Precision molecular spectroscopy in the near- and mid-infrared using frequency comb-based Fourier transform spectrometers

Adrian Hjältén

Department of Physics
Umeå University
2023
That’ll do pig,
that’ll do.
- Farmer Hoggett
Abstract

Absorption spectroscopy is a powerful scientific tool for non-invasive and remote sensing applications ranging from atmospheric monitoring to astrophysics. In spectroscopic detection schemes it is necessary to have spectral models for any molecular species to be detected or quantified. Such models are often based on spectroscopic measurements or at the very least require experimental validation. The experimental data need to be accurate in terms of absorption line positions and intensities, but should also cover as many absorption lines as possible, i.e. broadband measurements are highly desirable.

Fourier transform spectroscopy (FTS) based on optical frequency combs (OFCs) can supply laboratory data that meet these requirements. OFCs provide a broad optical bandwidth and high spectral brightness, and also revolutionized our ability to measure optical frequencies, which had a profound impact on the frequency accuracy of spectroscopic measurements. The combination of OFCs and FTS, using the recently developed sub-nominal resolution technique, allows for measuring broadband absorption spectra with very high resolution, and a frequency accuracy provided by the OFCs. The aim of the work in this thesis was to expand the application of sub-nominal OFC-FTS to provide the much needed high-accuracy data for validation and development of spectroscopic databases of molecules relevant for a wide range of sensing application.

We developed a spectrometer to target the strong molecular absorption bands in the mid-infrared using two OFC sources based on difference frequency generation (DFG) emitting in the 3 µm and 8 µm wavelength ranges. We measured the spectra of iodomethane, CH₃I, and dibromomethane, CH₂Br₂, around 3 µm, fitted Hamiltonian models to several bands using the PGOPHER software, and reported molecular constants. For CH₃I we improved on previous models, while for CH₂Br₂ we presented a new interpretation of the spectrum. We also reported the first assessments of line intensities of CH₃I performed using multispectral fitting. At 8 µm, we implemented OFC-FTS based on a fiber-based compact DFG OFC source and measured low pressure spectra of nitrous oxide, N₂O, methane, CH₄, and formaldehyde, H₂CO. After the frequency accuracy was confirmed by excellent agreement with an earlier accurate study of N₂O, we compiled extensive line lists for CH₄ and H₂CO containing hundreds of transition frequencies with a precision improved by one order of magnitude compared to previously available data, and also reported line intensities for most transitions. For CH₄ the new data were used to improve a global Hamiltonian model, while the H₂CO data were incorporated into an algorithm based on spectroscopic networks to yield better precision in predicted energy levels and transition frequencies.

We also further developed a recent implementation of double resonance (DR) spectroscopy where optical pumping by a continuous-wave laser was used to populate selected vibrational energy levels of CH₄ not populated at room temperature, and a near-infrared OFC probed sub-Doppler transitions from the pumped states. Such measurements are necessary to validate theoretical predictions of transitions between excited vibrational levels that are relevant for high-temperature environments such as the atmospheres of hot celestial objects. We reported an improved
measurement setup using a new pump laser, new enhancement cavity with an updated OFC-cavity locking scheme, and measured transitions between more highly excited rotational levels than was previously reported. The higher rotational excitation lead to a larger number of DR transitions, which could be readily detected in the broadband high-resolution OFC-FTS spectra. We retrieved parameters of 88 lines of which we could assign 79 to theoretically predicted transitions. We found systematic frequency discrepancies with the predictions, that had not been observed earlier for lower rotational levels.

These implementations of sub-nominal OFC-FTS thus provided highly accurate line lists and improved spectral models of absorption bands of several molecules in the universally important mid-infrared region, as well as the first detection of 88 transitions between excited vibrational states of CH₄ relevant for high-temperature environments. We demonstrated the high potential of these techniques for collecting large amounts of accurate spectroscopic data, that further the scope of applicability of molecular spectroscopy.

Fouriertransformsspektroskopi (FTS) baserad på optiska frekvenskammar (OFC) kan leverera laboratoriedata som möter dessa kriterier. Frekvenskammar tillhandahåller bred optisk bandvidd och hög spektral intensitet, och revolutionerade även våra möjligheter att mäta optiska frekvenser, vilket starkt påverkade spektroskopiska mätningars frekvensnoggrannhet. Kombinationen av frekvenskammar och FTS (OFC-FTS) tillsammans med den nyligen utvecklade sub-nominella upplösningstechniken tillåter bredbandsmätningar av absorptionsspektra med mycket hög upplösning, och en frekvensnoggrannhet given av frekvenskammen. Syftet med arbetena i denna avhandling var att vidga tillämpningarna av sub-nominell OFC-FTS för att erhålla den noggranna datan nödvändig för validering och vidareutveckling av spektroskopiska databaser över molekyler relevanta för en rad avkänningsstillämpningar.

Vi utvecklade en spektrometer för att mäta starka mellaninfraröda molekylär absorptionsband, med hjälp av två frekvenskammljuskällor baserade på differensfrekvensgeneration (DFG) vid våglängder kring 3 µm och 8 µm. Vi mätte spektra av jodmetan, CH₃I, dibrommetan, CH₂Br₂, kring 3 µm, anpassade Hamiltonianmodeller till flera band med hjälp av mjukvaran PGOPHER, och rapporterade molekylär parametrar. För CH₃I förbättrades vi tidigare modeller, medans vi för CH₂Br₂ presenterade en ny tolkning av spektrumet. Vi rapporterade också de första uppskattningarna av linjintensiteter av CH₃I erhållna genom multispektrumanpassningar. Vid 8 µm tillämpade vi OFC-FTS med en fiberbaserad kompakt DFG-frekvenskam och mätte lågtryckspektra av dikväveoxid, N₂O, metan, CH₄, och formaldehyd, H₂CO. Efter att frekvensnoggrannheten bekräftats av utmärkt överensstämmelse med en tidigare studie av N₂O, sammanställde vi omfattande linjelistor av CH₄ och H₂CO bestående av hundratals övergångsfrekvenser med en tiofald förbättring i precision jämfört med dittills tillgänglig data, och rapporterade även linjintensiteter för de flesta övergångar. För CH₄ användes den nya datan till att förbättra en global Hamiltoniananpassning, medans formaldehyddenatan matades in i en algoritm baserad på spektroskopiska nätverk, vilket ledde till bättre precision i beräknade energinivåer och övergångsfrekvenser.

Vi vidareutvecklade även en tidigare tillämpning av dubbelresonansspektroskopi, där optisk pumpning med en monokromatisk laser användes för att fylla utvalda energinivåer för CH₄ som är tomma vid rumstemperatur, och en närinfraröd frekvenskam mätte sub-Dopplerövergångar från de fyllda nivåerna. Sådana mätningar är nödvändiga för att validera teoretiska förutsägelser av övergångar mellan exzitierade vibrationsnivåer som är relevanta för miljöer med hög temperatur såsom

Dessa tillämpningar av sub-nominell OFC-FTS bidrog därmed med mycket noggranna linjelistor och förbättrade modeller av absorptionsband för flera molekyler i det allmänt viktiga mellaninfraröda området, tillsammans med den första detektering av 88 övergångar mellan exciterade vibrationsnivåer av CH$_3$$_4$ relevanta för miljöer med hög temperatur. Vi demonstrerade dessa teknikers höga potential för insamling av omfångsrik och noggrann spektroskopisk data, som vidgar tillämpningsmöjligheterna av molekylspektroskopi.
Abbreviations

BS beam splitter
CRDS cavity ring-down spectroscopy
CW continuous wave
DCS dual-comb spectroscopy
DDS direct digital synthesizer
DFG difference frequency generation
DR double resonance
FFT fast Fourier transform
FSR free spectral range
FWHM full width at half maximum
FTIR Fourier transform infrared spectroscopy
FTS Fourier transform spectrometer / spectroscopy
GS ground state
HNLF highly non-linear fiber
HVOC halogenated volatile organic compound
HWHM half width at half maximum
IDFG intra-pulse DFG
IR infrared
MIR mid-infrared
MS-SF microstructured silica fiber
NIR near-infrared
OFC optical frequency comb
OFC-FTS optical frequency comb Fourier transform spectroscopy
OODR optical-optical double-resonance spectroscopy
OPD optical path difference
OP-GaP orientation patterned gallium phosphide
OPO optical parameteric oscillator
PD  photo-diode
PDH  Pound-Drever-Hall
PPLN  periodically poled lithium niobate
PZT  piezo-electric transducer
RF  radio frequency
QCL  quantum cascade laser
SN  spectroscopic network
SNR  signal-to-noise ratio
UV  ultraviolet
VIPA  virtually imaged phase array
## Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_{12}$</td>
<td>Einstein coefficient of absorption</td>
</tr>
<tr>
<td>$B_{21}$</td>
<td>Einstein coefficient of stimulated emission</td>
</tr>
<tr>
<td>$E$</td>
<td>energy</td>
</tr>
<tr>
<td>$E$</td>
<td>electric field of electromagnetic wave</td>
</tr>
<tr>
<td>$E_0$</td>
<td>electric field amplitude</td>
</tr>
<tr>
<td>$F$</td>
<td>cavity finesse</td>
</tr>
<tr>
<td>$\hat{H}_N$</td>
<td>Hamiltonian operator of nuclear motion</td>
</tr>
<tr>
<td>$I_B$</td>
<td>intensity transmitted through empty sample container</td>
</tr>
<tr>
<td>$I_0$</td>
<td>incident intensity</td>
</tr>
<tr>
<td>$I_T$</td>
<td>intensity transmitted through absorbing sample</td>
</tr>
<tr>
<td>$J$</td>
<td>rotational quantum number</td>
</tr>
<tr>
<td>$L$</td>
<td>interaction length with sample</td>
</tr>
<tr>
<td>$L_c$</td>
<td>enhancement cavity length</td>
</tr>
<tr>
<td>$N$</td>
<td>molecular number density</td>
</tr>
<tr>
<td>$N^*_l$</td>
<td>ideal fractional number of sampling points of interferogram</td>
</tr>
<tr>
<td>$N_l$</td>
<td>number of sampling points of interferogram</td>
</tr>
<tr>
<td>$N_{steps}$</td>
<td>number of repetition rate steps</td>
</tr>
<tr>
<td>$P$</td>
<td>pressure</td>
</tr>
<tr>
<td>$S$</td>
<td>integrated molecular line intensity</td>
</tr>
<tr>
<td>$T$</td>
<td>temperature</td>
</tr>
<tr>
<td>$\tilde{T}$</td>
<td>complex electric field transmission function</td>
</tr>
<tr>
<td>$T_c$</td>
<td>cavity intensity transmission function</td>
</tr>
<tr>
<td>$Z$</td>
<td>partition function</td>
</tr>
<tr>
<td>$c$</td>
<td>speed of light in vacuum</td>
</tr>
<tr>
<td>$c_{rel}$</td>
<td>relative concentration of absorbing species</td>
</tr>
<tr>
<td>$f_{ceo}$</td>
<td>carrier envelope offset frequency</td>
</tr>
<tr>
<td>$f_{rep}$</td>
<td>repetition rate</td>
</tr>
</tbody>
</table>
\( g_i \)  statistical weight of energy level \( i \)
\( h \)  Planck’s constant
\( k_B \)  Boltzmann’s constant
\( m \)  molecular mass
\( n \)  index of comb modes and cavity modes
\( n_0 \)  Loschmidt number
\( n_b \)  background refractive index
\( p \)  pressure broadening coefficient
\( q \)  number of sampling points per reference laser wavelength
\( r \)  mirror reflection coefficient
\( t \)  mirror transmission coefficient
\( \tilde{w} \)  Faddeeva error function
\( \alpha \)  absorption coefficient
\( \Gamma_D \)  Doppler half-width at half maximum
\( \Gamma_L \)  Lorentzian half-width at half maximum
\( \Gamma_c \)  cavity mode full-width at half maximum
\( \gamma_{\text{air}} \)  air pressure broadening coefficient
\( \gamma_{\text{self}} \)  self pressure broadening coefficient
\( \Delta l \)  optical path difference
\( \Delta l_{\text{max}} \)  optical path difference scanning range
\( \Delta \Phi \)  comb-cavity phase offset
\( \delta l \)  optical path difference domain point spacing
\( \delta \nu \)  frequency domain point spacing
\( \delta_{\text{air}} \)  air pressure shift coefficient
\( \epsilon \)  fractional error in \( N_l \)
\( \epsilon_0 \)  vacuum permittivity
\( \lambda \)  wavelength
\( \lambda'_{\text{ref}} \)  assumed effective reference laser wavelength
\( \lambda_{ref} \) true effective reference laser wavelength
\( \bar{\nu} \) wavenumber
\( \bar{\nu}_c \) line center wavenumber
\( \nu \) frequency
\( \nu_{FTS} \) frequency scale of FTS spectrum
\( \nu_{OFC} \) frequency scale of comb modes
\( \sigma \) absorption cross section
\( \Phi_m \) mirror dispersion phase shift
\( \phi \) dispersion coefficient
\( \chi_{abs}^{D} \) Gaussian absorption line shape
\( \chi_{abs}^{L} \) Lorentzian absorption line shape
\( \chi_{abs}^{V} \) Voigt absorption line shape
\( \chi_{disp}^{D} \) Gaussian dispersion line shape
\( \chi_{disp}^{L} \) Lorentzian dispersion line shape
\( \chi_{disp}^{V} \) Voigt dispersion line shape
Publications

Paper I
Line positions and intensities of the $\nu_4$ band of methyl iodide using mid-infrared optical frequency comb Fourier transform spectroscopy
Ibrahim Sadiek, Adrian Hjältén, Francisco Senna Vieira, Chuang Lu, Michael Stuhr, Aleksandra Foltynowicz.

Paper II
Optical frequency comb Fourier transform spectroscopy of $^{14}\text{N}_2^{16}\text{O}$ at 7.8 $\mu$m
Adrian Hjältén, Matthias Germann, Karol Krzempek, Arkadiusz Hudzikowski, Aleksander Gluszek, Dorota Tomaszewska, Grzegorz Soboń, Aleksandra Foltynowicz.

Paper III
A methane line list with sub-MHz accuracy in the 1250 to 1380 cm$^{-1}$ range from optical frequency comb Fourier transform spectroscopy
Journal of Quantitative Spectroscopy & Radiative Transfer, 288, 108252 (2022)

Paper IV
Optical frequency comb-based measurements and the revisited assignment of high-resolution spectra of CH$_2$Br$_2$ in the 2960 to 3120 cm$^{-1}$ region
Ibrahim Sadiek, Adrian Hjältén, Frances C. Roberts, Julia H. Lehman, Aleksandra Foltynowicz.
Physical Chemistry Chemical Physics, 25, 8743–8754 (2023)

Paper V
Line positions and intensities of the $\nu_1$ band of $^{12}\text{CH}_3\text{I}$ using mid-infrared optical frequency comb Fourier transform spectroscopy
Adrian Hjältén, Aleksandra Foltynowicz, Ibrahim Sadiek.
Paper VI
Optical frequency comb Fourier transform spectroscopy of formaldehyde in the 1250 to 1390 cm$^{-1}$ range: experimental line list and improved MARVEL analysis
Journal of Quantitative Spectroscopy & Radiative Transfer, 312, 108782 (2024)

Paper VII
Accurate measurement and assignment of high rotational energy levels in the 9150 - 9370 cm$^{-1}$ range of methane using optical frequency comb double-resonance spectroscopy
Adrian Hjältén, Vinicius Silva de Oliveira, Andrea Rosina, Isak Silander, Lucile Rutkowski, Grzegorz Soboń, Kevin K. Lehmann, Aleksandra Foltynowicz.
In manuscript.
Contents

1 Introduction

2 Absorption and dispersion of light 7
   2.1 Rotational-vibrational transitions 7
   2.2 The thermal population distribution 9
   2.3 The Lambert-Beer law 10
   2.4 Line shapes and broadening mechanisms 12
      2.4.1 The Gaussian line profile - Doppler broadening 13
      2.4.2 Lorentzian line profiles 14
      2.4.3 The Voigt profile 15

3 Absorption spectroscopy, methods and concepts 17
   3.1 Enhancing the absorption signal 17
      3.1.1 Multipass cells 18
      3.1.2 The Fabry-Pérot enhancement cavity 18
   3.2 Sub-Doppler techniques and optical pumping 20
      3.2.1 Optical-optical double-resonance spectroscopy 21
   3.3 Broadband absorption spectra 23

4 Spectroscopic databases and models 25
   4.1 The HITRAN database 25
   4.2 Theoretical databases 26
   4.3 The MARVEL algorithm 26

5 Optical frequency combs 29
   5.1 Different types of OFCs 29
   5.2 Mode-locked OFCs in the time and frequency domains 30
   5.3 Non-linear frequency conversion 31
      5.3.1 Principles of non-linear conversion 31
      5.3.2 Difference frequency generation with OFCs 32
   5.4 Locking an OFC to an enhancement cavity 34

6 Optical frequency comb Fourier transform
   spectroscopy with sub-nominal resolution 37
   6.1 The Michelson interferometer 37
   6.2 Fourier transform spectroscopy with OFCs 39
   6.3 The sub-nominal resolution technique 41
      6.3.1 The instrumental line shape of an OFC-FTS spectrum 41
      6.3.2 Matching the OFC and FTS frequency scales 41
      6.3.3 Spectral interleaving 44
      6.3.4 Reference wavelength calibration 46
      6.3.5 Correcting for non-linear mapping of OFC frequencies 48
   6.4 Balanced and auto-balanced detection 49
7 MIR high-precision optical frequency comb
Fourier transform spectroscopy

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Experimental setup</td>
<td>52</td>
</tr>
<tr>
<td>7.1.1</td>
<td>The MIR FTS</td>
<td>53</td>
</tr>
<tr>
<td>7.1.2</td>
<td>3 µm OFC source</td>
<td>53</td>
</tr>
<tr>
<td>7.1.3</td>
<td>8 µm OFC source</td>
<td>54</td>
</tr>
<tr>
<td>7.2</td>
<td>Measurements and analysis methods</td>
<td>55</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Spectral acquisition and treatment</td>
<td>55</td>
</tr>
<tr>
<td>7.2.2</td>
<td>The PGOPHER software</td>
<td>56</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Line fitting</td>
<td>56</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Multispectral fitting</td>
<td>57</td>
</tr>
<tr>
<td>7.3</td>
<td>High-precision spectroscopy at 3 µm</td>
<td>58</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Iodomethane CH₃I</td>
<td>58</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Dibromomethane CH₂Br₂</td>
<td>60</td>
</tr>
<tr>
<td>7.3.3</td>
<td>Conclusions</td>
<td>61</td>
</tr>
<tr>
<td>7.4</td>
<td>High-precision spectroscopy at 8 µm</td>
<td>62</td>
</tr>
<tr>
<td>7.4.1</td>
<td>Nitrous oxide N₂O</td>
<td>62</td>
</tr>
<tr>
<td>7.4.2</td>
<td>Methane CH₄</td>
<td>63</td>
</tr>
<tr>
<td>7.4.3</td>
<td>Formaldehyde H₂CO</td>
<td>66</td>
</tr>
<tr>
<td>7.4.4</td>
<td>Conclusions</td>
<td>68</td>
</tr>
</tbody>
</table>

8 Optical frequency comb optical-optical double-resonance spectroscopy

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Experimental setup</td>
<td>70</td>
</tr>
<tr>
<td>8.2</td>
<td>Measurements and analysis methods</td>
<td>72</td>
</tr>
<tr>
<td>8.2.1</td>
<td>Spectrum acquisition and treatment</td>
<td>72</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Line fitting</td>
<td>73</td>
</tr>
<tr>
<td>8.2.3</td>
<td>Combination differences</td>
<td>74</td>
</tr>
<tr>
<td>8.3</td>
<td>OFC OODR spectroscopy of CH₄</td>
<td>76</td>
</tr>
</tbody>
</table>

9 Conclusions

10 Summary of papers

11 Acknowledgements
1 Introduction

Spectroscopy is the study of interactions between electromagnetic radiation (often referred to simply as light from here on) and matter. During the last two centuries, it has evolved into one of the most powerful tools for detecting and investigating atoms and molecules, the building blocks of the world around us. In gas absorption spectroscopy, one measures the light transmitted through a gas sample as a function of optical frequency $\nu$. The photon energy $E$ depends on the frequency as $E = h\nu$ where $h$ is Planck’s constant, and when this energy coincides with the difference between the discrete energy levels of atoms or molecules in the gas, the photons can be absorbed through an optical transition. This leads to a decrease in the transmitted intensity at the corresponding frequency, which can be detected as an absorption line. The collection of absorption lines due to the transitions of an atomic or molecular species makes up its absorption spectrum. The appearance of an absorption spectrum is determined by the internal structure of the atoms or molecules. Absorption spectroscopy can therefore give information about atoms and molecules and the quantum mechanical processes governing their behavior, but can also provide quantitative data about what substances are present in a gas sample and in what amounts.

Molecules are of vast importance, both for life on Earth, but also for understanding conditions and processes elsewhere in the universe. Molecular absorption spectroscopy is used for studying the composition of the Earth’s atmosphere, e.g. the natural cycling of gas species relevant to the climate. It can be used to monitor the levels of harmful substances emitted from industry or leaking into the air at work places, but it can also provide information about distant celestial objects such as exo-planets, and is a crucial tool in e.g. the search for signs of extra-terrestrial life.

Molecules possess vibrational and rotational energy states, and transitions between these lie in the infrared (IR) and microwave regions, respectively. The IR spectrum is further divided into different sub-regions, and here we will distinguish between the near-infrared (NIR) and mid-infrared (MIR). The exact definitions of these regions vary somewhat, but roughly, the NIR is considered to cover wavelengths from 800 nm to 2500 nm, where the MIR begins, stretching up to wavelengths of 20 - 50 $\mu$m.

The collection of transitions between the rotational sub-levels of two different vibrational states gives rise to a vibrational absorption band. Absorption transitions of fundamental bands, occurring to the first excited vibrational level, appear in the MIR, and are typically much stronger than the NIR overtone and combination bands, and hence, the former region is particularly important for molecular spectroscopy applications. Another important factor for many terrestrial and ground-based astrophysical applications is the overall transparency of the atmosphere. Background absorption of atmospheric molecular species can be significant at some wavelengths, and severely interfere with spectroscopic detection schemes. Most of the background absorption is due to water vapor but also to trace gases such as CO$_2$. However, there are spectral regions called atmospheric windows, where the atmosphere is more transparent, and these are therefore of particular
interest. One such window occurs around 3 µm - 5 µm, and a wider one spans between 8 µm and 14 µm.

In order to extract information from measured spectra for various applications, accurate model spectra are required. Comparison of an absorption measurement to spectral models for various molecules allows for identifying the presence of molecular species and quantifying their amounts. Spectral models can be calculated from databases containing the parameters describing the absorption spectra, primarily the frequencies and intensities of individual absorption lines. Inaccurate line parameters hamper the analysis of absorption measurements and can lead to incorrect conclusions. Line parameters can be calculated theoretically based on certain previous knowledge and assumptions regarding the molecular structure. By comparison with laboratory measurements, the models can be optimized in an iterative process to improve their accuracy. This gradual refinement calls for increasingly more accurate experimental data, and hence, steadily advancing experimental techniques pave the way for new spectroscopic applications. Providing the necessary laboratory data for improving spectral models is the main focus of this thesis.

The fundamental laboratory setup for absorption spectroscopy is displayed in Figure 1.1. It consists of a light source, a gas sample and a detection system. The absorption signal increases with the interaction length of the light beam with the sample, and the light source can be monochromatic or broadband. Before lasers, typically thermal sources were used, which emitted broadband light.

![Figure 1.1: The concept of absorption spectroscopy. A light beam is passed through a gas sample and the transmission measured with a detection system.](image)

With a broadband light source, the detection system must incorporate a spectrometer that resolves the individual spectral components of the transmitted light. Spectrometers can be dispersive, where the frequencies of the light are spatially spread out before detection [1]. This can be achieved by a simple glass prism or by a diffraction grating. These were the first kinds of spectrometers which were used to identify absorption lines in sunlight in the 19th century, though more evolved forms of dispersive spectrometers are still in use today.

Another type of spectrometer is the Fourier transform spectrometer (FTS). It does not disperse the light spatially but relies on the interference between two phase-shifted beams from the same light source. In a mechanical FTS, the phase shift is obtained by physically varying the path length difference between the beams. The spectrum is obtained by taking the Fourier transform of the recorded interferogram. The use of FTS with thermal sources is known as Fourier transform infrared spectroscopy (FTIR) [2]. While FTIR measurements provide very broad optical bandwidth, the nominal resolution of the FTS is limited by the length
of the interferogram, and hence by how much the path length difference can be scanned. Very high resolution thus requires very bulky interferometers, which in practice limits the nominal resolution of FTSs to tens of MHz.

The invention of lasers [3], was an important breakthrough for spectroscopy. Lasers provide much higher spectral brightness than thermal sources and laser beams are easier to collimate and couple through the measurement setup, allowing for longer interaction lengths with the absorbing gas. The interaction length can be increased by passing the beam multiple times through the sample in a multipass cell, and narrowband lasers can also be effectively coupled into a resonant mode of an enhancement cavity. All of these factors serve to increase the measurement sensitivity. The high intensity of lasers also made optical pumping possible, where intense laser light is used to populate excited molecular energy levels not populated at thermal equilibrium conditions [1]. This allows for probing otherwise not detectable transitions from these excited levels. Monochromatic lasers, whose wavelength can be stabilized and tuned also eliminated the need for spectrometers. This yielded improved spectral resolution as the laser linewidth can be orders of magnitude narrower than the maximum resolution of spectrometers. However, the tunability of narrowband lasers typically only allows for measuring a few absorption lines, which does not compare to the spectral coverage of broadband sources which can cover entire absorption bands.

The next leap forward came with the invention of optical frequency combs (OFCs), which are lasers emitting a large number of equidistant narrow frequency modes that can be stabilized and known with high precision. The first application of OFCs was in frequency metrology [4, 5], but their potential in spectroscopy was also soon realized as they allowed for very accurate stabilization of monochromatic lasers [6]. In addition, OFCs combine the high brightness of lasers with the broad optical bandwidth of thermal sources and are hence a suitable substitute for the latter. Initial implementations of OFCs replacing the thermal sources in traditional FTIR allowed for broadband measurements with great reductions in acquisition time [7]. However, the spectral resolution was still the nominal resolution of the FTS. Furthermore, the truncation of the interferograms causes instrumental effects that can broaden or distort absorption lines. A way around these problems was provided by the sub-nominal resolution sampling interleaving technique [8, 9], in which the instrumental distortions are largely avoided by matching the spectral sampling points to the OFC mode frequencies. The comb mode grid is then scanned stepwise over the absorption features and by interleaving spectra acquired at each step one can reach resolutions far beyond the FTS nominal resolution, where the ultimate limit is given by the comb mode width [10], which can be on the kHz level.

This combination of high spectral resolution and absolute frequency accuracy allows for very accurate measurements of absorption line parameters, particularly transition frequencies. Spectroscopic database parameters for many molecules in the MIR region are still largely based on traditional FTIR measurements, which limits the accuracy of spectral models, and MIR data for some important molecules are still missing from major databases such as HITRAN [11]. Much work is required to improve and expand the scope of the databases. In this thesis, we aim
at contributing to this process by further developing the technique of optical frequency comb Fourier transform spectroscopy (OFC-FTS) with sub-nominal resolution, and measuring highly accurate broadband molecular spectra for validating and improving spectroscopic databases.

We firmly establish the technique in the MIR wavelength region, by down-converting NIR OFC light with difference frequency generation (DFG). We developed a MIR FTS spectrometer compatible with two OFC sources operating within the atmospheric windows at 3 µm and 8 µm. In Papers I, IV and V we measured broadband spectra of iodomethane, CH₃I, and dibromomethane, CH₂Br₂, using the 3 µm system, and performed spectral simulations and fitting to retrieve molecular parameters. We improved on previous spectral models for CH₃I and a part of this model was included in the 2020 update of the HITRAN database. For CH₂Br₂, we introduced a new analysis of the spectrum and reported the obtained model parameters. Around 8 µm, we first targeted nitrous oxide, N₂O, in Paper II which served to validate the performance of the spectrometer in a spectral region where not much work had been done with OFCs. In Papers III and VI, we measured multiple low pressure spectra of methane, CH₄, and formaldehyde, H₂CO, and retrieved hundreds of transition frequencies with much lower uncertainties than previously reported. We also used the experimental data to improve theoretical models.

Another suitable application of sub-nominal OFC-FTS is in double-resonance (DR) spectroscopy. In DR measurements, a strong pump laser is used to optically pump selected molecular energy levels. Using a second probe laser, one can then detect transitions from these levels, that would be difficult to observe at thermal equilibrium conditions. For the CH₄ molecule, extensive theoretical predictions of such transitions exist but due to the difficulty in detecting them, they are largely not experimentally validated. Optical pumping by a monochromatic laser makes the DR transitions very narrow and the broadband high-resolution measurements performed using an OFC probe laser and the sub-nominal FTS technique are excellent for the search for previously undetected transitions, as was demonstrated in [12–14]. In Paper VII we demonstrated an improved setup for DR OFC-FTS spectroscopy of CH₄. We pumped more highly excited rotational states than what was done previously, which increased the number of DR transitions observed from these states, thus making more full use of the wide optical bandwidth of the probe. 79 experimental absorption lines could be matched to theoretically predicted transitions, and a comparison with the predictions revealed systematic frequency discrepancies not observed for lower rotational states.

The overall structure of the thesis is as follows. Chapter 2 gives a theoretical background of the interaction of light with molecules, describing the processes giving rise to absorption spectra, and determining the intensity and shape of absorption lines. Chapter 3 then summarizes some concepts, spectroscopic methods and experimental equipment relevant for the thesis, while in Chapter 4, we review different types of spectroscopic databases and how their data are obtained. We also briefly describe the particular databases we used for modeling our spectra and comparisons to our experimental data. Chapter 5 provides an overview of OFCs in general and of mode-locked OFCs in particular, and we also discuss some meth-
ods for converting NIR OFCs to MIR wavelengths. Chapter 6 then thoroughly reviews the sub-nominal OFC-FTS measurement technique used throughout the thesis. Finally, the results of the MIR OFC-FTS spectroscopy, and the DR OFC-FTS spectroscopy of CH$_4$ are summarized in Chapters 7 and 8, respectively.
2 Absorption and dispersion of light

In this chapter, we briefly describe the origin of vibrational-rotational spectra of molecules, and the factors determining the frequencies and intensities of transitions. We introduce the formalism for describing the effect on the electric field and the intensity of light caused by a transition, and introduce some key quantities used later on in the thesis. Finally, we look at different mechanisms that broaden transitions, the resulting shapes of the observed absorption lines and how these can be modeled mathematically.

2.1 Rotational-vibrational transitions

The atomic nuclei of a molecule can oscillate around their equilibrium positions and the molecule as a whole can rotate in space. In the Born-Oppenheimer approximation [15], the vibrational and rotational motion can be separated from the motion of the electrons in the atoms due to the large difference in mass between electrons and nuclei, and consequently the different time-scales of their motion. The allowed discrete rotational-vibrational energy states are then solutions to the Schrödinger equation using the Hamiltonian operator of nuclear motion

$$\hat{H}_N \psi_i(r_\alpha) = E_i \psi_i(r_\alpha),$$  \hspace{1cm} (2.1)

where $r_\alpha$ denotes the nuclear coordinates, $\psi_i(r_\alpha)$ is the wavefunction of the $i^{th}$ eigenstate of $\hat{H}_N$ and $E_i$ is the corresponding eigenenergy [15]. Rather than $i$, the eigenfunctions and energy levels are typically labeled with several numbers that describe the vibrational and rotational state of the molecule. Figure 2.1(a) shows a graphical depiction of a rotational-vibrational energy level structure in the simplest case, e.g. for a linear molecule. At the bottom is the vibrational ground state (GS), and the three lowest excited states of a vibrational mode $\nu_1$ are denoted as $\nu_1$, $2\nu_1$ and $3\nu_1$. The separation between rotational levels is much smaller than between vibrational levels, and each vibrational level splits into rotational sub-levels labeled by the quantum number $J$, which denotes the total angular momentum of the molecules. Photons with an energy equal to the difference between two energy levels can be absorbed, thereby exciting the molecule to the higher energy state. Such transitions between rotational levels of the same vibrational state occur at microwave wavelengths, while transitions between vibrational states appear in the IR region. The set of transitions occurring between the rotational sub-levels of two particular vibrational states is referred to as an absorption band. Transitions from the vibrational ground-state to the first excited state form fundamental bands, while transitions to more highly excited states are called overtones. Examples of such transitions are shown in Figure 2.1(a) as the orange and violet arrows respectively. Transitions originating from excited vibrational states are called hot bands, and one such transition of the $3\nu_1 - \nu_1$ band is depicted by the red arrow. Hot bands become more prominent at higher temperature, due to increased population in excited states (Section 2.2).

Quantum mechanically, the transitions occur due to coupling between the initial and final states induced by the electromagnetic radiation. The most dominant
Figure 2.1: (a) Schematic of a rotational-vibrational energy level structure, showing the vibrational ground-state (GS) and the first three excited levels of a vibrational mode \( \nu_1 \). The vibrational levels are divided into rotational sub-levels labeled by the rotational quantum number \( J \). The vertical arrows indicate transitions of the \( \nu_1 \) fundamental band (orange), the \( 2\nu_1 \) overtone band (violet) and the \( 3\nu_1 - \nu_1 \) hot band (red). (b) The P(1), Q(1) and R(1) transitions of the overtone band, originating from \( J = 1 \) with \( \Delta J = -1, 0 \) and 1 respectively.

The mechanism is the electric dipole transition, where the strength of a transition between two states 1 and 2, (describing how likely it is to occur), is proportional to the square of the transition dipole moment

\[
\mu_{21}^2 = \langle \psi_2 | \mu | \psi_1 \rangle^2 ,
\]

where the electric dipole moment of the molecule, \( \mu \), is a function of the nuclear coordinates [15]. In particular, transitions only take place between states for which \( \mu_{21} \) is non-zero, giving rise to selection rules. One common selection rule concerns the change in the rotational quantum number in rotational-vibrational transitions, which is often \( \Delta J = \pm 1 \) or 0. These correspond to P, Q and R-branches of a vibrational band. The triplet of transitions originating from the ground state with \( J = 1 \) are shown Figure 2.1(b), where they are conventionally labeled P(1), Q(1) and R(1). Note however, that the Q-branch transitions can be forbidden for certain absorption bands. The transition dipole moment is the same for transitions in either direction, and hence the interacting light field can also induce emission of a photon. However, the likelihood of a transition depends also on the population and degeneracy of the involved states (see Section 2.2).

A molecule can have several modes of vibration, and the number increases with the number of nuclei. The modes are commonly labeled \( \nu_1, \nu_2, \nu_3 \) etc., and transitions that simultaneously excite more than one vibrational mode form so-called combination bands. For more complex molecules, more quantum numbers are often required to label the states. For example, non-linear molecules may
have different moment of inertia around their different axes, and then additional rotational quantum numbers $K_a$, $K_b$ and $K_c$ can be used, denoting the projection of the angular momentum along the molecular axes $a$, $b$ and $c$.

Selection rules imposed by Eq. 2.2, apply to the vibrational components of wavefunctions as well, and not all absorption bands are allowed. Their intensity also varies depending on how strongly they are coupled by the light field. Fundamental bands are often one or two orders of magnitude stronger than overtone and combination bands. The fundamental bands are generally found in the MIR region, while overtone and combination bands, which require higher photon energies (shorter wavelength), appear in the NIR. The presence of strong fundamental bands is the reason why the MIR is sometimes referred to as the fingerprint region, as it is particularly useful for detecting and distinguishing molecular species. Hot bands are generally weak, but can appear both in the NIR and MIR depending on the energy difference between the involved states.

2.2 The thermal population distribution

The rotational-vibrational levels of a molecule can be excited not only through absorption of radiation, but also through collisions between molecules. In thermal equilibrium, interactions between molecules distribute the population of energy levels such that the number density $N_i$ [molecules/cm$^3$] of a gas sample in an energy level labeled $i$ is given by the Boltzmann distribution

$$N_i = N \frac{g_i}{Z} e^{-E_i/k_B T},$$

(2.3)

where $N$ is the total number density of the gas [molecules/cm$^3$], $E_i$ is the energy [J] of level $i$, $k_B$ is Boltzmann’s constant [J/K], $g_i$ is the statistical weight, or degeneracy of level $i$, $T$ is the temperature [K], and $Z$ is the partition function

$$Z = \sum_i g_i e^{-E_i/k_B T},$$

(2.4)

where the sum is over all rotational-vibrational energy levels of the molecule. The partition function ensures that $N = \sum N_i$. The exponential function in Eq. 2.3 shows that the thermal population decreases in levels with higher energy. As alluded to in the previous section, higher temperature increases the population in more highly excited levels, leading to the appearance or enhancement of hot bands, and overall increasing the number of observable transitions, thus resulting in more congested absorption spectra.
2.3 The Lambert-Beer law

The real electric field of light at frequency $\nu$ [Hz] propagating in the $z$-direction (measured in [cm]) is described by

$$ E(\nu, z, t) = \tilde{E}(\nu, z, t) + \tilde{E}^*(\nu, z, t), $$ (2.5)

where $t$ is time [s], the star denotes the complex conjugate and the complex magnetic field is given by

$$ \tilde{E}(\nu, z, t) = \frac{1}{2} E_0(\nu)e^{i(2\pi\nu t - kz)}. $$ (2.6)

Here, $E_0(\nu)$ is the vector amplitude of the field [V/m] at frequency $\nu$, and $k = 2\pi\nu/c$ is the wave vector [cm$^{-1}$], where $c$ is the speed of light. Expressed in terms of wavenumber $\bar{\nu} = \nu/c$ [cm$^{-1}$], the complex electric field becomes

$$ \tilde{E}(\bar{\nu}, z, t) = \frac{1}{2} E_0(\bar{\nu})e^{2\pi\bar{\nu}i(ct - z)}. $$ (2.7)

The complex field transmitted a distance $L$ [cm] through an absorbing medium can be expressed as

$$ \tilde{E}_T(\bar{\nu}, z, t) = \tilde{T}(\bar{\nu})\tilde{E}(\bar{\nu}, z + L, t), $$ (2.8)

where

$$ \tilde{T}(\bar{\nu}) = e^{-\delta(\bar{\nu}) - i\varphi(\bar{\nu})}, $$ (2.9)

is the complex electric field transmission function. The presence of molecular transitions at specific wavenumbers subjects the electric field to an amplitude attenuation $\delta(\bar{\nu})$ and a phase shift $\varphi(\bar{\nu})$ that can be expressed as [16]

$$ \delta(\bar{\nu}) = \frac{SNL\chi^{abs}(\bar{\nu})}{2}, $$ (2.10)

$$ \varphi(\bar{\nu}) = \frac{SNL\chi^{disp}(\bar{\nu})}{2}. $$ (2.11)

Here, $S$ is the integrated molecular line intensity [cm$^{-1}$/molecule cm$^{-2}$]), and $\chi^{abs}(\bar{\nu})$ and $\chi^{disp}(\bar{\nu})$ are the area-normalized absorption line shape function [cm], and its dispersion counterpart [cm]. The intensity of the light that is transmitted through the medium then becomes

$$ I_T(\bar{\nu}) = I_0(\bar{\nu})|\tilde{T}(\bar{\nu})|^2 = I_0(\bar{\nu})e^{-2\delta(\bar{\nu})}, $$ (2.12)

where the initial intensity is given by $I_0(\bar{\nu}) = \frac{1}{2}c\epsilon_0|E_0(\bar{\nu})|^2$, where $\epsilon_0$ is the vacuum permittivity. By defining the absorption coefficient $\alpha(\bar{\nu})$ [cm$^{-1}$] as

$$ \alpha(\bar{\nu}) = SN\chi^{abs}(\bar{\nu}), $$ (2.13)
we obtain the well-known Lambert-Beer law for intensity of light transmitted through an absorbing medium

\[ I_T(\bar{\nu}) = I_0(\bar{\nu})e^{-\alpha(\bar{\nu})L}. \]  

(2.14)

Similarly, we define the dispersion coefficient \( \phi(\bar{\nu}) \) [cm\(^{-1}\)] as

\[ \phi(\bar{\nu}) = SN\chi^{\text{disp}}(\bar{\nu}). \]  

(2.15)

The number density, \( N \), can be expressed in terms of the gas pressure \( P \) [Torr], and temperature \( T \) [K], assuming an ideal gas [17]. The absorption and dispersion coefficients then take the form

\[ \alpha(\bar{\nu}) = S n_0 c_{\text{rel}} \frac{P T_S}{P_0 T} \chi^{\text{abs}}(\bar{\nu}), \]  

(2.16)

\[ \phi(\bar{\nu}) = S n_0 c_{\text{rel}} \frac{P T_S}{P_0 T} \chi^{\text{disp}}(\bar{\nu}), \]  

(2.17)

where \( n_0 = 2.68676 \cdot 10^{19} \) molecules/cm\(^3\) is the Loschmidt number, \( c_{\text{rel}} \) is the relative concentration of the absorbing species, \( P_0 = 760 \) Torr is atmospheric pressure and \( T_S = 273.15 \) K. By normalizing the absorption coefficient to the number density of the absorbing species one obtains the absorption cross section

\[ \sigma(\bar{\nu}) = \frac{\alpha(\bar{\nu})}{N}, \]  

(2.18)

which can be useful for comparing the magnitude of absorption measured at different sample densities.

For various reasons, the absorption and dispersion line shape functions \( \chi^{\text{abs}}(\bar{\nu}) \) and \( \chi^{\text{disp}}(\bar{\nu}) \) are not infinitely narrow but cover some range around the center wavenumber \( \bar{\nu}_{12} \) of a transition. A few types of line shapes and the processes giving rise to them will be covered in the next section.

The absorption coefficient of a transition integrated over all wavenumbers is

\[ \int_0^\infty \alpha(\bar{\nu})d\bar{\nu} = SN \int_0^\infty \chi^{\text{abs}}(\bar{\nu})d\bar{\nu} = SN, \]  

(2.19)

since \( \chi^{\text{abs}}(\bar{\nu}) \) is area-normalized. From Eqs. 2.13, 2.14 and 2.19 the total intensity absorbed per unit length integrated over a whole transition is found to be

\[ \frac{\delta I_T}{\delta L} = -SN I_T. \]  

(2.20)

Assuming that a transition occurs between energy levels 1 and 2 of a closed two-level system, this same quantity can be expressed in terms of the Einstein coefficients of absorption \( B_{12} \) and stimulated emission \( B_{21} \) [cm\(^3\)/(Js\(^2\))] as [17]

\[ \frac{\delta I_T}{\delta L} = -(N_1 B_{12} - N_2 B_{21}) \frac{h\bar{\nu}_{12}}{c} I_T, \]  

(2.21)
where \( N_1 \) and \( N_2 \) are the population densities in levels 1 and 2, \( h \) is Planck’s constant \([Js]\), and \( \bar{\nu}_{12} \) is the center wavenumber of the transition. The Einstein coefficients are related by

\[
g_1 B_{12} = g_2 B_{21}, \tag{2.22}
\]

where \( g_1 \) and \( g_2 \) are the statistical weights of levels 1 and 2. The populations in levels 1 and 2 are given by the Boltzmann distribution (Eq. 2.3), denoting their energies as \( E_1 \) and \( E_2 \). Using this together with Eqs. 2.22, Eq. 2.21 can be rewritten as

\[
\frac{\delta I_T}{\delta L} = -\frac{g_1 B_{12}}{Z} N e^{-E_1/k_B T} (1 - e^{-\Delta E/k_B T}) \frac{\hbar \bar{\nu}_{12}}{c} I_T, \tag{2.23}
\]

where \( \Delta E = E_2 - E_1 \) is the energy difference between the levels, and hence the photon energy of the transition. Comparing this to Eq. 2.20, the line intensity can be expressed as

\[
S = \frac{g_1 B_{12}}{Z} e^{-E_1/k_B T} (1 - e^{-\Delta E/k_B T}) \frac{\hbar \bar{\nu}_{12}}{c}. \tag{2.24}
\]

In this expression, \( B_{12} \) is particular to the absorbing transition and proportional to \( \mu_2^2 \) [15], but \( S \) also depends on \( E_1 \) and \( \Delta E \), as these determine the populations in the initial state, and the population difference between the initial and final states for a given temperature. Note however, that we have ignored effects due to redistribution of population to levels other than those involved in the transition.

### 2.4 Line shapes and broadening mechanisms

There are a number of mechanisms that broaden transitions detected in an absorption spectrum, and these are ultimately limiting the precision with which their frequencies can be determined. When probing a gas sample with a light beam, the thermal motion of the molecules gives rise to Doppler broadening since the light appears Doppler shifted to molecules moving relative to the source. Doppler broadening is an *inhomogeneous* broadening mechanism where the absorption probability is different for different molecules of the gas depending on their velocity, and the detected absorption line shape is the combined signal from all velocity groups. Doppler broadening gives rise to *Gaussian* line shapes and is often the dominating broadening mechanism at low gas pressures, in the so called Doppler limit.

The Doppler width can be reduced by limiting the velocity component of molecules parallel to the propagation of the light by cooling the sample or measuring in molecular beams [1]. It is also possible to surpass the Doppler limit by employing various sub-Doppler detection techniques where the absorption of a single velocity group is detected (see Section 3.2).

The frequency precision of sub-Doppler measurements is still limited by other forms of broadening, ultimately by the natural linewidth of the transition which stems from the finite lifetime of the energy states of the absorbing molecules. This fundamental natural lifetime broadening gives rise to a *Lorentzian* line shape. Additional forms of broadening yielding Lorentzian profiles that are often relevant are
power broadening induced by high laser power and collisional broadening caused by inter-molecular collisions. These are homogeneous broadening mechanisms, which (to first order) affect all molecules in the same way. Collisional broadening increases with pressure and is typically dominating over Doppler broadening at atmospheric pressure, and hence low pressure conditions are advantageous for precision spectroscopy. Low pressure also reduces shifts in transition frequencies arising at higher pressures due to interactions between molecules that alter their internals energy levels.

Retrieving spectroscopic parameters from measured absorption spectra typically requires modeling the line profiles of transitions and this section provides a brief mathematical description of the Gaussian and Lorentzian line profile, as well as their origins. Absorption profiles cannot always be modeled purely as either of these, and we will also discuss the Voigt profile which is required in cases where a Doppler broadened transition is also affected by non-negligible collisional broadening. There are additional higher order effects, for example due to the speed dependence of the collision rates, that require more sophisticated line shape models [18]. Such effects become relevant in very sensitive measurements, but for our purposes, the Voigt profile will be sufficient.

The amount of broadening is often quantified in terms of the linewidth of the line profile in question. The half width at half maximum (HWHM) is the detuning from the transition center frequency at which the absorption coefficient is reduced to half of the maximum, while the full width at half maximum (FWHM) is the double-sided width, i.e. twice the HWHM.

2.4.1 The Gaussian line profile - Doppler broadening

Gaussian absorption profiles occur due to the thermal motion of molecules in a gas sample. When the sample is probed by a light beam, different molecules have different velocity components \( v_z \) parallel to the propagation direction of the light. Because of the Doppler effect, the wavenumber of the incoming light appears shifted to higher or lower values for molecules moving towards or away from the light source respectively. In thermal equilibrium, the number of molecules with a given \( v_z \) follows a Maxwell-Boltzmann distribution, which is a Gaussian function centered around \( v_z = 0 \). This carries over into the frequency dependence of the absorption response of the sample which becomes a Gaussian centered around the wavenumber of the transition \( \tilde{\nu}_{12} [1] \). The area normalized absorption line shape function is then

\[
\chi_{abs}^{D}(\tilde{\nu}) = \frac{\sqrt{\ln 2}}{\sqrt{\pi} \Gamma_{D}} e^{-\frac{1}{2} \left( \frac{\tilde{\nu} - \tilde{\nu}_{12}}{\Gamma_{D}} \right)^2}.
\]

The linewidth (HWHM), \( \Gamma_{D} \) [cm\(^{-1}\)], of the Doppler broadened line shape is given by

\[
\Gamma_{D} = \tilde{\nu}_{12} \sqrt{\frac{2 \ln 2 k_{B} T}{mc^2}},
\]
where $m$ is the molecular mass [kg] of the absorbing species. The corresponding dispersion line shape takes the form [16]

$$
\chi^{\text{disp}}_D(\bar{\nu}) = -\frac{2\sqrt{\ln 2}}{\pi \Gamma_D} \int_0^x e^{-x^2} e^{s^2} ds,
$$

(2.27)

where $x = \sqrt{\ln 2(\bar{\nu} - \bar{\nu}_{12})}/\Gamma_D$. Note that the Doppler width $\Gamma_D$ is proportional to the transition wavenumber $\bar{\nu}_{12}$ and hence Doppler broadened transitions become narrower at longer wavelengths. Moreover, it increases with temperature, but is smaller for heavier molecules since their mean velocity is lower at a given temperature.

### 2.4.2 Lorentzian line profiles

The finite lifetime of the excited state involved in a transition expands the range of photon energies it can absorb. This is a manifestation of the energy-time uncertainty principle of quantum mechanics [19]. The resulting broadening of an absorption line is termed natural lifetime broadening and gives rise to a Lorentzian absorption profile [1, 15]. The absorption and dispersion line shape functions take the form

$$
\chi^{\text{abs}}_L(\bar{\nu}) = \frac{1}{\pi \Gamma^2_L + (\bar{\nu} - \bar{\nu}_{12})^2},
$$

(2.28)

$$
\chi^{\text{disp}}_L(\bar{\nu}) = -\frac{1}{\pi \Gamma^2_L + (\bar{\nu} - \bar{\nu}_{12})^2},
$$

(2.29)

where $\Gamma_L$ [cm$^{-1}$] is the Lorentzian HWHM. Higher gas pressure, corresponding to higher density of molecules, leads to an increased rate of collisions between them. During a molecular collision, the internal rotational-vibrational energy levels are shifted by the interaction and collisions also lower the lifetime of excited states. These effects scale with the gas pressure and give rise to a pressure shift of transition frequencies and collisional broadening or pressure broadening of absorption lines [1]. Pressure broadening also yields a Lorentzian line shape, and for a sample containing multiple molecular species the total Lorentzian width can be expressed as

$$
\Gamma_L = \Gamma_n + \sum_j \gamma_j p_j,
$$

(2.30)

where $\Gamma_n$ is the natural lifetime linewidth, and $\gamma_j$ and $p_j$ are the pressure broadening coefficients and partial pressures of the $j^{th}$ collision partner. The pressure shift can analogously be expressed as

$$
\Delta P = \sum_j \delta_j p_j,
$$

(2.31)

where $\delta_j$ is the pressure shift coefficient of the $j^{th}$ collision partner.
2.4.3 The Voigt profile

In the pressure broadened regime where collisional broadening is not negligible compared to Doppler broadening, both of these mechanisms must be accounted for. To a first order approximation, each velocity group displays the same Lorentzian profile and the resulting line shape is described by the Voigt profile which is a convolution of the Lorentzian and Gaussian profiles. Although the convolution cannot be calculated analytically, it can be expressed in terms of the Faddeeva error function \( \tilde{w}(z) \). The absorption and dispersion line shapes are obtained from its real and imaginary parts [16]

\[
\chi_V^{\text{abs}}(\tilde{\nu}) = \chi_0 \text{Re}[\tilde{w}(x + iy)],
\]

\[
\chi_V^{\text{disp}}(\tilde{\nu}) = -\chi_0 \text{Im}[\tilde{w}(x + iy)],
\]

where \( x = \sqrt{\ln 2(\tilde{\nu} - \tilde{\nu}_{12})/\Gamma_D}, \ y = \sqrt{\ln 2\Gamma_L/\Gamma_D} \) and \( \chi_0 = \sqrt{\ln 2}/\sqrt{\pi}\Gamma_D \). The error function \( \tilde{w}(x + iy) \) is then computed numerically. Figures 2.2(a)-(b) show examples of absorption and dispersion line shapes of the Voigt model plotted against the frequency detuning from the transition \( \Delta \nu = \nu - \nu_{12} \), where the linewidths \( \Gamma_D = \Gamma_L = 100 \text{ MHz} \).

![Figure 2.2](image)

**Figure 2.2:** Example of Voigt absorption \( \chi_V^{\text{abs}}(\nu) \) (a), and dispersion \( \chi_V^{\text{disp}}(\nu) \) (b) line shapes with \( \Gamma_D = \Gamma_L = 100 \text{ MHz} \).
3 Absorption spectroscopy, methods and concepts

The general idea of absorption spectroscopy is to measure the absorption by transitions as a function of optical frequency. We will use frequency $\nu$ [Hz] when referring to absorption linewidths and occasionally for describing e.g. frequency scales of spectrometers, though we will mostly use wavenumber $\tilde{\nu} = \nu/c$ [cm$^{-1}$] as unit in the measurements, while we describe different spectral regions, and wavelengths of lasers in terms of wavelength $\lambda = c/\nu$ [\mu m]. The two main metrics that characterize spectroscopic measurements are the spectral resolution, i.e. the smallest frequency difference that can be resolved, and the sensitivity, often expressed as the smallest amount of absorption that can be detected. A few approaches to improving the sensitivity of a measurement system are explained in this chapter, as well as the possibility of enhancing selected transitions through optical pumping. Ways of improving measurements of transition frequencies with sub-Doppler techniques will also be briefly discussed. We employed broadband light sources, where a spectrometer was used to analyse the light in the frequency domain. The spectral resolution is typically mainly determined by the spectrometer, and we will look at the general representation and analysis of broadband absorption spectra.

3.1 Enhancing the absorption signal

As seen from Eq. 2.14, directly measuring the absorption of a sample fundamentally amounts to detecting a reduction in transmitted light intensity at certain frequencies. The detected light intensity inevitably carries some amount of noise, that can stem from the light source itself (intensity noise) or the detector circuit (detector noise), but is also fundamentally limited by statistical variations in the number of photons reaching the detector in a given time interval (shot noise). It is desirable to make the absorption signal as strong as possible compared to the noise level, and a simple metric for evaluating this relation is the signal-to-noise ratio (SNR) of a measurement. In our work, we define this as the ratio of the peak absorption signal to the noise level without absorption. The SNR can be increased either by reducing the noise level or enhancing the absorption.

As is clear from Eq. 2.14, the absorption increases with the distance $L$ that the light propagates through the sample, and a longer interaction length hence improves SNR. The simplest type of sample container is a single-pass cell which the light beam passes through once. Since the length of such containers is often limited by practical considerations, two common ways of enhancing the interaction length while keeping the sample containers compact are multipass cells and enhancement cavities, which are briefly described in the next two sections. Note however, that the absorption signal cannot be increased indefinitely in this way, since for long enough path lengths, all light is absorbed at the transitions. Longer path lengths then only result in increasingly distorted absorption lines.

There are various approaches for reducing the impact of noise fluctuations. In addition to the universally applied method of averaging the detected signal over time, there are specific detection schemes aimed at noise reduction. One rather simple method is balanced detection (to be discussed in Section 6.4), where by
subtracting two signals which carry noise from the same source, one can detect differences between them while canceling the common noise. More advanced examples are frequency modulation techniques [20], where the frequency of the light is modulated and the detected signal demodulated at the modulation frequency. Absorption features give rise to intensity modulations that are picked up by the demodulation, and this serves to shift the detection to the modulation frequency, which can be hundreds of MHz, where the noise density is lower.

### 3.1.1 Multipass cells

In a multipass cell, the light beam is reflected multiple times back and forth through the sample. For a given container length, the interaction length is then enhanced by a factor roughly equal to the number of reflections, which can yield tens of meters of path length, though longer path lengths tend to lead to larger power losses due to a higher number of reflections.

Two main types of multipass cells are commonly used. The White cell [21], is made from three mirrors, and the number of reflections (and hence the path length) can be changed by adjusting the angles of the mirrors. A Herriott cell [22], consists of only two fixed mirrors with one or more holes machined in them for the input and output beams. This design is simpler and often more robust than the White cell.

Unlike enhancement cavities, the multipass cell is non-resonant with the light and therefore Beer’s law is valid, and the absorption coefficient can be directly calculated from the spectral data (see Section 3.3). The exact value of $L$ depends on the alignment through the cell, which can vary within a certain tolerance, and this causes some uncertainty that propagates to the absorption intensity.

### 3.1.2 The Fabry-Pérot enhancement cavity

The Fabry-Pérot cavity consists of two parallel partially transmitting mirrors $M_1$ and $M_2$, shown schematically in Figure 3.1(a). The mirrors are assumed to have the same intensity reflection and transmission coefficients $r$ and $t$ (where $r$ is close to 1) and are separated by the distance $L_c$. Light with intensity $I_0(\bar{\nu})$ is coupled in from the left. Neglecting the effect of dispersion, light with a wavelength $\lambda$ that satisfies the condition [23]

$$n\frac{\lambda}{2} = L_c,$$  \hspace{1cm} (3.1)

where $n$ is an integer, can resonate in the cavity, leading to an enhancement of the intensity. With every reflection at the mirrors, a small portion of the intensity is coupled out. Light of other wavelengths is prevented from entering the cavity by destructive interference, and is instead reflected back at mirror $M_1$. Hence, only certain optical frequencies, corresponding to the cavity modes, can resonate. They are separated by the free spectral range (FSR) [Hz], which is given by

$$\text{FSR} = \frac{c}{2L_c},$$ \hspace{1cm} (3.2)
The cavity mode structure is depicted in Figure 3.1(b). In real cavities however, the FSR becomes wavelength dependent due to intra-cavity dispersion. The condition for resonance can also be viewed in terms of the round trip phase shift of the electromagnetic wave, which should be $n2\pi$, and this phase shift is affected both by dispersion at the cavity mirrors, and by the refractive index of any medium inside the cavity.

![Figure 3.1](image)

Figure 3.1: (a) Schematic depiction of a Fabry-Pérot cavity consisting of two mirrors with reflection and transmission coefficients $r$ and $t$ respectively, separated by a distance $L_c$. Light with intensity $I_0(\nu)$ is impinging from the left and the transmitted intensity is given by $I_T(\nu)$. (b) A frequency domain representation of the resonant cavity modes separated by the FSR.

When an absorbing medium is present in the cavity, the transmitted light is affected not only by its dispersion but also its absorption. The dispersion is divided into one component stemming from the slowly varying background refractive index $n_b(\nu)$ of the medium as a whole, and the dispersion occurring around molecular transitions $\phi(\nu)$ (Eq. 2.17). The transmitted intensity can be expressed by $I_T(\nu) = T_c(\nu)I_0(\nu)$ where $T_c(\nu)$ is the intensity transmission function of the cavity given by [24]

$$T_c(\nu) = \frac{t^2e^{-\alpha(\nu)L_c}/(1 - re^{-\alpha(\nu)L_c})^2}{1 + 4re^{-\alpha(\nu)L_c} \sin^2 \left(2\pi\nu n_b(\nu)L_c + \frac{\Phi_m(\nu)}{2} + \frac{\phi(\nu)L_c}{2} \right)/(1 - re^{-\alpha(\nu)L_c})^2},$$

where $\Phi_m(\nu)$ contains the dispersion component of the mirrors. The peaks of the cavity modes (corresponding to a round trip phase shift of $n2\pi$) occur when the argument of the sine function is an integer of $\pi$.

A characteristic parameter of the cavity is the finesse $F$, which is the ratio of the FSR to the width (FWHM) of the cavity modes $\Gamma_c$

$$F = \frac{\text{FSR}}{\Gamma_c} = \pi \frac{\sqrt{r}}{1 - r}. \quad (3.4)$$

The mirror reflectivity is slightly wavelength dependent, which carries over into $F$. Assuming light that is on resonance with a cavity mode, that $r \approx 1$ and that the absorption and dispersion are small one can show that

$$\frac{I_0(\nu) - I_T(\nu)}{I_0(\nu)} \approx \frac{2F}{\pi} \alpha(\nu). \quad (3.5)$$
The absorption is enhanced by a factor of $2F/\pi$, where $F$ can reach several thousands. Enhancement cavities are hence very effective at increasing sensitivity, though in order to couple light effectively through them, one generally needs a laser that is frequency stabilized to one of the cavity modes. There are schemes that do not necessarily require this, such as cavity ring-down spectroscopy (CRDS) [1], where it is sufficient to couple in some light and then measure the characteristic decay time of the intensity trapped in the cavity [25, 26].

A commonly used way of stabilizing lasers to cavities is the Pound-Drever-Hall (PDH) method [27]. It relies on phase-modulating the light, thus generating sidebands around the laser frequency. Detecting the reflection from the input mirror and demodulating at the phase modulation frequency yields an error signal with odd symmetry with respect to the detuning of the laser frequency from a near-by cavity mode. By feeding back either to the FSR via the cavity mirror separation, or to the laser frequency, one can use this error signal to lock one cavity mode to the laser or the laser frequency to a cavity mode.

To achieve effective coupling into a cavity (and for Eq. 3.3 to be valid), it is also necessary to match the collimation parameters of the input beam to the fundamental transverse mode of the cavity, which is determined by the curvature of the mirrors [28].

Finally, the spectrum transmitted through a cavity cannot be easily converted to e.g. absorption coefficient using the Lambert-Beer law, but rather one has to model the transmission function (Eq. 3.3), with all its parameters, which makes the analysis more complicated than when using single-pass or multipass cells.

### 3.2 Sub-Doppler techniques and optical pumping

In high-resolution spectroscopic measurements, the precision of measured transition frequencies can be restricted by the linewidth of the transitions, where Doppler broadening becomes the limiting factor in the low-pressure regime. Apart from reducing Doppler broadening as touched upon earlier, there are detection schemes that avoid Doppler broadening by addressing a single velocity group of the molecules. These techniques were strongly aided by the invention of lasers that provide intense, narrowband light that can be tuned to molecular transitions. Sub-Doppler measurements are often realized with pump-probe schemes, where one velocity group is targeted with a strong narrow-linewidth pump laser and the molecular response is detected with a weaker probe beam, though sub-Doppler signals can also be obtained by using counter-propagating beams of the same laser. Examples include saturation techniques [29], which rely on saturating the absorption of the selected velocity group, polarization spectroscopy [30], where one detects the effect on the polarization of the probe beam induced by absorption of the pump beam, and two-photon spectroscopy [31], where two photons are simultaneously absorbed by the sample.

Another potential issue in absorption measurements, is that the intensity of observed absorption lines depends not only on the inherent strength of the transitions (transition dipole moments) but also reflects the thermal population distribution. Absorption lines can be difficult to detect, if the lower level of a targeted transi-
tion has low population or if the population difference between the levels is small (typical for long wavelength transitions), recall Eq. 2.24. Population in more highly excited levels can be increased by raising the sample temperature but this also populates more levels than the ones of interest, leading to the appearance of many more absorption lines in the spectrum, which complicates the analysis. For molecules with complex or overlapping absorption bands, spectral congestion can be a problem already at room temperature. A way to circumvent these problems is by optical pumping [1], where a strong laser tuned to a transition selectively transfers population from the lower energy level to the higher level.

In double-resonance (DR) spectroscopic techniques the pump laser is complemented by a probe light field that addresses transitions sharing either upper or lower state with the pump transition. The probe field can be in the radio-frequency (RF) or microwave region, and such techniques are referred to as optical-RF and optical-microwave double-resonance [1]. Microwave spectroscopy is frequently used for studying pure rotational transitions, but the population difference between their energy levels can be small. A pump laser can induce a larger population difference, that is further affected by probing the involved levels with the microwave source. For both optical-RF and optical-microwave DR spectroscopy, detecting the population change due to the probe on the transmitted pump intensity is advantageous as the more energetic pump photons are more easily detected than the low energy probe photons. Optical pumping also allows for populating excited vibrational levels and probing the rotational structure in the microwave. Another DR technique employs a second laser as probe which also addresses rotational-vibrational transitions. This optical-optical DR spectroscopy was used in this thesis and is described in more detail below.

3.2.1 Optical-optical double-resonance spectroscopy

In optical-optical DR (OODR) spectroscopy [1], the probe is another laser beam that detects the population transfer induced by the pump. If the linewidth of the pump laser is much narrower than the Doppler width, the population transfer occurs in only one velocity group. The depopulated lower level of the pumped velocity group is detected by the probe as sub-Doppler absorption dips called \textit{v-type} features on the Doppler broadened transitions sharing a lower state with the pumped transition. Similarly, probe transitions from the upper level of the pump transition display sub-Doppler enhancements of the absorption. These are referred to as \textit{ladder-type} transitions as the pumped energy state serves as an intermediate level for the probe transitions, where the latter might not be observable at thermal population. Ladder-type transitions allow for studying high energy levels essential for modeling high-temperature spectra, without the above mentioned complications due to high temperature such as strong congestion. A third kind of transition is the \textit{Λ-type}, which occurs as stimulated emission from the pumped level, though this type will not be dealt with in this thesis.

The exact frequencies of the observed OODR features depend on which velocity group is addressed by the pump. If the pump and probe beam paths are parallel in the sample and the pump frequency is tuned on resonance with the pump
transition (pumping the velocity group with $v_z = 0$), the $v$-types appear at the centers of the Doppler profiles and the ladder-type lines at the centers of their corresponding transitions. Figure 3.2(a) shows a simplified molecular energy level diagram depicting the four first levels of a vibrational mode further split into rotational sub-levels. The orange arrow indicates a transition of the fundamental band excited by the pump, and the blue arrows show possible examples of a $v$-type transition observed in the first overtone band and a hot band ladder-type transition. Figure 3.2(b) shows the sub-Doppler absorption profiles of the ladder-type and $v$-type features observed in transmission through the sample (see Section 3.3), where the pump is tuned to the center of the Doppler broadened line.

![Figure 3.2](image)

**Figure 3.2:** (a) The energy levels of a molecule with the transition induced by the pump laser (orange) and two possible transitions addressed by the probe (blue). The probe transitions are both of the ladder-type and $v$-type. (b) The absorption profiles of the ladder-type (top) and $v$-type transitions (bottom), where the latter appears on top of a Doppler broadened line.

Retrieving standard-temperature line intensities from ladder-type transitions is not straightforward, since the observed intensities depend on the (non-thermal) population in the intermediate state. This cannot be easily known since it depends on factors such as the degree of saturation of the pumped transition and the overlap between the pump and probe beams. However, assuming equal relaxation rates from the two states of the pump transition, the population in the pumped intermediate state equals the depletion in the lower state, that causes the $v$-type features. Quantitative information of the ladder-type intensities can then be obtained by normalizing them to the $v$-type intensities. Furthermore, the intensities are effected by the relative polarization of the pump and probe lasers [12, 32].
3.3 Broadband absorption spectra

Throughout this thesis, we probe gas samples with broadband light, and analyse the transmitted spectrum using a spectrometer. A graphical depiction of the spectrum (optical intensity vs wavenumber) of a broadband light source transmitted through a sample, where the Lambert-Beer law is valid, is shown in Figure 3.3(a). In the notation of the previous chapter, this corresponds to $I_T(\bar{\nu})$, where a number of absorption lines can be seen on top of the spectral emission profile, or spectral envelope, of the source $I_0(\bar{\nu})$. Note however, that the envelope shape in Figure 3.3(a) does not necessarily exactly correspond to the incident intensity $I_0(\bar{\nu})$, but rather to $I_B(\bar{\nu})$ which is the intensity transmitted through the empty sample container, thus being subject to any losses along the beam path common with $I_T(\bar{\nu})$. $I_B(\bar{\nu})$ is provided by a background spectrum measured without the absorber, and the transmission through the sample is calculated by normalizing $I_T(\bar{\nu})$ to $I_B(\bar{\nu})$, essentially yielding the intensity transmission function $|\tilde{T}(\bar{\nu})|^2$ shown in Figure 3.3(b). The absorption coefficient can then be calculated from Eq. 2.14 (replacing $I_0(\bar{\nu})$ by $I_B(\bar{\nu})$) as

$$\alpha(\bar{\nu}) = -\frac{1}{L} \ln \left( \frac{I_T(\bar{\nu})}{I_B(\bar{\nu})} \right).$$

(3.6)

This is plotted in Figure 3.3(c), and a zoom of a single absorption line is shown in Figure 3.3(d) with the sampling points marked by black dots. Here, the point spacing is sufficiently dense to clearly resolve the Gaussian shape of the line with approximately eight points per FWHM.

![Figure 3.3](image)

**Figure 3.3:** (a) The intensity of a broadband source $I_T(\bar{\nu})$ transmitted through a sample, with absorption lines appearing as dips on top of the background intensity profile $I_B(\bar{\nu})$. (b) The transmission through the sample obtained by normalizing the transmitted intensity to $I_B(\bar{\nu})$. (c) The absorption coefficient calculated using Eq. 3.6. (d) Zoom on one absorption line.

For very precise absorption measurements, the spectral sampling point spacing is a key aspect, as it needs to be smaller than the absorption lines one wishes to
study. However, it is not the only factor determining the spectral resolution. Instrumental effects in the spectrometers can smear out the spectral features, limiting the available resolution, and distort the shapes of absorption lines. An even more crucial issue is the absolute accuracy of the wavenumber scale, since it is the transition frequencies that provide information about separations between molecular energy levels, and evaluating these correctly is essential for building accurate spectral models. In many cases, spectrometers are frequency-calibrated by measuring absorption lines with very well known positions. However, optical frequency combs (Section 5) are a great asset for this purpose, as they provide a direct link between optical frequencies and highly accurate atomic clocks.

Some additional details on the analysis of absorption spectra with enhancement cavities (where Eq. 2.14 is not valid) are given in Section 5.4.
4 Spectroscopic databases and models

To be able to benefit from spectroscopic measurements for various applications one generally needs some reference or model to compare to. By measuring the absorption of a sample it is then possible to deduce e.g. what molecules it contains and in what concentrations. In its most basic form, the reference data can be raw spectra measured at well defined experimental conditions. Such data are available e.g. in the database from the Pacific Northwest National Laboratory (PNNL) [33], which provides spectra at atmospheric pressure and different temperatures. More versatile comparisons are possible with the help of line lists that provide parameters of individual transitions of molecules. This allows for simulating molecular spectra at the exact conditions of various experiments.

Molecular line lists can be composed in different ways. There are theoretical line lists that are calculated ab initio, from solving the Schrödinger equation for given assumed models of the internuclear potential energy and dipole moments. Such calculations can produce very extensive line lists containing billions of transitions [34, 35]. However, they need to be validated against experimental data and if necessary adjusted to improve their accuracy. There are also line lists compiled from experimental measurements, though in order to assign measured absorption lines to molecular transitions, comparison to theoretical models is often required. The models can be based on ab initio calculations but are then gradually refined by fitting them to experiments. Line lists based on empirical data are generally more accurate than their purely theoretical counterparts, but collecting the required spectroscopic data and refining the models is a laborious and time-consuming process. There are also experimental conditions that are challenging to measure at in the lab (e.g. high temperatures) and to model such spectra one might have to rely on theoretical line lists to a greater extent.

In our work, we make use of spectral databases for modeling absorption bands during spectral data treatment and for assigning observed absorption lines to predicted transitions, but also for comparing retrieved line parameters to existing line lists, with the ultimate aim of improving their accuracy.

4.1 The HITRAN database

One very frequently used database of molecular line parameters is HITRAN (high-resolution transmission molecular absorption database) [11], which is maintained by the HITRAN group at Harvard University. The data in HITRAN typically stem from spectral models based on empirical data. It currently contains parameters of over 50 molecules, and is mostly aimed at atmospheric applications, though it is still applicable for low pressure measurements. For simulating spectra at higher temperature, the HITEMP (high-temperature molecular spectroscopic database) [36] is provided in parallel. It contains additional data for absorption bands observable at high temperatures, though it currently covers only 8 molecules.

HITRAN provides a variety of parameters, and the most relevant for our purposes are those necessary for simulating absorption spectra using the Voigt profile (Sections 2.3 and 2.4).
• The transition center wavenumber in vacuum \( \bar{\nu}_c \) [cm\(^{-1}\)].

• The spectral line intensity \( S \) [cm\(^{-1}\)/(molecule cm\(^{-2}\))] at temperature \( T_{\text{ref}} = 296 \) K.

• The pressure broadening (HWHM) coefficient \( \gamma_{\text{self}} \) [cm\(^{-1}\)/atm] due to collisions of the absorbing species with itself.

• The pressure broadening (HWHM) coefficient \( \gamma_{\text{air}} \) [cm\(^{-1}\)/atm] due collisions between the absorbing species and molecules of air.

• The pressure shift due to air \( \delta_{\text{air}} \) [cm\(^{-1}\)/atm].

HITRAN is a very useful source of reference data, and it is in a state of continuous development. The scope of the database, in terms of number of molecules and covered spectral ranges, as well as the accuracy of already included parameters is being improved, and updates have been published typically every four years. These improvements require new experimental data and analysis to provide more refined spectral models.

In the scope of this thesis, we made extensive use of HITRAN as a reference source, but also contributed data to its latest update published in 2022 [11].

4.2 Theoretical databases

In this thesis, we used two well established predominantly ab initio-based spectroscopic databases aimed at astrophysical applications such as the study of exoplanets, and cool stars. The ExoMol project [37], based at University Collolleg of London, provides calculated line lists of many molecules for high-temperature simulations. The TheoReTS database (Theoretical Reims–Tomsk Spectral data) [38], as the acronym suggests a joint French-Russian endeavor. It contains data for molecules with four to six atoms.

We used ExoMol as reference in Paper VI where the necessary data was not available in HITRAN, and also refined their line list using our measured data. In Paper VII, TheoReTS provided predictions of multiply excited vibrational energy levels of CH\(_4\), which we could validate experimentally for the first time.

4.3 The MARVEL algorithm

One powerful method for producing spectral models from large bulks of experimental data, without using a Hamiltonian is the \textit{MARVEL} (measured active rotational–vibrational energy levels) procedure [39], developed as a joint project between the ExoMol group at University Colllege London and Eötvös University in Budapest. It relies on collecting assigned transitions observed in many different measurements into spectroscopic networks (SNs) that join all the rotational-vibrational energy levels involved in the transitions. Taking the experimental uncertainties into account, a least-square-fit is performed that results in self-consistent evaluations of the energy levels and the set of transitions connecting them, with soundly estimated uncertainties. The resulting \textit{MARVELized} line lists
can be much more accurate than pure ab initio predictions (as found in Paper VI). Because of the interconnected SNs, the experimental input data impact the MARVEL evaluations of a larger number of transitions than those actually observed.

We relied extensively on a MARVEL-based line list in Paper VI for assignment of the spectrum of H$_2$CO, and also used our data to update the MARVEL analysis and improve its precision, leading to an update of the ExoMol database.
5 Optical frequency combs

Optical frequency combs (OFCs) are lasers that emit spectra consisting of a large number of equidistant narrow frequency components similar to the structure of a comb, and these frequencies are accordingly referred to as comb modes or comb teeth. An OFC is said to be fully stabilized, if the frequencies of the comb modes are fixed over time, providing a kind of optical ruler of well defined frequencies. OFCs were originally developed as a link transferring the accuracy of atomic clocks into the optical region [40], but soon proved to be useful for spectroscopy as well. As direct light sources, OFCs offer a broad optical bandwidth, coupled with higher spectral brightness and better spatial and spectral coherence than the previously used thermal broadband sources. The comb structure can also be exploited to provide highly accurate frequency scales of spectroscopic measurements, both using OFCs as direct light sources, and for frequency referencing narrowband lasers. In addition, broadband cavity-enhanced measurements are possible, by matching the comb mode frequencies to those of the cavity modes. Here, we review the major types of OFCs that exist. We then focus on the time and frequency domain description of light emitted by an OFC from a mode-locked laser, and we will see how a NIR OFC can be converted to the MIR region. We shall also look at how OFCs can be combined with enhancement cavities for broadband cavity-enhanced measurements.

5.1 Different types of OFCs

The first OFCs were generated with mode-locked lasers [28], and this is often regarded as the standard type of OFC. Mode locking is a process by which a fixed phase relationship is established between a large number of resonant modes of the laser cavity. The phase locking arises as a consequence of the pulsed operation of mode-locked lasers. The formation of short pulses is often made to occur passively by some mechanism that increases the gain for high field strengths, though it can also be induced actively by modulating the gain of the laser [28]. Mode-locked OFCs are commonly based on e.g. Ti:Sapphire lasers emitting around 600 nm [41], Cr:ZnSe lasers around 2.4 µm [42], and various fiber lasers. Typical doping substances for the latter are Ytterbium (Yb) yielding lasing around 1 µm [43], Erbium (Er) around 1.5 µm [44], and Thulium (Tm) around 2 µm [45]. These OFCs are all well established technologies but only provide light in the NIR region.

OFCs based on quantum cascade lasers (QCLs) [46], are a promising alternative, as QCLs can lase directly in both the MIR and THz regions. Passive pulsed mode-locking is difficult to achieve with QCLs, but active mode-locking has been demonstrated by external modulation via the driving current or optical seeding [47, 48]. Spontaneous multi-mode operation can arise in QCLs as a result of four wave mixing, and a stable phase relationship between the modes has been demonstrated [49], though without pulse formation. Fully stabilizing a QCL OFC is possible but has required referencing to a mode-locked OFC [50]. QCL resonators are compact compared to traditional mode-locked lasers, but the resulting large mode separation can be inconvenient for spectroscopy as it limits the spectral reso-
olution. A certain tuning of the mode separation is possible by current modulation, and the resolution of spectroscopic measurements can be improved by scanning the comb modes over the spectral features [51].

There are also ways of generating OFCs or comb-like output spectra from narrowband seed lasers. In micro-resonator combs [52], the frequency of a CW laser injected into a micro-resonator undergoes four-wave mixing processes to populate other resonant modes, thus forming a comb structure. The resonators can be very compact, but the comb mode separation is generally large (several GHz), and obtaining broader bandwidth tends to require increasing the mode spacing.

Electro-optic combs (EOCs) are produced by phase-modulating a seed laser, yielding multiple sidebands, which provides very high tunability of the mode spacing [40], and stabilization of the comb modes has been demonstrated [53].

Nevertheless, mode-locked lasers are the most widely used sources of fully stabilized OFCs, due to the inherent phase coherence of the comb modes with well established schemes for detecting and controlling the parameters determining the comb structure (to be discussed below). The OFCs used in our work were also of this type, and the remaining discussion of this chapter will assume a traditional pulsed mode-locked OFC.

5.2 Mode-locked OFCs in the time and frequency domains

Mode-locked OFCs emit trains of pulses with a repetition rate $f_{\text{rep}}$ which is the inverse of the laser cavity round-trip time. The electric field of such a pulse train is displayed schematically in Figure 5.1(a). The pulses consist of short envelopes with typical lengths of tens to hundreds of fs, with a superimposed modulation at the carrier frequency. The carrier wave undergoes a phase shift of $\Delta\Phi_{ce}$ relative to the envelope between consecutive pulses, due to the difference between the phase and group velocities in the laser cavity.

In the frequency domain, the OFC spectrum consists of modes separated by precisely $f_{\text{rep}}$ and the frequency of the $n^{th}$ comb mode can be expressed as

$$\nu_n = f_{\text{ceo}} + nf_{\text{rep}},$$

(5.1)

where $f_{\text{ceo}}$ is the cavity-envelope-offset frequency which ranges from 0 to $f_{\text{rep}}$ and is related to the cavity-envelope phase shift by

$$f_{\text{ceo}} = \frac{\Delta\Phi_{ce}}{2\pi} f_{\text{rep}}.$$  

(5.2)

The frequency domain structure is displayed in Figure 5.1(b). Note that lasing occurs only at a certain range of the modes described by Eq. 5.1, though the spectral coverage can still be hundreds of cm$^{-1}$ or many tens of thousands of mode indices.

The repetition rate $f_{\text{rep}}$ lies in the radio frequency range (typically a few hundred MHz) and can be detected with a fast photo-detector due to the pulsed nature of the OFC. To measure $f_{\text{ceo}}$ one often uses the f-2f technique [41], where the OFC spectrum is frequency-doubled and the $f_{\text{ceo}}$ is detected by beating a part
Figure 5.1: (a) Time domain depiction of the electric field of a mode-locked OFC. The envelope is outlined with the dashed line while the solid curve shows the carrier frequency modulation. The carrier envelope phase shift relative to the first pulse is also indicated. (b) Frequency domain illustration of the comb mode frequencies described by Eq. 5.1 (black dashed vertical lines) and the actual spectrum of the comb (solid vertical lines).

of the frequency-doubled spectrum with an overlapping segment of the original spectrum. By implementing feedback loops acting on the laser cavity length and cavity dispersion, \( f_{\text{rep}} \) and \( f_{\text{CEO}} \) can be locked, yielding full stabilization of the comb modes.

5.3 Non-linear frequency conversion

To be able to study different molecules and absorption bands, it is sometimes desirable to convert the radiation of a light source to a different wavelength. For example, mode-locked laser technologies are mature in the NIR, but to reach the finger-print region in the MIR with these sources, one has to resort to such wavelength conversion. It can be accomplished by exploiting non-linear effects possessed by various materials. In the following, we briefly describe the basics of these processes, and in particular how to implement one of them with OFCs.

5.3.1 Principles of non-linear conversion

Non-linear effects arise when the optical properties of materials are altered by strong radiation fields [54], and typically lasers are required to achieve the necessary intensities. The result is a conversion of a part of the input light field to a different frequency, the simplest example being second harmonic generation, where light with twice the optical frequency of the input field is generated.

To convert light to longer wavelengths, one can use difference frequency generation (DFG). In this process, two input fields termed the pump with frequency \( \nu_p \), and the signal with frequency \( \nu_s \) are overlapped in a non-linear material resulting in the production of an idler wave with frequency \( \nu_i = \nu_p - \nu_s \). The pump has
higher power than the signal and part of this power is transferred to the signal and idler fields. However, maintaining this unidirectional transfer of power requires matching the phases of the fields in the non-linear material, which change with propagation distances due to the frequency dependence of the refractive index. Phase matching can be achieved over longer distances by exploiting birefringence of the material, if present. One can also resort to quasi-phase matching, where the crystal geometry of the material is inverted at a regular pattern, which compensates for the dephasing of the waves [54]. The period of the geometry inversion determines which wavelengths can be quasi-phase matched in the material.

The efficiency of the conversion process can be increased by placing the non-linear material in an optical cavity that enhances the light intensity inside it. Such a device is called an optical parametric oscillator (OPO). In a doubly resonant OPO, the cavity enhances both the signal and idler fields, while if only one of these can resonate, the OPO is singly resonant. Due to the cavity enhancement, only the pump beam is required as input, while the other fields arise through quantum fluctuations that are subsequently amplified at the resonant frequencies. Singly resonant OPOs require higher pump powers than doubly resonant OPOs, but continuous frequency tuning is not possible in the doubly resonant regime since both the idler and signal wavelengths must coincide with one of the resonant modes of the cavity [54].

NIR mode-locked OFCs can be converted to MIR wavelengths using both OPOs and DFG-based systems. Pulsed lasers have the advantage that the peak powers of the pulses are much higher than for a CW laser with a similar average power. OPO-based OFCs can provide high power and bandwidth [55, 56] and wide tunability [57, 58]. The cavity makes them slightly more complex than DFG-based OFCs, and the $f_{ceo}$ of the OPO idler cannot in general be easily known, even when the $f_{ceo}$ of the pump is stabilized. DFG-based systems do not suffer from this problem, and the implementation of DFG with OFCs will be summarized briefly below.

### 5.3.2 Difference frequency generation with OFCs

It is possible to use DFG to down-convert the whole spectrum of an OFC, described by Eq. 5.1. One way is by producing the difference frequencies between a CW pump with frequency $\nu_p$ and an OFC [59], yielding an idler spectrum given by

$$\nu_{i,n} = \nu_p - f_{ceo} - nf_{rep}.$$  \hspace{1cm} (5.3)

The pump can also be another OFC, in particular, the pump and signal can be produced by the same OFC, which removes the need to frequency-stabilize several lasers. With very broad OFC spectra, DFG conversion can occur between the low and high frequency ends of the same spectrum, referred to as intra-pulse DFG (IDFG), which can yield very broad idler spectra [60–62]. Alternatively a signal OFC can be obtained by shifting the pump OFC spectrum to longer wavelengths in a highly non-linear fiber (HNLF) [63–67]. The shifted OFC retains the same $f_{rep}$ and $f_{ceo}$, but the light shifts to lower mode indices $n$. In both this and IDFG, the $f_{ceo}$ is hence the same in the pump and signal OFCs, which means that it cancels out in the idler spectrum given by
\[ \nu_{i,n} = f_{ceo} + n_p f_{rep} - f_{ceo} - n_s f_{rep} = n f_{rep}, \]  

(5.4)

where the indices \( n_p \) and \( n_s \) refer to the pump and signal OFCs respectively. This can be a disadvantage if both degrees of freedom in Eq. 5.1 are required e.g. in order to lock the OFC to an enhancement cavity (see Section 5.4), but it also means that stabilizing only \( f_{rep} \) of the pump OFC is sufficient to obtain a fully stabilized idler OFC.

The basic DFG scheme used in our work, where the signal OFC is derived from the pump OFC in a HNLF is depicted in Figure 5.2. The pump and signal pulses need to overlap temporally in the non-linear material for the conversion process to work. The timing of the pulses varies with wavelength tuning due to dispersion in the optical components, and can also drift over time. To achieve and maintain temporal overlap, an adjustable delay line is added to the path of the pump beam. In some systems, pulse timing can be crucial for the performance [68], and automatic stabilization schemes are thus beneficial.

![Figure 5.2: Schematic depiction of a DFG system generated from a single OFC. The signal (green) is obtained by shifting the spectrum in a HNLF, and the pump (blue) is passed through a variable delay stage for achieving temporal overlap between the signal and pump pulses. The beams are overlapped in the non-linear material, producing the idler OFC beam (red). Note that the signal and pump beams remaining after the DFG process are not drawn.](image-url)
5.4 Locking an OFC to an enhancement cavity

The OFC mode structure strongly resembles that of the cavity modes of a Fabry-Pérot cavity. If the cavity FSR were independent of wavenumber and matched to the \( f_{\text{rep}} \) of the OFC, it would be possible to achieve simultaneous transmission of all comb modes through the cavity. However, due to dispersion, the FSR varies over the OFC spectrum and cannot be matched to \( f_{\text{rep}} \) over arbitrarily large wavenumber ranges, which often limits the OFC spectral bandwidth that can be transmitted through the cavity. This effect can be quantified by the phase terms in Eq. 3.3 (the argument of the sine function). In the absence of molecular transitions, the last term vanishes and the phase offset between the comb mode at wavenumber \( \tilde{\nu}_n \) from its nearest cavity mode is

\[
\Delta \Phi(\tilde{\nu}_n) = 2\pi \tilde{\nu}_n n_b(\tilde{\nu}_n) L_c + \frac{\Phi_m(\tilde{\nu}_n)}{2} \mod \pi, \tag{5.5}
\]

since the phase at the cavity mode resonances is equal to integers of \( \pi \). The main variation in FSR is due to the second term stemming from the mirror dispersion. Figure 5.3 shows a simulation of the effect of the FSR increasing with wavenumber, where a comb mode with index chosen as \( n = 0 \) is exactly on resonance with a cavity mode. Figure 5.3(a) shows graphically the positions of the comb modes (red solid lines), which are increasingly offset from the cavity modes (black dotted lines) further away from \( n = 0 \). The cavity transmission function \( T_c(\tilde{\nu}) \) (where the mode width is depicted as unrealistically large compared to the FSR) is shown in Figure 5.3(b) (black), together with the transmission of the comb modes (red sticks). Figure 5.3(c) shows the comb-cavity phase offset \( \Delta \Phi(\tilde{\nu}_n) \) as a fraction of \( \pi \), which is zero at \( n = 0 \) but gradually deviates from zero further from this comb mode.

![Figure 5.3](image)

**Figure 5.3:** (a) The relative positions of the comb modes (red solid lines) and the cavity modes (black dotted lines), where the comb modes are on resonance with the cavity modes around comb mode index \( n = 0 \). The cavity FSR increases towards higher indices. (b) The cavity transmission function \( T_c(\tilde{\nu}) \) (black) and the relative intensity of the transmitted comb modes (red sticks). (c) The comb-cavity offset phase \( \Delta \Phi(\tilde{\nu}_n) \) in units of \( \pi \).
When measuring cavity enhanced absorption spectra with OFCs, the dispersion component due to the molecular transitions $\phi(\bar{\nu})$ also affects the cavity mode frequencies. The dispersion line shape (shown for a Voigt profile in Figure 2.2(b), is asymmetric with respect to the transition center frequency, resulting in the cavity modes shifting in opposite direction on either side of the absorption line. Whenever $\Delta \Phi \neq 0$, this causes asymmetry in the measured line shapes. The effect of $\Delta \Phi(\bar{\nu})$ on the transmitted spectral envelope (Figure 5.3(b)) can be canceled by normalizing the absorption spectrum to the cavity transmitted spectrum without absorbing sample (similarly to the background normalization described in Section 3.3). However, the asymmetric line shapes remain, and $\Delta \Phi(\bar{\nu})$ also affects the normalized transmission on resonance with the transitions (even though $\phi(\bar{\nu}) = 0$ there), since the cavity transmission does not have the simple exponential dependence on absorption of the Lamber-Beer law. Figure 5.4(a) shows a simulation (based on Eq. 3.3) of a normalized cavity transmission when $\Delta \Phi = 0$ (black) and when $\Delta \Phi(\bar{\nu})$ is given by the function plotted in panel (b) (red). The absorption lines become asymmetric in opposite directions depending on the sign of $\Delta \Phi(\bar{\nu})$. However, provided that the dispersion is properly accounted for in the model, the spectrum can still be used to retrieve accurate line parameters.

![Picture](image.png)

**Figure 5.4:** (a) The normalized cavity transmission containing a series of absorption lines. The black trace shows the case when $\Delta \Phi = 0$. When $\Delta \Phi(\bar{\nu})$ is given by the curve in panel (b), the absorption lines become asymmetric as indicated by the red trace in panel (a).

Although it might not be possible to couple the entire OFC spectrum through the cavity, various locking schemes exist that allow for stabilizing parts of the OFC bandwidth to the cavity modes.

The PDH-locking technique (Section 3.1.2) can be applied directly to the light of an OFC, where the light of all comb modes is affected by the phase modulation. Since only a limited number of comb modes can be on resonance with their
corresponding cavity modes, a narrow spectral portion of the OFC light reflected from the cavity is selected for producing the PDH error signal. This is done by dispersing the reflected OFC beam and directing a part of the spectrum on a detector. The error signal is then used to maintain transmission of the OFC spectrum through the cavity around this locking point, where $\Delta \Phi(\bar{\nu}) = 0$ as around the comb mode $n = 0$ in Figure 5.3.

One can choose two locking points in the spectrum to ensure transmission of as much of the OFC spectrum as possible [69, 70]. The two PDH feedback loops require one degree of freedom each of either the OFC or the cavity to stabilize one to the other. The cavity FSR can be controlled via the mirror separation, while the OFC provides control of $f_{rep}$ and $f_{ceo}$. Therefore, at least $f_{rep}$ or $f_{ceo}$ must be free to respond to one of the PDH locks, and they cannot both be absolutely stabilized. Not having absolutely stabilized OFC frequencies is a disadvantage in precision spectroscopy applications, particularly when one needs the mode structure to be absolutely stable over time, e.g. for averaging purposes. Instead, one can implement a single PDH feedback loop to lock the cavity to a point in the OFC spectrum, while $f_{rep}$ and $f_{ceo}$ are absolutely stabilized to values optimizing the cavity transmission. This is fundamentally the approach followed in Paper VII (see Section 8.1).

Additional variations are also possible, such as locking the cavity to a CW laser that is in turn locked to a comb mode of an OFC [71, 72]. Such schemes do not require phase-modulating the OFC light (and thereby transmission of the OFC), and the cavity modes can be tuned simply by adjusting the frequency offset between the OFC and the CW laser. Hence, there is no need to change locking point in the OFC spectrum, which would require mechanically adjusting the spectral portion detected for the PDH error signal.
6 Optical frequency comb Fourier transform spectroscopy with sub-nominal resolution

This chapter describes the spectroscopic technique that is the basis for all the works in the thesis, namely the combination of OFCs with Fourier transform spectroscopy (the acronym FTS will be used both for Fourier transform spectrometers and Fourier transform spectroscopy). In Fourier transform spectroscopy, the spectral information of a light beam is extracted via the Fourier transform from the interference observed when part of the light is phase shifted by a variable amount in an interferometer [2]. Fourier transform infrared spectroscopy (FTIR) based on incoherent light sources has long been used for measuring the spectra of molecules, but the replacement of the thermal sources by OFCs provides substantial improvements. The higher coherence and spectral brightness of OFCs allows for higher SNR in shorter acquisition times. OFCs are easier to couple through an interferometer and allow for longer interaction lengths with the samples, and as discussed in Section 5.4, they can also be efficiently coupled through enhancement cavities. The narrow discrete comb modes can also be exploited to circumvent the effects of instrumental line shapes which result from the finite length of the recorded interferograms.

The spectrometers used in this thesis are based on the Michelson interferometer, whose principle of operation is described in Section 6.1. Then we will cover the basics of OFC-based FTS, followed by the sub-nominal technique which allows for avoiding the aforementioned instrumental effects and to surpass the spectral resolution inherent to the FTS which is limited by the length of the interferogram. We will also discuss some elements of the analysis of the specific measurements in the thesis.

6.1 The Michelson interferometer

A simple schematic depiction of a Michelson interferometer is shown in Figure 6.1. An incoming monochromatic laser beam with intensity $I_0$ is split by a beamsplitter (BS) into two beams which are reflected by two mirrors ($M_1$ and $M_2$) and recombined at the beamsplitter. Mirror $M_1$ is fixed while mirror $M_2$ can be scanned along the beam direction. At a detector at point A, the beams reflected from the two arms of the interferometer produce interference, that depends on the optical path difference (OPD) between the arms $\Delta l$, which here incorporates the refractive index of the medium inside the interferometer.

Assuming a 50-50 splitting ratio of the beamsplitter the intensity at the detector is given by [1]

$$I_A = \frac{1}{2} I_0 \left(1 + \cos \left(\frac{2\pi \nu_0 \Delta l}{c}\right)\right),$$  \hspace{1cm} (6.1)

where $\nu_0$ is the frequency of the laser light. If mirror $M_2$ is scanned, the change in $\Delta l$ gives rise to a sinusoidal modulation in $I_A$ with a periodicity that depends
Figure 6.1: Schematic of a Michelson interferometer. An incoming laser beam with intensity $I_0$ is split by a beamsplitter (BS) into two arms and reflected of mirrors (M$_1$ and M$_2$), where the latter can be translated along the beam direction. The two beams are recombined at the beamsplitter and a detector is placed at point A. The direction of incoming light is shown by red arrows and back-reflected light by black arrows.

on $\nu_0$. The output beam reaching point B is phase-shifted by $180^\circ$ compared to point A due to a different number of reflections and the intensity is given by

$$I_B = \frac{1}{2} I_0 \left( 1 - \cos \left( \frac{2\pi \nu_0 \Delta l}{c} \right) \right). \quad (6.2)$$

By taking the difference between the two signals one obtains

$$I = I_A - I_B = I_0 \cos \left( \frac{2\pi \nu_0 \Delta l}{c} \right). \quad (6.3)$$

This is the idea of balanced detection. By cancelling the offset in the signals, one gets rid of much of the intensity noise contained in $I_0$, while the amplitude of the interferogram is doubled. Note however that for optimal noise performance, the powers at points A and B need to be the same which is not generally the case in real interferometers due to e.g. unequal splitting ratio of the beamsplitter and losses occurring along the beam paths. The above description also assumes 100% contrast of the interferogram which is not achieved in practice.

Fourier transforming the interferogram in Eq. 6.3 using the fast Fourier transform (FFT) algorithm, yields the spectrum. The sampling points in the spectrum are spaced by $\delta \nu$, which is given by the length of the interferogram

$$\delta \nu = \frac{c}{\Delta l_{\text{max}}}, \quad (6.4)$$

where $\Delta l_{\text{max}}$ is the maximum change in OPD of the interferogram. This is referred to as the nominal resolution of the interferometer. Higher spectral resolution hence requires a longer scanning distance of the mirror, and this is the factor limiting the resolution of FTSs, as the instruments cannot be made arbitrarily large.
Another consequence of the truncation of the interferogram is an instrumental line shape (ILS) of the peak appearing in the spectrum at $\nu_0$. If the interferogram is simply truncated without applying any additional apodization function, the spectrum is convolved with an instrumental line shape function given by

$$g_{ILS}(\nu) = \frac{\Delta l_{\text{max}}}{c} \frac{\sin(\pi \nu \Delta l_{\text{max}}/c)}{\pi \nu \Delta l_{\text{max}}/c}.$$  \hspace{1cm} (6.5)

This is a normalized sinc function displayed graphically in Fig. 6.2. The shape of $g_{ILS}$ carries over to the detected frequency peak at $\nu_0$ which is broadened with ringing occurring at the wings. Note that the ILS function is zero when the detuning is an integer value of the nominal resolution $\delta \nu$.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure6.2}
\caption{The instrumental line shape function $g_{ILS}$ due to truncation of the interferogram, where the x-axis is in units of the nominal resolution $\delta \nu$. Zero crossings occur when $\nu$ is an integer times $\delta \nu$.}
\end{figure}

### 6.2 Fourier transform spectroscopy with OFCs

Here we describe the signals obtained in the OPD and spectral domains when using an FTS with an OFC light source (OFC-FTS). Since the light of an OFC consists of a large number of comb modes, the interference signal from the FTS (Eq. 6.3) becomes a sum over all frequency components

$$I = \sum_n I_n \cos \left( \frac{2\pi \nu_n \Delta l}{c} \right). \hspace{1cm} (6.6)$$

The superposition of all sine waves results in an interferogram consisting of a series of bursts occurring when $\Delta l_{\text{max}}$ is an integer value of $c/f_{\text{rep}}$, shown in red in Figure 6.3 with a zoom of one of the bursts revealing the modulation superimposed on it. Since the comb modes are separated by $f_{\text{rep}}$ in the spectral domain, resolving individual comb modes requires $\delta \nu < f_{\text{rep}}$ which corresponds to $\Delta l_{\text{max}} > c/f_{\text{rep}}$. Hence one needs to record an interferogram longer than the separation between bursts, and the resulting spectrum is shown at the top of Figure 6.3. This can be
done with mechanical FTSs [73–75], but is more often done in dual-comb spectroscopy (DCS) [76]. The latter is a variation of FTS where beat notes between the modes of two OFCs with slightly different $f_{rep}$ are detected. This gives rise to a time-domain interferogram similar to that in Figure 6.3, which is Fourier-transformed to yield the spectrum. However, the gaps between the comb modes do not contain any spectral information, and (as will be elaborated on in Section 6.3) it is sufficient to acquire an interferogram of length $\Delta l_{max} = c/f_{rep}$ thereby matching the spectral point spacing to the comb mode spacing $f_{rep}$, as in the spectrum at the bottom of Figure 6.3.

\[ \text{Figure 6.3: The train of bursts separated by } c/f_{rep} \text{ that make up the interferogram from an OFC (red), with a zoom of one of the bursts. Fourier transforming an interferogram segment spanning many burst results in a comb-mode resolved spectrum (top), while limiting the OPD range to } c/f_{rep} \text{ yields a spectrum with a nominal resolution matched to } f_{rep} \text{ (bottom), where the dots indicate sampling points.} \]

In order to calculate the correct frequency scale of the Fourier spectrum (by knowing $\delta \nu$) one needs to measure $\Delta l_{max}$. This is done by using a continuous wave (CW) laser with a known and stable wavelength that propagates through the FTS on a path parallel to the OFC beam. The sinusoidal interferogram of this reference laser serves as a calibration signal to determine $\Delta l_{max}$, but is also used to resample the OFC interferogram at equal OPD steps, which is necessary for the FFT.
6.3 The sub-nominal resolution technique

For retrieving highly-accurate frequencies of molecular transitions, it is generally desirable to measure in regimes with narrow absorption features, in the Doppler limit or even using sub-Doppler techniques. In the context of OFC-FTS, the absorption lines might frequently be narrower than the available nominal resolution of the FTS and the comb mode separation. In addition, unless great care is taken when sampling the spectra, the ILS stemming from interferogram truncation can distort the absorption lines, making them broader, and causing modulations on their wings that affect line parameters retrieved by standard line shape fits [9]. A way around these problems is provided by the sub-nominal resolution sampling-interleaving technique [8, 9]. This section summarizes the theory behind this technique and the various steps and operations it consists of.

6.3.1 The instrumental line shape of an OFC-FTS spectrum

All information contained in an OFC spectrum can be retrieved when \( \Delta l_{\text{max}} = c / f_{\text{rep}} \) since one then has precisely one sampling point per comb mode. The comb modes are negligibly narrow compared to the ILS function \( g_{\text{ILS}} \) (Fig. 6.2) and can be considered as Dirac \( \delta \)-functions. Convolving the spectrum with \( g_{\text{ILS}} \) then results in a sum of the sinc functions stemming from each comb mode. Fig. 6.4 shows a simulation where one of the comb modes (red vertical bars) with index \( n_{\text{abs}} \) is absorbed by an absorption line that is narrower than \( f_{\text{rep}} \). The ILS of the absorbed comb mode is shown as the red dotted line, while the sum of the ILS contributions from all comb modes results in the dashed red line at the top. Here, the nominal resolution \( \delta \nu = f_{\text{rep}} \), which means that the comb modes occur at the zero-crossings of the ILS contributions of their neighbors. If the spectrum is sampled such that the sampling points (black dots) are precisely on top of the comb modes, one can hence avoid the ILS and retrieve the correct intensity of each comb mode. As the black solid line shows, one obtains a flat baseline free from the underlying modulations, with reduced intensity on comb mode \( n_{\text{abs}} \) only.

In the following, we describe the process of achieving the situation in Figure 6.4, which requires not only matching the nominal resolution to \( f_{\text{rep}} \) but also shifting the sampling points to match them to the comb modes.

6.3.2 Matching the OFC and FTS frequency scales

Avoiding the ILS distortions hence requires matching two different frequency scales, one describing the comb mode frequencies and one marking the sampling point positions. The frequency scale of the comb modes is given by (Eq. 5.1)

\[
\nu_{\text{OFC}} = f_{\text{ceo}} + n f_{\text{rep}}. 
\]  

(6.7)

Similarly, the frequency scale obtained by default from the FFT is given by

\[
\nu_{\text{FTS}}^0 = n \delta \nu, 
\]  

(6.8)

where \( \delta \nu = c / \Delta l_{\text{max}} \) (Eq. 6.4). The first step is to choose \( \Delta l_{\text{max}} \) such that
Figure 6.4: Illustration of the ILS caused by an absorption line absorbing a single comb mode in an OFC-FTS spectrum. The comb modes are shown as red vertical lines. The mode with index $n_{abs}$ is absorbed by a transition. The dotted red line shows the ILS profile stemming from only this comb mode, while the red dashed line is the sum of the ILS of all comb modes. The black dots indicate sampling points which are matched to the comb mode positions. Thereby, the ILS distortions are avoided and the comb mode intensities are accurately retrieved.

$$\delta \nu = f_{rep}. \quad (6.9)$$

However, the interferogram consists of discrete sampling points with the OPD spacing $\delta l$ equal to a set fraction of the reference laser wavelength $\lambda_{ref}$. Hence it is in general not possible to satisfy Eq. 6.9 exactly, but the closest match occurs when

$$\Delta l_{max} = N_l \delta l, \quad (6.10)$$

where $N_l$ is the number of points in the interferogram given by

$$N_l = \text{round}\left(\frac{c}{f_{rep} \delta l}\right) = \text{round}(N_l^*), \quad (6.11)$$

and $N_l^*$ is the ideal unrounded number of sampling points. The dimensionless parameter $\epsilon$ is used to quantify the deviation between $N_l$ and $N_l^*$

$$N_l^* = (1 + \epsilon)N_l. \quad (6.12)$$

The maximum value of $\epsilon$ is $1/N_l$ [9], where $N_l$ is on the order of $10^6$. However, since the indices $n$ in Eqs. 6.7 and 6.8 are also of this order at the measured optical frequencies, the resulting offset between the sampling points and comb modes can become significant. Figure 6.5 shows a depiction of $\nu_{OFC}$ and $\nu_{FTS}^0$ as the red
and black dots respectively, where the difference in point spacings (slopes) and the $f_{\text{ceo}}$ offset have been strongly exaggerated for clarity. With $\delta \nu$ calculated as

$$\delta \nu = \frac{c}{\Delta l_{\text{max}}} = \frac{c}{N_l \delta l}$$

(6.13)

the offset between the two scales due to the rounding of $N^*$ at the comb mode with index $n_{\text{opt}}$ is

$$f_{\text{shift}} = n_{\text{opt}}(f_{\text{rep}} - \delta \nu) = -n_{\text{opt}} \epsilon f_{\text{rep}},$$

(6.14)

where the last step follows from Eq. 6.11 and 6.12. Then, a final FTS scale is produced by adding this shift together with any non-zero $f_{\text{ceo}}$

$$\nu_{\text{FTS}} = n \delta \nu + f_{\text{shift}} + f_{\text{ceo}}.$$  

(6.15)

To sample the spectrum at these frequencies, the interferogram is multiplied by an exponential function containing $f_{\text{shift}}$ and $f_{\text{ceo}}$ [9], prior to taking the FFT. The scale described by $\nu_{\text{FTS}}$ is indicated by the blue dots in Figure 6.5, and yields a local match to $\nu_{\text{OFC}}$ at comb index $n_{\text{opt}}$. The discrepancies between the two frequency scales often remain small enough to avoid discernible ILS distortions in a region measured around $n_{\text{opt}}$.

![Graphical depiction of the matching of the OFC and FTS frequency scales.](image)

**Figure 6.5:** Graphical depiction of the matching of the OFC and FTS frequency scales. The comb mode frequencies $\nu_{\text{OFC}}$ occur along the red dotted line, which is shifted by $f_{\text{ceo}}$ from the FTS sampling points $\nu_{\text{FTS}}^0$, indicated by the black dotted line. The sampling point spacing also differ slightly from $f_{\text{rep}}$, yielding different slopes of the curves. A vertical shift produces $\nu_{\text{FTS}}$ (blue dotted line) which is matched to $\nu_{\text{OFC}}$ around the comb with index $n_{\text{opt}}$. Hence, to match the sampling points to the comb modes, one first selects the number of OPD domain sampling points $N_l$ according to Eq. 6.11. The

43
interferogram is cut to this required length centered around the zero-OPD burst (Figure 6.3). Before calculating the FFT, the interferogram is multiplied by the exponential introducing the sampling point shift equal to $f_{ceo} + f_{shift}$ calculated according to Eqs. 6.13 and 6.14, where $n_{opt}$ is chosen as the comb mode index in the center of the analyzed spectral range.

However, the procedure described above relies on precise knowledge of $\delta l$ which is not available a priori, but can be determined from the data, as will be discussed in Section 6.3.4. In addition, for some measurements included in this thesis, instrumental effects in the FTS caused the comb modes to be mapped into the FTS spectrum with a wavelength-dependent spacing, resulting in a curvature of the red dotted line in Figure 6.5. Unlike the corrections introduced in Eq. 6.15, such effects cannot be calculated beforehand, and the way to correct for them is described in Section 6.3.5. First however, we will see how to surpass the nominal resolution of the FTS. In fact, the method for determining $\delta l$ fundamentally relies on the width of the absorption features being smaller than or comparable to $f_{rep}$, and hence additional procedures are needed to resolve them.

### 6.3.3 Spectral interleaving

The previous sections described how to obtain an OFC-FTS spectrum with minimized ILS distortions at a sample point spacing equal to the repetition rate $f_{rep}$ of the OFC. For precision spectroscopy however, one often needs to measure absorption features narrower than or comparable to $f_{rep}$, which for mode-locked OFCs is typically on the order of 100 MHz. The sampling point spacing required to resolve these features is provided by the interleaving step of the sub-nominal technique. Since a spectrum is sampled at the comb mode positions, one can acquire a set of spectra with the comb modes at different positions, perform the procedure described above for each spectrum and finally interleave them. This allows for surpassing the limit posed by $f_{rep}$ by orders of magnitude (see e.g. paper VII), which is far beyond the nominal resolution of any available FTS.

The comb modes can be scanned either by tuning the $f_{ceo}$ or $f_{rep}$, but for all the works in this thesis, $f_{rep}$-tuning was used. This is the only degree of freedom available when using inherently $f_{ceo}$-free DFG sources (Section 5.3.2), but also has the advantage that very small adjustments of $f_{rep}$ are necessary. This is because the resulting shift of the comb modes scales with the mode index $n$. When $f_{rep}$ is changed by an amount $\Delta f_{rep}$, the resulting shift in the $n^{th}$ comb mode in the optical domain is

$$\Delta \nu = n \Delta f_{rep},$$  \hspace{1cm} (6.16)

where $n$ is typically around $10^6$. Knowing the approximate mode number $n$ in the targeted spectral region and the width of the absorption lines of interest, one can thus choose a $\Delta f_{rep}$ to yield the required spectral point spacing. The number of steps $N_{steps}$ is chosen such that

$$N_{steps} n \Delta f_{rep} \approx f_{rep}.$$ \hspace{1cm} (6.17)
Hence, in the optical domain the separation between the comb modes is constant to within one part in $\sim 10^6$, but they are scanned over the absorption features. This is depicted in Figure 6.6, where the comb modes at a given $f_{\text{rep}}$ setting are shown as solid red vertical lines, and one mode is partly absorbed by an absorption line, shown as the black dashed line. The modes are scanned at increments of $\Delta \nu$ over seven positions (red dashed lines) to map out the absorption line with appropriate point spacing.

Figure 6.6: The separation ($f_{\text{rep}}$) between the comb modes (solid red lines) is larger than the width of an absorption line (black dashed line). By tuning $f_{\text{rep}}$ the comb modes are scanned at increments of $\Delta \nu$ (dashed red lines). For this increment size, seven steps are required to cover the gaps of one $f_{\text{rep}}$.

The general measurement procedure is thus to choose a set of repetition rate settings

$$f_{rep}^s = f_{rep}^0 + s \Delta f_{rep}, \quad (6.18)$$

where $s = 0, 1, \ldots, N_{\text{steps}}$, and record at total of $N_{\text{steps}} + 1$ spectra at these $f_{rep}$ values. Each spectrum is averaged a number of times to improve the SNR. During the analysis, the spectra are then sampled according to Section 6.3.2 using the respective $f_{rep}^s$, before they are interleaved to the final spectrum with a spectral point spacing of $\Delta \nu$.

The technique is versatile as the resolution can be adapted to a wide range of needs, and is in the ideal case limited only by the width of the comb modes, as was demonstrated in [10]. There are however, other effects that can limit the resolution. For example, due to instrumental effects not taken into account in the above model, the ILS function may be broader and have a different shape than the assumed sinc function. That can result in an instrumental broadening of measured absorption lines and reduce the maximum available resolution. To a varying extent, this was the case for Papers I - VI.
6.3.4 Reference wavelength calibration

Matching the FTS and OFC scales relies on choosing $\Delta l_{\text{max}}$ correctly and according to Eqs 6.10 and 6.11, this requires knowing $\delta l$, which is given by

$$\delta l = \frac{\lambda_{\text{ref}}}{q}, \quad (6.19)$$

where $\lambda_{\text{ref}}$ is the reference laser wavelength and $q$ is an integer indicating the number of sampling points per period of the reference laser interferogram. Note however, that $\lambda_{\text{ref}}$ depends on the refractive index, and unless the inside of the FTS is under vacuum, the refractive index is different for the reference laser and OFC, and may also vary with the ambient conditions in the lab. Also, the OFC and the reference beam are in practice not exactly parallel in the FTS and have different collimation, so a given mirror translation will correspond to slightly different changes in OPD for the two lasers. Hence, $\lambda_{\text{ref}}$ should be considered an effective value of the wavelength.

To minimize the ILS to below the noise level, $\delta \nu$ typically needs to be matched to $f_{\text{rep}}$ with a relative accuracy of at least $10^{-8}$. While the repetition rate can readily be known with such accuracy, the value of $\lambda_{\text{ref}}$ cannot, for the aforementioned reasons. Instead, it is determined from the measured spectrum. Assuming a reference wavelength value $\lambda'_{\text{ref}}$ in Eq. 6.19 to calculate $N_l$ (Eq. 6.11), and inserting this in Eq. 6.13 using Eq. 6.12 and the $\delta l$ given by the true effective reference wavelength $\lambda_{\text{ref}}$, one obtains

$$\delta \nu = f_{\text{rep}} (1 + \epsilon) \frac{\lambda'_{\text{ref}}}{\lambda_{\text{ref}}}, \quad (6.20)$$

This actual point spacing in general differs from the point spacing assumed in the analysis $\delta \nu' = f_{\text{rep}} (1 + \epsilon)$, used for calculating $f_{\text{shift}}$ (Eq. 6.14). From Eq. 6.15, the sampling point with index $n_{\text{opt}}$ then occurs at frequency

$$\nu_{\text{FTS}}^{\text{opt}} = f_{\text{CEO}} + n_{\text{opt}} \epsilon f_{\text{rep}} \left( \frac{\lambda'_{\text{ref}}}{\lambda_{\text{ref}}} - 1 \right) + f_{\text{rep}} n_{\text{opt}} \frac{\lambda'_{\text{ref}}}{\lambda_{\text{ref}}}, \quad (6.21)$$

and this reduces to the frequency of comb mode $n_{\text{opt}}$ given by

$$\nu_{\text{OFC}}^{\text{opt}} = f_{\text{CEO}} + n_{\text{opt}} f_{\text{rep}} \quad (6.22)$$

when $\lambda'_{\text{ref}} = \lambda_{\text{ref}}$. The effective reference wavelength can hence be determined by varying $\lambda'_{\text{ref}}$ in the analysis (which corresponds to varying the slope of the blue line in Figure 6.5). The ILS distortions around index $n_{\text{opt}}$ are minimized when $\lambda'_{\text{ref}}$ is equal to $\lambda_{\text{ref}}$ (within some tolerance determined by the SNR). To quantify the degree of ILS, one can look at the residuals of the fits to a selection of lines in the spectrum. Figure 6.7(a) shows a plot of the root-mean-square (black markers) of the residuals of the fit to a line of methane CH$_4$, using the Voigt model in spectra analyzed with $\lambda'_{\text{ref}}$ incrementally adjusted at 30 steps with a relative step size of $10^{-8}$. The insets show the fits at steps 9 and 30, where the latter displays strong ILS effects while the former yields flat residuals with no obvious modulation.
Figure 6.7: (a) The black markers show the dependence of fit residual root-mean-square (rms) of Voigt fits to a methane line on the reference laser wavelength $\lambda_{\text{ref}}'$ assumed in the analysis, which is adjusted with a relative step size of $10^{-8}$. The insets show the fits (blue) to the data (black) around the rms minimum (step 9), and when $\lambda_{\text{ref}}'$ is far from the optimum value (step 30). (b) The dependence on $\lambda_{\text{ref}}'$ of the line center frequency retrieved from the Voigt fits, relative to the optimum at step 9.

The optimum $\lambda_{\text{ref}}'$ for the line in Figure 6.7(a) occurs around step 9. Preferably though, the final value is obtained as an average of the optima found for several lines. Individual lines can be affected by baseline effects that may resemble ILS, and fitting a number of lines spread over the spectrum also shows whether there is some wavelength dependence in $\lambda_{\text{ref}}'$ which requires additional corrections (see Section 6.3.5). The uncertainty in determining $\lambda_{\text{ref}}$ propagates to the frequency scale of a measurement [9]. The uncertainty in transition frequencies retrieved by line fitting results from a combination of the sampling point uncertainties and the ILS effects due to imperfect matching of the OFC and FTS scales. The effect of the ILS on fitted line centers depends on the linewidths. For lines much narrower than $f_{\text{rep}}$, the ILS modulation appears far to the sides of the transition, while for broader lines it overlaps with the line profile and has a larger influence on the fit. For lines much broader than $f_{\text{rep}}$, ILS effects become very weak [9]. The uncertainty contribution of $\lambda_{\text{ref}}'$ on transition frequencies is thus best assessed by determining the uncertainty in $\lambda_{\text{ref}}'$, which can be evaluated either from a curve such as that in Figure 6.7(a), or from the spread in optima found for many lines.
and then evaluating the shift in retrieved center frequencies corresponding to this $\lambda'_{ref}$ uncertainty. Figure 6.7(b) shows the dependence of retrieved center frequency on $\lambda'_{ref}$ for the methane line in Figure 6.7(a).

Sometimes, the measured absorption features may be too broad to display clear ILS (e.g. Paper IV). In such cases, $\lambda'_{ref}$ can be optimized in alternative ways, such as comparing the positions of certain absorption lines to previously reported center frequencies. The frequency scale uncertainty is then limited by the uncertainty in these reference data.

### 6.3.5 Correcting for non-linear mapping of OFC frequencies

An OFC by definition contains a large range of frequency components and if they do not all propagate in the same way through the FTS, e.g. if there is a frequency dependence in the propagation direction (spatial chirp) or the collimation, then $\delta l$ can also vary with frequency. This results in the comb modes being mapped non-linearly into the FTS spectrum with a mode spacing that depends on wavelength. A procedure to deal with such situations is described in Paper III.

The non-linear mapping of the different frequency components results in the line of equidistant comb modes $\nu_{OFC}$ in Figure 6.7 to become curved. This OFC scale $\nu'_{OFC}$ is depicted graphically (curvature exaggerated) by the red dotted line in Figure 6.8.

**Figure 6.8:** Graphical depiction of the correction to non-linear mapping of the comb modes in the FTS spectrum. The comb modes (red dots) occur along a curved line. The black dots indicate the FTS sampling points for two different values of $\lambda'_{ref}$ which translate into different point spacings reflected by the different slopes of these lines. The sampling points of the black curves coincide with the comb modes with indices $n_A$ and $n_B$. The comb mode frequencies between $n_A$ and $n_B$ are approximated with the corrected FTS scale (blue dots).
The non-linearity of $\nu'_{OFC}$ manifests itself as a wavelength dependence of the optimum $\lambda'_{\text{ref}}$ value. Two different FTS scales $\nu^A_{FTS}$ and $\nu^B_{FTS}$ obtained with $\lambda'_{\text{ref}} = \lambda^A_{\text{ref}}$ and $\lambda'_{\text{ref}} = \lambda^B_{\text{ref}}$ are shown that intersect $\nu'_{OFC}$, and hence minimize the ILS, at comb modes with indices $n_A$ and $n_B$. According to Eq. 6.20, the slopes of these lines are proportional to $\lambda^A_{\text{ref}}$ and $\lambda^B_{\text{ref}}$ and from those one can calculate the $\lambda'_{\text{ref}}$ corresponding to the average slope of $\nu'_{OFC}$ between $n_A$ and $n_B$ with Eq. A.11 in Paper III. Using this $\lambda'_{\text{ref}}$ value together with an appropriate shift calculated with Eq. A.13 in Paper III, yields the corrected FTS scale $\nu^\text{corr}_{FTS}$ displayed by the blue dotted line in Figure 6.8, that is a much better match to $\nu'_{OFC}$ between $n_A$ and $n_B$.

The above procedure thus provides a way to match the FTS and OFC scales and retrieve an accurate spectrum even when the frequency components of the OFC are mapped non-linearly into the FTS spectrum. However, if the curvature of $\nu'_{OFC}$ is severe, it may be necessary to divide the measured wavelength range into segments and calculate different $\nu^\text{corr}_{FTS}$ for each segment (Paper III and VI).

### 6.4 Balanced and auto-balanced detection

As described in Section 6.1, the intensity noise of the OFC source can be reduced by detecting both out-of-phase outputs of the FTS and subtracting the signals, provided that the power levels on the detectors are properly balanced. This is referred to as balanced detection, and the balancing can either be performed by adjusting the optical power on the detectors or balancing the magnitude of the detected electrical signals. The state of optimum balancing can change from day to day e.g. due to changing conditions in the laboratory, or even while acquiring a spectrum due to changes in alignment caused by the moving mirror. In auto-balanced detection an electronic feedback-loop is used to automatically maintain optimum noise cancellation [77]. This convenient configuration simplifies operation and can lead to very substantial noise reduction in OFC-FTS measurements [69].

In this thesis, auto-balanced detection was used in the NIR OODR experiment and 3 $\mu$m regions (Papers I, IV, V and VII), while the measurements at 8 $\mu$m (Papers II, III and VI) were performed in a balanced configuration only.
MIR high-precision optical frequency comb
Fourier transform spectroscopy

Historically, most of the data for the important MIR region in spectroscopic databases were supplied by traditional FTIR measurements, which provided broad bandwidth but limited sensitivity and resolution. The introduction of OFCs as light sources brought about improvements in these aspects, and after the first demonstrations of OFC-FTS in the NIR [7, 78], the technique was soon expanded to the MIR.

Early MIR OFC-FTS implementations included detection of several molecular species in the 3 µm - 5 µm region using OPO-based sources [58, 79]. Cavity enhanced OFC-FTS was demonstrated for hydrogen peroxide, H₂O₂, around 3.8 µm [70], and for several molecules, including methane, CH₄, and formaldehyde, H₂CO, at 2.5 µm - 5 µm [80], in the latter study by placing the sample inside the OPO cavity. Initially, much focus was on improved sensitivity compared to traditional FTIR, and these works also demonstrated the potential for simultaneous detection of multiple molecular species.

Studies aimed at very high-resolution measurements and detailed analysis of molecular spectra also appeared. Measurement of rotationally resolved cold spectra, and accompanying Hamiltonian fits, were reported for several complex molecules such as naphthalene, C₁₀H₈, hexamethylenetetramine, C₆H₁₂N₄, and adamantane C₁₀H₁₆ using an OPO-based OFC at 3.3 µm [81], and of the Buckminsterfullerene C₆₀ using a DFG-based OFC at 8.5 µm [75]. The required high resolution was achieved by interleaving multiple comb mode-resolved FTS spectra with incrementally scanned mode positions.

MIR applications of dual-comb spectroscopy (DCS) also began, with measurements of CH₄ at 3.4 µm using DFG and OPO sources [82, 83]. In the former of these, low pressure measurements and absolutely stabilized OFCs allowed for reporting line center frequencies with sub-MHz uncertainties. More recently, DCS with spectral interleaving was performed on carbon disulfide CS₂ at 4.6 µm using an OPO source [84], which led to the introduction of this molecule into HITRAN. Similar interleaving schemes have lately been implemented using QCL OFCs at 8 µm for CH₄ [51, 85]. Interleaving solves the issue of the large mode-spacings (several GHz) of QCLs, though the frequency scale was still calibrated to reference absorption spectra in these works.

The sub-nominal OFC-FTS technique provides very high-resolution measurements without needing to resolve individual comb modes, and a frequency accuracy obtained directly from the probing OFC. In addition to a NIR spectrum of CO₂, the first demonstration of sub-nominal OFC-FTS included a measurement of carbon monoxide, CO, around 4.8 µm [8]. The focus there was mainly on the study of the line shape and pressure broadening of one CO line. Here, we applied sub-nominal OFC-FTS for measurements and analysis of entire MIR absorption bands of several molecules in the 3 µm and 8 µm regions. We begin this chapter by outlining the experimental setup, light sources, and analysis methods, before describing the results obtained for the different molecules in the two spectral ranges.
7.1 Experimental setup

The MIR setup is schematically depicted in Figure 7.1, and consisted of two OFCs, two multipass cells and a single FTS that could be used with either OFC source. The MIR OFCs were produced by DFG, yielding idlers around 3 µm and 8 µm, both with a degree of tunability. They will be briefly described in separate sections below. The 3 µm OFC was a home-built system [66] placed on a separate optical table and its light was passed through a single-mode ZBLAN fiber (Thorlabs P3-23Z-FC-5) to the table where the FTS and multipass cells were mounted. The 3 µm beam was coupled into a multipass cell with a 76 m path length (Aerodyne AMAC-76LW). The mostly fiber-based and more compact 8 µm OFC source [67] was placed on the FTS table and coupled through the other multipass cell with a path length of 10 m (Thorlabs, HC10L/M-M02). Due to the DFG process, the \( f_{ceo} \) of both MIR idler OFCs was zero, and thus we only stabilized the \( f_{rep} \) of the pump OFC oscillators. Both sources had repetition rates \( f_{rep} \) of 125 MHz which were detected by fast photo-detectors, and locked to tunable reference frequencies. For the 3 µm source the reference frequency was produced by a direct digital synthesizer (DDS) while for the 8 µm source, a low noise RF-generator was used. Both the DDS and the RF-generator were referenced to a GPS-disciplined Rb clock with a relative stability of \( 10^{-11} \) in 1 s.

![Figure 7.1: Schematic depiction of the MIR OFC-FTS system, comprised of two OFC sources, two multipass cells and an FTS. ABD: auto-balanced detector, BD: balanced detector, BS: beamsplitter, PD: photo-diode, RR: retro-reflector.](image)
7.1.1 The MIR FTS

The FTS was based on the Michelson interferometer and differed from the simplified description in Section 6.1, in that the beams in both arms were reflected back from two retro-reflectors (Edmund Optics 49-675) mounted facing in opposite directions on a translation stage. The use of retro-reflectors simplifies the alignment since the reflected beam is always parallel to the incoming beam, and also displaced horizontally and vertically which allows for easily separating the FTS output beams from the input beam. The length of the translation stage was 75 cm, but since in this configuration the length of both arms was changed in opposite directions as the stage moved, the maximum change in OPD was four times larger, yielding a maximum nominal resolution of \( \sim 100 \text{ MHz} \). The OFC beams were detected in a balanced/auto-balanced configuration and two different sets of detectors were used for the two sources. The two out-of-phase outputs could be directed at either detector pair by flipping a set of mirrors. For the 3 \( \mu \text{m} \) light we used an integrated auto-balanced HgCdTe detector (VIGO System, PVI-4TE-6). The light was focused on the detector diodes by two lenses. The 8 \( \mu \text{m} \) beams were detected in a balanced configuration by two separate HgCdTe detectors (VIGO Systems, PVI-4TE-10.6-1x1) and the detector outputs were subsequently subtracted in a low-noise preamplifier (Stanford Research Systems SR560). In addition to a focusing lens, an adjustable pin-hole was placed in front of each detector to balance the powers and thereby optimize the noise cancelling and improve the contrast of the interference.

A continuous wave (CW) reference laser with \( \lambda_{\text{ref}} = 1563 \text{ nm} \) (long term stability measured to 4 MHz) was coupled to the FTS along a path parallel to the OFC beams, but offset diagonally on the retro-reflector. One of its output beams was detected by a photo-diode and the interferogram was used to calibrate the OPD scale of the OFC interferograms. Both the OFC and the CW interferograms were acquired with a digital oscilloscope (National Instruments, PCI-5922) and a home-written LABVIEW program. For the measurements, the OPD scanning range was set slightly longer than that required to obtain a nominal resolution of 125 MHz, and one interferogram was acquired in \( \sim 3 \text{ s} \).

The FTS was enclosed in a plexi-glas cover with an opening for the free-space OFC beams, while the CW reference laser was coupled in through a fiber. In addition to preventing disturbances due to air currents, the enclosure allowed for purging the spectrometer with dry air in order to reduce interference by atmospheric \( \text{H}_2\text{O} \) absorption, which could cause baseline problems and sometimes interfered with the \( \lambda_{\text{ref}}' \) optimization through the dispersion near \( \text{H}_2\text{O} \) transitions.

7.1.2 3 \( \mu \text{m} \) OFC source

The OFC used for spectroscopy at 3 \( \mu \text{m} \) was a home-built DFG system that was first described in [66], and it is schematically depicted Figure 1 of Paper I where a few modifications were introduced. The pump and the signal for the DFG process were derived from a mode-locked Yb-doped fiber laser (Menlo Systems HP, Orange) generating an OFC around 1 \( \mu \text{m} \) with an \( f_{\text{rep}} \) of 125 MHz, that
was amplified to a power of $\sim 2$ W. A part of the light was passed through a HNLF which produced a soliton Raman-shifted to longer wavelengths. The soliton served as the signal and its center wavelength was tuned by adjusting the power into the HNLF. The pump beam was provided by the remaining 1 µm light passed through a variable delay stage. Fast control of the delay was provided by a piezo-electric transducer (PZT) to which a retro-reflector was glued, while slow adjustments were made with a translation stage holding the retro-reflector mount. The pump and the signal beams were overlapped in a Mg-doped periodically poled lithium niobate (PPLN) crystal (MOP01-0.5-3, Covesion Ltd.) with nine spatially separated poling periods ranging from 28 µm to 32 µm, mounted in a temperature controlled oven. The idler could be tuned between 2.7 µm and 4.2 µm by tuning the soliton wavelength and choosing an appropriate poling period and crystal temperature.

The noise level in the idler can depend rather crucially on maintaining the temporal overlap between the pulses [68], which was stabilized using the method described in [68] and Paper I.

The 3 µm source provided ample and convenient wavelength tuning and considerable optical bandwidth (up to 400 cm$^{-1}$) and MIR power (>100 mW). However, the free-space design had the drawback of requiring regular re-alignment and being relatively bulky.

### 7.1.3 8 µm OFC source

The 8 µm DFG system was designed and built by the Laser & Fiber Electronics Group at Wrocław University of Technology and is described in [67], with additional details about its integration into our spectroscopy setup given in Paper II. It had the particular feature of being mostly fiber-based and hence alignment-free and compact. A schematic of the source is shown in Figure 1 of [67]. All fiber components and electronics required for producing the pump and signal were contained in a table-top dual wavelength source. The pump was a 1.56 µm OFC with an $f_{\text{rep}}$ of 125 MHz generated in an Er-doped fiber oscillator, and the signal was obtained by shifting part of this light towards 2 µm, in a photonic crystal fiber. The pump and signal were coupled out of a common fiber into free space in an external module which housed an orientation-patterned gallium-phosphate (OP-GaP) crystal (BAE Systems), in a temperature stabilized oven.

As described in [67], the source offered significant wavelength tunability, though this required changing between crystals with different poling periods, which affected the alignment and collimation of the MIR beam. In our work we used poling periods of 58 µm and 60 µm. These provided 90 - 100 cm$^{-1}$ of optical bandwidth centered at 1330 cm$^{-1}$ and 1290 cm$^{-1}$ respectively.

The temporal overlap of the pulses was optimized with the help of an adjustable in-fiber delay line. The delay was not stabilized since long term drifts in the delay were rather minor and the measurement noise was dominated by the detector noise in the 8 µm region.

The 8 µm system was unique in combining a wavelength region that is rare for OFCs with compactness and simplicity, being alignment-free, though with more
involved wavelength tuning and more modest MIR power (∼1 mW), compared to the 3 μm source.

7.2 Measurements and analysis methods

Here we describe the general measurement procedure and methods for producing the high-resolution MIR absorption spectra, as well as for their analysis. The latter to some extent depended on the degree of overlap between the measured absorption lines.

In several works, we used the PGOPHER software (Section 7.2.2) to retrieve molecular constants of different vibrational bands. This was possible also in congested spectra where the individual peaks of at least strong lines could be identified using the built-in peak finding function of the software, even though neighboring line profiles were strongly overlapping.

In spectra where individual line profiles could be resolved, we fitted line shape models to these to retrieve their line parameters. Either we applied the fits to one spectrum at a time, which we refer to simply as line fitting, or to the same lines observed in several spectra measured at different conditions, so called multispectral fitting. In too congested spectra, these methods were not applicable, but we could still report the absorption cross section (Section 2.3) of the whole spectrum.

7.2.1 Spectral acquisition and treatment

We measured most MIR molecular spectra at low pressure (sub-mbar to a few mbar) near the Doppler limit, though a few spectra (papers I and V) were acquired at up to 10 mbar. Depending on the absorption linewidth, the number of \( f_{rep} \) steps \( N_{steps} \) was chosen to yield sampling point spacings of 10-15 MHz in the interleaved spectra, corresponding to about 8 sampling points per FWHM. In addition to the \( N_{steps} + 1 \) absorption spectra of the gas sample, we always acquired one background spectrum at the first \( f_{rep} \) step with the multipass cell evacuated. A single background spectrum was sufficient despite the difference in sampling points, since the narrowest features in the background were typically \( H_2O \) lines from the atmosphere with linewidths (∼GHz) much larger than the comb mode spacing. For averaging, we recorded a number (125-400) of interferograms of both the background and \( N_{steps} + 1 \) sample spectra. The interferograms were sampled as described in Section 6.3, Fourier transformed and then averaged. With the intensity distribution of the OFC (the background spectrum) given by \( I_B(\tilde{\nu}) \) and the transmission through the sample given by \( I_T(\tilde{\nu}) \), we calculated the absorption coefficient from Eq. 3.6.

Some residual baseline structure generally remained after background normalization, caused mostly by temporal drifts in the OFC spectral envelope. Prior to interleaving, we therefore modeled the baseline as a polynomial and a series of sine-terms. These were typically fitted together with a model of the absorption spectrum calculated based on data from HITRAN (Papers II and III) or ExoMol (Paper VI). In cases where no spectral model was available, the absorption features were simply masked out (Papers I and V) when fitting the baseline, or
the contour of very congested spectra was modeled using one spectrum where the baseline structure was found to be negligible (Paper IV). The fitted model of the polynomial and sine-terms was subtracted from the data to eliminate the overall baseline structure. However, some local distortions sometimes remained at the positions of strong H₂O background lines from the ambient air.

### 7.2.2 The PGOPHER software

**PGOPHER** is a software for simulating and fitting molecular spectra [86], developed by the late Colin Western at the University of Bristol. It is capable of modeling linear molecules, symmetric and asymmetric top molecules using Hamiltonians with the appropriate forms. Several vibrational bands can be simulated simultaneously and couplings between them can be accounted for by introducing interaction terms. **PGOPHER** is an effective tool for bridging the gap between pure calculations and empirical measurements, allowing for refining ab initio molecular constants to yield better agreement with experiments. Other software packages exist for this purpose but **PGOPHER** is particularly user friendly, as it possesses a graphical interface that displays the simulated bands together with the experimental data that it is compared to. It provides a wide range of functionality, including fitting of individual line positions, line intensities or just the overall contour of a spectrum in cases where the rotational structure is not fully resolved. We used the approach of fitting individual line positions, in Papers I, IV and V. We assigned observed lines to simulated transitions based on visual comparison. By fitting the simulated transition frequencies to the measured line positions, we obtained a gradually more refined set of molecular parameters.

### 7.2.3 Line fitting

Having retrieved the absorption coefficient \( \alpha(\tilde{\nu}) \) in the measured spectral range, individual absorption lines can be modeled using Eq. 2.16. For retrieving parameters of individual transitions in Papers II, III and VI, we fitted a line shape model to the spectra line by line, using the Voigt profile for \( \chi^{abs}(\tilde{\nu}) \) in a home-written MATLAB routine based on the Levenberg-Marquardt algorithm. The fit parameters were the line center wavenumber \( \tilde{\nu}_c \), line intensity \( S \) and the Lorentzian width \( \Gamma_L \), while the Gaussian width \( \Gamma_D \) was fixed to its theoretical value (Section 2.4).

The MIR spectra were all affected by some degree of instrumental broadening. However, we found that this could be accounted for in line fitting by \( \Gamma_L \), with a maximum error of 5% in line intensities. Furthermore, since the broadening was symmetric, it did not affect the retrieved line positions.

In most of the absorption spectra measured, absorption lines frequently partly overlapped. In such cases, we modeled them as a sum of two or more Voigt shapes with their own set of parameters. However, lines that were separated by less than one Doppler FWHM were considered non-resolved (in analogy with the Rayleigh criterion [1]) and hence left out of fitting. Where weak overlapping lines were predicted by HITRAN (Papers II and III) or ExoMol (Paper VI), but their SNR was expected to be too low for fitting, these lines were included in the fit windows.
but with their parameters fixed to the predicted values. Figure 7.2 shows an example of a measured line of N$_2$O from Paper II (black dots) and a Voigt line fit (red line) with fit residuals in the lower panel.

![Figure 7.2: A measured absorption line of N$_2$O shown as black dots and the corresponding Voigt fit as the red line. The fit residuals are shown in the lower panel.](image)

### 7.2.4 Multispectral fitting

In multispectral fitting [87], the line fit is simultaneously applied to the same line (or set of overlapping lines) in several spectra measured at different conditions. The absorption coefficient $\alpha(\bar{\nu})$ is proportional to the pressure $P$ and sample concentration $c_{rel}$ (Eq. 2.16) and a Voigt profile can thus be fitted to several spectra with different values of these parameters, retrieving e.g. one value of the line strength $S$ from all data sets. Also, the Lorentzian width depends on pressure (Eq. 2.30) and so multispectral fitting can be used to directly fit the pressure broadening coefficient $\gamma_i$. Fitting to multiple spectra increases the reliability in the retrieved parameters since they stem from larger data sets. Particularly in cases where certain measurement conditions are difficult to control or measure, multispectral fitting provides extra robustness to such uncertainties compared to using a single spectrum where they would directly propagate to the uncertainty in retrieved parameters.

We retrieved intensities of individual lines using multispectral fitting in Papers I and V, where the gas sample was produced by evaporation of a liquid substance, which made the sample purity (i.e. $c_{rel}$) difficult to determine. We used another adaptation of the Levenberg-Marquardt routine, where in addition to the line intensities $S$, we fit the center frequencies $\bar{\nu}_c$ in the Voigt model, while $\Gamma_D$ was calculated from theory and $\Gamma_L$ was fixed to an average value, since it was dominated by the instrumental broadening. An example of a multispectral fit to a line of iodomethane, CH$_3$I, from Paper I is shown in Figure 7.3, where the data (black) measured at four different sample partial pressures and the corresponding fits (red, blue, purple and turquoise) are shown vertically offset for clarity.
7.3 High-precision spectroscopy at 3 µm

We used the 3 µm OFC source to study two molecules, iodomethane, CH\textsubscript{3}I, and dibromomethane, CH\textsubscript{2}Br\textsubscript{2}, in the atmospheric window at 3 µm - 5 µm. These molecules belong to the so called halogenated volatile organic compounds (HVOCs) that play important roles in atmospheric chemistry as well as in some industrial and agricultural applications, e.g. as fumigant agents [88], and as by-products of water sterilization [89]. Both compounds are harmful to humans, and a strong motivation for spectroscopic studies of these molecules is the development of optical detection techniques for e.g. exposure limit monitoring in work places. Line parameters were not available in HITRAN for neither molecule around 3 µm.

7.3.1 Iodomethane CH\textsubscript{3}I

Figure 7.4(a) shows a high-resolution OFC-FTS spectrum of CH\textsubscript{3}I acquired in a single measurement spanning almost 400 cm\textsuperscript{-1}. It contains the $\nu_1$ (symmetric stretch) and $\nu_4$ (asymmetric stretch) fundamental bands as well as a series of overlapping overtone and combination bands to the left. Zooms of each region are shown in panels (b)-(c). Of the fundamental bands, $\nu_1$ can be of interest due to its high intensity, though it is relatively congested, while more distinct isolated lines can be found in the $\nu_4$ band. Spectroscopic models of the fundamental bands existed [90, 91], but they were based on conventional FTIR measurements with a resolution larger than the Doppler width. There were no available data on line intensities of neither band.
Figure 7.4: (a) A broadband high-resolution spectrum of pure CH₃I measured at 0.03 mbar. (b)-(d) Zoomed in spectra showing the dense rotational-vibrational structures of the different bands. Spectra in panels (b) and (d) were measured at a pressure of 0.11 mbar to increase the absorption signal due to the lower intensity compared to the ν₁ band in panel (c).

**Paper I** was the first demonstration of the sub-nominal technique using the 3 µm OFC source, and focused on the analysis of the ν₄ band of CH₃I. We fitted a Hamiltonian of the ν₄ band and an overlapping hot band to the measured spectrum using PGOPHER. Figure 7.5 shows a comparison between the experiment (black) and the fitted model of the two bands (blue and red), where the model is inverted for clarity. The new fit yielded a clear reduction in residuals compared to previous fits to FTIR data (see Figure 6 of Paper I). We reported the molecular constants of this Hamiltonian model, and it served as the basis for an update in the 2020-version of HITRAN [11], which introduced line parameters of the two analysed bands.

We expanded the analysis to the region around the ν₁ band in **Paper V** where this band was modeled together with the ν₄ band as a common level system coupled by perturbations, and an analogous analysis was applied to two hot bands overlapping with the fundamental bands. This more thorough treatment made maximum use of the wide optical bandwidth of the OFC-FTS measurements. Again, we observed a clear reduction in the Hamiltonian fit residuals compared to the previous treatment of Ref. [91]. We reported molecular constants for all bands included in the analysis.

In Papers I and V we also reported CH₃I line intensities of this region for the first time. We retrieved these by a multispectral fit applied to a selection of lines measured at four different pressures. The relatively congested nature of the bands, particularly ν₁, limited the number of lines suitable for fitting. However
the experimental line intensities for the \( \nu_4 \) region proved useful for the adaption into HITRAN as they provided a basis for scaling the PGOPHER intensities of the remaining transitions [11].

![Figure 7.5](image)

**Figure 7.5:** (a) Measured absorption spectrum of 0.11 mbar of pure CH\(_3\)I (black) and simulations of the \( \nu_4 \) band (blue) and the \( \nu_3 + \nu_4 - \nu_3 \) hot band (red) fitted using PGOPHER. Panel (b) is a zoom showing the presence of a significant number of isolated lines as well as a more congested Q-branch.

### 7.3.2 Dibromomethane CH\(_2\)Br\(_2\)

In **Paper IV** we performed measurements and analysis of dibromomethane, CH\(_2\)Br\(_2\). It was a collaboration with the research group of Julia Lehman at the Universities of Leeds and Birmingham, where they were using an OFC together with a VIPA spectrometer [92], to study the same molecule. The targeted spectral region around 3080 cm\(^{-1}\) was dominated by the \( \nu_1 \) (symmetric stretch) and \( \nu_6 \) (asymmetric stretch) fundamental bands, but as for CH\(_3\)I, there were also overlapping hot bands. The congestion of the spectrum was increased further by the existence of two bromine isotopes, \(^{79}\)Br and \(^{81}\)Br, with nearly equal abundance. The spectral region had been studied before using cavity ring-down spectroscopy (CRDS) with a CW laser [26], where the observed structure of the \( \nu_6 \) band consisting
of superimposed progressions of similar absorption features were assumed to be stemming from the three isotopologues of the molecule CH$_2^{79}$Br$_2$, CH$_2^{79}$Br$^{81}$Br and CH$_2^{81}$Br$_2$. In Paper IV we re-interpreted these progressions as a series of hot bands, and this was more consistent with the observed intensities of the features and isotopic shifts predicted for the molecule. Figure 7.6(a) shows the measured FTS spectrum of the $\nu_6$ band in black and the inverted overall simulation in blue, while Figure 7.6(b) shows a zoom displaying the substructures assigned to the $\nu_6$ band and a series of hot bands. We reported the molecular constants for these bands as well as for $\nu_4$ and its first two overtones that were provided by the Hamiltonian fit.

![Figure 7.6:](image)

**Figure 7.6:** (a) Absorption cross-section measured using the FTS spectrometer (black) together with the overall simulations (blue, inverted), involving twelve vibrational states. (b) Zoomed-in window showing different clusters of transitions corresponding to different bands.

Due to the high degree of congestion of the spectrum, individual line parameters could not be retrieved e.g. by multispectral fitting. However, the reported absorption cross section measured with the FTS and VIPA measurements were in good agreement, and the sample density of the FTS measurement was validated by checking the linearity of a series of measurements at various CH$_2$Br$_2$ partial pressures (Figures 2 and 3 in Paper IV).

### 7.3.3 Conclusions

Papers I, IV and V demonstrated the high potential of sub-nominal OFC-FTS for molecular spectroscopy in the MIR. The high resolution and accurate frequency scale proved useful for improving spectral simulations of CH$_3$I and CH$_2$Br$_2$ and contributed to expanding the scope of the HITRAN database. The high resolution also made line intensity retrieval possible even in the congested spectrum of CH$_3$I.
7.4 High-precision spectroscopy at 8 µm

At 8 µm, we are entering into the greater infrared atmospheric window at 8 - 14 µm. The high potential of spectroscopic applications at these wavelengths together with the fact that OFCs working around 8 µm are still on the pioneering stage, made it a hot target for sub-nominal OFC-FTS. Here, we studied three rather different molecules, nitrous oxide, N₂O, methane, CH₄, and formaldehyde, H₂CO. The first one of these provided a suitable target for proof-of-concept measurements, while the full potential of the technique for high-precision spectroscopy was exploited for the remaining two molecules with their more complicated absorption spectra.

7.4.1 Nitrous oxide N₂O

As a first demonstration of the 8 µm OFC-FTS setup in Paper II, we selected nitrous oxide, N₂O, as the target molecule. It is a linear molecule with regular uncongested absorption bands that are relatively easy to model. We measured the ν₁ fundamental band and the ν₁ + ν₂ − ν₂ hot band, that both occur around 1280 cm⁻¹. Figure 7.7 shows the spectrum of these two bands, where the hot band lines are weaker by about a factor of ten and slightly shifted to higher wavenumbers.

![Figure 7.7: The spectrum of 2.98% N₂O diluted in N₂ measured at a pressure of 0.02 mbar.](image)

The combination of 8 µm OFC-FTS with the sub-nominal method allowed us to retrieve line centers of both vibrational bands with sub-MHz uncertainties. At the time, OFCs had been utilized to study N₂O in this region in one set of studies [93–95], but only for referencing a CW QCL laser. The accuracy of our measurements was confirmed by excellent agreement in observed line centers of the ν₁ band with one of these previous OFC-referenced studies [94]. Figure 7.8(a) shows a comparison of the ν₁ band line centers experimentally retrieved in Paper II (black markers) and Ref. [94] (red markers) to those in HITRAN, which also indicates systematic inaccuracies in the latter. We fitted Hamiltonian models using a Levenberg-Marquardt routine in MATLAB to both bands, which resulted in flat residuals. Figure 7.8(b) compares the results of the Hamiltonian fit to the measured ν₁-band line centers of Paper II and an analogous fit from [94]. The line centers of the ν₁ + ν₂ − ν₂ band showed better agreement with HITRAN, but
the Hamiltonian fit allowed for improving the precision in the band center by one order of magnitude compared to the fit serving as a basis for the HITRAN data.

![Figure 7.8](image-url)

**Figure 7.8:** (a) The line positions of the \( \nu_1 \) fundamental band from Paper II (black markers) and from AlSaif et al. [94] (red markers) relative to those from HITRAN. (b) The \( \nu_1 \)-band line positions obtained by line-by-line fitting (black markers) relative to the simulation based on the Hamiltonian fit from Paper II. The red markers indicate line positions of the Hamiltonian fit in [94], relative to the simulation from Paper II.

### 7.4.2 Methane CH\(_4\)

Having verified the performance of the system in Paper II, we shifted focus to spectroscopy of methane, CH\(_4\), in Paper III. It is a molecule of wide relevance, being a potent greenhouse gas present in the Earth’s atmosphere [96, 97], that has also attracted attention in the realms of astrophysics and study of exoplanets [98, 99], though much of the available CH\(_4\) data in the 8 \( \mu \)m region is still based on traditional FTIR measurements [100, 101]. The energy level structure of CH\(_4\) with near coincidences between wavenumbers of the vibrational modes and their multiples gives rise to congested and irregular absorption spectra consisting of overlapping bands, which calls for accurate models. Due to extensive interactions between the vibrational modes, building the models requires large sets of data of transitions of several bands.

Using both OP-GaP crystals (58 \( \mu \)m and 60 \( \mu \)m poling periods) we probed the spectral range 1250 cm\(^{-1}\) - 1380 cm\(^{-1}\), measuring at a series of pressures and concentrations to cover transitions spanning four orders of magnitude in intensity.
This spectral window contains the $\nu_4$ fundamental band (anti-symmetric bend) overlapping with several hot bands. Figure 7.9(a) shows a spectrum of CH$_4$ diluted to 5% concentration in N$_2$ and measured at 0.04 mbar, while Figure 7.9(b) shows a spectrum of pure CH$_4$ measured at 0.49 mbar, where the now saturated lines of panel (a) are colored in gray.

Figure 7.9: (a) The spectrum of 5% CH$_4$ diluted in N$_2$ at a total pressure of 0.04 mbar. The data stem from two measurements with different spectral coverage, stitched together at 1320 cm$^{-1}$. (b) The spectrum from two measurements of pure CH$_4$ at 0.49 mbar, stitched together at 1320 cm$^{-1}$. The analyzed lines are shown in black, while the gray lines depict the stronger lines shown in the upper panel, which are saturated at these conditions. The negative excursions in the baseline are distortions caused by H$_2$O lines in the ambient air.

We retrieved line centers of over 800 lines, and using HITRAN as reference, we assigned them to the $\nu_4$, $2\nu_4 - \nu_4$ and $\nu_4 + \nu_2 - \nu_2$ bands of the main isotopologue $^{12}$CH$_4$ as well as the $\nu_4$ band of $^{13}$CH$_4$. We compiled this data into a line list, with line center uncertainties (1σ) ranging from 200 kHz to 2.3 MHz. We also reported line intensities for most of these transitions, omitting only those measured on gas samples diluted with N$_2$ due to high uncertainty in the CH$_4$ concentrations.

Comparisons of retrieved line centers to HITRAN yielded various degrees of discrepancies for different bands but they were always significant, i.e. exceeding the experimental uncertainties (see Figure 3 in Paper III). The line positions were included into a global model of the CH$_4$ spectrum based on the polyad formalism [102], that was performed by the group of Vincent Boudon at the University of Burgundy. Most of the transitions replaced previous data, while 40 transitions were included for the first time. Figure 7.10(a) shows the residuals for the $^{12}$CH$_4$ $\nu_4$ band of the global fit, and 7.10(b) is a zoom around the new OFC-FTS data.
marked by the red crosses. The clear reduction in both the experimental errorbars and the scatter of the residuals for these transitions compared to the remaining absorption (blue circles) and emission data (green diamonds) is easily appreciated.

Figure 7.10: (a) The residuals of the global fit to the $^{12}$CH$_4$ $\nu_4$-band line centers obtained in Paper III (red crosses) as well as from previous absorption (blue circles) and emission (green diamonds) measurements. (b) Detailed view of the spectral range including the lines reported in Paper III, where the previous emission measurements have been omitted for clarity. The reduced uncertainty of the OFC-FTS measurements is evident.

The CH$_4$ measurements was also where the wavenumber dependence of the optimum reference wavelength $\lambda_{ref}$ became significant. It became even more pronounced after switching to the crystal with 58 $\mu$m poling period. The crystal change affected both the collimation and alignment of the MIR beam, and the effect on $\lambda_{ref}$ can likely be traced to either of these properties. This called for the development of the correction procedure described in Appendix A of Paper III and outlined here in Section 6.3.5. It proved successful at achieving the OFC to FTS scale matching with a single value of $\lambda'_{ref}$ over a larger spectral range than would
otherwise have been possible, though after the crystal change, the correction still had to be applied segmentwise. The reliability of the method was confirmed by the cancellation of the wavelength dependence of $\lambda'_{\text{ref}}$ as exemplified by Figure A.3 of Paper III, and by agreement in retrieved line centers at the intersections between different spectral segments.

### 7.4.3 Formaldehyde $\text{H}_2\text{CO}$

In Papers II and III we established that our 8 $\mu$m OFC-FTS measurements can contribute to improving the accuracy of the HITRAN database. Data for formaldehyde, $\text{H}_2\text{CO}$, is missing from HITRAN in the region from 6 $\mu$m to 20 $\mu$m, which made it the target of Paper VI. Formaldehyde bears relevance for our everyday life, being a common pollutant, known for its adverse health effects [103]. It can also be of interest e.g. in the study of exoplanets as a potential biomarker [104].

Previous absorption measurements of $\text{H}_2\text{CO}$ around 8 $\mu$m have been performed mainly with conventional FTIR and previous studies have also been aimed at creating Hamiltonian models of the molecule [105, 106]. This is a challenging task since, similarly to $\text{CH}_4$, the spectrum of $\text{H}_2\text{CO}$ is complex and congested with several overlapping bands that are strongly coupled to each other. High-accuracy broadband spectra provided by OFC-FTS are therefore a valuable addition to the bulk of data, facilitating Hamiltonian fitting and contributing to filling the present gap in HITRAN.

Formaldehyde is difficult to store in gas form and all samples had to be synthesized on site from solid paraformaldehyde shortly before the measurements, though they could be stored for some time at cryogenic temperatures, as described in Paper VI.

We performed measurements between 1250 cm$^{-1}$ and 1380 cm$^{-1}$ (using the crystal with 58 $\mu$m poling period) at different $\text{H}_2\text{CO}$ partial pressures and retrieved over 1000 line centers and intensities. Figure 7.11 shows a spectrum of $\text{H}_2\text{CO}$ measured at 0.02 mbar (a), and 0.58 mbar (b). Again, lines saturated at the higher pressure are colored gray for clarity.

Here the observed transitions could not be assigned using HITRAN as for the previous two molecules due to the gap in the database. Instead we relied mostly on a MARVELized (Section 4.3) line list of $\text{H}_2\text{CO}$ [107], which proved accurate enough to enable unambiguous assignments, though it only contained 638 of the measured transitions. However, fitting an effective Hamiltonian in PGOPHER to the MARVEL data provided additional predictions that could be matched to more lines observed in the experiment.

In total, we assigned 747 measured lines to the $\nu_3$, $\nu_4$ and $\nu_6$ fundamental bands, as well as four hot bands, and reported their experimental line centers with 1$\sigma$ uncertainties from sub-MHz to a few MHz, as well as intensities. Analysis of the intensities was somewhat challenging due to uncertainties in sample concentrations stemming from the synthesis process and the instability of gas-phase $\text{H}_2\text{CO}$ that caused the density of the sample to change after introducing it into the multipass cell. The varying sample density could be monitored as drifts in the gas pressure, and these drifts were found to be negligible for the measurement at
the highest pressure. We used the line intensities retrieved from this measurement as a standard to scale line intensities from the remaining measurements. This was possible due to a certain overlap in observed absorption lines between the series of measurements. We conservatively estimated the $1\sigma$ uncertainty in line intensities to 15%.

Although the agreement of the experimental line centers with the MARVEL line list was sufficient for assignment purposes, a closer comparison (see Figure 3 in Paper VI) revealed both random and systematic discrepancies that were outside the experimental uncertainties. This indicated that the accuracy of the line list could be improved. The MARVEL optimization was repeated by our collaborators in the ExoMol group at University College London, including the new 747 experimental line centers. This resulted in an updated MARVEL line list containing 82 new energy levels and over 5000 new transitions. Considering that the accuracy of the previous MARVEL list was found to be quite good, the addition of these transitions represents a significant improvement. Figure 7.12(a) shows the changes in energy levels resulting from rerunning the MARVEL algorithm with the new data. Apart from the energy levels directly probed in the experiment (black markers), inclusion of the OFC-FTS data also affects a large number of other levels (gray markers) through the spectroscopic network. The uncertainties in the directly probed levels were reduced by up to two orders of magnitude, but significant reductions occurred for other levels as well, as seen in Figure 7.12(b), displaying the uncertainty ratios before and after including the OFC-FTS data.
Figure 7.12: (a) The difference between the energy levels before and after revising the MARVEL analysis with the line positions from Paper VI. The revised energies of the levels directly probed in the OFC-FTS measurements (black markers) propagate via spectroscopic networks to a multitude of other levels (gray markers). (b) The ratios of energy uncertainties before and after revising the MARVEL analysis with the line positions from Paper VI. Uncertainties of levels directly probed (black markers) are reduced by up to two orders of magnitude, while many other levels (gray markers) underwent substantial uncertainty reductions because of their links to probed levels through spectroscopic networks.

7.4.4 Conclusions

Papers II, III and VI served to establish 8 µm OFC-FTS as a useful tool for providing highly accurate experimental center frequencies of large numbers of transitions. These data were successfully used to improve models, in even larger spectral windows than those covered by the measurements, and can contribute to improving the accuracy and the spectral coverage of spectroscopic databases in the MIR.
8 Optical frequency comb optical-optical double-resonance spectroscopy

OODR spectroscopy allows for probing transitions and energy levels crucial for modeling high-temperature spectra. Here, the molecule of interest was again methane, CH$_4$, which is relevant in this context due to its role in combustion processes and its presence on the class of so-called hot-Jupiter exoplanets [98]. Calculated spectral parameters for methane at the relevant temperatures are available both from the TheoReTS [108], and ExoMol [109] databases (Section 4.2). Until recently, experimental verification of the predicted highly excited energy levels was still lacking.

OODR spectroscopy of CH$_4$ had been performed before using nano-second pump and probe lasers reaching the 3$\nu_3$ region [110–112], and with tunable diode lasers [113] probing the 2$\nu_3$ region. These methods suffered from limited resolution and spectral coverage respectively. The broad bandwidth of OFCs would be an advantage in the search for previously undetected transitions for which the accuracy of predictions is unverified. OODR spectroscopy of CH$_4$ using a CW pump laser and an OFC probe laser was recently realized [12, 13], pumping the $\nu_3$ fundamental band and detecting 36 ladder-type transitions from the $\nu_3$ vibrational state to the 3$\nu_3$ region. This verified the accuracy of TheoReTS while the agreement with ExoMol was found to be significantly worse. The sample was contained in a single-pass cell that was cooled with liquid N$_2$ to increase the population in the lower states of the pump transitions. The setup was later modified by introducing an enhancement cavity, removing the need for cooling, while improving the sensitivity [14]. Another advantage of the cavity is that the probe is propagating in both directions relative to the pump beam in the cavity. This improves the frequency accuracy of the OODR transitions since a detuning of the pump wavelength from the center of the pumped transition leads to the probe addressing slightly different velocity groups for the two propagation directions. This manifests itself as a broadening of the OODR transitions, rather than a frequency shift, as in the case where the probe makes a single pass through the sample.

Here, we upgraded the setup with a new pump laser with narrower linewidth and a new enhancement cavity for the CH$_4$ sample. The new cavity had mirrors with lower dispersion than previously, and the FSR was now matched to $f_{rep}$ as opposed to previously when every third comb mode was resonant with the cavity. These improvements served to increase the optical bandwidth and power of the probe spectrum transmitted through the sample. We also implemented a different OFC-cavity locking scheme that simplified averaging. The upper energy levels probed in the previous OODR OFC work on CH$_4$ had angular momentum quantum numbers $J \leq 4$. We now extended the probed $J$-number range, by pumping $\nu_3$ transitions with lower rotational quantum numbers $J = 7$. 
8.1 Experimental setup

The experimental setup is shown schematically in Figure 8.1. Its main components were a MIR CW pump laser, an NIR OFC probe laser that was absolutely stabilized and locked to a cavity containing a CH$_4$ gas sample, and an FTS used for recording the spectrum of the probe transmitted through the cavity.

![Figure 8.1: Schematic of the OODR OFC-FTS setup, divided into three main parts; An absolutely stabilized OFC probe locked to an enhancement cavity, a CW pump laser stabilized to a CH$_4$ transition and an FTS for analyzing the probe spectrum. ABD: auto-balanced detector, BS: beamsplitter, DM: dichroic mirror, FD: fast detector, MS-SF: micro-structured silica fiber, PD: photo-diode, PZT: piezo-electric transducer, RR: retro-reflector.]

The probe was an Er:fiber OFC (Menlo Systems, FC1500-250-WG) emitting around 1.5 µm with an $f_{rep} = 250$ MHz, that was passed through a micro-structured silica fiber (MS-SF) where it was shifted to longer wavelengths. By adjusting the input power to the MS-SF, the probe center wavenumber was tuned between 5900 cm$^{-1}$ and 6000 cm$^{-1}$ depending on the measurement. The OFC light then passed through an optical circulator used for PDH locking (see below) and was mode-matched and coupled to a 60 cm long Fabry-Pérot cavity with an FSR matched to the $f_{rep}$ of the OFC. Tunability of the cavity length was provided by a PZT at one cavity mirror. To simplify long term averaging by avoiding drifts of the comb mode frequencies, we used a comb-cavity locking scheme where both $f_{ceo}$ and $f_{rep}$ were absolutely stabilized (compare discussion in Section 5.4). The comb modes were locked to the cavity modes using the PDH method (Section 5.4) at only one locking point. The PDH error signal was obtained from detecting a part of the cavity reflection that was picked off with the optical circulator, and then dispersed by a grating. The PDH lock was implemented feeding back to $f_{rep}$. 
while the $f_{ee0}$ was measured with a f-2f interferometer and separately stabilized to a DDS-generated reference frequency that optimized the coverage and shape of the transmitted spectrum. Absolute stabilization and stepping of $f_{rep}$ was achieved by detecting it with a fast detector and feeding back to the Fabry-Pérot cavity length to maintain $f_{rep}$ at a fixed reference frequency synthesized by a second DDS. This reference frequency was incrementally scanned during the measurements to tune $f_{rep}$ for spectral interleaving (see Section 8.2.1). Both DDS units were referenced to a GPS-disciplined Rb clock (the same as in Section 7.1).

The OFC beam transmitted through the cavity was coupled through a fiber to the FTS. The FTS differed from the MIR FTS (Section 7.1.1) in that the translation stage was shorter (35 cm) since the required nominal resolution was 250 MHz. The CW reference laser was a frequency stabilized HeNe laser with $\lambda_{ref} = 633$ nm. The CW laser was detected by a photo-diode and the OFC by a home-built auto-balanced detector. The OPD was scanned slightly more than the $\Delta l_{\text{max}}$ required for a nominal resolution of 250 MHz, and the acquisition time of one interferogram was $\sim 3$ s.

The pump laser was generated by a singly resonant OPO seeded by a narrow linewidth diode laser (CW-OPO, TOPTICA, TOPO). The OPO provided up to 2 W of power and was tunable between 1.5 $\mu$m and 4 $\mu$m. In order to stabilize the pump wavelength to a particular CH$_4$ transition, a few tens of mW of the pump beam was split off and passed to a 30 cm long gas cell containing CH$_4$. The beam made two passes through the cell, giving rise to a Lamb-dip [1] in the back-reflected light when it was on resonance with a CH$_4$ transition. Part of the reflection was detected with a fast detector. The pump wavelength was locked to the Lamb-dip using an error signal produced by modulating the frequency of the diode laser at 23 MHz and demodulating the detected absorption signal at the same frequency [20].

The remaining part of the pump beam was overlapped with the OFC beam using a dichroic mirror and coupled to the cavity. For optimum beam overlap between the pump and probe, the collimation of the pump was adjusted to have the same beam waist position and Rayleigh range [28] as the probe. The cavity transmission at the pump wavelength was $\sim 95\%$ and it essentially made a single pass through the sample.

The relative polarization angle between the pump and probe was set to $54.7^\circ$, in order to eliminate polarization dependent effects on the intensities of ladder-type transitions with different $\Delta J$. The pump beam could be blocked by a shutter to record background spectra without the OODR transitions.
8.2 Measurements and analysis methods

In the OODR experiment, the NIR probe spectra contain both a background consisting of Doppler broadened overtone and combination band transitions originating from the ground vibrational state, and the sub-Doppler OODR features in the form of ladder-type hot band transitions, and v-type dips appearing on certain overtone/combination band lines that share their lower state with the pumped transition (recall Figure 3.2). The required spectral resolution is determined by the width of the OODR transitions and fitting these require different models than the Voigt profile used previously. Below we describe the methodology particular to the acquisition and analysis of the OODR measurements.

8.2.1 Spectrum acquisition and treatment

All OODR spectra were measured on a sample of pure CH\textsubscript{4} at a pressure of 200 mTorr. To resolve the sub-Doppler OODR features with a HWHM around 10 MHz, the sampling point spacing in the interleaved spectra had to be much smaller than in the Doppler broadened MIR measurements. The number of \( f_{\text{rep}} \) steps, \( N_{\text{steps}} \), was hence 129 corresponding to steps of \(~2\) MHz in the optical domain. Here, the Doppler broadened CH\textsubscript{4} lines in the background had linewidths (HWHM \(~270\) MHz) comparable to \( f_{\text{rep}} \), and the spectral envelope of the OFC drifted relatively quickly. Therefore, one background spectrum was recorded for each \( f_{\text{rep}} \) step. We scanned \( f_{\text{rep}} \) over the series of 130 values (\( N_{\text{steps}} \) + 1), recording one background and one OODR spectrum consecutively at each step, by closing and opening the shutter. For averaging, we repeated the \( f_{\text{rep}} \) step sequence and acquisition series a number (8-10) of times in alternating directions.

We treated the background and OODR spectra with the sub-nominal routine (Section 6.3). For each \( f_{\text{rep}} \) step, we removed the baseline remaining after normalization with a method similar to the cepstral analysis described in [114], where baseline structures are retrieved from the OPD domain by taking the FFT of the region around zero-OPD. Figure 8.2 shows the interleaved OODR (pump un-blocked) and background (pump blocked) spectra separately. Both spectra contain the Doppler broadened background due to overtone and combination bands, but the pump laser induces the appearance of a sub-Doppler ladder-type transition in the OODR spectrum around 5844.16 cm\(^{-1}\).
Figure 8.2: Interleaved CH₄ spectrum of the OODR (red) and the background (blue). The OODR spectrum displays a ladder-type transitions at 5844.16 cm⁻¹, in addition to the Doppler broadened background lines.

For studying the ladder-types, we normalized the OODR spectra to the background spectra. For initial detection of ladder-type transitions, we interleaved the 130 spectra in the whole measured spectral range. For line fitting, we then interleaved individual spectral segments around the identified ladder-type lines, since the baseline correction was more effective in smaller windows.

Quantitative comparisons of the ladder-type intensities to predictions can be done by normalizing them to the v-type intensities (recall Section 3.2.1). The overtone and combination band transitions displaying v-types could be determined from HITRAN as those originating from the same state as the pump transitions. Spectral segments were interleaved for these, similarly as for the ladder-type transitions. Here we did not do background normalization, as this would eliminate the Doppler broadened lines (recall Figure 3.2(b)).

8.2.2 Line fitting

The transmission through the cavity in the OODR (optically pumped) spectra was modeled with Eq. 3.3 where, apart from the sub-Doppler OODR transitions, the spectrum also contains the absorption and dispersion due to the Doppler-broadened background, which was modeled from HITRAN. To account for the background normalization, we divided the OODR transmission by the transmission calculated with only the Doppler-broadened background.

For each detected ladder-type line, we fitted the model described above in 2 GHz wide segments around the line. We modeled the ladder-type transition as a sum of a Lorentzian and a broader Gaussian with a common center frequency. The Gaussian component was due to redistribution of the pumped population into other velocity groups through elastic collisions. We floated the ladder-type center frequency, the widths of the Lorentzian and Gaussian components, the integrated absorption coefficient of the Lorentzian (Eq. 2.19), as well as the comb-cavity
phase offset $\Delta \Phi$ at the position of the ladder-type. For lines with lower SNR, where the Gaussian component was not clearly above the noise level, we fixed the Gaussian width and the ratio of the integrated absorption of the Gaussian to that of the Lorentzian. The values of the Gaussian width and proportionality constant for the Gaussian-to-Lorentzian integrated absorption were found as the mean values for a set of strong ladder-type lines. Figure 8.3(a) shows a measured ladder-type transition (black) and the fit (red), with residuals in the bottom panel.

The v-type features appeared at the bottom of Doppler broadened overtone and combination band lines (recall figure 3.2). We used the same model for the v-type dip as for the ladder-type lines, only inverted, and modeled the Doppler-broadened transition as a Gaussian with the theoretically calculated Doppler width. Here, we fixed the widths of the v-type components to the mean values of the ladder-type lines, leaving only a common center frequency of the v-type and Doppler broadened transition as well as their individual integrated absorptions as fitting parameters. Figure 8.3(b) shows a measured transition displaying a v-type (black) and the fit (red). A zoom around the v-type feature is shown in the inset, with residuals in the lower panel.

![Figure 8.3:](image)

Figure 8.3: (a) A measured ladder-type line (black) and fit (red) with residuals at the bottom. (b) A measured Doppler broadened line displaying a v-type (black) and a fit (red). The inset shows a zoom around the v-type feature with residuals.

8.2.3 Combination differences

In an OODR experiment, it is possible to reach the same final states through ladder-type transitions from different pumped intermediate states. Here, we used a pump laser around 3000 cm$^{-1}$ and a probe around 5900 cm$^{-1}$ - 6000 cm$^{-1}$ and possible combinations of pump and probe transitions are shown schematically in Figure 8.4. Three intermediate states are pumped through P, Q and R-transitions originating from the same $J'' = 7$ level of the vibrational ground state. The populated intermediate levels thus have $J' = 6 - 8$, and the final states addressed
by the probe for the three pump transitions cover $J$-numbers from 5 to 9. As can be appreciated from the figure, final states with $J = 7$ can be reached from all three pumped levels, while final states with $J = 6$ and $J = 8$ can be reached for the $(P,Q)$ and $(Q,R)$ pairs of pump transitions respectively.

**Figure 8.4:** Illustration of combination differences observed in an OODR experiment. Three intermediate levels are pumped through $P(7)$ (orange solid arrow), $Q(7)$ (orange dashed arrow) and $R(7)$ (orange dotted arrow) transitions from the vibrational ground state with $J'' = 7$. Probe transitions from the intermediate states with $J' = 6$ (red solid arrows), $J' = 7$ (red dashed arrows) and $J' = 8$ (red dotted arrows) then reach final levels with $J = 5 - 9$, where levels with $J = 6 - 8$ are reached for more than one pump transition.

Probing the same upper levels from different lower levels is an example of combination differences, and the probed transition wavenumbers depend only on the differences in the lower energy levels. In this case, it is also useful for assigning observed ladder-type transitions, since if the pumped intermediate energy levels are well known, one can calculate the energies of the probed final states. Identifying final states common between different pump transitions then in principle allows for deducing the $J$-number of these states, as shown in Figure 8.4. However, the possibility that certain probe transitions are not observed due to e.g. low intensity or overlap with other spectral features means that some caution must be exercised when assigning final $J$-numbers. Nevertheless, combination differences can impose restrictions on the possible final state $J$-numbers which are useful for line assignment.
8.3 OFC OODR spectroscopy of CH$_4$

In Paper VII we pumped three transitions P(7,A$_2$), Q(7,A$_2$) and R(7,A$_2$) of the $\nu_3$ fundamental band of CH$_4$ originating from the same lower state. For each pump transition, we probed the sample with the OFC in regions spanning 140 cm$^{-1}$ - 200 cm$^{-1}$ depending on the measurement. We adjusted the center of the OFC spectrum in each measurement, to probe hot band transitions reaching final states common between the measurements. For CH$_4$, more highly excited rotational levels possess larger numbers of sub-levels leading to a larger number of observed ladder-type transitions.

Using line fitting, we retrieved line parameters of a total of 88 ladder-type hot band transitions from the three measurements, with line center uncertainties (1$\sigma$) ranging between 360 kHz and 2 MHz. A comparison of these to the predictions from TheoReTS is shown in Figure 8.5(a)-(c) where the positions and integrated absorption of ladder-type transitions observed for the three pump transitions are shown as black vertical lines, and the TheoReTS predictions are plotted inverted in red. The predicted intensities are arbitrarily scaled to match the measurements. We calculated the final state energies plotted on the upper x-axis by summing the ladder-type transition frequencies to the energies of the pumped intermediate states. The latter we calculated from transition frequencies of the $\nu_3$ band published in [115, 116], and a lower state energy value of the pumped transitions obtained directly from the authors of these publications.

![Figure 8.5](image)

**Figure 8.5:** (a) – (c) The ladder-type transitions (black) detected when pumping the P(7,A$_2$), Q(7,A$_2$) and R(7,A$_2$) transitions compared to the predictions from TheoReTS (red), where the predicted intensities are scaled to match the experiment and plotted inverted.

We determined 19 final states common between two or more measurements, by identifying groups of probed final state energies with a mutual agreement three orders of magnitude smaller than the minimum energy difference between any group and the nearest observed level. For all but one of these combination differences, the two or three final state energies all agreed with their weighted mean to within their respective 3$\sigma$ uncertainties. Based on the common upper levels,
the \( J \)-numbers of 45 observed transitions could be restricted according to Section 8.2.3.

We could assign 79 of the 88 hot band transitions to TheoReTS [38], by using the restrictions in \( J \)-numbers imposed by the combination differences, and observed patterns in center frequencies and intensities common between the measurements and predictions. Figure 8.6(a) shows a comparison of the retrieved center frequencies of the assigned ladder-type transitions to those predicted by TheoReTS. The experimental uncertainties are negligible on the displayed scale. The scatter in the discrepancies is comparable to what was seen for lower \( J \)-numbers [12–14], but here there is a systematic offset between the experiment and predictions, which was not observed previously. This indicates the presence of systematic errors in the predictions for higher rotational levels.

![Figure 8.6](image_url)

**Figure 8.6:** (a) The differences between the retrieved center wavenumbers of the assigned ladder-type transitions observed for the three pump transitions and the predictions from TheoReTS. (b) The ratios of the experimental to predicted normalized intensities for the three pump transitions.

To be able to compare the intensities of the hot band transitions to the TheoReTS predictions, we fitted one \( v \)-type feature in each measurement. The intensity of the \( v \)-type dips were assumed to be proportional to the population transferred to the intermediate \( \nu_3 \) states, as discussed in Section 3.2.1. We calculated normalized intensities \( I_N \) of the ladder-type transitions by normalizing their retrieved integrated absorption to the integrated absorption of the fitted \( v \)-type feature of the respective measurements. We calculated the theoretical normalized intensities from data obtained from HITRAN for the transitions displaying the \( v \)-types, and
from TheoReTS for the ladder-type transitions. The ratios of the experimental normalized intensities to the calculated ones are shown in figure 8.6(b). These are centered around 1, as expected but the discrepancies are still well outside the experimental uncertainties.
9 Conclusions

We have demonstrated new implementations of OFC-FTS with sub-nominal resolution, in order to provide highly accurate spectroscopic data necessary for validation and development of molecular spectroscopic databases. The broadband high-resolution absorption spectra allowed for compiling extensive line lists, but also for fitting Hamiltonian models and retrieving molecular parameters.

Around 3 µm, we targeted the HVOCs iodomethane, CH$_3$I, and dibromomethane, CH$_2$Br$_2$. The broadband data covering several bands of each molecule proved very suitable for spectral PGOPHER simulations. We demonstrated the spectrometer setup for CH$_3$I in Paper I and expanded the analysis in Paper V to make full use of the broad spectral coverage of the measurements. The spectral resolution was improved compared to previous work, and the results of the Hamiltonian fit yielded significantly reduced residuals compared to the previously reported models. The high resolution also allowed for applying multispectral fits to selected absorption lines, and reporting line intensities of these for the first time. For CH$_2$Br$_2$ we presented new measurements in Paper IV with improved resolution compared to before, as well as a new more self-consistent interpretation of the observed absorption spectrum. The spectrum was too congested for fitting of individual absorption lines, but the measured absorption cross section agreed well with that measured by our collaborators using an OFC and VIPA spectrometer.

In Paper II we demonstrated an OFC-FTS system which is noteworthy not least because of the compact and practical OFC source emitting in the hard-to-access 8 µm region. The performance of the setup was validated by good agreement in retrieved line center frequencies with another recent highly accurate work on nitrous oxide, N$_2$O.

We exploited this potential for precision spectroscopy in Papers III and VI where we studied methane, CH$_4$ and formaldehyde, H$_2$CO. While the spectra of these molecules were complex, they were less congested than for the HVOCs and we therefore aimed at retrieving line parameters for as many transitions as possible. We measured at a range of sample partial pressures and compiled line lists containing center frequencies and intensities for many hundreds of transitions of both molecules. The precision in line center frequencies was higher by about one order of magnitude compared to previously available data, with 1σ uncertainties as low as ~ 100 kHz. The frequency accuracy of the measurements was further supported by the flat residuals observed for our CH$_4$ data when it was included in a global Hamiltonian model. For H$_2$CO, we validated the accuracy of the line list produced by the MARVEL procedure, and our data was also used for improving the MARVEL line list in terms of precision and number of transitions.

In Paper VII, we improved on a previously demonstrated setup for OODR spectroscopy using a MIR CW pump laser and an NIR OFC probe laser. The new enhancement cavity allowed for transmitting a broader bandwidth and higher power of the OFC, and an updated OFC-cavity locking scheme simplified long-term averaging. We targeted more highly rotationally excited states than was reported previously, pumping three transitions of the P, Q, and R-branches of the $\nu_3$ band originating from the same state with $J = 7$. In the three measurements,
we experimentally detected 88 hot band transitions to final states with $J = 5 - 9$ for the first time. The possible upper level $J$-numbers could be restricted for 45 lines with common final states, and this together with correspondences in center frequencies and intensities allowed for assigning 79 observed lines to transitions predicted by the TheoReTS database. Unlike what was reported previously for transitions with lower $J$-numbers, we found a systematic offset between observed and predicted transition frequencies.

Our work in the MIR has contributed accurate data important in the quest for perfecting spectroscopic models. The Hamiltonian fit to CH$_3$I in Paper I already proved useful for expanding the scope of HITRAN. Apart from their use in the development of databases, the extensive line lists produced in the 8 µm region can be used as direct reference data in developing and validating other MIR spectroscopic techniques [117]. Moreover, we have firmly established sub-nominal OFC-FTS as a platform for future work in the important MIR region. The study of CH$_4$ and H$_2$CO could be expanded in the range accessible by the 8 µm source, and the updated MARVEL analysis of H$_2$CO demonstrated that new data can have an impact in a wider spectral range than that measured, by virtue of the spectroscopic networks. Also, numerous other molecules e.g. ethane C$_2$H$_6$, ethylene C$_2$H$_4$ and ammonia NH$_3$ would be of interest here. Further possible developments of the setup could be detectors with lower noise at 8 µm that would reduce acquisition times.

The NIR OODR OFC-FTS system employed in our work on CH$_4$ proved highly effective at detecting large numbers of OODR transitions. Although theoretical predictions of the measured hot band transitions are less accurate than for e.g. the MIR fundamental bands, assignment of measured transitions is possible by exploiting combination differences observed between measurements at different pump transitions. The large amounts of experimental data obtained with this setup has a great potential for further improving spectral models for molecules relevant for high-temperature applications. Also, implementing OODR spectroscopy using either a 3 µm or 8 µm OFC probe would be an intriguing next step.
10 Summary of papers

Paper I
Line positions and intensities of the $\nu_4$ band of methyl iodide using mid-infrared optical frequency comb Fourier transform spectroscopy
Ibrahim Sadiek, Adrian Hjältén, Francisco Senna Vieira, Chuang Lu, Michael Stuhr, Aleksandra Foltynowicz.

We implemented the sub-nominal method in the MIR using a 3 $\mu$m DFG-based OFC source. We measured high-resolution spectra of iodomethane CH$_3$I between 2800 cm$^{-1}$ and 3160 cm$^{-1}$ covering several vibrational bands, and achieved a clear improvement in spectral resolution compared to previously available data.

We fitted a Hamiltonian model of the $\nu_4$ fundamental band and one hot band to the data using PGOPHER, and reported updated molecular constants. The model was the basis for an update in HITRAN2020 which added data for these bands. We also retrieved line intensities of selected $\nu_4$ transitions, which had not been reported previously.

I contributed to setting up the spectrometer for use with the 3 $\mu$m OFC source and measuring the spectra. I was responsible for interleaving the spectra and fitting the line intensities, and wrote a part of the manuscript.

Paper II
Optical frequency comb Fourier transform spectroscopy of $^{14}$N$_2^{16}$O at 7.8 $\mu$m
Adrian Hjältén, Matthias Germann, Karol Krzempek, Arkadiusz Hudzikowski, Aleksander Głuszek, Dorota Tomaszewska, Grzegorz Soboń, Aleksandra Foltynowicz.

We demonstrated OFC-FTS with sub-nominal resolution based on a DFG OFC source operating in the important spectral region around 8 $\mu$m. We measured the $\nu_1$ band of N$_2$O and the $\nu_1 + \nu_2 - \nu_3$ hot band, and retrieved line center frequencies with 1$\sigma$ uncertainties on the order of 100 kHz, which was a significant improvement compared to HITRAN.

The line center accuracy was confirmed by good agreement with another highly accurate measurement. Fitting a Hamiltonian to the two vibrational bands also allowed for improving the precision in previously reported band centers.

I participated in adapting and optimizing the spectrometer for 8 $\mu$m and making the measurements. I interleaved the spectral data, retrieved the line positions, fitted the Hamiltonian models and wrote parts of the manuscript.

Paper III
A methane line list with sub-MHz accuracy in the 1250 to 1380 cm$^{-1}$ range from optical frequency comb Fourier transform spectroscopy

81
Here, we measured a series of spectra of CH$_4$ at different pressures around 8 µm and retrieved center frequencies of 843 lines and line intensities of 678 lines belonging to several bands and two isotopologues.

The line center uncertainties were smaller than for the previously available data, and adding the retrieved line centers to a larger data set used for a global Hamiltonian fit revealed a significant reduction in residuals for the OFC-FTS data.

My contribution consisted of participating in the measurements, interleaving the spectra, fitting the line centers and intensities, and writing parts of the paper. I also developed the method used for handling the wavenumber dependence of the optimum reference laser wavelength.

**Paper IV**

*Optical frequency comb-based measurements and the revisited assignment of high-resolution spectra of CH$_2$Br$_2$ in the 2960 to 3120 cm$^{-1}$ region*

Ibrahim Sadiek, Adrian Hjältén, Frances C. Roberts, Julia H. Lehman, Aleksandra Foltynowicz.

*Physical Chemistry Chemical Physics, 25, 8743–8754 (2023)*

We used the 3 µm DFG OFC to measure the spectrum of CH$_2$Br$_2$ and made a combined analysis with measurements made using an OFC and a VIPA spectrometer.

The measurements covered several overlapping vibrational bands, and the absorption cross-sections retrieved with the two methods were in good agreement.

The region around the $\nu_6$ band was simulated in PGOPHER with a revised interpretation of the spectrum, where observed progressions of features overlapping with $\nu_6$ were ascribed to hot bands rather than different isotopologues of the molecule.

I took part in the OFC-FTS measurements, performed the spectral interleaving and comparison of the absorption cross-sections. I also wrote parts of the manuscript.

**Paper V**

*Line positions and intensities of the $\nu_1$ band of $^{12}$CH$_3$I using mid-infrared optical frequency comb Fourier transform spectroscopy*

Adrian Hjältén, Aleksandra Foltynowicz, Ibrahim Sadiek.


We expanded the analysis of the spectrum presented in Paper I to cover the region around the $\nu_1$ band. We simulated the $\nu_1$ band coupled with the $\nu_4$ band in a common system, to obtain a more complete model than previously. An analogous approach was applied to two hot bands. We reported molecular constants for the involved bands, and also line intensities for a selection of lines in the $\nu_1$ region.

I interleaved the spectra, fitted the line intensities, and wrote the corresponding sections of the paper.
Paper VI
Optical frequency comb Fourier transform spectroscopy of formaldehyde in the 1250 to 1390 cm$^{-1}$ range: experimental line list and improved MARVEL analysis
Journal of Quantitative Spectroscopy & Radiative Transfer, 312, 108782 (2024)

Using the 8 µm DFG source, we measured the spectrum of H$_2$CO, which we synthesized onsite. We fitted lines of several bands between 1250 cm$^{-1}$ and 1390 cm$^{-1}$ and assigned 747 transitions. We reported their center frequencies and line intensities and the data served to confirm the accuracy of a line list produced with the MARVEL algorithm. Furthermore, we had our experimental line list added to the MARVEL data pool, and rerunning the MARVEL procedure yielded a large number of new MARVELized transitions and reduced uncertainties of many energy levels.

I contributed to the sample synthesis and the measurements. I interleaved the spectra and wrote a part of the manuscript.

Paper VII
Accurate measurement and assignment of high rotational energy levels in the 9150 - 9370 cm$^{-1}$ range of methane using optical frequency comb double-resonance spectroscopy
Adrian Hjältén, Vinicius Silva de Oliveira, Andrea Rosina, Isak Silander, Lucile Rutkowski, Grzegorz Soboń, Kevin K. Lehmann, Aleksandra Foltynowicz.
In manuscript.

We demonstrated an improved setup for OFC-based OODR spectroscopy with sub-nominal FTS, and continued the work on measuring hot band transitions of CH$_4$, relevant for high-temperature applications.

We made three measurements pumping three different transitions of the $\nu_3$ band. We detected 88 ladder-type lines, retrieved their line parameters, and could assign 79 of them to predicted hot band transitions. Comparison to predictions revealed systematic discrepancies in line center frequencies, not observed for transitions with lower rotational energy levels.

I was involved in upgrading the measurement setup, selecting the pump transitions, and making the measurements. I was responsible for line fitting and assigning the hot band transitions to TheoReTS.
11 Acknowledgements

First of all I want to thank my supervisor Aleksandra Foltynowicz for having confidence in me and giving me the opportunity to embark on this journey, and for guiding me every step of the way. I must also acknowledge my co-supervisor Ove Axner for sharing his cheerfulness and knowledge of the field of optics and spectroscopy, and my reference person László Veisz for his keen record-keeping and tracking of my progress.

I want to thank my nearest partners in the group, Ibrahim Sadiek, Matthias Germann, Vinicius Silva de Oliveira and Andrea Rosina, for all their efforts in our scientific endeavors, and for not losing me abroad. I thank Isak Silander for his useful input to all discussions and capable engineering skills, for fixing the things I broke and building the things I couldn’t. I also acknowledge past group members, Alexandra Johansson, Anna Filipsson, Francisco Senna Vieira and Chuang Lu who brought me into the local frequency comb community and contributed to these publications in one way or another, as well as our external collaborators.

I am grateful to the administrative staff, Katarina Hassler for answering any question I ever had, Nils Blix for fixing all computer related problems (even when self-inflicted) and Jörgen Eriksson for the help with other practical matters (sorry for using so much tape).

A big recognition is also due to my fellow scientists at the lunch table, you helped more than you think.

Finally, I thank my parents and other relatives for their support and their relentless, though exaggerated, belief in me and my abilities.
References


96. IPCC A et al. Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change 2013;1535.
105. Perrin A, Keller F, and Flaud JM. New analysis of the $\nu_2$, $\nu_3$, $\nu_4$, and $\nu_6$ bands of formaldehyde H$_2^{12}$C$^{16}$O line positions and intensities in the 5–10 $\mu$m spectral region. J. Mol. Spectrosc. 2003;221:192–8.
106. Tchana FK, Perrin A, and Lacome N. New analysis of the $\nu_2$ band of formaldehyde (H$_2^{12}$C$^{16}$O): Line positions for the $\nu_2$, $\nu_3$, $\nu_4$ and $\nu_6$ interacting bands. J. Mol. Spectrosc. 2007;245:141–4.


113. Okubo S, Inaba H, Okuda S, and Sasada H. Frequency measurements of the 2\nu\textsubscript{3}A\textsubscript{1} − \nu\textsubscript{3} band transitions of methane in comb-referenced infrared-infrared double-resonance spectroscopy. Phys. Rev. A 2021;103:022809.


