the electrical probe moves simultaneously with the contacted laser chip. A hexapod parallel-kinematics precision alignment system (PI F-206.S HexAlign\textsuperscript{TM}) with 6 degrees of freedom is used to control the motion of the vacuum chuck, which positions the QCL chips with respect to the collimation lens. It also performs an auto-alignment for each of the QCL chips in the external cavity. This requires the
position of the chip as precise as $\sim 1 \, \mu m$ for a strong EC feedback. Compared to the traditional serial kinematics systems (system built by several stacked translation stages for multiple degrees of motion), F-206.S system has only one moving platform, which makes it be able to provide higher accuracy of positioning capacity due to the lower error accumulation. The output laser beam is permanently aligned with the FTIR. Since the external cavity is always fixed during chip test, the beam path providing the maximum cavity feedback and lasing intensity will not change for different chips, we do not need to realign the FTIR with every chip. Once the EC and the gain chip are aligned, the cavity loss is minimized and the output power of the EC-QCL is maximized, indicating the optimized chip position. The laser is operated in a pulsed mode and the optical signal detected in the FTIR is demodulated by a lock-in amplifier synchronized with the laser pulse repetition rate. The demodulated signal is proportional to the optical power, which is utilized by the F-206.S stage controller as a feedback to perform automatic EC alignment. The control software for the automatic test is designed such that once the first laser and the last laser on the bar are aligned with the EC, their position are saved and coordinates of all the other lasers on the bar are calculated and used for further coarse translation by the F-206.S stage. Before the actual characterization of each laser an auto-alignment procedure is executed to find the optimum EC coupling. Once the EC coupling is optimized the FTIR spectral characterization and LI curve measurement at each grating angle are recorded automatically. A software interface for operating the system is shown in Fig. D.1 in Appendix D.
6.3 QCL chip performance testing

6.3.1 Tuning range vs. threshold current

The system was used to test several QCL wafer bars (developed and fabricated by Corning Inc) lasing at different wavelengths. Firstly the system was used to automatically test the EC tuning spectrum and LI curves measurement performance of 40 QCL chips on a bar emitting at \( \sim 3.5 \ \mu m \). The temperature of the bar is kept at at \( 15 \ ^\circ C \) and the driving current is kept constant while testing the tuning range of the chips.

Fig. 6.3 shows the test results of a selected chip (#28) on a laser bar. Each colored line indicates an FTIR spectral measured at an equally spaced grating angle. An average side-mode suppression ratio (SMSR) of \( \sim 20 \) dB is shown by Fig. 6.4 with y-axis in Fig. 6.3 converted into logarithm scale. The black dots in Fig. 6.3 indicate the lasing threshold by analyzing the LI curves obtained at each of grating angles.

A Littrow configured EC-QCL can be analyzed as two adjacent coupled cavities. One is the chip FP cavity formed by the front and back facets of the chip; another is the EC cavity formed by the diffraction grating. Cavity losses of the two cavities can be written as:

\[
\alpha_{FP} = \alpha_w + \frac{1}{2L} \ln \left( \frac{1}{R_F R_B} \right), \tag{6.2}
\]
\[
\alpha_{EC} = \alpha_w + \frac{1}{2L} \ln \left( \frac{1}{R_{EC} R_B} \right), \tag{6.3}
\]

where \( \alpha_{FP} \) and \( \alpha_{EC} \) are FP cavity loss and EC cavity loss respectively, \( \alpha_w \) is the waveguide loss, \( L \) is the length of the waveguide, \( R_F \) and \( R_B \) are the reflectivity on the chip’s front facet and back facet, respectively. \( R_{EC} \) is the effective reflectivity combining the chip’s front facet and the grating feedback. \( R_{EC} \) in Eq. 6.3 is greater than \( R_F \) in Eq. 6.2 because there is an additional feedback from the grating. Thus, the EC cavity loss \( \alpha_{EC} \) is less than the FP cavity loss \( \alpha_{FP} \), leading to a lower lasing
Figure 6.3: EC spectral tunability and threshold current spectrum for chip#28 on a bar operating at 3.5 µm.

Figure 6.4: Logarithm plot of the same spectrum as above. Side mode suppression ratio ∼20dB.
Figure 6.5: Schematic diagram of the gain and loss in a Littrow configured EC-QCL. Less threshold current is needed at the center of the gain profile because of the maximum differential gain. Maximum tuning range is achieved once the gain peak meets the FP cavity loss.

Thus, the threshold vs. grating angle plot provides a convenient way for coarse estimation of the EC-QCL tuning range. In addition, since the lasing wavelength corresponding the Littrow (grating) angle can be calculated based on Eq. 6.1, it provides a possibility of evaluation of the output wavelength without the need of an expensive and time consuming spectral characterization instrument (e.g. FTIR, wavemeter etc.).

6.3.2 Selecting the best gain chips for EC laser

Subsequently the system was used to characterize the laser chips which were optimized for ultra-broadband operation. Fig. 6.6 shows the tuning range of ~344 cm⁻¹ was
achieved for chip #22. This large tunability proves that the system is reliable for test of QCL chips with ultra-broadband design. Unfortunately, for such a wide tuning range (> 15% of the central wavelength) a chromatic dispersion of the germanium collimating lens starts playing an important role. The refractive index of germanium, which is considered as a material with minimum chromatic dispersion, changes by ∼0.15% within the tuning range which is sufficient to change the focal length of the lens. In the auto-alignment mode, the system does not optimize the focal distance (collimation). Thus for such a wide tuning range manual focus length adjustment is needed to perform optimized measurements in the entire wavelength tuning range. When a manual lens position compensation is applied a wider tuning range of ∼450 cm\(^{-1}\) (from 2140 cm\(^{-1}\) to 2590 cm\(^{-1}\)) was achieved for the same chip.

The bars fabricated by Corning consist of 40 lasers with 5 different waveguide widths. They repeat the same design sequence every 5-lasers (i.e. lasers #1, #6,
Figure 6.7: Tuning spectrum of the 3.5 \( \mu m \) gain chips with waveguide width of 5 \( \mu m \) on the wafer bar. Chip numbers are indicated at each graph, followed the number, “AR” means the front facet of the chip is applied by anti-reflection (AR) coating; ”ARHR” means besides AR coating, a high-reflection (HR) coating is applied to the back facet.

#11, ...#36 have a same design and the narrowest waveguide width, and lasers #5, #10, #15, #40 have the widest waveguide width.). With such a design one is able to evaluate the homogeneity of the process by analyzing 8 similar lasers inside a laser bar. Due to the process inhomogeneity, coating non-uniformity and other defects inside the waveguides differences in the cavity loss and the facet reflectivity will be observed, which result into variation of threshold current and the EC-QCL tuning range. The highly efficient chip testing capacity allows us to analyze large volume of the lasers in a relative short time, therefore the QCL chips with the most appropriate wavelength and tuning range can be quickly selected. Fig. 6.7 shows the FTIR test results for 6 gain chips with the exactly same design. Chip #30 and #40 did not lase during the test. To increase the coupling efficiency, an AR coating is applied on the
front facet of the bar facing to the EC. Half of the bar’s rear facet (chip #21 to #40) is HR coated for comparison. Fully characterizing all the 40 chips is an 8 hour process, which includes the automatic EC coupling, the FTIR spectrum acquisition and the LI curve recording. To compare the tunability of the chips directly, the threshold vs. grating angle data of all the lasers are plotted in Fig. 6.8. Obviously, chip #25 and #35 have the greatest tuning range and chip #15 has the lowest tuning range.

With all the data acquired by the system, the tunability of each laser chip can be determined. Therefore we can pre-select the best chips for EC-QCL applications at very early stage of chip manufacturing process.
6.3.3 Estimating the AR coating quality

AR coating is a key element in an EC laser configuration. Low reflectivity allows for a large wavelength tuning range and mode-hop free fine tuning [136]. With this system, if the external cavity is off-aligned on purpose, all the laser chips will lase at FP mode as the grating is not able to provide enough feedback to the gain medium. Since the threshold current is determined by the AR/HR coating and waveguide loss, the AR coating uniformity within the bar can be examined by statistically studying the electrical properties of the chips. The threshold current is related to the cavity loss and facet reflection by the following lasing conditions:

\[ I_{th}^{UCUC} = C \cdot \left( \alpha_w + \frac{1}{2L} \ln \left( \frac{1}{R_{UC}R_{UC}} \right) \right), \]  \hspace{1cm} (6.4)

\[ I_{th}^{UCAR} = C \cdot \left( \alpha_w + \frac{1}{2L} \ln \left( \frac{1}{R_{UC}R_{AR}} \right) \right), \]  \hspace{1cm} (6.5)

\[ I_{th}^{HRAR} = C \cdot \left( \alpha_w + \frac{1}{2L} \ln \left( \frac{1}{R_{HR}R_{AR}} \right) \right), \]  \hspace{1cm} (6.6)

where \( I_{th}^{UCUC}, I_{th}^{UCAR}, I_{th}^{HRAR} \) are the threshold current of FP mode measured when the chip is processed under the stage of uncoated, AR coated, and HR/AR coated, respectively. \( C \) is a constant coefficient related to the gain at the center wavelength. \( R_{UC}, R_{AR} \) and \( R_{HR} \) are the reflectivity of the un-coated, AR coated and HR coated cavity facet, respectively. \( R_{UC} \) can be calculated by \( R_{UC} = (n-1)^2/(n+1)^2 \), which is 27% assuming \( n \approx 3.15 \) for the waveguide material. \( R_{HR} \approx 95\% \) is obtained from the measurement of the witness sample after HR coating being deposited. Waveguide length \( L = 6 \text{ mm} \). Now, only \( C, \alpha_w \) and \( R_{AR} \) are unknown parameters. Solving Eq. 6.4 to Eq. 6.6 one can get:
\[ R_{AR} = \exp \left( \ln(R_{UC}) + [\ln(R_{UC}) - \ln(R_{HR})] \cdot \frac{I_{th}^{UCUC} - I_{th}^{UCAR}}{I_{th}^{HRAR} - I_{th}^{UCAR}} \right) \]

\[ = \exp \left( \ln(R_{UC}) + \frac{\ln(R_{UC}) - \ln(R_{HR})}{1 - \frac{I_{th}^{HRAR} - I_{th}^{UCUC}}{I_{th}^{UCAR} - I_{th}^{UCUC}}} \right) \]  

(6.7)

Table 6.1: Threshold current of 5µm width chips with different coating stages.

<table>
<thead>
<tr>
<th>Chip No.</th>
<th>( I_{th}^{UCUC} )</th>
<th>( I_{th}^{UCAR} )</th>
<th>( I_{th}^{HRAR} )</th>
<th>( I_{th}^{UCAR} - I_{th}^{UCUC} )</th>
<th>( I_{th}^{HRAR} - I_{th}^{UCUC} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.838</td>
<td>1.012</td>
<td>N/A</td>
<td>0.174</td>
<td>N/A</td>
</tr>
<tr>
<td>10</td>
<td>0.802</td>
<td>0.980</td>
<td>N/A</td>
<td>0.178</td>
<td>N/A</td>
</tr>
<tr>
<td>15</td>
<td>0.794</td>
<td>1.275</td>
<td>N/A</td>
<td>0.481</td>
<td>N/A</td>
</tr>
<tr>
<td>20</td>
<td>0.762</td>
<td>0.893</td>
<td>N/A</td>
<td>0.131</td>
<td>N/A</td>
</tr>
<tr>
<td>25</td>
<td>0.754</td>
<td>N/A</td>
<td>0.815</td>
<td>N/A</td>
<td>0.061</td>
</tr>
<tr>
<td>35</td>
<td>0.759</td>
<td>N/A</td>
<td>0.811</td>
<td>N/A</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Eq. 6.7 shows that the AR coating reflectivity is only related to the ratio \( \gamma \) of the threshold current difference between different coating stages.

\[ \gamma = \frac{I_{th}^{HRAR} - I_{th}^{UCUC}}{I_{th}^{UCAR} - I_{th}^{UCUC}} \]  

(6.8)

Table 6.1 lists the measured threshold current \( I_{th}^{UCUC} \), \( I_{th}^{UCAR} \), \( I_{th}^{HRAR} \) when the chips are processed under different coating stages. Due to the coating distribution, only chip #5, #10, #15, #20 are available to calculate \( I_{th}^{UCAR} - I_{th}^{UCUC} \), and chip #25, #35 are available to calculate \( I_{th}^{HRAR} - I_{th}^{UCUC} \). Substituting all the possible values of ratio \( \gamma \) by substituting the data in Table 6.1 into Eq. 6.7, one can estimate a range of \( R_{AR} \), with maximum of ~6.51% and minimum of ~2.52%.
Table 6.2: Threshold current of 4 µm width chips with different coating stages.

<table>
<thead>
<tr>
<th>Chip No.</th>
<th>$I_{UCUC}^{th}$</th>
<th>$I_{UCAR}^{th}$</th>
<th>$I_{HRAR}^{th}$</th>
<th>$I_{UCAR}^{th} - I_{UCUC}^{th}$</th>
<th>$I_{HRAR}^{th} - I_{UCUC}^{th}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0.694</td>
<td>0.847</td>
<td>N/A</td>
<td>0.153</td>
<td>N/A</td>
</tr>
<tr>
<td>14</td>
<td>0.666</td>
<td>0.811</td>
<td>N/A</td>
<td>0.145</td>
<td>N/A</td>
</tr>
<tr>
<td>19</td>
<td>0.618</td>
<td>0.764</td>
<td>N/A</td>
<td>0.146</td>
<td>N/A</td>
</tr>
<tr>
<td>29</td>
<td>0.694</td>
<td>N/A</td>
<td>0.746</td>
<td>N/A</td>
<td>0.052</td>
</tr>
<tr>
<td>34</td>
<td>0.621</td>
<td>N/A</td>
<td>0.674</td>
<td>N/A</td>
<td>0.053</td>
</tr>
</tbody>
</table>

Please note in this coarse estimation the $I_{UCAR}^{th} - I_{UCUC}^{th}$ and $I_{HRAR}^{th} - I_{UCUC}^{th}$ are obtained from different chips and the $R_{UC}$ and $R_{HR}$ are assumed as the constants. An error propagation (see Appendix D) shows that it is reasonable to have such approximation and the variations of these values have little influence on the $R_{AR}$.

Table 6.2 shows another set of threshold current measured for the chips with 4 µm waveguide width. With the similar calculation one can derive that the $R_{AR}$ falls in the region of 3.69% - 3.96% which is smaller than the one estimated based on the data from Table 6.1. For this bar we did not have the witness sample to verify the actual AR coating reflectivity. However, the results are consistent with the previous chips and the method developed here (although at very early stage) can give a simple and coarse estimation on the range of the AR coating reflectivity. It can be clearly seen from the calculations that less variation of the ratio $\gamma$ results in higher accuracy of the estimation result.

Compared to the conventional methods that require each laser to be separately mounted and analyzed at each fabrication step (i.e. after HR and AR coating) \cite{181,136,111}. This new method determines the coating quality by measuring threshold
current from several similar laser chips at different coating stages on a laser bar, allowing a significant improvement on the estimation of the coating uniformity.

6.4 Chapter summary

Characterization and optimization of the EC-QCL gain media parameters, such as characterization of the tuning range, electrical properties and AR coating of the laser, play one of the most important roles in the EC-QCL design and manufacturing. In this chapter we demonstrated an automatic EC laser chip testing and characterization system that is capable of performing fully automatic characterization of the spectral and electrical parameters of the EC-QCL chips. The system developed provides a direct and efficient way of selecting the best gain chips in EC configurations without the need for dicing the laser bar into chips and custom mounting for EC-QCL operation. As a result, the system greatly reduces the labor and time cost required to select good QC chips for EC-QCL applications. Several different QCL gain chips developed and fabricated by Corning have been characterized using the system. A statistical analysis of the fabrication process which provides useful information to the active gain region and coating designers/growers has been developed and demonstrated. This system provides a direct and efficient way of selecting the best gain chips which will be used in the future spectroscopy systems. Moreover it is applicable to any EC laser test as long as the optical components of the EC are optimized for a specific wavelength.
Chapter 7

Multi-Heterodyne Spectroscopy based on Fabry-Perot Quantum Cascade Lasers

Many trace-gas sensing applications required for environmental monitoring, medical diagnostics or homeland security would benefit from broadband spectroscopy techniques which can be implemented using a simple, robust, compact and inexpensive optical system. In this chapter a new method of performing broadband mid-IR spectroscopy with conventional, free-running, CW FP-QCLs is demonstrated. The measurement technique is inspired by the elegant dual-comb spectroscopy approach [187, 188, 189, 190], in which the optical spectrum probed by one comb source is down-converted to the RF domain through an optical multi-heterodyne process using a second frequency comb as the reference LO. Both a multi-mode spectral retrieval as well as a high-resolution (~15 MHz) spectral scan capability show a great potential for development of high performance FP lasers based spectrometers for chemical sensing.
7.1 Background and motivation

Despite the wide tunability EC-QCLs are rather complex opto-mechanical systems that are vibration-sensitive, which makes construction of robust transportable systems difficult. A few attempts to address this issue have been made by several research groups constructing fully integrated arrays of DFB-QCLs covering broad spectral range of frequencies and containing no moving parts \[191, 192\]. An interesting alternative to QCLs are optical frequency combs that combine broad spectral coverage with high spectral resolution and unlike the DFB-QCL arrays deliver their multi-mode radiation in a single laser beam \[187, 193\]. Several mid-IR frequency comb sources have been reported, but due to the lack of appropriate materials most of them rely on indirect generation schemes based on nonlinear frequency conversion from near-IR \[194\] or on four-wave mixing process by pumping a micro-resonator \[195\]. These approaches suffer from some fundamental limitations such as low output efficiency, restricted wavelength availability, or their need for complex, expensive near-IR mode-locked pump lasers. Recently an impressive progress towards mode-locked frequency combs based on electrically pumped QCL chips has been made \[196, 197, 126\]. Hugi et. al. \[126\] reported a phase-locking mechanism between modes in a free-running CW FP-QCL as an effect of four-wave-mixing process in the fast QCL gain medium. Some evidence of active mode locking has also been observed in QCLs with RF modulation of the injection current tuned to the round trip frequency \[126\]. Despite these recent findings, ultra-short mode-locked pulsed operation in QCLs has not been yet fully demonstrated. However it should be noted that spectroscopic measurements commonly required for gas phase samples do not usually require the high precision provided by ultra-stable frequency comb sources (<kHz comb mode line-width). In such applications spectral resolution at the level of ~0.001 cm\(^{-1}\) (or 30 MHz) is in most cases sufficient. With this in mind the phase-locking mechanism in QCLs
shows a real potential to provide all-solid-state, electrically-controlled, broadband, high-resolution mid-IR spectrometers.

### 7.2 Experimental set-up

The configuration of the experimental set-up of the FP-QCL multi-heterodyne spectrometer is shown in Fig. 7.1. A copper sub-mount, which contained two 1.23 mm long QCLs with different ridge widths (3 µm and 5 µm) cleaved from the same wafer, was mounted inside evacuated laser housing. The QCLs operating at 8.4 µm developed by Diehl et. al. [198] (used previously in Ref. [111]) employed a conventional double-phonon resonance design of the active region. The laser housing is equipped with a thermoelectric cooler (TEC) controlling the operating temperature of both
lasers and allows providing separate bias current to each of the two laser ridges. A single anti-reflection coated (3-12 \( \mu \text{m} \)) \( f/0.6 \) germanium aspheric lens with 24 mm-diameter (Optical Solutions, model: 4682) is used to collimate both beams. The QCLs are separated by 0.5 mm, which results in slightly different propagation direction of the collimated beams. The beams are spatially separated and one beam is used as the reference local oscillator beam (referred to as the LO beam) and the other is used to probe the sample (referred to as the sample). The two beams are recombined on a ZnSe wedge used as a beamsplitter (BS) and focused onto a high-speed mercury cadmium telluride (MCT) photodetector integrated with a \( \sim1 \text{ GHz} \) bandwidth pre-amplifier (VIGO system S.A. detector model: PVI-3TE-10.6; pre-amplifier model: VPAC-1000F). The output signal is then analyzed by RF spectrum analyzers (Tektronics RSA6106A and/or RSA6120). To perform spectrum analysis in the optical domain the beams can be re-directed into an FTIR spectrometer (Nicolet 8700) using a flip mirror placed after the beam splitter (not shown in the figure). The bias currents of the two chips are separately controlled by two low noise current supplies (Wavelength Electronics QCL500 and QCL1000).

The principle of the multi-heterodyne spectroscopy based on two FP-QCLs is schematically shown in Fig. 7.2. The FSR of the signal QCL (FSR\(_1\)) is slightly different from that of the LO-QCL (FSR\(_2\)). As shown in Fig. 7.2 the heterodyne beat between the two combs of FP modes produces a series of beat notes in the RF domain. Due to difference in FSRs the frequency of each beat note is shifted by a \( \Delta f = \text{FSR}_2 - \text{FSR}_1 \). With this Vernier-like effect every heterodyne signal observed in the RF domain can be clearly assigned to an optical signal corresponding to a particular FP mode of the signal QCL. All heterodyne beat notes carry information about amplitude and phase of the optical waves corresponding to specific FP modes. This opens a way to variety of spectroscopic techniques including absorption [54] and dispersion [199] spectroscopy methods that could be performed with this FP-QCL.
Figure 7.2: A schematic operation principle of the multi-heterodyne spectroscopy with FP-QCLs. A series of heterodyne beat notes is produced by two FP-QCLs with slightly different FSR (the optical and the RF comb structures are not to scale).

multi-heterodyne technique. In practical implementation, the number of beat notes that can be simultaneously observed on a photodetector is limited by the photodetector bandwidth, by \( \Delta f \), and by the line-width of each FP mode.

Despite the same physical length of the QCL chips and the same gain material, the two FP lasers used in this experiment had different ridge widths, different turn-on voltages, and different threshold currents (affecting the internal chip temperature), which resulted in slightly shifted mode generation profiles and different FSRs. The spectra of the two QCLs measured by the FTIR are shown in Fig. 7.3a. The spectral overlap of the broadband FP mode structures generated by the two lasers is marked with a green box in Fig. 7.3a inset. The approximate mode spacing for both lasers is equal to \( \sim 1.25 \text{ cm}^{-1} \), and the actual difference in FSRs (\( \Delta f \)) cannot be clearly resolved with the FTIR. In order to generate beat notes within the bandwidth of the photodetector (1 GHz) the bias currents of the two QCLs are carefully adjusted so that their modes observed with the FTIR are nearly overlapped. As shown in Fig. 7.3b, four strong heterodyne signals can be clearly observed in the RF domain by
Figure 7.3: a, FTIR spectra of the signal- and the LO-QCL output radiation are shown in blue and red respectively. The spectral overlap is indicated by the green box in the inset. b, An RF spectrum of the heterodyne signal acquired by an RF spectrum analyzer.
mixing the signal and LO radiation with the optical spectra shown in Fig. 7.3a. The beat-notes are spaced by $\Delta f \sim 230$ MHz, which indicates the difference in the FSRs of the two FP-QCLs used in this experiment. The beat note below 200 MHz was also visible, but had large interferences from laser excess noise with $1/f$ characteristics; therefore it is not shown and has not been used in this work. Another weak beat signal at $\sim 1275$ MHz was outside the photodetector bandwidth, and has not been analyzed here. Nonetheless it clearly indicates that faster photodetectors (e.g. QWIP [200, 201]) or smaller $\Delta f$ could be used to increase the number of simultaneously measured beat notes in the RF domain, and thus increase the instantaneous optical bandwidth provided by the instrument. The QCL FSR can be changed within rather limited range using temperature- or current-tuning. The absolute frequency of the modes can be tuned with -612.8 MHz/mA rate for the signal QCL and with -690.9 MHz/mA rate for the LO QCL. Provided the operation frequency of $\sim 35.4$ THz and an approximate FSR of 37.5 GHz (1.25 cm$^{-1}$) this frequency tuning rates correspond to only $\sim-649$ kHz/mA tuning rate for FSR$_1$ and $\sim-732$ kHz/mA tuning rate for FSR$_2$. A temperature tuning coefficient of the mode frequencies of $\sim-3.38$ GHz/$^{\circ}$C was observed from FTIR measurements. This corresponds to FSR tuning rate of $\sim-3.58$ MHz/$^{\circ}$C. The actual difference in thermal tuning rates for both QCLs was estimated from observation of RF beat note shift and yielded the tuning rate for the signal QCL modes to be 296.6 MHz/$^{\circ}$C lower than that of the LO-QCL. Based on these observations the FSR tuning rates are very small, thus a decrease in $\Delta f$ through current- or temperature-tuning would not be practical. Therefore in order to achieve larger optical bandwidth of a multi-heterodyne system based on FP-QCLs, a proper selection of FP-QCLs with smaller $\Delta f$ must be performed up-front. In this work a proof of concept experiments have been carried out using the four available beat notes shown in Fig. 7.3b, which are sufficient to demonstrate multi-mode and high resolution spectral capabilities of the system.
Figure 7.4: a, A single RF spectrum with 100kHz resolution bandwidth (RBW) indicated a full-width at half-maximum of the beat note to be $\sim 170$ kHz. b, Center frequencies recorded over time for two different beat notes at $\sim 397$ MHz (upper) and at $\sim 860$ MHz (lower). High correlation between the graphs indicates strong phase-lock between the different longitudinal modes in FP-QCLs. (RBW=100 kHz).
The RF spectrum in Fig. 7.3b was recorded with 1 min averaging, which yielded a beat note full width at half maximum (FWHM) of ∼15 MHz. Single frequency QCLs are known to have very narrow intrinsic line-width of several hundred hertz [202], but in case of FP-QCLs both the intensity as well as phase noise might be higher due to various internal mode-competition effects. By setting the RF spectrum analyzer acquisition bandwidth to 100 kHz and measuring a single RF scan we could estimate an instantaneous line-width of the beat note to be ∼170 kHz (see Fig. 7.4a). This suggests that the instantaneous line-width of the FP mode is much narrower. Therefore the results with 1 min averaging indicate that there is strong frequency noise in the free running FP-QCLs. With an assumption that both FP-QCLs have similar un-correlated phase noise, the FP mode line-width in a single laser can be estimated at ∼10.6 MHz level. Given the measured QCL current tuning slope of ∼650 MHz/mA, this would correspond to a variation in the bias current of ±8 µA. Since this is much greater than the current noise (∼0.6 µA_{rms}) specified for our QCL driver, we attribute these FP mode frequency fluctuations to be an intrinsic phase noise of the FP-QCLs rather than the technical noise.

In order to investigate if the phase noise of one FP mode is correlated with phase noise of another FP mode of the same laser we have performed simultaneous measurements of two different beat notes belonging to different laser FP modes. Fig. 7.4b shows time traces of the instantaneous center frequencies determined for the two heterodyne beat notes at ∼397 MHz and ∼860 MHz respectively and recorded continuously over 30 seconds. Both center frequencies shift randomly over time yielding 4.23 MHz standard deviation calculated for the 30 s window, however high correlation between the two traces is clearly noticeable. In fact, the frequency difference of these two peaks is always a constant of 463.4 MHz (2×FSR), which indicates that the frequencies of the two participating FP modes in each laser are phase-locked to each other with accuracy exceeding the resolution of the RF spectrum analyzer (100
kHz). Frequency locking of the longitudinal modes has been observed in near-infrared (near-IR) InGaAsP FP-lasers, and has been attributed to a four-wave mixing mechanism within the active region [203]. Recently Hugi et al. observed narrow line-width (8.8 Hz) of the intermode beat signal in a mid-IR FP-QCL with specially designed broadband gain structure providing low group-velocity dispersion. This has also been attributed to a phase-lock mechanism based on nonlinear parametric process of four-wave mixing [126]. Our measurement of a conventional FP-QCL presented in Fig. 7.4b also supports the conclusion that FP-QCL longitudinal modes are not lasing independently but exhibit a phase-locked operation. This is also supported by recent theoretical studies of QCL dynamics by Dikmelik and Khurgin [204]. Therefore it might be possible to stabilize the longitudinal modes of both QCLs to a fixed frequency reference (e.g. a molecular transition or a high fitness cavity mode), and their intrinsic phase-lock mechanism should allow broad-band IR spectroscopy with ultra-high spectral resolution limited only by their intrinsic narrow line-width [202]. However it should be noted that the average long-term line-width of 15 MHz (∼0.0005 cm⁻¹) recorded for the beat note between two free-running FP-QCLs can easily be used to perform broadband mid-IR spectroscopy with resolution of < 0.001 cm⁻¹, which is sufficient for a wide range of applications.

### 7.3 Multi-heterodyne spectroscopy

First we investigated the potential of performing broad-band mid-IR spectroscopy using all four beat notes accessible in the current prototype spectrometer. An FTIR spectrum shown in the upper panel of Fig. 7.5a depicts four modes of the signal QCL that produce four RF beat notes within the electrical bandwidth of the detector (see lower panel of Fig. 7.5a) by heterodyning with the LO-QCL modes. The modes cover the spectral range from 1174 cm⁻¹ to 1179 cm⁻¹. The RF beat notes at ∼530 MHz,
Figure 7.5: a, A comparison of FTIR spectrum of the relevant FP modes emitted by the signal QCL (upper) and an RF spectrum of heterodyne beat notes (lower) with (red) and without (black) NH$_3$ absorption. b, Sample transmission measured for 1% (blue dots) and 50% (green dots) NH$_3$ in N$_2$ mixture (1 atm pressure, 15 cm path-length) is compared with HITRAN simulations shown by the colored lines (error bars represent experimental uncertainty observed during the beat note RF power measurement).
∼760 MHz, ∼990 MHz, and ∼1220 MHz correspond to the signal laser FP modes at ∼1174.4 cm\(^{-1}\), ∼1175.6 cm\(^{-1}\), ∼1176.8 cm\(^{-1}\), and ∼1178.1 cm\(^{-1}\) respectively. When the 15 cm long sample gas cell in the signal arm of the spectrometer is filled with pure ammonia (NH\(_3\)) gas, the longitudinal mode close to 1176.8 cm\(^{-1}\) is fully absorbed as indicated by the red FTIR spectrum in Fig. 7.5a. The beat note at ∼990 MHz in RF domain, generated by the mode at 1176.8 cm\(^{-1}\), is also suppressed, which clearly enables measurement of optical spectrum down-converted to the RF domain. To demonstrate multi-heterodyne measurement of a broad absorption profile, two different gas mixtures of 1% NH\(_3\) in N\(_2\) (producing ∼0.5 cm\(^{-1}\) wide absorption feature) and 50% NH\(_3\) in N\(_2\) (producing >2 cm\(^{-1}\) wide absorption feature with up to ∼100% absorption) were used in the experiment. To calculate the transmission of the sample the power of the heterodyne beat notes were first measured with the gas cell evacuated and then with the gas mixture at pressure of 1 atm introduced into the cell. The ratio of the beat note power with NH\(_3\) absorption to the power without absorption was used to calculate the transmission of the NH\(_3\) mixture at wavelengths corresponding to the four longitudinal modes. The measurement results are shown in Fig. 7.5b in blue for 1% mixture and in green for 50% NH\(_3\) in N\(_2\) mixture. The measured transmission agrees quite well with the simulated transmission profile based on HITRAN database (shown by the black and red lines for 1% and 50% mixture respectively).

To explore high-resolution capability of the prototype spectrometer we have performed spectroscopy of a single R11 N\(_2\)O transition at 1178.328 cm\(^{-1}\) which belongs to the 2\(\nu_2\) N\(_2\)O ro-vibrational band and could be targeted within the range of spectrometer coverage. The temperature and current of both FP-QCLs were first adjusted such that one of the signal QCL modes coincides with the target N\(_2\)O transition, and one of the LO-QCL modes creates a beatnote within the photodetector bandwidth (in this case at ∼750 MHz). Then the frequency of the signal QCL mode was scanned
Figure 7.6: Measuring low pressure N$_2$O absorption line with high resolution. One of the FP modes of the signal QCL is current tuned across the N$_2$O transition around 1178.328 cm$^{-1}$. A corresponding beatnote is measured in the RF domain and the N$_2$O absorption is calculated based on measurements with and without the N$_2$O gas in the sample cell. The black dots show the absorption data acquired for 25% N$_2$O in N$_2$ gas mixture at 10 Torr (15 cm optical path); The blue dots show the absorption data acquired for pure N$_2$O gas at 40 Torr indicating a wider line-width due to collision broadening. HITRAN simulations of the N2O absorption at 10 Torr and 40 Torr are shown by the colored lines. (Error bars represent experimental uncertainty observed during the beatnote RF power measurement).
over the absorption line by varying the laser bias current from 149.4 mA to 149.0 mA. The generated beatnote was also tuned in frequency, and the power of the beatnote was recorded during the scan. If the signal QCL and the LO QCL are well isolated and the LO frequency is kept constant, the change in the beatnote frequency can be directly linked to the change in optical frequency of the FP mode. However in our prototype spectrometer the QCL chips were located on the same substrate and a current tuning of the signal laser caused a slight influence on the frequency of the LO-QCL modes. This cross-talk of 78 MHz/mA was first characterized using FTIR spectra and an appropriate correction was applied to the frequency values in Fig. 7.6. The absorption spectrum was calculated in the post-processing using data collected for an evacuated cell and for a cell containing the N$_2$O sample. The data points in Fig. 7.6 represent the measured results for 25% mixture of N$_2$O in N$_2$ at two different pressures (10 Torr and 40 Torr). The measured data agree well with the spectrum simulated for the same conditions using HITRAN database (shown by the colored traces).

7.4 Chapter summary

We have demonstrated a compact spectrometer based on multi-heterodyne measurement performed with two free running, CW mid-IR FP-QCLs emitting at $\sim 8.5 \mu$m. We have verified the existence of the phase-lock mechanism between longitudinal modes in the FP-QCLs by observing the instantaneous frequencies of two heterodyne beatnotes generated by different FP modes. The heterodyne signals in RF domain show FWHM of $\sim 15$ MHz, which provides optical frequency resolution that is much higher than that of conventional tabletop FTIR spectrometers. Broadband spectroscopy of wide absorption feature of NH$_3$ with spectral sampling interval determined by the FP mode spacing of $\sim 1.25$ cm$^{-1}$ was demonstrated using simultaneous
measurement of all four heterodyne beat notes available in this prototype system. High-resolution capability of the multi-heterodyne spectrometer was demonstrated by resolving a narrow absorption feature (∼100 MHz) of N₂O at 1178.328 cm⁻¹. Due to technical limitations only one heterodyne beatnote could be recorded in this operation mode that uses fine-tuning of optical frequency through the laser bias current. However technologically feasible improvements in the data acquisition and computation system should allow parallel processing of all available beatnotes, which should enable the capability of a high-resolution broadband spectroscopy using all the available heterodyne beatnotes. In the current set-up, due to high difference between QCL FSRs and due to limitations in the MCT photodetector bandwidth, only four beatnotes corresponding to a spectral coverage of ∼4 cm⁻¹ could be collected simultaneously. Wider spectrum coverage can be achieved with high speed photodetectors (e.g. QWIP with bandwidths up to 100 GHz [201]), or with two FP-QCL chips with smaller FSR difference. In summary, we demonstrated a robust dual-FP-QCL multi-heterodyne spectrometer configuration, that is all electrically driven, contains no moving parts (high opto-mechanical stability), uses intrinsically small foot-print semiconductor devices, and shows a great potential for development of simple and compact broadband spectrometers that are able to access the fundamental molecular absorption bands in the mid-IR with high resolution and speed.
Chapter 8

Conclusion

Spectroscopic gas sensing has drawn enormous interest since it has many applications including environmental research, combustion studies, industry process monitoring, medical diagnostics and homeland security. Mid-infrared spectroscopy has the potential to realize the most selective and sensitive gas detection because many gas molecules have their distinct and strong ro-vibrational transitions at this electrical magnetic field wavelength region (also known as molecular “fingerprint” region). Although traditional mid-infrared spectrometers such as the FTIRs have been extensively used for broadband spectral analysis of chemical species, the instrument’s large physical dimension, long data-processing times along with the limitations in spectral resolution due to the using of incoherent IR light source are the main shortcomings of this spectroscopic technology. The invention of room temperature, high power, continuous wave operation QCLs opens the door for the development of novel spectroscopic technologies with high resolution, fast access to the molecular transitions in the mid-infrared band. This thesis has focused, to a significant part, on the development of various mid-infrared spectroscopic technologies enabled by QCL sources.
A major goal of this dissertation was to develop the most powerful biogenic NO sensing instrumentation with the unique capability of simultaneous detection of all NO isotopes for medical diagnostics. I have developed new methods that can improve the sensitivity of the FRS technique that provides ultra-sensitive and selective detection of gas phase free radical species (e.g. NO, NO$_2$, O$_2$ and •OH radical etc.). Record performance that is close to the fundamental shot noise limit has been achieved. The theoretical studies of Faraday rotation angle based on the parameters such as molecular integrated line intensity and line-width, g-factors, magnetic field strength, sample pressure have been carried out. Based on this analysis the ultimate sensitivity (limited by fundamental shot-noise) to the polarization rotation for an FRS system has been derived and can be used as a figure of merit to compare the performance achieved by different FRS techniques.

The most common practice to improve spectroscopic system performance is to increase in the signal strength by utilizing multi-pass cells. A DFB-QCL based FRS system integrated with a compact Herriott multi-pass cell was developed and characterized. It was found that in multi-pass cell arrangement there is optimum number of passes providing best FRS sensitivity. By measuring the NO transition line $R(17/2)f$ at 1906.73 cm$^{-1}$, the system has achieved NO sensitivity of 3.2 ppbv/Hz$^{1/2}$ with noise equivalent polarization rotation angle of $3.9 \times 10^{-8}$ rad/Hz$^{1/2}$.

Suppression of the laser noise has been proposed as the most viable strategy to approach the ultimate shot-noise limited performance. A new technique that provides significant noise reduction by employing heterodyne detection of the FRS signal has been developed. H-FRS effectively shifts the FRS signal detection to RF range where the QCL noise is significantly suppressed. The developed system is cryogen-free and operates at 5.24 $\mu$m. The sensitivity to the polarization rotation was improved by almost two orders of magnitude as compared to the previous conventional AC-FRS techniques.
system for NO detection, exhibiting a total noise of only 3.7 times higher than the fundamental quantum shot noise.

With the H-FRS is usually limited by complex optical interferometer, the H-FRS sensing system is typically operated under laboratory conditions. To overcome this issue a new technique utilizing DM-FRS has been developed. With 45 cm optical path and magnetic field of 158 Gauss, the NO detection sensitivity (1σ) of 0.53 ppbv/Hz\(^{1/2}\) and 3.72 ppbv/Hz\(^{1/2}\) for \(^{15}\)NO and \(^{14}\)NO were achieved, respectively, which translates into noise-equivalent polarization rotation angle of 7.11 × 10\(^{-9}\) rad/Hz\(^{1/2}\) and total noise only 2 times higher than the quantum shot-noise. A prototype DM-FRS instrument was built as a transportable instrument and was used in clinical setting to perform the first, proof-of-principle studies of NO metabolism in human body. Additionally when using multiple laser sources modulated at different carrier frequencies this system can be further improved to enable simultaneous concentration monitoring of multiple paramagnetic species by encoding molecular concentration signal at different carrier frequencies. This makes DM-FRS a very promising technique for a next generation NO isotope spectrometers for applications in environmental and medical studies.

Development of compact, robust and inexpensive spectrometers with broadband, high resolution capability has many important applications such as monitoring multiple trace gas species whose absorption lines are far apart, or detecting large molecules with unresolvable absorption lines in a wide frequency range. The EC-QCL is a single mode mid-IR laser source that gives access to the entire gain spectrum of the laser and offers large tuning range achievable for a single semiconductor laser chip. Characterization and optimization of the EC-QCL gain media parameters, such as characterization of the tuning range, electrical properties and AR coating of the laser, play one of the most important roles in the EC-QCL design and manufacture. However, the low efficiency of the conventional EC-QCL chip manufacturing method limited the
production process. Here, an automatic EC laser chip testing and characterization system has been demonstrated. The system is capable of automatic characterization of the threshold current, emission spectrum and electrical parameters of the EC-QCL chips. The system developed provides a direct and efficient way of selecting the best gain chips in EC configurations without the need for lengthy and cost-ineffective chip mounting steps. The system allows for statistical analysis of the fabrication process and its uniformity, which provides useful information to the active gain region and coating designers, growers or to improve the wafer fabrication process.

To address the need for high-resolution broadband spectrometers a novel spectrometer based on multi-heterodyne spectroscopy of two free running, CW FP-QCLs has been proposed and developed. The phase-lock mechanism between longitudinal modes in the FP-QCLs allows for this new spectrometer operation. Broadband spectrum retrieval of NH$_3$ absorption feature from 1174 cm$^{-1}$ to 1179 cm$^{-1}$ using simultaneous measurement of all four heterodyne beat notes available in this prototype system was demonstrated. The system has a potential to cover a broader spectral range if high speed photodetectors (e.g. QWIP) or FP-QCL sources with smaller FSR-difference were used. High-resolution spectrum retrieval of a $\sim$100 MHz narrow absorption feature of N$_2$O at 1178.328 cm$^{-1}$ was demonstrated by using one heterodyne beat note through precise current-tuning of the laser frequency. The heterodyne beat note in RF domain shows an FWHM of $\sim$15 MHz, which results in an optical frequency resolution that is much higher than that of conventional tabletop FTIR spectrometers. With further technologically feasible improvements in the data acquisition and computation system, parallel processing of all available beat notes should be straight forward, which will enable high-resolution broadband spectroscopy applications. The all electrically driven and robust dual-FP-QCL multi-heterodyne spectrometer has shown a great potential for the development of broadband mid-IR spectrometers with insuperable resolution and speed in terms of its size and cost.
Appendix A

A bit more about Zeeman effect

Magnetic dipole moment for an orbital electron

The magnetic dipole moment $\mu_{lz}$ of an electron traveling in a circular orbit in a plane perpendicular to $z$-axis is related to its $z$-axis angular orbital momentum $L_z$ via

$$\mu_{lz} = \frac{e}{2m} L_z, \quad (A.1)$$

where $e$ and $m$ are the charge and static mass of an electron.

On the other hand, the magnetic dipole moment caused by the electron spin $S_z$ can be expressed by

$$\mu_{sz} = g_s \frac{e}{2m} S_z, \quad (A.2)$$

where $g_s \approx 2.0023$ (from quantum field theory). Therefore the $z$-component total magnetic dipole moment $\mu_{Jz}$ is related to the total $z$-component angular momentum $J_z$

$$\mu_{Jz} = g_J \frac{e}{2m} J_z, \quad (A.3)$$
where $g_J$ is called Lande’ g-factor, which is given by L-S coupling theory

$$g_J = 1 + \frac{J(J + 1) + S(S + 1) - L(L + 1)}{2J(J + 1)} \quad (A.4)$$

Please note if $J = L$ and $S = 0$, then $g_J = 1$, Eq. A.3 becomes Eq. A.1 if $J = S$ and $L = 0$, then $g_J = 2 = g_s$, Eq. A.3 becomes Eq. A.2 For the other values of $S$ and $L$, Eq. A.4 gives the intermediate value for $g_J$.

**Zeeman effect for hydrogen atom**

Zeeman effect occurs because the interaction of the magnetic dipole moment $\mu_{Jz}$ with the external magnetic field $B$ slightly shifts the energy levels of the atom by the amount

$$\Delta E = -\mu \cdot B \quad (A.5)$$

For simplicity let us assume that the magnetic field is applied on z-axis. Substituting Eq. A.3 into Eq. A.8 we get

$$\Delta E = -g_J \frac{e}{2m} J_z B \quad (A.6)$$

The z-component angular momentum $J_z$ has the quantized eigenvalue of $m_j \hbar$, where $m_j = -j, -j + 1, -j + 2, ..., j - 1, j$. Thus

$$\Delta E = -g_J \frac{e}{2m} m_j \hbar B = -g_J \mu_B m_j B \quad (A.7)$$

where constant $\mu_B = \frac{e \hbar}{2m} = 5.78 \times 10^{-5} eV/T$ is called Bohr magneton. When an electron makes transition from an excited state to the ground state, it emits light with energy $\Delta h \nu$ equals to the difference between the two states. The electron transitions are determined by selection rules $p = m'_j - m''_j = 0, \pm 1$. If the transition is between
the levels with the same $m_j (p = 0)$ then the photon energy keeps unshifted. If the transition corresponds to $p = \pm 1$ then the photon energy is shifted by

$$\Delta h\nu = (g'_j m'_j - g''_j m''_j)\mu_B B$$ \hspace{1cm} (A.8)

where prime refers to the initial state and double-prime refers to the ground state. This is exactly the Eq. 2.5 shown in the Chapter 2. For total spin $S = 0$ states where the spin does not contribute to the total angular momentum, Landé’ g-factor $g_J = 1$ (see Eq. A.4) and Eq. A.8 reduces to $\Delta h\nu = 0, \pm \mu_B B$. This is called “normal Zeeman effect” as only three split transitions can be observed when the magnetic field is on. When electron spin is included, there is a greater variety of splitting patterns because energy splitting $\Delta h\nu$ is different for different Landé’ g-factors.

For example in the presence of spin-orbit interaction the hydrogen atom has Lyman-alpha transitions $2P_{1/2} \rightarrow 1S_{1/2}$ and $2P_{3/2} \rightarrow 1S_{1/2}$. When the magnetic field is on, the ground state $1S_{1/2}$ splits into two energy states (with $m_j = 1/2$ and $-1/2$), the excited state $2P_{1/2}$ splits into two states (with $m_j = 1/2$ and $-1/2$) and $2P_{3/2}$ splits into four states (with $m_j = 3/2, 1/2, -1/2$ and $-3/2$) as shown by Fig. A.1.
The Lande’ g-factor for the three levels are (calculated by Eq. A.4): $g_{J} = 2$ for $1S_{1/2}(J = 1/2, S = 1/2, L = 0)$; $g_{J} = 2/3$ for $2P_{1/2}(J = 1/2, S = -1/2, L = 1)$ and $g_{J} = 4/3$ for $2P_{3/2}(J = 3/2, S = 1/2, L = 1)$. The arrows on the right show the transitions allowed by the selection rules. Using Eq. A.8 one can get there are 10 different splitting levels with energy shift $\Delta h\nu = \pm 1/3\mu_{B}B, \pm 2/3\mu_{B}B, \pm 1\mu_{B}B, \pm 4/3\mu_{B}B$ and $\pm 5/3\mu_{B}B$. This is so-called “anomalous Zeeman effect”.
Appendix B

H-FRS implementation details

B.1 H-FRS customized photodetector

The photodetector for H-FRS experiment is a specially customized MCT photodetector with high speed, high saturation power especially suitable for heterodyne detection at peak response wavelength \(\sim 5.3 \, \mu \text{m}\). The whole detector module includes a low NEP, two stage TE cooled, photovoltaic MCT photodetector unit, followed by a DC coupled amplifier as the front-end, and an AC coupled amplifier as the second amplification stage. The area of the IR active element is only \(0.05 \times 0.05 \, \text{mm}^2\). This helps to reduce the dark noise and parasited capacitance so that the frequency response improves. Optical immersion of the IR active element to high refraction microlenses was used to improve the collection efficiency for non-coherent beam (e.g. solar radiation). However, since both of the signal and local oscillator (LO) are spatial coherent laser beam, we selected non-immersed lens device in order to obtain a relatively uniform optical intensity on the IR active surface which helps to avoid saturation of the spot area if an immersion lens were used. Although AC coupled pre-amp could be useful to suppress large pedestal due to the dark and LO generated current, DC coupling provides better operation conditions for detector, especially in our case, DC
coupling is necessary for adding user adjustable bias voltage to the detector to provide best frequency response. So the pre-amplifier for this application was configured that the first stage is a DC coupled amplifier with relatively low transimpedance for large dynamic range, tunable detector bias voltage for the best frequency performance, and tunable photocurrent compensation for reducing dark and LO-generated DC current offset. A DC monitor output after the first pre-amp is used for measuring the optical power arriving on the detector and adjusting photocurrent compensation when user wants to increase the dynamic range. At the second stage, an AC coupled amplifier with fixed gain (\( \sim 9.5 \times \)) is used to further enhance the gain. It is also equipped with output offset adjustment to cancel the DC voltage and avoid saturation of the second amplifier. Test indicates the detector has high cut-off frequency of \( \sim 59 \text{ MHz} \), NEP \( \sim 1 \times 10^{-11} W/Hz^{1/2} \) (at \( \lambda = 5.2 \mu \text{m} \) and \( f = 30 \text{MHz} \)). A high dynamic range with saturation power up to \( \sim 1.2 \text{ mW} \) was measured for this photodetector.

\section*{B.2 H-FRS system photo}

Fig. B.1 shows the actual optical setup photo for the H-FRS system.
Figure B.1: A photo showing the H-FRS setup.
Appendix C

DM-FRS implementation details

C.1 DM-FRS signal spectrum calculation

%% This Matlab program calculate the different harmonic signal spectrum based on
%% Fourier expansion. Ref. [171]
clear all

close all

load VoigtTemp.dat; % load AC-FRS spectrum normalized to peak. FWHM~42
v = [1 : length(VoigtTemp)]; % x axis
theta = linspace(0,pi,100); % theta integral variable
d_theta = theta(2) - theta(1);
H = zeros(1 , length(VoigtTemp)); % buffer
w = 49;
R(w) = w/42; % Ratio of modulation depth/FWHM
x_index = floor(w*cos(theta)); % x index
for i = w+1 : length(VoigtTemp)-w-1
temp = VoigtTemp(i + x_index).*cos(2*theta)*d_theta; % integral function

H(i) = 2/(pi)*sum(temp); % integral
end

harm_max(w) = abs(min(H)); % get amplitude
figure(1)
hold on
plot(VoigtTemp,'b.');
plot(H,'r.');
figure(2)
plot(R,harm_max,'g.');

C.2 The Allan variance and the Allan deviation

Created for timekeeping, the Allan variance plot is an effective way to characterize
the relative frequency instability of oscillation systems (e.g. clocks, LC oscillators
and lasers etc.) on an averaging time \( \tau \). Mathematically, Allan variance is expressed
as \( \sigma_y^2(\tau) \). The square root of Allan variance, Allan deviation, is expressed by \( \sigma_y(\tau) \).
The Allan variance is defined by [205]:

\[
\sigma_y^2(\tau) = \frac{1}{2} \langle (\bar{y}_{n+1}^\tau - \bar{y}_n^\tau) \rangle, \tag{C.1}
\]

where \( \bar{y}_n^\tau \) is the \( n^{th} \) average of the measured frequency data over the averaging time \( \tau \). From Eq. [C.1] it can be concluded that Allan variance is a measurement of
the frequency drift between the consecutive averaging time intervals. By choosing
different averaging times, many Allan deviation values can be calculated from the
same data set [206]. Conventionally a log-log plot is made where the x-axis is the
averaging time \( (\tau) \) and the y-axis is the calculated relative drift \( (\sigma) \). For an oscillator with white frequency noise performance, the greater the averaging time \( \tau \), the less frequency difference or drift can be observed.

For the gas sensing systems (e.g. H-FRS, DM-FRS etc.), delivering reliable measured signal without significant drift in a long-term running is required. Therefore, instead of using frequency, sensor output signal was used to calculate Allan deviation. The Allan deviation plot (Fig. 5.9) is a powerful tool to diagnostic the baseline drift and background noise compositions of the DM-FRS system. A Matlab program published on the Internet Ref. [207] was used to calculate the standard Allan deviation of a time domain signal as shown by the upper panel of Fig. 5.9. The Matlab code below was used to perform data reading and “allan()” function call:

```matlab
clear all
data = load('file.dat'); % file.dat has the format of first two columns as the signal data acquired from Zurich Instrument lock-in amplifier, and the third column as the timing data
CH1 = data(:,1);
CH2 = data(:,2);
time = data(:,3);
array.freq = CH1-CH2;
array.rate = 1;
tau = 1:1:5000;
allan(array, tau);
```
C.3 Signal spectrum comparison between DM-FRS and conventional AC-FRS

Comparison of the spectrum acquired with conventional AC-FRS and 2\textsuperscript{nd} harmonic of DM-FRS is shown by Fig. C.1. Both spectra were acquired under the same experimental conditions except optimum analyzer angle $\theta_{\text{opt}}$=1.5° in (a) and $\theta_{\text{opt}}$=3.4° in (b). Experimental conditions are: sample gas 2 ppmv NO in N\textsubscript{2} mixture; Since a different DFB-QCL was used, the target transition used for this comparison is NO R(17/2)f at 1906.73 cm\textsuperscript{-1}, which is different from the one targeted by the sensor; QCL output power $P_0$=12 mW; magnetic field $B$=100 G; pressure = 30 Torr; effective path-length $L$=15 cm; $f_0$=3.26 kHz; $f_1$=50 kHz. By comparing the spectra shown in Fig. C.1 it is obvious that the DM-FRS’s SNR is improved by $\sim$3 times, and its
signal baseline is free of offset, allowing much better long-term performance without signal drift.

C.4 DM-FRS photos

A top view of the DM-FRS sensor optical set-up is shown in the Fig. C.2. Fig. C.3 shows the author is demonstrating how to use the DM-FRS sensor at Cleveland Clinic.
C.5 Purge system for the liquid sample measurement

The purge system for converting nitrate and nitrite to NO is shown in Fig. C.4 [208, 209]. The chemical conversion produces both NO isotopes converted from the metabolites (mainly nitrate and nitrite) in human urine and blood. The NO gas is removed from the reaction solution by scrubbing with helium carrier gas and is subsequently detected by the DM-FRS sensor.
Figure C.4: Equipment for converting nitrite and nitrate in liquid samples to nitric oxide [208].
Appendix D

Implementation details of EC-QCL studies

D.1 EC-QCL bar tester software interface

The user interface of the software to control the EC-QCL testing platform is shown in Fig. D.1.

Figure D.1: Software interface for EC-QCL testing platform.
D.2 Error propagation for AR coating formula

Assuming the refractive index $n$ has a variation region of $n = 3.1 \sim 3.3$, which results in a variation region of $26.23\% \sim 28.61\%$ for $R_{UC}$. By substituting $R_{HR} = 95\%$, and a typical ratio of $\gamma = (I_{HR AR}^{th} - I_{UC UC}^{th})/(I_{UC AR}^{th} - I_{UC UC}^{th}) = 0.3$ into Eq. 6.7, one can calculate that $R_{AR}$ falls in a region of $4.17\% \sim 5.15\%$. On the other hand, assuming $R_{HR}$ varies between $90\%$ and $100\%$. By substituting $R_{UC} = 27\%$ and the same threshold change ratio into Eq. 6.7, one can calculate that $R_{AR}$ falls in a region of $4.16\% \sim 4.83\%$. The influence of the $R_{UC}$ or $R_{HR}$ variations on the $R_{AR}$ does not exceed $1\%$, which is acceptable to our coarse estimation.
Bibliography


168


[175] N. M. Grob and R. A. Dweik, “Exhaled nitric oxide in asthma from diagnosis, to monitoring, to screening: Are we there yet?”


[204] Y. Dikmelik and J. Khurgin, “private communication, Department of Material Science and Engineering, Johns Hopkins University, 102 Maryland Hall, 3400 N. Charles Street, Baltimore, Maryland 21216.”

[205] This introduction is based on the information from wikipedia and Dr. allan’s website, online access:. [Online]. Available: http://en.wikipedia.org/wiki/Allan_variance


