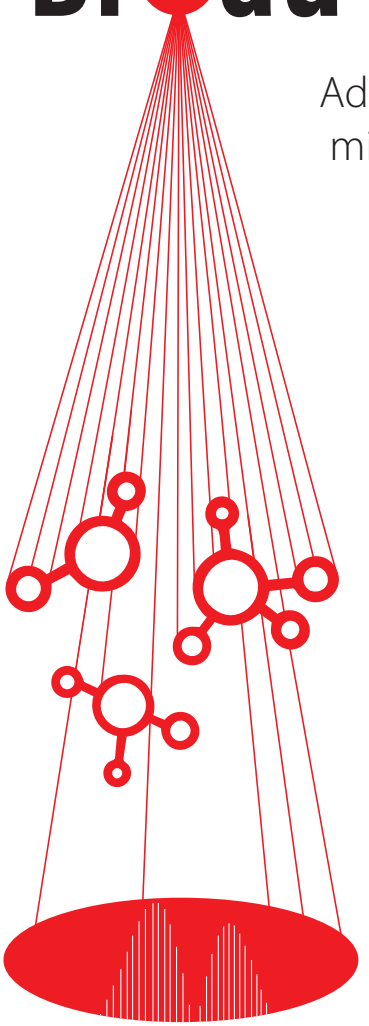
A Broad Perspective

Advancing gas-sensing applications of mid-infrared supercontinuum sources

Roderik Krebbers



Institute for Molecules and Materials

RADBOUD UNIVERSITY PRESS

Radboud Dissertation Series

A Broad Perspective

Advancing gas-sensing applications of mid-infrared supercontinuum sources

Roderik Krebbers, MSc.

A Broad Perspective: Advancing gas-sensing applications of mid-infrared supercontinuum sources
Roderik Krebbers

Radboud Dissertation Series
ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS
Postbus 9100, 6500 HA Nijmegen, The Netherlands
www.radbouduniversitypress.nl

Design: Proefschrift AIO | Annelies Lips
Cover: Proefschrift AIO | Guntra Laivacuma
Printing: DPN Rikken/Pumbo

ISBN: 9789465150932
DOI: 10.54195/9789465150932
Free download at: https://doi.org/10.54195/9789465150932

RADBOUD
UNIVERSITY
PRESS

A Broad Perspective

Advancing gas-sensing applications of mid-infrared supercontinuum sources

Proefschrift ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.M. Sanders,

volgens besluit van het college voor promoties

in het openbaar te verdedigen op

woensdag 24 september 2025

om 14.30 uur precies

door

Roderik Krebbers

geboren op 15 februari 1998

te Deventer

Promotor:
dr. S.M. Cristescu

Copromotor:
dr. A. Khodabakhsh

Manuscriptcommissie:
Prof. dr. J. Oomens
Prof. dr. R.A.H. Engeln (Technische Universiteit Eindhoven)
Prof. dr. M.W. Beijersbergen (Universiteit Leiden)

A Broad Perspective

Advancing gas-sensing applications of mid-infrared supercontinuum sources

Dissertation to obtain the degree of doctor

from Radboud University Nijmegen

on the authority of the Rector Magnificus prof. dr. J.M. Sanders,

according to the decision of the Doctorate Board

to be defended in public on

Wednesday, September 24, 2025

at 2.30 pm

by

Roderik Krebbers

born on February 15, 1998

in Deventer (the Netherlands)

Supervisor:
dr. S.M. Cristescu

Co-supervisor:
dr. A. Khodabakhsh

Manuscript Committee:
Prof. dr. J. Oomens
Prof. dr. R.A.H. Engeln (Eindhoven University of Technology)
Prof. dr. M.W. Beijersbergen (Leiden University)

Table of contents

Summary

Gas sensing through infrared absorption spectroscopy is a beautiful example of how physical and chemical theory and technical know-how can be used for real-life applications. It is built upon the physics and physical chemistry of the interaction of light with gas molecules and on the fundamental ro-vibrational structure of these molecules. Through it, concrete information on the different molecules present in a gas sample can be derived, such as the concentrations of these molecules, or the pressure and temperature of the sample. Therefore, spectroscopy finds applications in a wide range of fields, from environmental science, microbiology, ecology, or human health to plasma chemistry. This thesis explores the applications and potential impact of gas sensing using absorption spectroscopy with novel mid-infrared supercontinuum sources and Fourier-transform spectrometers.

Infrared supercontinuum sources have emerged over the past years as novel light sources, ideal for spectroscopy. They combine properties of lasers, such as spatial coherence and high spectral brightness with a broad spectral coverage, previously only seen in thermal sources. Combined with the renowned technique of Fourier transform spectroscopy, a measurement system arises that shows great promise for measurements of gas samples.

The current societal demands for innovative action in various domains, such as greenhouse gas emission reduction, circular economy, and exposure to a polluted environment call for innovative solutions. The required gas-sensing approaches need to be field-deployable and widely applicable. This thesis demonstrates how Fourier transform spectroscopy with novel mid-infrared supercontinuum sources can be used within these relevant domains. The **Introduction** starts with the concepts of broadband infrared absorption spectroscopy and a technical background on the two different types of supercontinuum sources used within the scope of this thesis – fiber-based and difference-frequency generation-based systems, respectively. Following the introduction, the thesis consists of two parts, of which one focuses on the applications of fiber-based systems and the other on the applications of difference-frequency generation-based systems.

An important domain of applications for spectroscopic research is the analysis of chemical reactions. As the chemical industry shifts towards electrified and circular chemical processes, methane is becoming a major factor in the challenge of closing the carbon loop. Valorization of methane is needed to achieve net zero CO_2 emissions. Plasma-catalytic conversion of methane to ethylene is an example of an innovative and economically viable process for methane valorization. However, optimization of these conversions requires fundamental insights into chemical reactions in which broadband infrared absorption spectroscopy can serve as an

essential tool. **Part A** focuses on the application of fiber-based supercontinuum sources to investigate plasma-based conversions. Optical spectroscopic methods can be used to quantify the products of different carbon-based chemical reactions (**Chapter 4**). Supercontinuum sources demonstrate to be an excellent tool for such studies, as their broad spectral coverage enables unambiguous identification and quantification of many gaseous products. Furthermore, they possess the ability to detect different molecular species simultaneously. This provides a broader picture of the reaction pathways and enables optimization of the reaction towards a favored product. Methane and carbon dioxide are typical waste products of agriculture and industry and can be used as feedstock for plasma reactors to create valuable chemical compounds. Therefore, the potential of plasma-driven chemical reactions goes beyond the reduction of the carbon footprint of the chemical industry and provides a way to up-cycle waste products of other sectors towards a circular economy.

In addition to studies on reaction products, a more fundamental understanding of the plasmas can be gained through *in-situ* studies. Inside the plasma, essential discharge parameters, such as temperature and pressure, can be traced and the formation and dissociation of compounds can be observed. Insights into the reaction mechanisms contribute to optimizing energy efficiency and reaction selectivity of the plasma-driven conversions. In **Chapter 5**, methane plasmas were studied by probing the plasma with a spatially coherent, fiber-based supercontinuum source. The broadband supercontinuum spectrum traces the breakdown of the methane molecules and the formation of the reaction products. Moreover, it adds additional insights into the dynamics inside the plasma, where the molecules are no longer in thermal equilibrium. Monitoring the full ro-vibrational bands of molecules, the vibrational and rotational excitation inside the plasma is assessed. The characteristics of supercontinuum sources make them excellent tools for these types of *in-situ* studies, which is demonstrated for the first time in this chapter. Upon publication, the work was selected as *Editor's Choice Article Highlight* and cover image of the issue.

The versatility in multi-species gas detection, both in the type of molecular species detected, as well as the investigated applications is further shown in **Part B**. Here, a new type of supercontinuum source is used: intra-pulse difference-frequency generation (IDFG)-based systems. The first-of-its-kind commercially available IDFG-based system emerged during the course of this PhD project and possesses unprecedented spectral coverage of 2 – 11.5 µm, ultralow intensity noise, and high brightness (3 W output power). These characteristics enable further advances in the simultaneous detection of a wide range of molecular species, as shown in **Chapter 6**. This work is the first spectroscopic study of simultaneous gas sensing over the full spectral coverage of 2 – 11.5 µm for IDFG-based systems. An

application of this ability is shown in the detection of the outflow of a bioreactor with ammonia-oxidizing organisms. Simultaneous detection of all nitrogen oxides present in the samples enables closing of the nitrogen mass balance. In appreciation of this merit, the associated paper was selected as *Editors' pick*.

The wide applicability of supercontinuum-based Fourier transform spectroscopy is demonstrated with the field deployment for the open-path measurements in **Chapter 7**. Field-deployable techniques are essential to provide insights into greenhouse gases and polluting gases in the outdoor environment. Supercontinuum sources provide a flexible platform to detect a wide variety of these gases simultaneously. Furthermore, the open-path optical arrangement enables measurements in environments with heterogeneous emissions and probing emission points from a distance. The coherent open-path spectroscopy (COPS) system was deployed over the aeration tank of a wastewater treatment plant to monitor its emissions, which was a noteworthy milestone as it represented the first out-of-laboratory field-deployment of an IDFG-based supercontinuum source. The atmospheric concentrations of nitrous oxide, ammonia, and carbon monoxide were detected, and clear emission patterns were retrieved in the concentrations of methane and carbon dioxide. This first demonstrator system paves the way for future emission studies in the outdoor environment, where the COPS system can be used flexibly in different types of studies, ranging from greenhouse gas to environmentally polluting (nitrogen-based) gas emissions.

With the extensive spectral coverage, high resolution and high sensitivity of the measurement system, data analysis becomes increasingly more challenging. A wealth of spectroscopic information is hidden in the fast amount of absorbance features of complex mixtures. The intensities of the absorption lines are spread over a large dynamic range and overlapping absorption features introduce further needs to improve spectral data analysis. Current analytical approaches are centered around quantification of known, identified absorption features. Therefore, an effort is needed to retrieve currently hidden information from complex spectra, such as the identification of unknown molecular species in a mixture. **Chapter 8** demonstrates an approach to enhance the algorithms converting the absorbance spectra to concentrations of the gas species present. Options to assess the precision and accuracy of data processing for a specific system are discussed, and a workflow to optimize retrieving concentrations of gases in complex mixtures is proposed. This workflow leads to a systematic approach to evaluate and optimize the performance of the data analysis prior to application to measured data. In this way, models can be developed with reliable performance in identifying compounds that have not been measured before.

Measurements of polluting gases in the outdoor environment, for example through coherent open-path spectroscopy, provide insights into the compounds people are exposed to. However, the effects of such exposure remain unknown. Analysis of exhaled (human) breath is a useful tool to measure the effects of exposure in humans non-invasively and to provide insights into their health conditions. Breath contains many small volatile compounds that can serve as biomarkers, indicative of a certain condition. Studying the metabolomic profile comprehensively through exhaled breath analysis is also known as breathomics. Laser-based infrared absorption spectroscopic methods have mostly focused on the detection of a single or a few gases tailored to a specific condition or disease. Mass spectrometry-based methods, on the other hand, are considered the golden standard for breath analysis due to their versatility in measuring many compounds simultaneously, typically untargeted. The application of supercontinuum-based Fourier transform spectroscopy in exhaled breath analysis is introduced in **Chapter 9** as a new approach. This technique unites the sensitivity and selectivity for small gas molecules, known from laser-based spectroscopic methods, with the versatility arising from detecting a broad range of gases simultaneously, known from mass spectrometry systems. Here, a demonstrator system is shown, including the standardized breath sample collection for on-line measurements of relevant metabolites of known biochemical pathways. Previous chapters provided insights into the detection capabilities of supercontinuum-based systems for numerous polluting and greenhouse gases. In contrast, in this chapter, molecular compounds with metabolic relevance were retrieved from human breath samples and related to several conditions imposed on the (healthy) participants of the study.

Samenvatting

Gasdetectie met infraroodspectroscopie is een prachtig voorbeeld van de manier waarop fysische en chemische theorie en technische knowhow kan worden gebruikt voor real-life toepassingen. Het fundament ligt in de natuurkunde en de fysische chemie van de interactie van licht met gasmoleculen en op de fundamentele rovibrationele structuur van deze moleculen. Hiermee kan concrete informatie worden afgeleid over de verschillende moleculen die aanwezig zijn in een gasmonster, zoals hun concentratie, of de druk en temperatuur van het monster. Op deze manier vindt spectroscopie zijn toepassingen in velerlei gebieden, variërend van milieukunde, microbiologie, ecologie, gezondheidszorg, tot plasmachemie. Deze dissertatie onderzoekt de toepassingen en potentiële impact van gasdetectie door middel van absorptiespectroscopie met innovatieve (middel-infrarood)-supercontinuümlichtbronnen en Fourier-transformspectrometers.

Supercontinuümlichtbronnen in het infraroodspectrum zijn de afgelopen jaren opgekomen als nieuwe, ideale lichtbronnen voor spectroscopie. Ze beschikken over een combinatie van eigenschappen die enerzijds overeenkomen met lasers, zoals de ruimtelijke coherentie en hoge spectrale stralingssterkte, en anderzijds met thermische lichtbronnen, zoals het brede spectrale bereik. In combinatie met de gevestigde Fourier-transformspectrometers ontstaat er zo een meetsysteem met grote potentie voor analyses van gasmonsters.

De huidige maatschappelijke roep om innovatie in verscheidene domeinen, zoals broeikasgasreductie, circulaire economie, en de blootstelling aan een vervuilde leefomgeving, vraagt om innovatieve oplossingen. De benodigde gasdetectiemethoden moeten in het veld inzetbaar zijn en breed toegepast kunnen worden. Deze dissertatie demonstreert hoe Fourier-transformspectroscopie met innovatieve supercontinuümlichtbronnen toepasbaar is in deze domeinen. De **Introductie** begint met de concepten van breedbandige infraroodspectroscopie en een technische achtergrond in de twee typen supercontinuümlichtbronnen die gebruikt worden in deze dissertatie – systemen gebaseerd op optische vezels en systemen gebaseerd op *difference-frequency generation*. Daaropvolgend bestaat de dissertatie uit twee delen, waarvan een zich richt op de vezelgebaseerde systemen en de ander op *difference-frequency generation*-gebaseerde systemen.

Een belangrijk domein binnen de toepassingen van spectroscopisch onderzoek is de analyse van chemische reacties. Nu de chemische industrie in transitie is naar geëlektrificeerde en circulaire chemische processen, begint methaan een belangrijke factor te worden in de uitdaging de koolstofcyclus te sluiten. Valorisatie van methaan is nodig om per saldo CO_2-uitstootvrij te worden. Plasmakatalytische methaan-etheenomzettingen zijn voorbeelden van

innovatieve en economisch haalbare processen binnen de valorisatie van methaan. Echter vereist de optimalisatie van deze omzettingen fundamentele inzichten in de chemische reacties. Breedbandige infraroodspectroscopie kan hierin een essentieel hulpmiddel zijn. **Deel A** richt zich op de toepassing van vezelgebaseerde supercontinuümlichtbronnen in onderzoek naar plasmagestuurde omzettingen. Optische methoden kunnen ingezet worden om de reactieproducten van verschillende koolstofgebaseerde chemische reacties te kwantificeren (**Hoofdstuk 4**). Supercontinuümlichtbronnen zijn excellente hulpmiddelen voor dit type studies, omdat het brede spectrale bereik eenduidige identificatie en kwantificatie van vele (gasfase) reactieproducten mogelijk maakt. Daarnaast is het mogelijk verschillende moleculaire stoffen gelijktijdig te detecteren. Dit geeft een breed beeld van de reactiemechanismen en helpt in de optimalisatie van een reactie richting een gewenst reactieproduct. Methaan en koolstofdioxide zijn typische rest- en bijproducten van de landboud en industrie en kunnen worden gebruikt als grondstof om waardevolle chemische producten te maken in plasmareactoren. Op die manier gaat het potentieel van plasmareacties zelfs verder dan de reductie van de CO_2-voetafdruk van de chemische industrie en biedt het een pad naar upcycling van restproducten van andere sectoren voor een circulaire economie.

Naast de studies op het gebied van reactieproducten kan een meer fundamenteel begrip van plasma's worden verworven door middel van *in-situ* studies. In het plasma kunnen essentiële ontladingsparameters als de temperatuur en druk worden gemonitord en de dissociatie van moleculen geobserveerd. Inzichten in de reactiemechanismen dragen bij aan de optimalisatie van de energie-efficiëntie en reactieselectiviteit van plasmagestuurde omzettingen. In **Hoofdstuk 5** zijn methaanplasma's bestudeerd door de ruimtelijk coherente lichtbundel van de vezelgebaseerde supercontinuümlichtbron door het plasma te sturen. Het breedbandige supercontinuümspectrum monitort de afbraak van methaanmoleculen en de formatie van de reactieproducten. Daarnaast biedt het aanvullende inzichten in de dynamiek in het plasma, waarin de moleculen zich niet langer in thermisch equilibrium bevinden. Door de volledige rovibrationele band van de moleculen te observeren kan de vibrationele en rotationele excitatie in het plasma worden beoordeeld. De karakteristieken van de supercontinuümlichtbronnen lenen zich uitstekend voor dit type *in-situ* studies, wat voor het eerst wordt aangetoond in dit hoofdstuk. De publicatie van deze studie is verkozen tot *Editor's Choice Article Highlight* en coverafbeelding van de uitgave.

De veelzijdigheid in detectie van meerdere soorten gassen, zowel wat betreft de verschillende gassen die gedetecteerd kunnen worden als de mogelijke toepassingen, komt verder aan bod in **Deel B**. Hier wordt een nieuw type supercontinuümlichtbron geïntroduceerd: *intra-pulse difference-frequency*

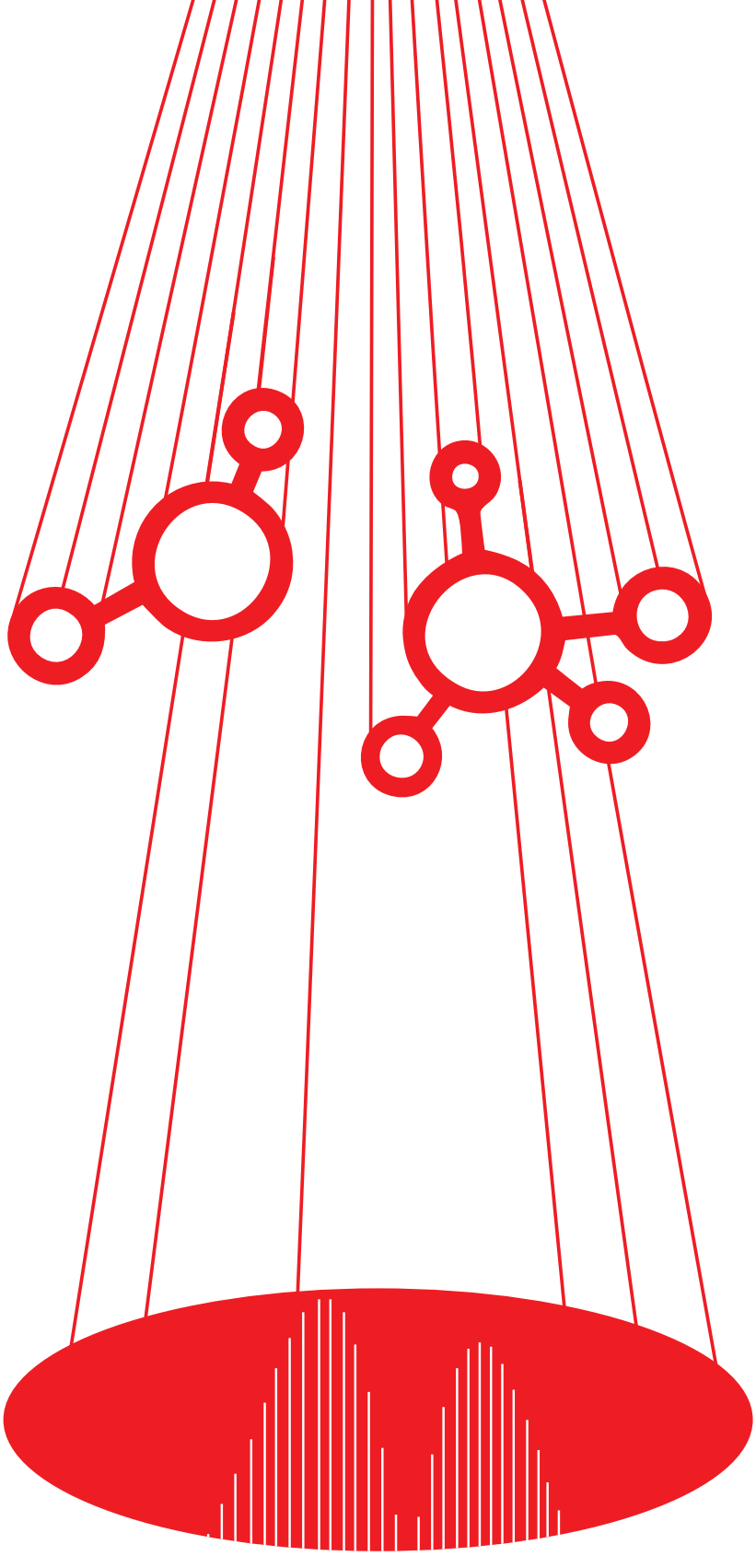
generation (IDFG)-gebaseerde systemen. Dit unieke, IDFG-gebaseerde systeem kwam voor het eerst commercieel beschikbaar in de loop van het PhD project en beschikt over een ongeëvenaard spectraal bereik van 2 – 11.5 µm, een ultralage intensiteitsruis, en een hoge spectrale stralingssterkte (3 W vermogen). Deze karakteristieken drijven de verdere ontwikkelingen in het simultaan detecteren van een groot aantal moleculaire gassen, zoals aangetoond is in **Hoofdstuk 6**. Dit is de eerste spectroscopische studie naar het gelijktijdig detecteren van gassen over het volledige spectrale bereik van 2 – 11.5 µm met IDFG-gebaseerde systemen. Een toepassing van deze eigenschap is het monitoren van het effluent van een bioreactor met ammonia-oxiderende organismen. Door gelijktijdig alle stikstofoxiden te meten in de monsters is het mogelijk om een sluitende stikstofbalans op te stellen. Als blijk van waardering van deze prestaties, is de bijbehorende publicatie geselecteerd als *Editors' pick*.

De wijde toepasbaarheid van supercontinuüm Fourier-transformspectroscopie blijkt ook uit de veldmetingen in het kader van de openpadmetingen van **Hoofdstuk 7**. Technieken die in het veld toegepast kunnen worden zijn essentieel voor het verkrijgen van inzichten in broeikas- en vervuilende gasemissies in de buitenlucht. Supercontinuümlichtbronnen bieden een flexibel platform voor de simultane detectie van verscheidende gassen. Daarnaast biedt de openpadopstelling de mogelijkheid tot metingen in omgevingen met heterogene emissies en het meten van emissiepunten vanaf afstand. Het *coherent open-path spectroscopy* (COPS) systeem is ingezet voor emissiemonitoring bij een beluchtingstank van een rioolwaterzuiveringsinstallatie: een noemenswaardige mijlpaal, aangezien het de eerste veldmetingen buiten het laboratorium betrof voor een IDFG-gebaseerde supercontinuümlichtbron. De atmosferische concentraties van lachgas, ammoniak, en koolmonoxide konden worden gemonitord, evenals duidelijke emissiepatronen in de concentraties van methaan en koolstofdioxide. Dit eerste demonstratiesysteem legt de basis voor toekomstige studies naar emissies in de buitenlucht, waarbij het COPS-systeem flexibel ingezet kan worden in verschillende typen studies, variërend van broeikasgasemissies tot aan milieuverontreinigende (stikstof-gebaseerde) gasemissies.

Met het uitgebreide spectrale bereik, de hoge resolutie en sensitiviteit van het meetsysteem wordt data-analyse een steeds grotere uitdaging. Een rijkdom aan spectroscopische informatie ligt verborgen in de gigantische hoeveelheid absorptielijnen van complexe mengsels. Dat de intensiteit van deze lijnen is verspreid over een breed dynamisch bereik en lijnen van verschillende moleculen regelmatig overlappen, maken verbeterde data-analyse des te noodzakelijker. De huidige analytische methoden richten zich op kwantificatie van reeds bekende, geïdentificeerde absorptielijnen. Daarom is het voor de identificatie

van onverwachte moleculaire gassen in een mengsel nodig om beter om te gaan met momenteel ongebruikte, verborgen informatie in de complexe spectra. **Hoofdstuk 8** toont een aanpak om de algoritmen voor het omzetten van spectra naar concentraties van gassen te verbeteren. Opties om de precisie en accuraatheid van de dataverwerking te beoordelen voor een specifiek meetsysteem worden besproken en een workflow om de verkregen concentraties te optimaliseren wordt geïntroduceerd. Deze workflow leidt tot een systematische aanpak voor de evaluatie en optimalisatie van de data-analyse nog voordat deze wordt toegepast op gemeten spectra. Op die manier kunnen betrouwbare modellen worden ontwikkeld voor de identificatie van niet eerder gemeten gassen.

Metingen van milieuverontreinigende gassen in de buitenlucht, bijvoorbeeld met behulp van *coherent open-path spectroscopy*, brengen inzicht in de gassen waar men aan wordt blootgesteld. De effecten van zulke blootstelling blijven echter onbekend. De analyse van ademmonsters is een bruikbaar hulpmiddel voor het niet-invasief meten van de effecten van blootstelling bij mensen en geeft inzichten in hun gezondheidstoestand. Adem bevat vele kleine vluchtige stoffen die kunnen dienen als biomarkers die indicatief zijn voor een bepaalde toestand. Het bestuderen van het gehele metabolische profiel door middel van ademanalyse staat ook bekend als *breathomics*. Infraroodspectroscopische methoden met lasers hebben zich tot dusver voornamelijk gericht op de detectie van een enkel of enkele gas(sen), specifiek voor een bepaalde toestand of ziekte. Daarentegen worden massaspectrometrische methoden gezien als de gouden standaard voor ademanalyse, juist vanwege hun veelzijdigheid in het gelijktijdig meten van vele soorten gassen, maar doen dit over het algemeen zonder specifieke selectiviteit. De toepassing van supercontinuüm-gebaseerde Fourier-transformspectroscopie in ademanalyse wordt in **Hoofdstuk 9** geïntroduceerd als een nieuwe aanpak. Deze techniek brengt de sensitiviteit en selectiviteit voor kleine gasmoleculen, bekend van de laser-gebaseerde spectroscopische methoden, samen met veelzijdigheid vanuit het simultaan detecteren van een breed scala aan gassen, zoals in massaspectrometrische methoden. In dit hoofdstuk tonen we een demonstratiesysteem, inclusief gestandaardiseerde adembemonstering, voor directe metingen van relevante metabolieten van bekende biochemische processen. De voorgaande hoofdstukken boden inzichten in de detectiemogelijkheden van supercontinuümgebaseerde systemen voor verscheidene milieuverontreinigende en broeikasgassen. Dit hoofdstuk, daarentegen, richt zich op moleculaire stoffen met metabolische relevantie die in ademmonsters gevonden kunnen worden en gerelateerd kunnen wordenaan verschillende interventies bij de (gezonde) proefpersonen van de studie.



Chapter 1

Introduction

This thesis explores spectroscopic gas-sensing applications that revolve around novel mid-infrared supercontinuum sources. As these light sources are maturing, novel fields of applications open up, involving out-of-laboratory experiments in environmental monitoring, human health, or plasma chemistry. The central theme in this thesis is the development of a flexible spectroscopic platform, which consists of a spectrometer and can be equipped with different supercontinuum light sources. In the following chapters, a general introduction to infrared absorption spectroscopy for gas sensing will be given, followed by an introduction to the various elements required in our spectroscopic platform, such as the supercontinuum source (Chapter 2), and the spectrometer (Chapter 3.1). The primary focus of the research presented in Chapters 4 to 9 is to assess the performance of the system with the various supercontinuum sources for these applications in gas sensing.

1.1 Infrared absorption spectroscopy

Absorption spectroscopy is an elegant way to determine the concentration of a compound in a medium. When we dissolve a molecule in water and the water starts to turn a certain color, we intuitively understand that the stronger or darker the color gets, the higher the concentration of the molecule in the water: more light is absorbed. However, in the gas phase, the density of molecules is much lower, and thus, the absorption of light per volume is much lower. Consequently, there are hardly any examples of gas mixtures with a certain color. However, while we cannot see the molecules' concentration in the visible spectrum, they absorb strongly in the infrared spectrum, leading to a well-detectable signal[1].

Besides the strong absorption of gaseous molecules in the infrared, it is also particularly useful to perform absorption spectroscopy for gas sensing in this spectral region as each molecule exhibits a unique absorption pattern. The infrared absorption of molecules is caused by rovibrational transitions. At resonant frequencies, when the frequency of the radiation (i.e., the energy of the photons) matches the energy of the vibrational transitions of a molecule, that molecule can absorb the photons. This can be detected as a distinct absorption peak at that specific frequency. The molecular structure and the number of atoms in the molecule, N, determine the number of vibrational modes of the molecule: it can be calculated as $3N - 5$ for linear molecules, and $3N - 6$ for nonlinear molecules. A simple, diatomic molecule, such as carbon monoxide, has one bond and consequently one fundamental vibrational mode. A slightly bigger, triatomic

1. The strong infrared absorption of small gaseous molecules, particularly CO_2, CH_4, N_2O, and H_2O, also explains their role in the greenhouse effect.

molecule, such as carbon dioxide, has two bonds and consequently already 4 fundamental vibrational modes. However, some of these modes are degenerate, resulting in three distinct fundamental vibrational bands. Not all vibrational modes are observable in the infrared spectrum as this requires a vibrationally induced dipole moment in the molecule. For example, in the bond of a nitrogen molecule (N_2), no dipole can be induced, and thus this molecule does not have a vibrational absorption band in the infrared. Contrarily, as shown in Figure 1.1, methane, which does not have a dipole due to the symmetry in its molecular structure, has both vibrational modes that induce a dipole in the molecule and thus are IR active (such as the ν_3 mode), as well as vibrational modes that do not cause a change in the dipole (such as the ν_3 mode) and thus are not visible in the absorption spectrum.

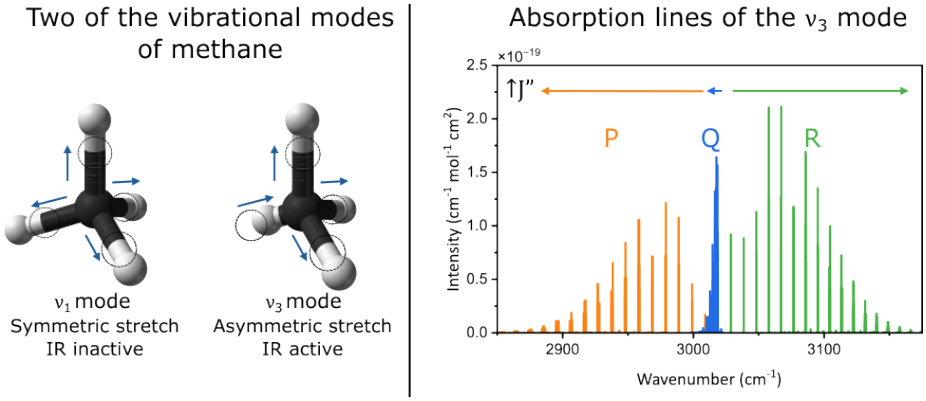


Figure 1.1: Left: Examples of the vibrational modes of methane: only vibrational modes that cause a change in dipole lead to IR absorption, such as the ν_3 mode. Right: Spectral lines within the fundamental ν_3 band of methane, with P-, Q-, R-branches indicated. The arrows show the directions of increasing lower-state rotational quantum number (J'') within one branch.

A vibrational mode does not consist of just a single absorption line, as the coupling of rotational and vibrational energies in the molecule causes the transition to be not merely vibrational. Depending on the symmetry of the molecule and the vibrational transition, the selection rules demand the change in rotational quantum number from the lower (J'') to the higher (J') state to be $\Delta J = J' - J'' = 0 \pm 1$ or $\Delta J = \pm 1$, during the vibrational transition $\Delta\nu = \pm 1$. This means that each vibrational mode, in spectroscopic terms: *bands*, consists of separate branches of transitions for $\Delta J = -1$, $\Delta J = 0$, and $\Delta J = +1$, which are called the P-, Q-, and R-branch, respectively. As the $\Delta J = -1$ transitions demand slightly less and the $\Delta J = +1$ slightly more energy, the P- and R-branches appear at slightly lower and higher wavenumbers than the Q-branch, respectively. Finally, each branch consists of separate absorption

lines. Within each branch, each line relates to a distinct lower-state rotational level. The difference in transition energy between the individual peaks within the same branch arises from the anharmonicity of the vibrational modes.

Particularly relevant is that the molecules in the gas phase maintain all their degrees of freedom, in contrast to solid and liquid phase molecules. A familiar feature of the spectra from classical infrared spectrometers for liquid or solid samples is their broad spectral profiles: these spectra can at best be used to distinguish molecules with different functional groups. Gas-phase molecules, on the other hand, have minimal interactions with their neighboring molecules. Thus, the absorption line broadening is limited, resulting in absorption spectra with very distinct patterns of narrow absorption lines. The exact broadening of absorption lines, such as due to the pressure broadening and the instrumental resolution, and the resulting line shape will be part of Chapter 3.2.

The presence of a certain molecule in a gas sample can be detected and quantified by measuring the absorption spectrum of a certain molecule – that is, measuring the absorption from the rovibrational transitions as a function of wavenumber or wavelength of the applied light. To detect a gas molecule successfully, high sensitivity and selectivity is required. High sensitivity can be achieved via the strong absorption of the vibrational levels in the infrared region, while high selectivity can be achieved via the distinct patterns of the various absorption bands of the different molecules. For this reason, infrared spectroscopy has been the pinnacle of applied gas-sensing techniques for many decades (1–3).

In classical infrared absorption spectroscopy with thermal lamps, the fundamental elements needed are the light source, the (gas) sample interaction, and the spectrometer (Figure 1.2). These elements are also reflected within the novel systems presented in this thesis. In the subsequent chapters, the novel type of light source used in this thesis, the supercontinuum source (Chapter 2), and the spectrometer (Chapter 3.1) will be reviewed. The interaction with the gas sample can occur in many ways, from sampling the gas in a dedicated gas cell, to *in-situ* in environments such as reactors or directly in the atmosphere. These variations in interacting with the gas sample will come back throughout this dissertation.

The first step towards retrieving an absorption spectrum is the illumination of the sample. Thermal light sources are traditionally used for this and are still in most commercial Fourier-transform infrared (FTIR) systems. While thermal light sources are reliable sources, with low noise and a broad emission profile, their applicability for gas sensing applications is limited. This is due to the incoherent nature of the emitted light, making long interaction path lengths with the gas sample harder to achieve, and their low spectral brightness in the infrared. This low brightness is a fundamental limitation of the thermal light source being a black body radiator: the

output power is governed by the temperature of the source (according to Stefan-Boltzmann law), but a higher temperature also shifts the peak intensity of the light emitted to shorter wavelengths, thus away of the infrared (according to Wien's law).

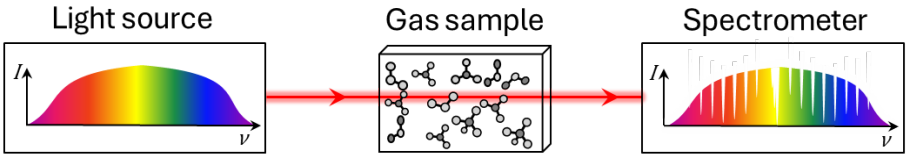


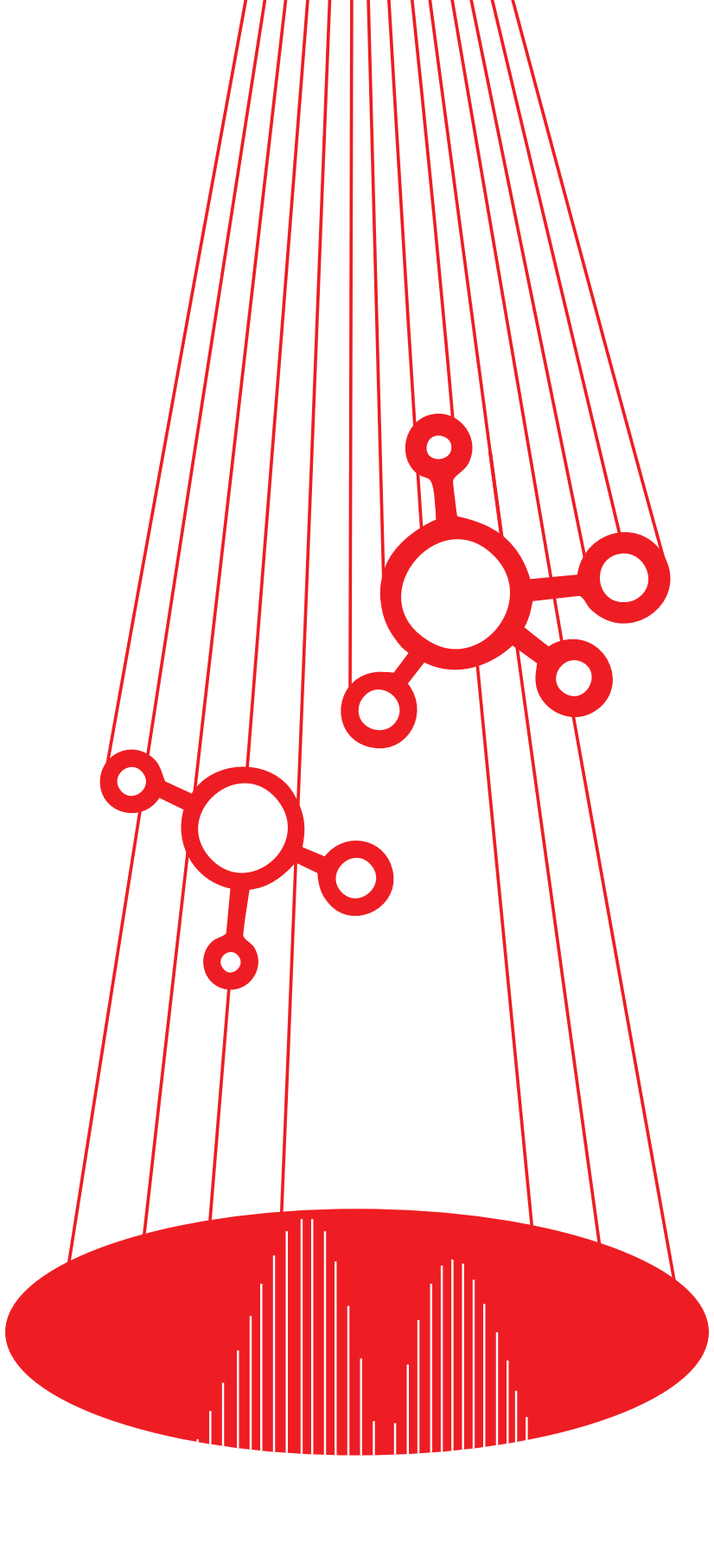
Figure 1.2: The essential elements of infrared absorption spectroscopy: a light source, the interaction with the gas sample, and the spectrometer.

An alternative to thermal sources is a laser-based source, which has been the standard in the gas-sensing community for multiple decades (4). Infrared lasers deliver a high spectral brightness in the infrared, possess spatial and temporal coherence, and can have excellent stability for low-noise operation. These characteristics enable sensitive detection of molecular species in complex gas mixtures as they allow for interaction with the gas sample over long interaction path lengths. Lasers usually have a narrow-band spectrum. This could be used as an advantage to remove the need for a spectrometer. However, a narrow spectral coverage also limits the number of molecular species that can be measured simultaneously. Furthermore, it compromises the identification of molecular species with broader spectral profiles and complicates the simultaneous detection of multiple gas mixture parameters, such as pressure or separate vibrational and rotational temperature distributions in non-thermal media.

Therefore, from the perspective of infrared absorption spectroscopy, it is ideal to have a light source with a broad range in the spectral domain, while maintaining the (spatial) coherence known from lasers, and a low overall noise profile. A significant effort within the spectroscopic community is being made to broaden the coverage of the laser sources. Two main directions can be identified in this field. Firstly, the ability to sweep the narrowband emission of the laser over a broader coverage. For example, quantum cascade lasers with an external cavity (EC-QCLs) are known to be able to sweep the narrowband emission of the laser over a range as wide as 400 cm^{-1}, allowing for broadband spectral coverage (5). On the other hand, there are efforts to broaden the spectral coverage of the emitted light itself, which results in an instantaneous broad spectral coverage without the need for sweeping over a certain spectral range. This does, however, re-emphasize the need for a spectrometer. The spectrum of emitted laser light is broadened using nonlinear effects in the propagating media, in a process called supercontinuum generation.

Sources based on supercontinuum generation, so-called supercontinuum sources, are the principal light sources used for spectroscopy in this dissertation. We will explore two approaches to building an infrared supercontinuum source: supercontinuum generation in nonlinear fibers, and supercontinuum generation using difference-frequency generation (DFG) in crystals.

1



Chapter 2

Light sources based on supercontinuum generation

Over the past years, commercial, compact, high-power, and broadband mid-infrared supercontinuum light sources have emerged on the market. Their spatially coherent beams deliver a brightness that is 2-6 orders of magnitude higher than thermal light sources or even synchrotron radiation (6,7) (see Figure 2.1).

A supercontinuum source is a light source exploiting the effects of nonlinear processes to accomplish a broadened spectrum. To understand what a supercontinuum source is, and get an insight into its strengths and weaknesses, the fundamentals of these sources will be briefly reviewed. The essential building blocks for a supercontinuum light source are the (often pulsed) pump laser and a nonlinear material. The type of nonlinear medium used defines the two categories of supercontinuum sources discussed here: the most frequently used supercontinuum sources that employ nonlinear fibers relying on $\chi^{(3)}$-nonlinear effects, and the emerging crystal-based supercontinuum sources, relying on $\chi^{(2)}$-nonlinear effects. Understanding the fundamental differences of $\chi^{(2)}$- and $\chi^{(3)}$-based effects, the current state-of-the-art of both types of sources will be reviewed, resulting in an overview of some of the (commercially) available supercontinuum sources.

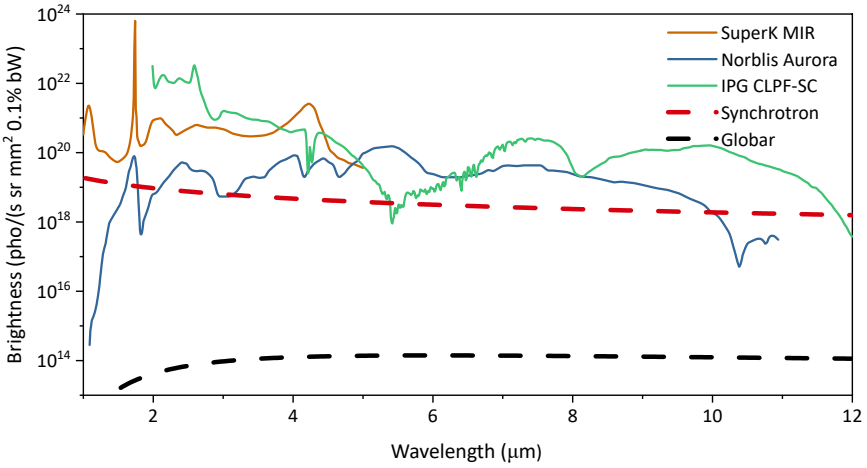


Figure 2.1: Spectral brightness of various supercontinuum sources versus traditional broadband light sources. Data adapted from (8,9)

2.1 Nonlinear effects

Most optical interactions of light waves in a medium can be described as linear interactions. These phenomena, which have mostly been known for centuries, such as refraction, dispersion, or absorption of light, can be observed for any light wave, with arbitrary intensity. With the introduction of lasers, an intensity of the electric

field of the light beam could be reached at which the light interacts nonlinearly with the polarization density of a medium, leading to new observable phenomena. The relationship between the polarization density of the material (P) and the electric field of the light (E) can be described as a Taylor series expansion as

$$P(t) = \varepsilon_0(\chi^{(1)}E(t) + \chi^{(2)}E^2(t) + \chi^{(3)}E^3(t) + \cdots) \quad (2.1)$$

in which ε_0 is the permittivity of free space and $\chi^{(n)}$ the n-th order susceptibility of the medium to the electric field. For sufficiently small values of E, the term $\chi^{(1)}E$ dominates – this is the linear regime. However, eventually, at a strong electric field, the first two nonlinear terms, $\chi^{(2)}E^2(t)$ and $\chi^{(3)}E^3(t)$, also come into play.

While 2nd-order effects may seem most significant, 3rd-order nonlinearities are relevant in materials with *inversion symmetry*. This includes crystals that possess a center of symmetry and for which flipping their coordinates through the center of the crystal, the lattice structure remains identical. Similarly, isotropic materials, such as glass, have the same structure in all directions. In such cases, the $\chi^{(2)}$ nonlinear term poses a problem: as both polarization and the electric field are vectors, inverting these terms changes the direction of the vector. However, in the 2nd-order term of equation (2.1), reverting the direction of the electric field vector does **not** result in an inversion of the polarization: $\chi^{(2)}E^2 \rightarrow \chi^{(2)}(-E)^2 = \chi^{(2)}E^2$. In other words, the relation between the vectors of the polarization and electric field would not be inversion symmetric. Therefore, in an inversion symmetric material, even order nonlinear terms are forbidden and the 3rd-order term will be dominant.

For this reason, nonlinear materials can be split into two classes: those in which $\chi^{(2)}$-nonlinear effects are dominant and those that depend on $\chi^{(3)}$-nonlinear effects. Both are used to create supercontinuum spectra that can be used for infrared spectroscopy, with their specific characteristics.

2.2 $\chi^{(3)}$-dominant supercontinuum generation

The majority of the commercially available supercontinuum-generating systems consists of $\chi^{(3)}$-based nonlinearities. Typically, these systems are equipped with crystal or glass fibers. The spectral broadening is caused by a combination of different $\chi^{(3)}$-based processes, of which a good overview is provided by Eslami Jahromi et al. (6), and Sylvestre et al. (10).

There is an entire class of different $\chi^{(3)}$ nonlinear effects that interact, compete with, and amplify each other. It is difficult to isolate one specific effect in a $\chi^{(3)}$ material. Generally, the combination and mixing of multiple effects finally result in spectral broadening. In combination with their stochastic nature, these fiber-based

supercontinuum sources tend to lose temporal coherence and introduce large pulse-to-pulse fluctuations in their intensity and phase.

One central $\chi^{(3)}$ effect, which is the root cause for many of the processes used in spectral broadening and supercontinuum generation, is the optical Kerr effect, which states that the refractive index (n) is intensity-dependent for strong light pulses with intensity I:

$$n(I) = n_0 + n_2 \cdot I. \tag{2.2}$$

This intuitively means that for a pulsed laser, the high-intensity peak of the pulse experiences a higher refractive index and thus propagates more slowly through the medium than the lower-intensity head and tail of the pulse.

The process is referred to as self-phase modulation (SPM). With the difference in propagation velocity, the pulse becomes chirped, with a stretched, thus red-shifted, head, and a compressed, thus blue-shifted, tail. Through this mechanism, self-phase modulation broadens the spectrum, and the broadening, i.e., the change in frequency, scales with the change in intensity.

Pulses encounter group-velocity dispersion (GVD) when propagating through the fiber. The GVD is a function of the refractive index of the material and the central wavelength of the pulse: a wavelength-dependent refractive index. Each material has a wavelength region below the zero-dispersion wavelength (λ_D), where the GVD > 0 and normal dispersion occurs: different wavelengths propagate at different velocities, and thus the pulse broadens. Consequently, this also enforces a similar chirping as we saw due to SPM: lower frequencies travel faster than higher frequencies, resulting in a positive chirp. At wavelengths longer than λ_D, the GVD < 0 and the opposite is true: negative chirping occurs, where lower frequencies travel slower than the shorter frequencies. This is referred to as the anomalous dispersion regime.

Carefully selecting the central wavelength of the input pulse within this anomalous dispersion regime of the GVD, the negative chirp can counterbalance the positive chirp from the SPM, effectively counterbalancing intensity-dependent and frequency-dependent refractive indexes. The resulting propagating pulse is a soliton: a wave packet that travels undistorted in shape over long distances.

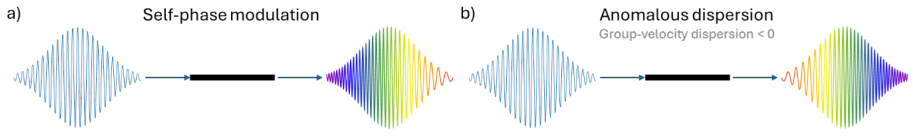


Figure 2.2: The effect of self-phase modulation (**a**) and anomalous dispersion (**b**), causing chirping and spectral broadening on a pulse traveling through a fiber.

Playing with the soliton dynamics throughout the fiber is the main parameter to engineer the desired supercontinuum spectrum. The two central approaches for generating a supercontinuum from the initial, generated solitons are soliton fission and modulation instability.

Solution fission is the primary mechanism with short-pulsed laser systems (fs). With soliton fission, higher-order solitons are compressed, causing fission into fundamental solitons. Higher-order dispersion and stimulated Raman scattering perturb the solitons, causing, for example, the emission of dispersive waves in the visible spectrum. The stimulated Raman scattering drives spectral broadening towards the infrared spectrum through soliton self-frequency shifts. Inside the soliton pulse, the shorter wavelength components act as a pump for stimulated Raman scattering of the longer wavelength, and thus, as the soliton propagates through the fiber, its spectrum gradually shifts into the infrared (11).

Noise-seeded modulation instability is the primary mechanism for laser systems with long pulses (pico- to nanoseconds). The input pulse envelope is broken into a train of sub-picosecond pulses with random amplitude and duration, thus being a major cause for noise amplification in supercontinuum generation. These individual pulses undergo similar soliton dynamics as the short-pulsed systems, with stimulated Raman scattering and corresponding soliton self-frequency shifting being the major effects for extending the spectrum (far) into the infrared.

The generation of supercontinua over its propagation through a fiber can be simulated using the generalized nonlinear Schrödinger equation (GNLSE). In Figure 2.3, an example is given of the generation of an infrared supercontinuum through a ZBLAN fiber (12). The fiber is pumped with laser pulses at 1.5 µm. Over the propagation distance, many solitons form through modulation instability. In this simulation, soliton self-frequency shift is the main mechanism driving the solitons further into the infrared.

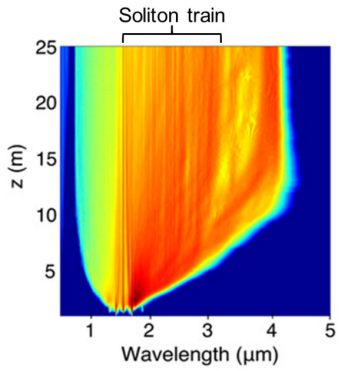


Figure 2.3: Supercontinuum generation through the length of a ZBLAN fiber (z), simulated using the GNLSE. Adapted from (12)

A relevant note is that these dynamics hold specifically for pulsed systems for supercontinuum generation towards the infrared. Other types of supercontinuum-generating systems, for example, all-normal dispersion (ANDi) fibers, exclusively use the normal dispersion region where GVD > 0, and do not rely on soliton formation. These systems only rely on SPM and optical wave-breaking for spectral broadening and provide all-coherent supercontinuum spectra. However, they cannot (yet) deliver the spectral bandwidth and coverage deep into the infrared as the soliton-based systems (10).

2.2.1 Fiber-based supercontinuum sources

Exploiting the $\chi^{(3)}$-based effects in nonlinear fibers, increasingly more systems are available revolving around these types of fibers. An overview of state-of-the-art (commercial) supercontinuum systems used in spectroscopic applications is provided by Zorin et al. (13). Two major types of fibers used in the systems for supercontinuum generation in the infrared are based on chalcogenides or ZBLAN. ZBLAN is a fluoride glass composed of a mix of zirconium fluoride, barium fluoride, lanthanum fluoride, aluminum fluoride, and sodium fluoride (hence ZBLAN). As it is commercially readily available and a stable material, it is used in most commercial mid-infrared supercontinuum sources. However, the high attenuation of the material for wavelengths longer than 4 µm limits the spectral range to this wavelength.

Chalcogenide fibers are a promising candidate for spectral coverage deeper into the infrared. Chalcogenides are glasses that contain chalcogen elements of the period table, such as sulfur, selenium, or tellurium. They possess much broader transmission windows, which, depending on their exact composition, extend up to ~20 µm. Furthermore, their nonlinearities are up to 1000 times stronger than silica fibers (8,14). However, the number of commercially available chalcogenide fibers suitable to incorporate in systems for supercontinuum generation is still limited (10,15).

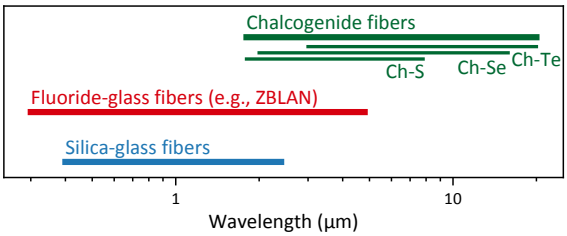


Figure 2.4: Transmission window of various glass fibers used for SC generation (10,16).

Fiber-based systems can be directly pumped into their anomalous dispersion regime if the required pump laser systems are available. For many nonlinear

fibers, however, this is not possible or desirable. Particularly for chalcogenide fibers, the zero-dispersion wavelength (λ_D) often lies deep into the infrared. For example, bulk As_2Se_3 glass has a λ_D at 7.4 µm (8), for which it is difficult to find a commercially available high-power laser to pump the anomalous dispersion regime. Some notable demonstrations of chalcogenide fibers pumped directly into their anomalous dispersion regime include Lemière et al. (17), who created a supercontinuum spanning 1.7 – 18 µm using a GaSeTe fiber, and Zhao et al. (18), who achieved a 2 – 16 µm supercontinuum using a GeTeAgI fiber. However, both systems rely on an optical parametric amplifier (OPA) laser source to deliver a pump pulse with a wavelength of 8.15 µm and 7 µm, respectively, to pump the chalcogenide fiber. These bulky, expensive OPA systems and the low power of the generated supercontinuum (usually at best ~1 mW (10,19)) limit the applicability of direct-pumping-based supercontinuum sources.

Consequently, most fiber-based laser systems use a cascaded design, in which multiple different fibers are combined. Particularly all-fiber-based systems are used, which use commercially available fiber lasers at telecommunication wavelengths (1550 nm). A cascade of silica and fluoride fibers, which have transmission windows up to 2.4 µm and 4.8 µm, respectively (see Figure 2.4), gradually shifts the wavelength to the infrared, from which it is possible to access the anomalous dispersion regimes of the chalcogenide fibers. For example, Bizot et al. used an all-fiber cascaded design consisting of silica, ZBLAN, InF_3, and GeSeTe chalcogenide fibers to achieve a 3-mW supercontinuum spectrum covering the range 3 – 8 µm (20).

Another example of a state-of-the-art infrared supercontinuum source using a cascaded design was developed by Woyessa et al. (21). This specific source is used in the research presented in Chapter 4. It delivers a spectral coverage from 1.5 – 10.5 µm with an average power output of >80 mW. As shown in Figure 2.5, it consists of an all-fiber design, starting with a 3 MHz pump laser with a wavelength of 1.56 µm. The light propagates through a cascade of fibers, with the thulium-doped and germanium/thulium-doped silica fibers extending the wavelength to ~2.75 µm. The ZBLAN consequently extends the spectrum up to 4.4 µm, which is sufficient for pumping the chalcogenide fibers: a tandem of an arsenic trisulfide fiber, which usually has a λ_D around 5 µm (22) and pushes the wavelength coverage up to 7 µm, and an arsenic triselenide fiber. The coverage up to 7 µm is sufficient to reach the λ_D of arsenic triselenide, which is usually between 6 and 7 µm (8,22), finally yielding a total wavelength coverage of up to 10.5 µm. A wide variety of light coupling techniques is used to couple light through these different fibers. Fusion splicing of the fibers would be ideal from the perspective of optimal transmission between the fibers. However, this is difficult to achieve due to the different melting

points of the various compositions of the fibers, which is why, for example, free space coupling is used between the ZBLAN and As_2S_3 fibers.

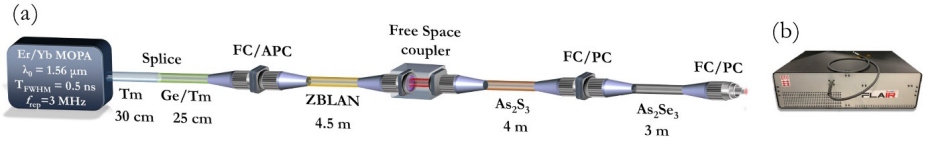


Figure 2.5: Architecture of the chalcogenide-based infrared supercontinuum source by Woyessa et al., reused from (21). FC/APC: ferrule connector with angled physical contact, FC/PC: ferrule connector with physical contact.

An overview of currently commercially available fiber-based (mid-)infrared supercontinuum sources can be found in Table 2.1. More comprehensive overviews of fiber-based infrared supercontinuum sources demonstrated in a laboratory context, but not (yet) available commercially, can be found in references (10,19).

Table 2.1: Overview of commercially available fiber-based infrared supercontinuum sources.

Manufacturer and model	*Type*	*Wavelength coverage*	*Power*	*Repetition rate*	*Ref.*
Fluoride-based sources:					
Leukos Electro MIR 4.8	Cascaded, ending in InF_3	0.9 – 4.8 µm	600 mW	250 kHz	(13)
Thorlabs SC4500	Cascaded, ending in InF_3	1.3 - 4.5 µm	300 mW	50 MHz	(13)
NKT Photonics SuperK MIR, now: NORBLIS Aurora 5	Cascaded, ending in ZBLAN	1.4 – 4.1 µm	500 mW	2.5 MHz	(13,23)
Chalcogenide-based sources:					
NORBLIS Aurora	Cascaded, ending in As_2Se_3	1.5 – 10.5 µm	>80 mW	3 MHz	(21,23)

2.3 $\chi^{(2)}$-dominant supercontinuum generation

In case the nonlinear medium is not inversion symmetric, $\chi^{(2)}$ nonlinear effects are enabled and expected to dominate the (much weaker) $\chi^{(3)}$ effects. Important $\chi^{(2)}$ effects involve frequency-mixing processes such as second-harmonic generation, sum-frequency generation, optical parametric oscillation, and difference-frequency generation (DFG). As shown in Figure 2.6, frequency-mixing consists of processes in which (at least) two photons interact inside the nonlinear medium to create a photon with a different frequency. DFG is the process in which the resulting photon is the difference between the two input frequencies: $\omega_{DFG} = |\omega_1 - \omega_2|$.

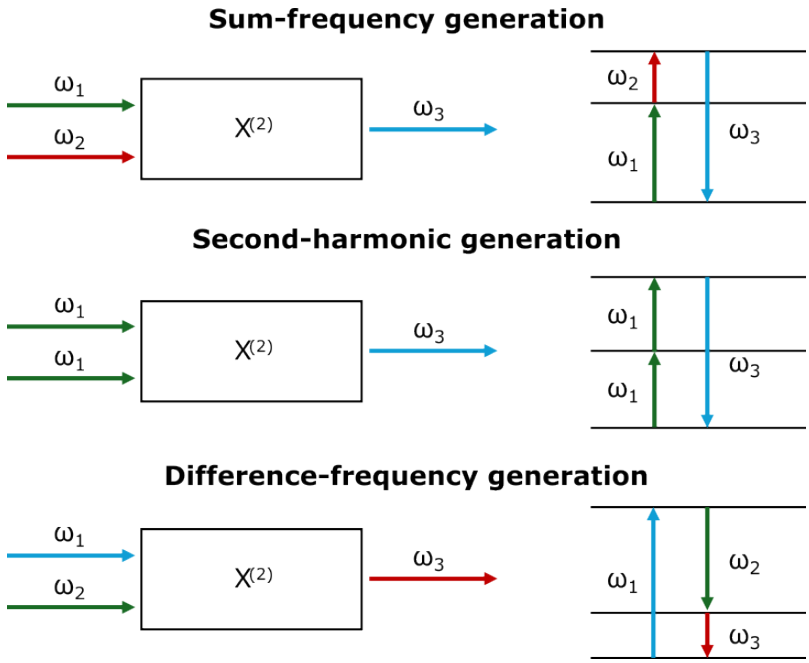


Figure 2.6: Schematic of three different three-wave mixing processes, with in- and outcoming waves depicted (left) and related energy diagram (right).

A specific variant of difference-frequency generation is intrapulse DFG (IDFG). This process is, mostly within the THz community, also referred to as optical rectification (24,25). In this case, the DFG process does not rely on two distinct input laser beams, but the DFG signal is generated from two distinct frequency components (e.g., ω and ω') of the same laser beam: $\omega_{IDFG} = |\omega - \omega'|$. As these components are close to each other in frequency, the resulting IDFG signal has a much lower frequency. Near-infrared pulses can thus be efficiently converted to signals deep into the infrared. For IDFG to occur, a spectrally broad input beam of sufficient power is required, but it reduces constraints on phase-lock between the two input beams, as the two components from the same beam are inherently locked. Since the IDFG signal is generated from a spectrally broadened input pulse, the IDFG signal is a sum of many different combinations of frequency components, leading to a broad spectrum (26).

In contrast to $\chi^{(3)}$ nonlinear effects, $\chi^{(2)}$-based effects can be steered towards a preferred process. The wavelengths of the light source used, in combination with the crystal, determine which process takes place. Most importantly, the selectiveness of $\chi^{(2)}$ effects is the result of constraints like phase-matching. In its essence, phase-matching is constructive interference of the radiation generated at different depths in the nonlinear crystal.

Perfect phase-matching is difficult to achieve, as due to dispersion photons with different wavelengths travel at different velocities through the crystal. As a result of this different velocity between the input waves (ω_1 and ω_2) and the generated wave (ω_3), the constructive interference of generated photons with frequency ω_3 is lost and the law of momentum conservation (the momentum of ω_1, ω_2, and ω_3 is to obey $\vec{k}_3 = \vec{k}_1 + \vec{k}_2$) is broken.

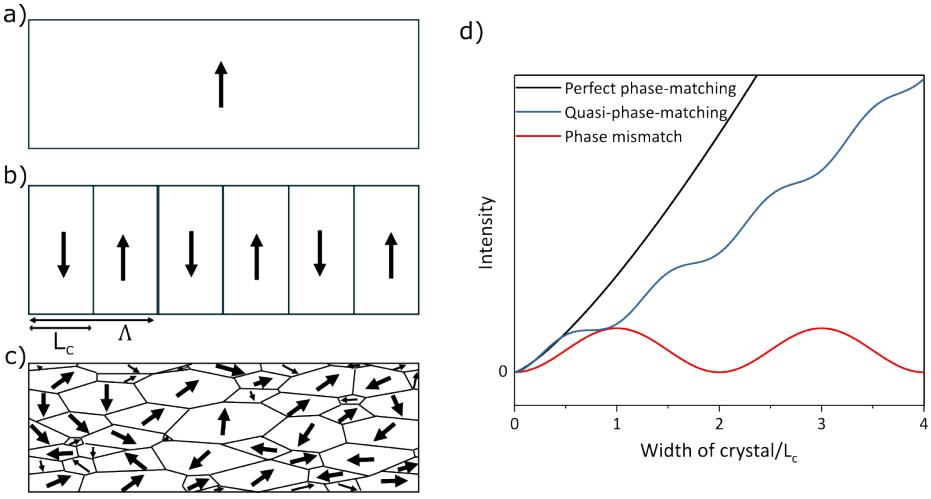


Figure 2.7: The principle of quasi-phase-matching. A crystal with a single orientation (**a**) will result in a phase mismatch of the generated wave. In a periodically poled (**b**) or polycrystalline (**c**) crystal, quasi-phase-matching can be achieved. The effective intensity of the generated wave is shown in **d**) for perfect phase-matching, quasi-phase-matching, and phase-mismatching conditions.

The maximum length after which the input and generated light walk out of phase would dictate the maximum length of the crystal. This maximum can be derived from the quasi-phase-matching (QPM) length Λ, which is defined as the distance over which a 2π phase difference is realized. To overcome this limitation and allow for longer coherence in the crystal, a new crystal could be introduced after each $\frac{1}{2}\Lambda = L_c$ with an opposite direction of $\chi^{(2)}$, effectively resetting or compensating for the phase mismatch. This is known as quasi-phase-matching (QPM) through periodic poling. Early attempts to achieve QPM were made by slicing the crystal into very thin layers of L_c , and inverting the orientation of each layer when putting the layers back together. Currently, the common approach is to apply a periodic charge during crystal growth to induce the periodic poling within a crystal (27). The L_c is specific to the wavelengths of the input and generated waves, and thus the poling period provides a regime for which only specific wavelengths will not walk out of phase. This results in a crystal highly specific for the generation of a signal with a specific, mostly narrowband, wavelength.

An alternative approach is random quasi-phase-matching (RQPM)(28). Rather than the organized, alternating structure of the periodically poled material, the RQPM crystal consists of a random polycrystalline structure with a broad distribution of grain sizes and orientations: it is a multitude of microscopic single-crystal grains. Therefore, RQPM materials do not allow for phase matching with a specific narrowband wavelength but rather allow for phase matching of a broader spectral coverage. Even though the loss in specificity comes with a loss in conversion efficiency, it opens the opportunity to generate broadband, supercontinuum spectra (26). Steering for different distributions of grain sizes and orientations during the fabrication of the crystal makes the crystal more favorable to a specific nonlinear process and wavelength coverage.

2.3.1 Crystal-based supercontinuum sources

The crystals used are made from various materials that have different transmission windows and parameters but share a high nonlinear coefficient. An overview is provided in Figure 2.8. There is a wide variety of infrared light sources using these $\chi^{(2)}$ nonlinear crystals. For the sake of brevity, some recent examples of sources with particularly broadband mid-infrared coverage are highlighted here.

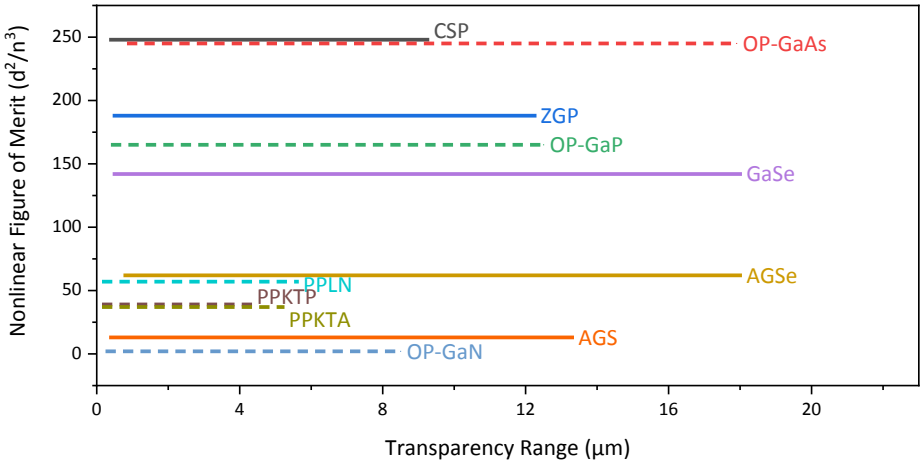


Figure 2.8: Examples of different crystals used in $\chi^{(2)}$ processes, with their transparency range and nonlinear coefficient indicated, based on (29,30). Crystals supporting QPM through periodic poling (PP) or orientation patterning (OP) are dashed. CSP: $CdSiP_2$, ZGP: $ZnGeP_2$, AGSe: $AgGaSe_2$, LN: $LiNbO_3$ (lithium niobate), KTP: $K[TiO]PO_4$ (potassium titanyl phosphate), KTA: $K[TiO]AsO_4$ (potassium titanyl arsenate), AGS: $AgGaS_2$.

Orientation-patterned gallium phosphide (OP-GaP) crystals are, together with periodically poled lithium niobate (PPLN), the most frequently used crystals for

achieving (periodically poled) quasi-phase matching (30). Alternatively, OP-GaAs crystals can be used. For example, Khodabakhsh et al. produced a system covering the spectrum from 3 – 5.4 µm using an optical parametric oscillator (OPO) with an OP-GaAs crystal for the detection of CH_4 and NO (31). Muraviev et al. later configured a similar system using an OPO with an OP-GaAs crystal with ~70 mW (32). OPO shares similarities with DFG but places the crystal in a cavity in which the ω_2, here called 'signal', is also generated by the $\chi^{(2)}$ interaction and kept resonating in the cavity.

An OP-GaP crystal can also be used for DFG, as shown by Timmers et al. (33), who used it to create a broadband spectrum spanning 4 – 12 µm with 0.25 mW of power. These very broadband characteristics of the (intra-pulse) DFG process are recently implemented in a commercial laser system (34), equipped with a zinc gallium phosphide (ZGP) crystal. This system achieves a spectral coverage of 3.5 – 11.5 µm at a power of ~20 mW.

The IDFG-based system used in Chapters 6 up to 9 is a commercial laser system also based on ZGP by Vasilyev et al. (see Figure 2.9) (35,36). In the systems of Timmers et al. (33) and Salem et al. (34), the spectral broadening and redshift of the input beam prior to the DFG stage is done using nonlinear fibers. In contrast, Vasilyev et al. use a master oscillator – power amplifier using Cr:ZnS crystals. As a result, the input power of the DFG in the system of Vasilyev et al. is ~ 3 W, roughly an order of magnitude stronger than the power delivered through the nonlinear fibers. Consequently, the resulting power output of the IDFG-produced beam in the spectral range 6 - 12 µm (~300 mW) is an order of magnitude stronger as well.

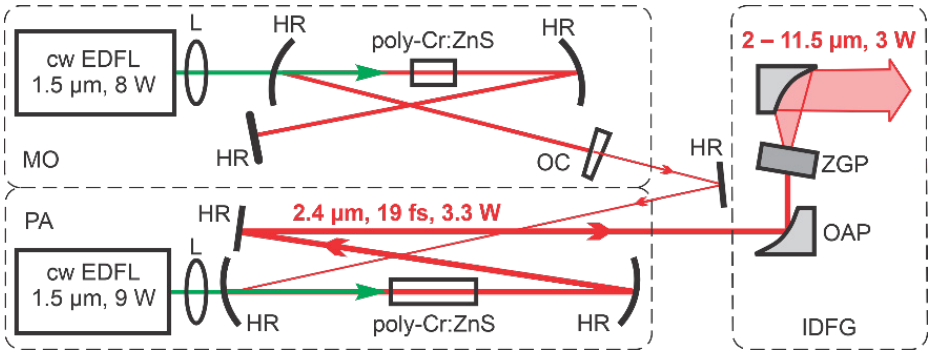
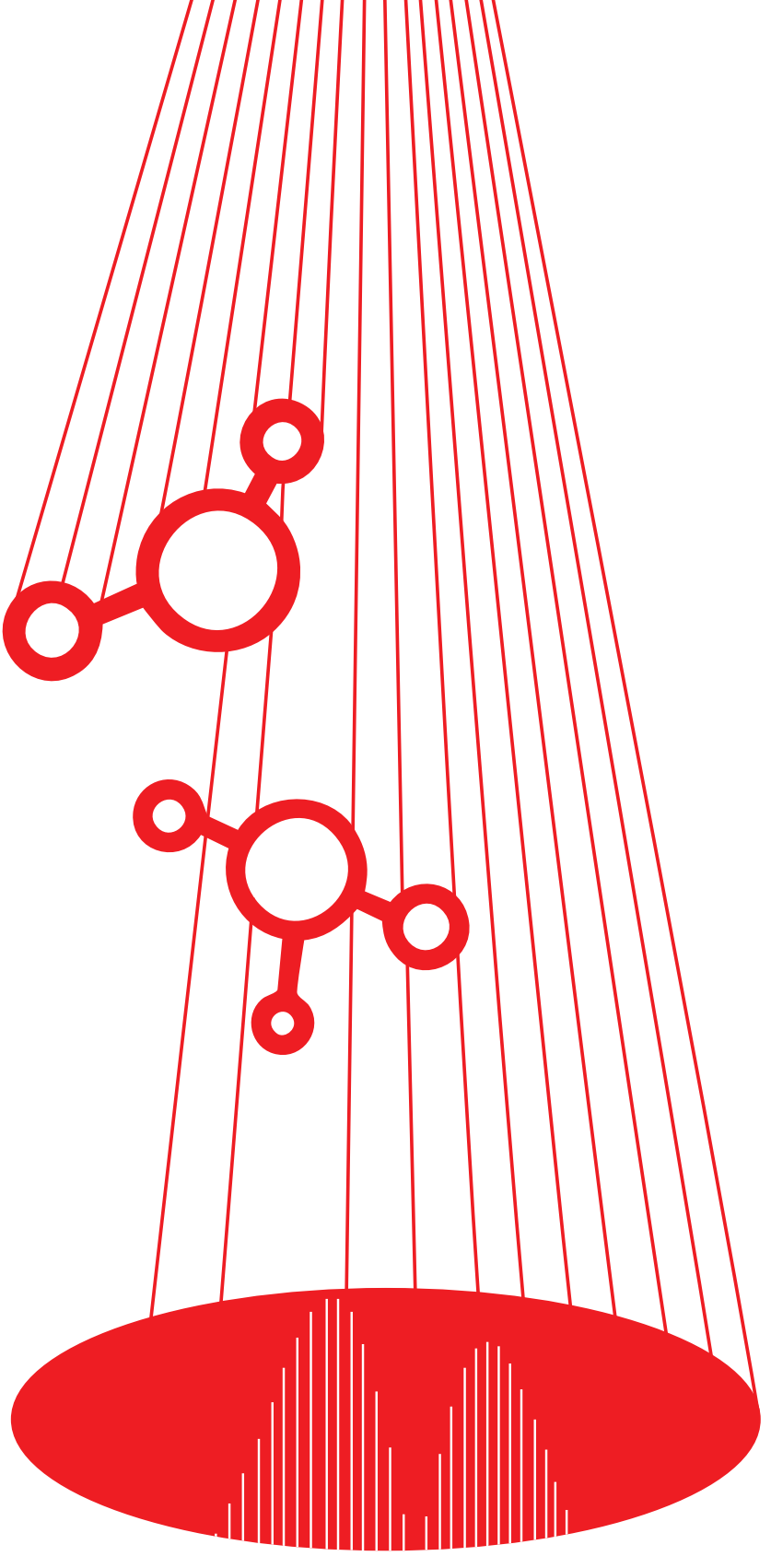


Figure 2.9: Schematic structure of the IDFG source used in chapters 6 up to 9, consisting of a MO-PA structure pumping the IDFG stage with 3.3 W, resulting in a total power output of 3 W of which 0.3 W in the range 6-11.5 µm.



Chapter 3

Trace gas sensing with broadband infrared supercontinuum sources

Having covered the concept of infrared light absorption in gas samples and introduced the broadband infrared supercontinuum sources, the spectroscopic system is completed with the spectrometer. Subsequently, it can be shown how the retrieved absorbance spectra are modeled to identify the presence of certain compounds and quantify their concentration. Finally, relevant parameters and metrics to assess the spectroscopic performance of the system for (trace) gas sensing can be given.

3.1 Fourier-transform spectroscopy

The spectrometer is used to unravel the spectrum of the detected light. Spectrometers can be classified predominantly into two types: dispersive and interferometric spectrometers. The most common dispersive spectrometer is equipped with a grating as the dispersive element to spatially separate different wavelengths of the light source. In recent years, grating spectrometers have been used in combination with supercontinuum sources (37,38). Here, grating-based spectrometers exhibited fast measurements of the spectrum (up to tens of Hz), at the cost of a poor spectral resolution (~90 up to 240 GHz) and a limited spectral coverage. For specific applications, such as the detection of larger (bio)molecules with a broad spectral profile that does not rely on a high spectral resolution and that can be detected in a specific spectral window, grating-based spectrometers can be the preferred option.

For their potential for broad spectral coverage and high spectral resolution, interferometric spectrometer types are often the preferred option. Furthermore, they have the potential to be inherently frequency-calibrated (which will be explained below). Interferometric spectrometers depend on the interference of light to create a pattern – an interference pattern or interferogram – that can be resolved to a spectrum via Fourier analysis. Two main advantages of Fourier-transform spectrometers over grating-based spectrometers are Fellgett's and Jacquinot's advantage (39).

Fellgett's advantage arises from the fact that the interferometer is measuring the total light intensity of the whole spectrum constantly. This contrasts with the grating spectrometer, which disperses the light, leaving the detector to only measure the intensity of the light of a specific wavelength at a time. This is an advantage for the interferometer when the measurement is detector-noise limited. It yields an improvement in the SNR that scales with , in which is the number of spectral elements in the spectrum, of the interferometer measurement compared to a grating-based measurement at the same resolution and measurement time. An important note is that this advantage is not fully exploited when other types of

noise are dominant, such as relative intensity noise: a highly relevant noise source when working with (fiber-based) supercontinuum sources.

The spectral resolution of a Fourier-transform spectrometer depends solely on the scan length of the arms of the interferometer. This is known as Jacquinot's advantage, as in contrast, dispersive spectrometers require optical elements such as gratings or slits, whose apertures always pose trade-offs between the resolution and the optical throughput.

Different interferometric spectrometer designs exist, such as Mach-Zehnder interferometers (40), often used to detect differences in refractive index or phase shifts between samples, wavefront-division spectrometers (41) or Michelson interferometers. Most interferometers depend on delay lines or mechanically moving arms to create a path length difference. A notable exception is dual frequency comb spectroscopy (42), which creates an interferogram through the heterodyne beating of two frequency comb lasers with a slight difference in their emitted wavelengths (through a small difference in their repetition rate frequency). While dual frequency comb systems are attractive for their advantages – which they share with other interferometer designs – such as broad spectral coverage and high spectral resolution, they also overcome the largest disadvantage of these interferometers: they can have much faster measurement times as they do not need any mechanically moving components to produce the interferogram. On the other hand, dual frequency comb systems are limited to be used with two coherent frequency combs, with associated higher costs (as usually two lasers are required) and added complexity, as both systems need to be stabilized and coherent with each other.

3.1.1 The Michelson interferometer

The Michelson-type interferometer is a classical and broadly applied spectrometer. While its initial purpose in 1887 was to prove the existence of ether (in which it famously failed (43)) it can serve as a very reliable method to detect a path length difference between two "arms" of the measurement, finding applications even in the detection of gravitational waves in the LIGO experiment (44). Moreover, reversing the logic, for a known path length difference in the two arms of the experiment, the retrieved interferogram can be used to determine the wavelength of the light source, which is exactly the purpose of the spectrometer.

A basic example of the Michelson interferometer can be seen in Figure 3.1.a. When used as a spectrometer, this system is referred to as a Fourier-transform spectrometer (FTS). In case the interferogram of a narrowband light source is detected over the displaced distance of the moving mirror, this will appear as a sinusoidal signal (Figure 3.1.b). As a result, the Fourier transform of this sinusoidal signal retrieves the narrow spectrum of this light source.

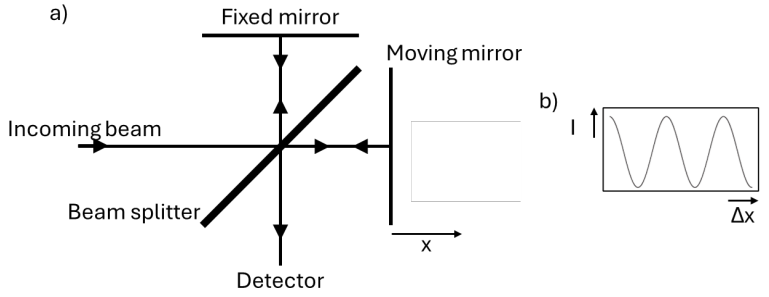


Figure 3.1: a) Basic layout of a Michelson interferometer. As the moving mirror translates over a distance (x), it creates an optical path difference (Δx) between the two arms of the beam. **b**) The interferogram is measured at the detector as a function of the optical path difference.

3.1.2 A Fourier-transform spectrometer for supercontinuum-based spectroscopy

When broadband light is supplied to the FTS, this can be pictured as the superposition of individual light beams of distinct wavelengths (Figure 3.2.a). Each of these beams, traveling through the interferometer, creates a sinusoidal interferogram. However, due to the difference in wavelength, their constructive and destructive interference occurs at different displacements (Δx). Consequently, this constructive and destructive interference of the different light beams cancel each other out, except for the optical path difference being – almost – zero (Δx = 0, also referred to as the zero path difference, ZPD), where all beams have constructive interference. For this reason, the interferogram of a broadband light beam appears as a burst, centered around the ZPD (Figure 3.2.b).

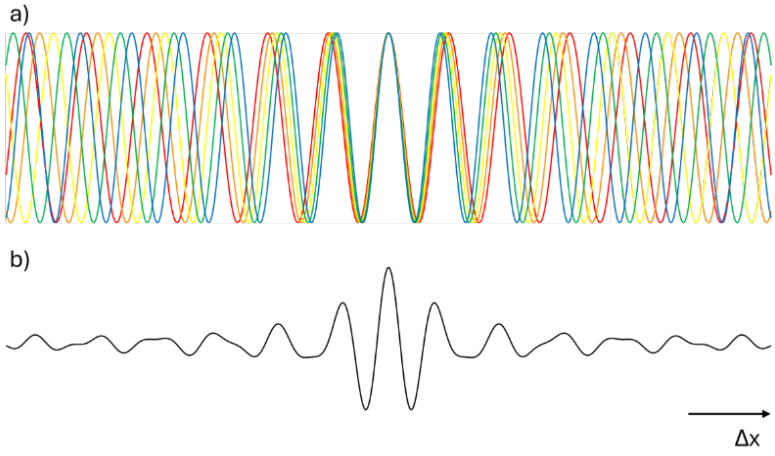


Figure 3.2: The interferogram of a broadband light source, pictured as individual components with distinct wavelengths (**a**) and the resulting, detected interferogram "burst" (**b**).

The design of the Fourier-transform spectrometer central to the work presented in this thesis is shown in Figure 3.3. The design and development of this spectrometer, tailored to be used with broadband infrared supercontinuum sources, is described in the paper by Abbas et al.(23).

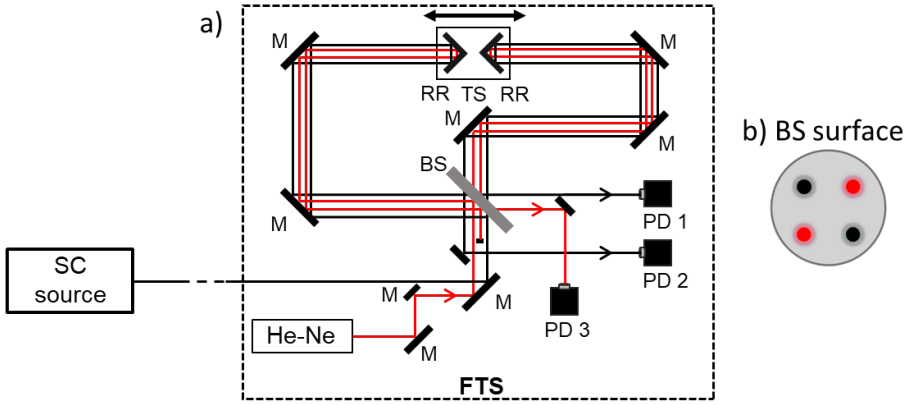


Figure 3.3: a) Schematic of the Fourier-transform spectrometer used, with **b**) placement of in and outcoming beams of the He-Ne laser (red) and supercontinuum source (black) on the beam splitter. M: Mirror, BS: Beam splitter, RR: Retroreflector, TS: translation stage, PD: Photodetector.

The incoming SC beam is directed to a ZnSe beam splitter, which is wedged to minimize etalons and optimized for the infrared spectral range of 1 – 12 µm (BSW711, Thorlabs). The incoming SC beam is split into two arms, which both propagate through a series of mirrors towards hollow retroreflectors that are mounted back-to-back on a moving translation stage. Retroreflectors are used, as they reflect the beam under the same angle, but with slight x - y translation. Therefore, and because of the coherent and small beam size of the light sources used, the returning beam does not overlap with the initial beam upon recombination on the beam splitter (Figure 3.3.b). Mounting both retroreflectors on the translation stage doubles the optical path difference (OPD) compared to the classical interferometer design with one moving arm. Moving the stage 1 cm consequently results in a 2 cm change in optical path length per measurement arm (as the beam propagates to and from the retroreflector), creating a total OPD of 4 cm for both arms together. The spectral resolution can be calculated as the reciprocal of the OPD, thus, a 1 cm translation of the stage results in a spectral resolution of 0.25 cm^{-1}.

To have precise knowledge of the absolute wavelengths of the detected SC beam, the exact displacement (OPD, Δx) of the scanning arms of the interferometer should be known. With the photodetectors detecting the interferogram as a function of time, and the translation stage being subject to the laws of physics and

having a moment of inertia, there is a need for careful calibration and conversion of the time-dependent interferogram to an Δx-dependent interferogram. Therefore, the He-Ne beam is sent in parallel to the SC beam through the interferometer and detected by a separate photodetector. As the He-Ne has a known, stable, narrowband wavelength, its detected sinusoidal interferogram over time has a period that corresponds to a specific step size in the OPD domain. Resampling the SC interferogram in the time domain to a sampling step size of a period of the He-Ne interferogram converts the time-domain SC interferogram to a well-calibrated SC interferogram in the OPD-domain – in units of, e.g., cm (see Figure 3.4). The Fourier-transform of this interferogram yields a spectrum in the reciprocal of these units, e.g., cm^{-1}, also known as wavenumbers.

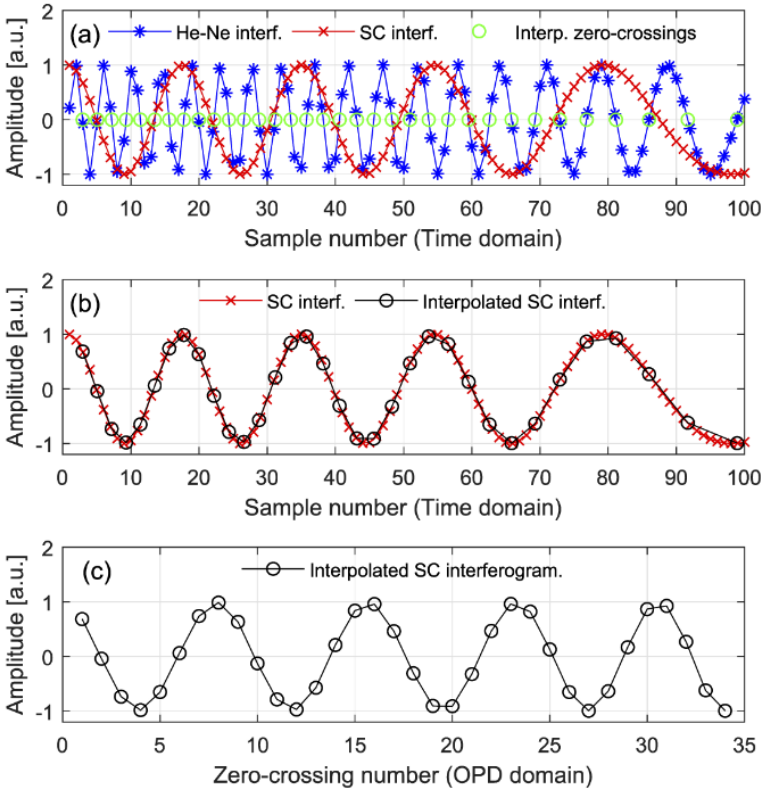


Figure 3.4: Visualization of the resampling of the SC interferogram. **a)** A sinusoidal He-Ne interferogram (blue) and SC interferogram (red) are sampled over time. The zero-crossings of the He-Ne (green) are taken to get a Δx-dependent time interval. **b)** The SC interferogram is resampled to the zero-crossings of the He-Ne, resulting in a SC interferogram sampled over Δx (**c**). Figure reused from (23).

Another advantage of separating the returning beam from the retroreflector from the incoming beam on the beam splitter, is that the recombined beam can be captured by a photodetector for both directions it travels in. Importantly, these two beams, both a recombination of the light previously split into two arms of the interferometer, have opposite phases.

Therefore, as the two out-of-phase interferograms are detected by a separate photodetector, the subtraction of these two signals results in a doubling of the amplitude of the detected interferogram. Moreover, any common mode noise of the two signals is still in phase and will cancel. Specifically (fiber-based) SC sources are known to have a large relative intensity noise (RIN). This arises, for example, from supercontinuum generation through modulation instability processes (see Chapter 2.2). Such noise sources appear as common mode noise in the interferogram signal, and thus, balanced detection is an effective way to reduce the RIN typical to (fiber-based) SC sources.

3.2 Modeling absorption spectra

The spectra acquired with the FTS described in the previous chapter are, after the Fourier-transform of the interferogram, intensity spectra (I). According to the Beer-Lambert law, after the absorption of parts of this light intensity by the molecules in the gas sample, the intensity spectrum can be described as

$$I_T(\tilde{\nu}) = I_0 e^{-A(\tilde{\nu})}. \tag{3.1}$$

Here, the absorbance (A) is a function of the wavenumber ($\tilde{\nu}$), thus yielding an absorbance spectrum ($A(\tilde{\nu})$). The absorbance spectrum can be derived by measuring a background spectrum (I_0) in which the gas sample was not present and a spectrum of the gas sample (I_T), as

$$A(\tilde{\nu}) = -\ln\left(\frac{I_0}{I_T}\right). \tag{3.2}$$

This logarithmic scale is used to ensure that the absorbance depends linearly on the concentration of the molecule in the gas sample (enclosed in the number density (N)) as

$$A(\tilde{\nu}) = \alpha L = S_{ij} N L \chi(\tilde{\nu}). \tag{3.3}$$

In this formula, the absorbance (unitless) over wavenumber is described in terms of the absorption coefficient (α, in cm^{-1}) and the optical interaction path length (L, in cm). The absorption coefficient is the product of the spectral line intensity

(S_{ij}, in cm^{-1}/(molecule·cm^{-2}), or simply cm/molecule), which is a standardized parameter to describe the intensity of the absorption line corresponding to the transition of state i to j at a temperature of $T_{\text{ref}} =$ 296 K, the number density (N, in molecules/cm^3), and the line shape ($\chi(\tilde{\nu})$, in cm) over wavenumber.

The number density depends on the conditions of the gas. For an ideal gas, according to the ideal gas law, $PV = nk_BT$, a standardized number can be found for the number of molecules (n) in 1 cm^3 at standard conditions (T_0= 271.15 K, P_0 = 1 atm). This number, the Loschmidt constant (n_0) is

$$n_0 = (P_0)/(k_BT_0) = 2.686 \cdot 10^{19}\ \text{molecules/cm}^3, \tag{3.4}$$

in which k_B is the Boltzmann constant. To calculate the number density from the Loschmidt constant for the specific gas species in a gas sample at specific pressure and temperature would then be

$$N = c\frac{n_0T_0}{P_0}\frac{P}{T} \tag{3.5}$$

when also the concentration of the gas in the sample (c, as volume fraction) is taken into account. The number density is now dependent on the concentration (c), pressure (P), and temperature (T).

Substituting equation (3.5) into equation (3.3) results in a formula for the absorbance as

$$A(\tilde{\nu}) = S_{ij}c\frac{n_0T_0}{P_0}\frac{P}{T}L\chi(\tilde{\nu}). \tag{3.6}$$

The line shape of the features in the absorbance spectrum, $\chi(\tilde{\nu})$, is introduced through two mechanisms. The conditions of the gas and the resolution parameters of the measurement instrument. Therefore, the total line shape can be described as the product of these two independent factors, as $\chi(\tilde{\nu}) = \chi_{\text{gas}}(\tilde{\nu}) \cdot \chi_{\text{instr.}}(\tilde{\nu})$ and inserted in equation (3.6):

$$A(\tilde{\nu}) = S_{ij}\frac{n_0T_0}{P_0}\frac{Pc\chi_{\text{gas}}(\tilde{\nu})}{T}L\chi_{\text{instr.}}(\tilde{\nu}). \tag{3.7}$$

Interestingly, this has become a formula that can be used for modeling the absorbance spectrum of a certain gas compound, in which each term has a distinct origin or role in the equation. The first term, S_{ij}, is the information retrieved from the HITRAN database (45), specific to a certain rovibrational transition for a single molecule at standardized conditions. The second term, $\frac{n_0T_0}{P_0}$, is a constant, accounting for the number of molecules in a volume. The third term, $\frac{Pc\chi_{\text{gas}}(\tilde{\nu})}{T}$,

describes conditions of the gas sample, and the fourth term, $L\,\chi_{\text{instr.}}(\tilde{\nu})$, a condition of the measurement instrument.

3.2.1 Temperature-dependence of the spectral line intensity

As mentioned, the spectral line intensity describes the intensity of a certain absorption line at standardized conditions of a temperature of $T_{\text{ref}} = 296\text{K}$. This dependency on temperature relates to the population distribution of the molecule over its lower state energy levels. For example, at a lower temperature, more molecules will be in the vibrational ground state than at higher temperatures. Consequently, the observed intensity of the absorption lines corresponding to a lower energy state in the vibrational ground state will be stronger at lower temperatures. To adjust the line intensity for a different temperature, it can, according to (46), be recalculated as:

$$S_{ij}(T) = S_{ij}(T_{\text{ref}})\frac{Q(T_{\text{ref}})}{Q(T)}\frac{\exp\left(-\frac{c_2 E''}{T}\right)\left[1-\exp\left(-\frac{c_2 \nu_{ij}}{T}\right)\right]}{\exp\left(-\frac{c_2 E''}{T_{\text{ref}}}\right)\left[1-\exp\left(-\frac{c_2 \nu_{ij}}{T_{\text{ref}}}\right)\right]}. \tag{3.8}$$

Here, c_2 is the second radiative constant, $c_2 = {}^{hc}/_{k_B}$, i.e., Planck's constant and the speed of light, divided by the Boltzmann constant. Q is the total internal partition function of the molecule, and E'' and ν_{ij} the lower state energy and center wavenumber of the transition associated with S_{ij}. This formula consists of three terms adjusting the line intensity: the first term accounting for the change in the total internal partition function, which is provided in the HITRAN database, the second for the change in the Boltzmann distribution, and the third for the effect of stimulated emission. For mid-infrared transitions – roughly in the wavenumber range 1000-3000 cm^{-1} – and relatively low temperatures (up to several hundred K), the stimulated emission term is negligible, resulting in the spectral line intensity predominantly influenced by the Boltzmann distribution. An example is provided in Figure 3.5, in which the absorbance spectrum is simulated for the R-branch of the ν_3 mode of CH_4 for a temperature of 250 K (black) and 450 K (blue, inverted). Using equation (3.8), the population is distributed over the states belonging to the different rotational quantum numbers (J'').

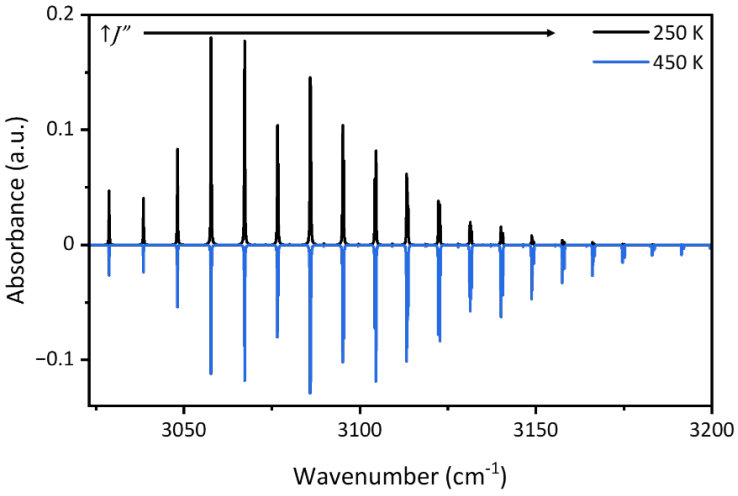


Figure 3.5: Simulated absorbance spectrum of the R-branch of the ν_3 mode of CH_4 for a temperature of 250 K (in black) and 450 K (in blue, inverted). At higher temperatures, the Boltzmann distribution over the rotational levels shifts to higher states.

An important assumption for this equation is that the system is in thermal equilibrium. That is, that the molecules have a Boltzmann distribution over the energy states. Under normal conditions, such as outdoor, environmental measurements, this holds. But various examples exist where this assumption does not hold, such as in jet-cooled experiments (47) or discharge experiments (48). For example, in the work presented in Chapter 5, a modified version of equation (3.8) was used to account for discrepancies between the vibrational and rotational energy distributions.

3.2.2 Line shape functions

The observed width of the absorption lines depends on the broadening of the absorption line in the gas and the measurement instrument. In the gas phase, two effects dominate the broadening of the line shape: the Doppler effect and molecular collisions. One can imagine these two effects by picturing the molecules moving freely through their medium in all directions: some molecules will propagate in a direction opposite to that of the laser beam, causing a blueshift in the wavelength they perceive, while other molecules move along the direction of the laser beam, resulting in a perceived redshift. With the molecules moving randomly in arbitrary directions, this leads to a Gaussian distribution of the absorption around its center frequency. The broadening scales with the velocity of the molecules, i.e., a higher temperature equals a higher translational energy, resulting in a broader

absorption profile. This type of broadening, resulting from the molecules moving in different directions, is thus referred to as inhomogeneous broadening or Doppler broadening.

Furthermore, the gas molecules continuously collide with each other. During a collision, the molecules can exchange energy, changing their energy state. More frequent collisions thus mean that the lifetime of the energy states is shorter. The probability of collision depends on the number of molecules inside the volume, which, following the ideal gas law, is analogous to the pressure inside the gas. In other words, a higher pressure reduces the lifetime of the energy states of the molecules. Following Heisenberg's uncertainty principle, a shorter lifetime corresponds to a higher uncertainty in energy, thus the width of the absorption line is larger. Since all molecules experience an identical effect of the collisions on their lifetime, this broadening is referred to as homogeneous broadening, or pressure or collision broadening. Together with the natural, inherent (and pressure-independent) lifetime of the energy state, it can be described by a Lorentzian profile.

The homogeneous and inhomogeneous broadening can be described together as a line shape with a Voigt profile. The Voigt profile is a convolution of the Gaussian and Lorentzian distributions and is used to account for the line shape $\chi_{\mathrm{gas}}(\tilde{\nu})$ in the absorbance spectrum, considering the aforementioned effects. The Voigt profile is widely used in spectroscopic applications. However, spectroscopy demanding extremely high spectral resolutions and accuracy in the line shape, such as the research field of precision spectroscopy, goes beyond the Voigt profile. The Hartmann-Tran profile has become the standard model when the highest possible accuracy in line shape is needed, but the seven different parameters needed to compute it, such as to account for the effect of collision-induced velocity changes and the speed dependence of the collisional broadening, limit the number of molecules that can currently be calculated with it (49,50).

The broadening caused by the measurement instrument ($\chi_{\mathrm{instr.}}(\tilde{\nu})$) is also referred to as the instrument line shape (ILS). An important difference between these two broadening parameters is that the line shape from the gas parameters applies to an individual absorption line, while the instrumental line shape applies to the spectrum as a whole.

In a Fourier-transform spectrometer, the spectrum is acquired by measuring the interferogram over a finite distance: in most cases by physically moving a translation stage. Picturing a narrowband light source with a sinusoidal interferogram, a finite number of periods of this sine is recorded. Upon resolving its spectrum, there is a limit to the spectral resolution. However, the bigger the interferogram, the more periods of the sinewave are recorded, and thus the higher the precision in determining its frequency.

The width of the spectral line is inversely related to the length of the interferogram as $\Delta\tilde{\nu} = {}^{1}/_{x}$. The interferogram with finite width can be seen as a rectangular, or boxcar, function applied to the theoretically infinite interferogram. The line shape associated with a rectangular function is a sinc function. Consequently, all absorption features with a spectral width narrower than the resolution of the measurement instrument will get broadened with a distinct sinc line shape.

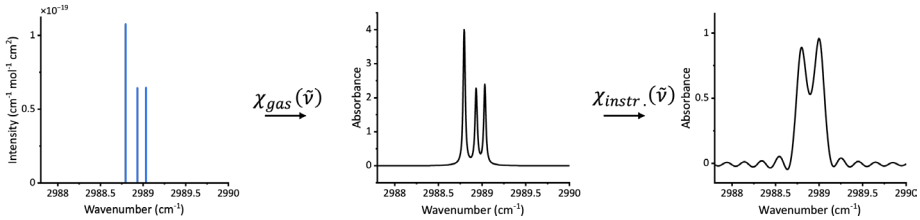


Figure 3.6: The effect of the line shape on the detected absorbance signal. The $\chi_{gas}(\tilde{\nu})$ is calculated with a Voigt profile for a temperature of 298 K and a pressure of 200 mbar. The $\chi_{instr.}(\tilde{\nu})$ is calculated with a sinc profile for an instrumental resolution of 0.1 cm^{-1}.

3.3 Quantifying spectroscopic performance

There are several ways to assess the performance of a measurement instrument. The goal is two-fold: to give an indication of the performance of the system for a certain application, and to be able to compare different (types of) systems. As for the latter application, these values often are theoretical limits, and in practical applications, the achieved detection limit is worse than the estimations in these metrics. This is the result of measurement effects, such as overlapping absorbance features of other compounds in the gas matrix. A way to make more reliable sensitivity estimations for a specific use case would be to use a test environment with realistic simulated spectra and apply the data analysis procedure to these spectra to determine the accuracy and precision of the data analysis. This is investigated further in the work presented in Chapter 8.

The limit of detection of the system is a result of noise in the spectrum, obscuring (the fit to) the absorption lines. One common type of noise is white noise: a frequency-independent noise (thus adding an equal amount of noise over the entire spectrum), mostly arising from thermal noise in the detectors or shot noise. When white noise is dominating the system, averaging can be used to reduce the noise and increase the sensitivity. The reduction of white noise over time by averaging scales with ${}^{1}/_{\sqrt{\tau}}$ for an averaging time of τ.

3.3.1 Sensitivity

The simplest way to assess the sensitivity of a system is to measure a gas with a stable, known concentration with the system. Assuming accurate detection of the stable gas sample, measuring and averaging more spectra of the gas will reduce the noise in the spectrum and improve the retrieved concentration. The standard deviation of the retrieved concentration of multiple measurements gives an idea of the precision of the system and thus the sensitivity of the system. Usually, the units associated with this figure are in volume fraction when associated with a specific measurement time, e.g., 2 ppm for a 1-second measurement.

When comparing systems with different measurement speeds, though, it is relevant to incorporate the time in this figure. Assuming the sensitivity is white-noise limited, the appropriate unit is volume fraction times the square root of time, e.g., $\mathrm{ppm} \cdot \sqrt{\mathrm{s}}$ or equivalently $\mathrm{ppm}/\sqrt{\mathrm{Hz}}$.

Spectroscopic systems with a variable optical interaction path length, such as open-path systems (see Chapter 7), have another variable to consider in their assessment of the sensitivity. The sensitivity of the systems depends on the (variable) path length, as from equation (3.3) $A \propto L$. Consequently, sensitivities for such systems can be expressed as path-integrated concentrations rather than concentrations. Path-integrated concentrations, in units of volume fraction times path length, e.g., ppm·m, express the volume fraction of molecules for a one-meter column of gas. That is, 1000 ppm·m equals the absorbance signal of both a concentration of 1000 ppm in 1 meter path length and 1 ppm in 1 km path length. This enables comparison of open-path systems with different optical interaction path lengths but is also often used in investigations of samples with inhomogeneous concentration distributions, such as when detecting leaks (51) or emissions (52), or for applications in which the path length cannot be determined with certainty, such as LiDAR-based sensing (53,54).

3.3.2 Noise-equivalent absorption sensitivity

A method to quantify the sensitivity of the system based on its noise and independent of the availability of a stable, calibrated gas standard is noise-equivalent absorption. In the absence of a gas sample, the absorption spectrum should be free of absorption peaks and dominated by noise. In case the spectrum is dominated by white noise, a common way to characterize the noise and make it comparable across different systems is the noise-equivalent absorption sensitivity (NEAS). The goal of the NEAS is to relate the noise of the measurement system to its sensitivity in the application of gas sensing and provide a figure that can compare different systems, e.g., fast versus slow detection systems, or broadband versus narrowband systems. The NEAS is defined as

$$\mathrm{NEAS} = \frac{\sigma\sqrt{2T}}{L\sqrt{M}}, \tag{3.9}$$

with the standard deviation of the noise in the *absorption* spectrum (σ), the measurement time (T), the optical interaction path length (L), and the number of spectral elements (M), i.e., the bandwidth of the spectrum divided by the spectral resolution, $(\nu_{\mathrm{end}} - \nu_{\mathrm{start}})/\Delta\nu$. The NEAS, with units of $\mathrm{cm}^{-1}\,\mathrm{Hz}^{-1/2}$ spectral elements^{-1}, thus yields a value to compare different spectroscopic systems, accounting for noise, resolution, spectral coverage, and measurement speed, but does not give a tangible number for the performance in detecting a specific gas.

3.3.3 Noise-equivalent concentrations and Allan-Werle deviation

The absorption-free spectra used for the NEAS can also be used to fit the spectra of a desired gas compound. These fits yield concentrations representing the uncertainty introduced by the noise: noise-equivalent concentrations (NECs). These NECs can be used to estimate the sensitivity of the system. An important remark is that these NECs only account for sensitivity limitations due to instrument noise, without considering effects due to gas composition, such as overlap with absorption lines of other gas species, or other effects.

The NECs can be used to build an Allan-Werle deviation plot. The concept of Allan deviation was originally developed to study the stability of (atomic) clocks (55). Similar to enhancing the sensitivity of investigating gas samples, the frequency of an oscillator used as a reference for a clock can be determined more precisely by measuring more oscillations (i.e., measuring longer). The Allan deviation determines the frequency stability of an oscillator by plotting the deviation against the averaging time. The central formula to calculate the Allan deviation is

$$\sigma_{\mathrm{Allan}}(\tau) = \sqrt{\frac{1}{2(M-1)}\sum_{i=1}^{M-1}\left(A_{i+1}(\tau) - A_i(\tau)\right)^2}, \tag{3.10}$$

in which the standard deviation is determined for the averaged NECs (A) with an averaging time of τ. M is the number of averaged NECs (A) at this averaging time.

Peter Werle translated the concept of Allan deviation to gas-sensing instruments (56), in which the NECs are used to give a detection sensitivity as a function of averaging time. An example of a typical Allan-Werle deviation plot is given in Figure 3.7. As long as the measurement instrument is white-noise dominated, the sensitivity follows the expected trend of $\sigma \propto 1/\sqrt{\tau}$. However, a certain limit is reached after which long-term drifts in the system start to dominate.

Therefore, the Allan-Werle deviation plot is particularly useful to determine the long-term stability of a system and show the optimal sensitivity and averaging time.

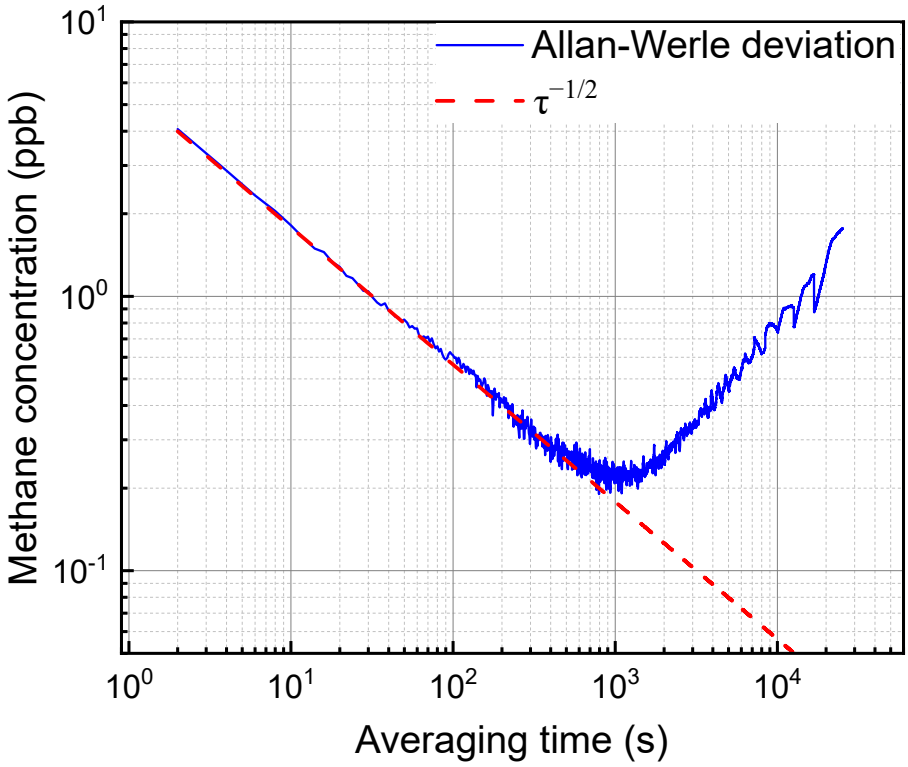


Figure 3.7: A typical Allan-Werle deviation plot, showing the detection sensitivity of methane improving as $\tau^{-1/2}$, which is expected for white-noise reduction. An optimum of 0.2 ppb is reached at an averaging time of 10^3 seconds, after which long-term instrument drifts become dominant. Data adapted from (57).

References

1. Tittel FK, Richter D, Fried A. Mid-Infrared Laser Applications in Spectroscopy. In: Sorokina IT, Vodopyanov KL, editors. Solid-State Mid-Infrared Laser Sources. Berlin, Heidelberg: Springer Berlin Heidelberg; 2003. p. 458–529. (Topics in Applied Physics; vol. 89).
2. Picqué N, Hänsch TW. Frequency comb spectroscopy. Nat Photon. 2019;13(3):146–157.
3. Hollas JM. Modern spectroscopy. 2nd ed. Chichester; New York: Wiley; 1992. 407 p.
4. Griffiths PR, de Haseth JA. Fourier Transform Infrared Spectrometry: Second Edition. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2007. XV p.
5. Hugi A, Maulini R, Faist J. External cavity quantum cascade laser. Semicond Sci Technol. 2010;25(8):083001.
6. Jahromi KE, Khodabakhsh A, Cristescu SM, Harren FJM. Applications of Supercontinuum Sources for Trace Gas Sensing. In: Meyers RA, editor. Encyclopedia of Analytical Chemistry. 1st ed. Wiley; 2024. p. 1–26.
7. Petersen CR, Moselund PM, Huot L, Hooper L, Bang O. Towards a table-top synchrotron based on supercontinuum generation. Infrared Phys Technol. 2018;91:182–6.
8. Petersen CR, Møller U, Kubat I, Zhou B, Dupont S, Ramsay J, et al. Mid-infrared supercontinuum covering the 1.4-13.3 µm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre. Nat Photon. 2014;8(11):830–834.
9. Vasilyev S, Smolski V, Moskalev I, Peppers J, Mirov M, Barnakov Y, et al. Multi-octave infrared femtosecond continuum generation in Cr:ZnS-GaSe and Cr:ZnS-ZGP tandems. In: Schunemann PG, Schepler KL, editors. Nonlinear Frequency Generation and Conversion: Materials and Devices XIX. San Francisco, United States: SPIE; 2020. p. 6.
10. Sylvestre T, Genier E, Ghosh AN, Bowen P, Genty G, Troles J, et al. Recent advances in supercontinuum generation in specialty optical fibers [Invited]. J Opt Soc Am B. 2021;38(12):F90.
11. Mitschke FM, Mollenauer LF. Discovery of the soliton self-frequency shift. Opt Lett. 1986;11(10):659.
12. Venck S, St-Hilaire F, Brilland L, Ghosh AN, Chahal R, Caillaud C, et al. 2–10 µm Mid-Infrared Fiber-Based Supercontinuum Laser Source: Experiment and Simulation. Laser & Photonics Reviews. 2020;14(6):2000011.
13. Zorin I, Gattinger P, Ebner A, Brandstetter M. Advances in mid-infrared spectroscopy enabled by supercontinuum laser sources. Opt Express. 2022;30(4):5222–5254.
14. Slusher RE, Lenz G, Hodelin J, Sanghera J, Shaw LB, Aggarwal ID. Large Raman gain and nonlinear phase shifts in high-purity As_2Se_3 chalcogenide fibers. J Opt Soc Am B. 2004;21(6):1146.
15. Wang Y, Dai S. Mid-infrared supercontinuum generation in chalcogenide glass fibers: a brief review. PhotoniX. 2021;2(1):9.
16. Saini TS, Sinha RK. Mid-infrared supercontinuum generation in soft-glass specialty optical fibers: A review. Progress in Quantum Electronics. 2021;78:100342.
17. Lemière A, Bizot R, Désévédavy F, Gadret G, Jules JC, Mathey P, et al. 1.7–18 µm mid-infrared supercontinuum generation in a dispersion-engineered step-index chalcogenide fiber. Results in Physics. 2021;26:104397.
18. Zhao Z, Wu B, Wang X, Pan Z, Liu Z, Zhang P, et al. Mid-infrared supercontinuum covering 2.0–16 µm in a low-loss telluride single-mode fiber. Laser & Photonics Reviews. 2017;11(2):1700005.
19. Dai S, Wang Y, Peng X, Zhang P, Wang X, Xu Y. A Review of Mid-Infrared Supercontinuum Generation in Chalcogenide Glass Fibers. Applied Sciences. 2018;8(5):707.

20. Bizot R, Tiliouine I, Désévédavy F, Gadret G, Strutynski C, Serrano E, et al. All-fiber supercontinuum absorption spectroscopy for mid-infrared gas sensing. APL Photonics. 2024;9(11):111303.
21. Woyessa G, Kwarkye K, Dasa MK, Petersen CR, Sidharthan R, Chen S, et al. Power stable 1.5–10.5 μm cascaded mid-infrared supercontinuum laser without thulium amplifier. Opt Lett. 2021;46(5):1129–1132.
22. El-Amraoui M, Gadret G, Jules JC, Fatome J, Fortier C, Désévédavy F, et al. Microstructured chalcogenide optical fibers from As_2S_3 glass: towards new IR broadband sources. Opt Express. 2010;18(25):26655.
23. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
24. Hoffmann MC, Yeh KL, Hebling J, Nelson KA. Efficient terahertz generation by optical rectification at 1035 nm. Opt Express. 2007;15(18):11706.
25. Kadlec F, Kužel P, Coutaz JL. Optical rectification at metal surfaces. Opt Lett. 2004;29(22):2674.
26. Zhang J, Fritsch K, Wang Q, Krausz F, Mak KF, Pronin O. Intra-pulse difference-frequency generation of mid-infrared (27–20 μm) by random quasi-phase-matching. Opt Lett. 2019;44(12):2986.
27. Boyd RW. Nonlinear optics. 3rd ed. Burlington, MA: Academic Press; 2008. 1 p.
28. Baudrier-Raybaut M, Haïdar R, Kupecek Ph, Lemasson Ph, Rosencher E. Random quasi-phase-matching in bulk polycrystalline isotropic nonlinear materials. Nature. 2004;432(7015):374–6.
29. Harren FJM. New coherent sources for mid infrared spectroscopic applications. In: OSA High-brightness Sources and Light-driven Interactions Congress 2020 (EUVXRAY, HILAS, MICS). Washington, DC: Optica Publishing Group; 2020. p. MTh2C.1.
30. Shumakova V, Heckl OH. A short guide to recent developments in laser-based gas phase spectroscopy, applications, and tools. APL Photonics. 2024;9(1):010803.
31. Khodabakhsh A, Ramaiah-Badarla V, Rutkowski L, Johansson AC, Lee KF, Jiang J, et al. Fourier transform and Vernier spectroscopy using an optical frequency comb at 3–54 μm. Opt Lett. 2016;41(11):2541.
32. Muraviev AV, Smolski VO, Loparo ZE, Vodopyanov KL. Massively parallel sensing of trace molecules and their isotopologues with broadband subharmonic mid-infrared frequency combs. Nat Photon. 2018;12(4):209–14.
33. Timmers H, Kowligy A, Lind A, Cruz FC, Nader N, Silfies M, et al. Molecular fingerprinting with bright, broadband infrared frequency combs. Optica. 2018;5(6):727.
34. Salem R, Backus S, Liu D, Wan C, Domingue S, Kirchner M, et al. 5-11 μm Supercontinuum Generation Using Cascaded Nonlinearities in Fluoride Fiber and ZGP. In: Conference on Lasers and Electro-Optics. San Jose, California: Optica Publishing Group; 2021. p. STh1L.5.
35. Vasilyev S, Moskalev IS, Smolski VO, Peppers JM, Mirov M, Muraviev AV, et al. Super-octave longwave mid-infrared coherent transients produced by optical rectification of few-cycle 25-μm pulses. Optica. 2019;6(1):111–114.
36. Vasilyev S, Muraviev A, Konnov D, Mirov M, Smolski V, Moskalev I, et al. Longwave infrared (6.6–11.4 μm) dual-comb spectroscopy with 240,000 comb-mode-resolved data points at video rate. Opt Lett. 2023;48(9):2273–6.
37. Jahromi KE, Pan Q, Høgstedt L, Friis SMM, Khodabakhsh A, Moselund PM, et al. Mid-infrared supercontinuum-based upconversion detection for trace gas sensing. Opt Express. 2019;27(17):24469.

38. Jahromi KE, Pan Q, Khodabakhsh A, Sikkens C, Assman P, Cristescu SM, et al. A Broadband Mid-Infrared Trace Gas Sensor Using Supercontinuum Light Source: Applications for Real-Time Quality Control for Fruit Storage. Sensors. 2019;19(10):2334.
39. Griffiths PR, De Haseth JA. Fourier transform infrared spectrometry. 2nd ed. Hoboken, N.J: Wiley-Interscience; 2007. (Chemical analysis).
40. Zetie KP, Adams SF, Tocknell RM. How does a Mach-Zehnder interferometer work? Phys Educ. 2000;35(1):46–8.
41. Zorin I, Gattinger P, Ricchiuti G, Lendl B, Heise B, Brandstetter M. All-mirror wavefront division interferometer for Fourier transform spectrometry across multiple spectral ranges. Opt Express. 2025;33(1):867.
42. Coddington I, Newbury N, Swann W. Dual-comb spectroscopy. Optica. 2016;3(4):414.
43. Giacomo P. The Michelson interferometer. Mikrochim Acta. 1987;93(1–6):19–31.
44. The LIGO Scientific Collaboration, Aasi J, Abbott BP, Abbott R, Abbott T, Abernathy MR, et al. Advanced LIGO. Class Quantum Grav. 2015;32(7):074001.
45. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf. 2022;277:107949.
46. Šimečková M, Jacquemart D, Rothman LS, Gamache RR, Goldman A. Einstein A-coefficients and statistical weights for molecular absorption transitions in the HITRAN database. J Quant Spectrosc Radiat Transf. 2006;98(1):130–155.
47. Perot S, Lecomte J, Suas-David N, Rutkowski L, Rey M, Kassi S, et al. Jet-cooled ethylene cavity ring-down spectroscopy between 5880 and 6200 cm^{-1}. Journal of Quantitative Spectroscopy and Radiative Transfer. 2024;324:109065.
48. Sadiek I, Fleisher AJ, Hayden J, Huang X, Hugi A, Engeln R, et al. Dual-comb spectroscopy of ammonia formation in non-thermal plasmas. Commun Chem. 2024;7(1):110.
49. Tennyson J, Bernath PF, Campargue A, Császár AG, Daumont L, Gamache RR, et al. Recommended isolated-line profile for representing high-resolution spectroscopic transitions (IUPAC Technical Report). Pure and Applied Chemistry. 2014;86(12):1931–43.
50. Ngo NH, Lisak D, Tran H, Hartmann JM. An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. Journal of Quantitative Spectroscopy and Radiative Transfer. 2013;129:89–100.
51. Soskind MG, Li NP, Moore DP, Chen Y, Wendt LP, McSpiritt J, et al. Stationary and drone-assisted methane plume localization with dispersion spectroscopy. Remote Sensing of Environment. 2023;289:113513.
52. Rodrigo PJ, Larsen HE, Ashik AS, Vechi NT, Kissas K, Fredenslund AM, et al. Fast horizontal radial plume mapping of N_2O using open-path absorption spectroscopy with a quantum-cascade laser. Atmospheric Environment. 2024;120510.
53. Cadiou E, Mammez D, Dherbecourt JB, Gorju G, Pelon J, Melkonian JM, et al. Atmospheric boundary layer CO_2 remote sensing with a direct detection LIDAR instrument based on a widely tunable optical parametric source. Opt Lett. 2017;42(20):4044.
54. Robinson R, Gardiner T, Innocenti F, Woods P, Coleman M. Infrared differential absorption Lidar (DIAL) measurements of hydrocarbon emissions. J Environ Monit. 2011;13(8):2213.
55. Allan DW. Statistics of atomic frequency standards. Proc IEEE. 1966;54(2):221–30.
56. Werle P, Mücke R, Slemr F. The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS). Applied Physics B Photophysics and Laser Chemistry. 1993;57(2):131–9.

57. Krebbers R, Van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, et al. Ultra-broadband spectroscopy using a 2–11.5 μm IDFG-based supercontinuum source. Opt Express. 2024;32(8):14506–14520.

Part A

Spectroscopy applications with fiber-based SC sources

Introduction

Fourier-transform spectroscopy with fiber-based infrared supercontinuum sources[2]

The capabilities of the fiber-based SC sources have been explored in the subsequent chapters for two sources. The first source has a spectral coverage in the infrared region from 1.4 to 4.1 µm, which contains the functional group region covering strong absorption features related to hydrocarbons (1). The source uses a cascaded fiber design terminating in a ZBLAN fiber. It was commercially available from NKT Photonics (SuperK MIR) and provides a power output of ~450 mW at a repetition rate of 2.5 MHz.

The second source was made available as a prototype, developed by Woyessa et al. (2). Its cascaded design contains a ZBLAN fiber, which is free space-coupled to a tandem of two chalcogenide fibers, delivering a spectral coverage of 1.5 - 10.5 µm (see Figure 2.5, Chapter 2). The repetition rate of this source is 3 MHz with a pulse duration of 0.5 ns, delivering ~86 mW of power.

Both sources were used in identical systems, consisting of a multipass Herriott-type absorption gas cell (HC30L, Thorlabs), and the previously introduced Fourier-transform spectrometer (FTS, see Figure 3.3). In the case of the chalcogenide fiber-based source, the FTS and surrounding system is purged with dry nitrogen gas to minimize the absorption of water in the range 5.5 – 7.5 µm.

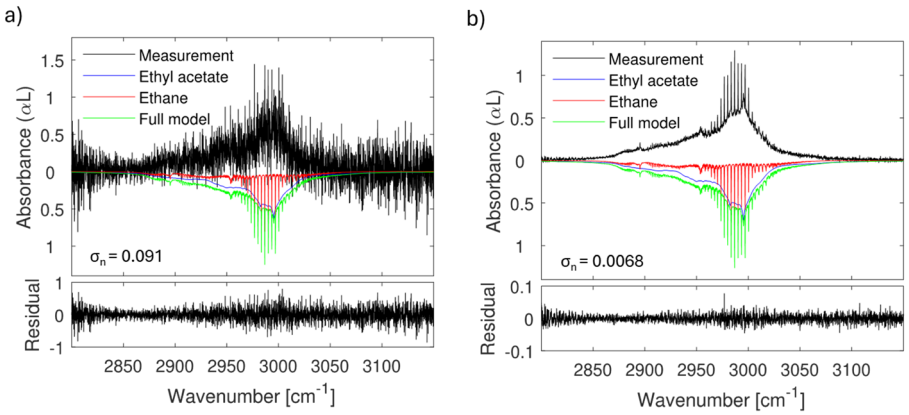


Figure A.1: Measured absorbance spectrum of 25.0 ± 0.7 ppm ethyl acetate and 5.0 ± 0.1 ppm ethane with the ZBLAN source (in black) using a single detector in the FTS **(a)** and using balanced detection **(b)**, resulting in a ~13 improvement in SNR. Reference spectra from the PNNL database (3) are shown in color, inverted.

[2] This chapter contains content from the following article: Abbas MA, Eslami Jahromi K, Nematollahi M, Krebbers R, Liu N, Woyessa G, Bang O, Huot L, Harren FJM, Khodabakhsh, A. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.

As a proof-of-concept of both systems for accurate gas detection, gas samples were measured of known composition. To demonstrate the effect of balanced detection in the FTS, a sample of a mixture of 25.0 ± 0.7 ppm ethyl acetate and 5.0 ± 0.1 ppm ethane in nitrogen gas was studied using the NKT Photonics ZBLAN-based source. The mixture was kept in the gas cell at 1 bar and the FTS was set to a spectral resolution of 3 GHz. The FTS measures the spectrum in 1.9 seconds. The results shown in Figure A.1 are averaged 500x (total measurement time: ~15 minutes). The reduction of relative intensity noise (RIN) in the spectrum using balanced detection can be seen by comparing Figure A.1.a and Figure A.1.b, where the standard deviation of the noise improved with a factor 13.

The spectrum of the chalcogenide source extends up to the molecular fingerprint region, which is the spectral region containing strong molecular absorption features beyond the strong water absorption window of 5.5 – 7.5 µm (1). Thus, this source is particularly useful for detecting molecules without strong absorption features in the functional group region. A sample of 30.0 ± 0.9 ppm SO_2 in nitrogen was measured at a spectral resolution of 3 GHz, averaging 4000 spectra in ~130 minutes (Figure A.2). Despite purging the system with dry nitrogen gas, some strong water absorption features interfere with the spectrum. A fit of the measurement to reference spectra constructed using the HITRAN database (4), results in a retrieved concentration of SO_2 of 29.4 ± 0.8 ppm.

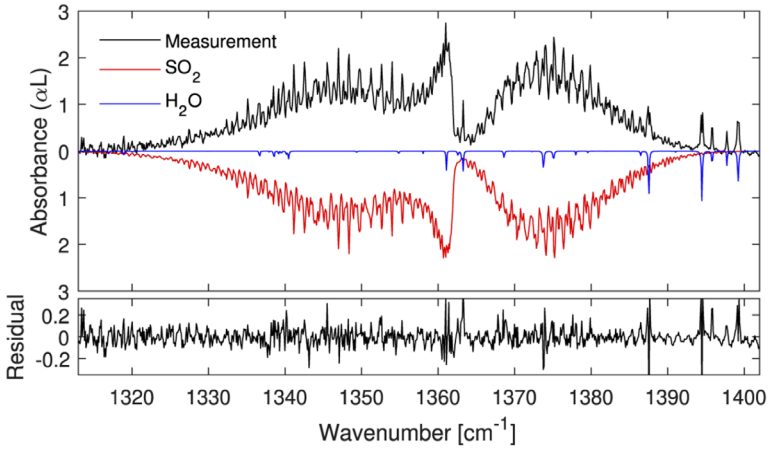


Figure A.2: Absorbance spectrum of 30.0 ± 0.9 ppm SO2, measured with the chalcogenide source (in black). Reference spectra from the HITRAN database (4) are shown in color, inverted.

The performance of the two sources in the spectroscopic setup with the FTS can be quantified and put into perspective with other systems with the noise-equivalent absorption sensitivity (NEAS) (see Table A.1). The NEAS indicates the sensitivity of the system in combination with the measurement speed, spectral resolution, and spectral width, but does not account for the exact spectral coverage of the system and its suitability to detect specific molecular species and its multi-species detection capabilities. The fiber-based sources currently yield worse NEAS than frequency comb-based systems. However, particularly for the chalcogenide-based source, the broadband spectral coverage, as well as its specific coverage in windows that are not being covered by the other sources, should be considered: depending on the application, these advances in spectral coverage outweigh the disadvantage of a worse NEAS.

Table A.1: Comparison of the fiber-based SC FTS systems to other broadband MIR spectroscopy systems.

Spectroscopy method	Reference/Note	Spectral coverage (µm)	Spectral resolution (GHz)	NEAS (cm^{-1} $Hz^{-1/2}$ per spectral elem.)
Fiber-based SC source with FTS	This work, ZBLAN source	1.4 – 4.1	1	$9.9 \cdot 10^{-7}$
	This work, Chalcogenide source	1.5 – 10.5	1	$5.3 \cdot 10^{-6}$
Frequency comb with FTS	(5), OPO pumped by fs-fiber laser	3.2 – 3.5 or 4.8 – 5.3	0.25[a]	$5.7 \cdot 10^{-8}$ or $3.1 \cdot 10^{-8}$
Dual-comb spectroscopy	(6), OPOs pumped by fs-fiber laser	3.1 – 5.5	0.115[a]	$8.5 \cdot 10^{-9}$
	(7,8), DFGs pumped by fs-fiber laser	2.6 – 5.2	0.2[a]	$\sim 2 \cdot 10^{-9}$
	(9), QCL-based frequency combs	8.1 – 8.6	10[a]	$\sim 8 \cdot 10^{-7}$
FTIR	(10), Bruker IRcube FTIR	2.5 – 5	30	$\sim 2 \cdot 10^{-8}$

[a]No spectral interleaving was considered

Plasma studies

The drive for sustainable, climate-neutral industrial processes requires an alternative to currently fossil fuel-driven chemical processes. Electric discharges, or plasmas, provide a solution to electrify industrial processes, replacing fossil fuels with sustainable energy (11,12). To enable and accelerate the adoption of plasma-driven chemical processes in industry, there is an urgent need for spectroscopic tools to monitor plasma reactions and gain insights into the fundamental processes in the plasma to optimize chemical conversions.

Plasmas are a fascinating medium for spectroscopists. The contents of a plasma can be very complex and consist of, among others, ions, radicals, and free electrons. Therefore, the plasma can be highly reactive. There can be a non-equilibrium between electronic, vibrational, and translational energy distributions (13,14), which can result in chemical reaction rates that are significantly different from traditional thermal chemical reactions (15). This results in rich absorbance spectra, containing many molecular species, in non-ambient conditions.

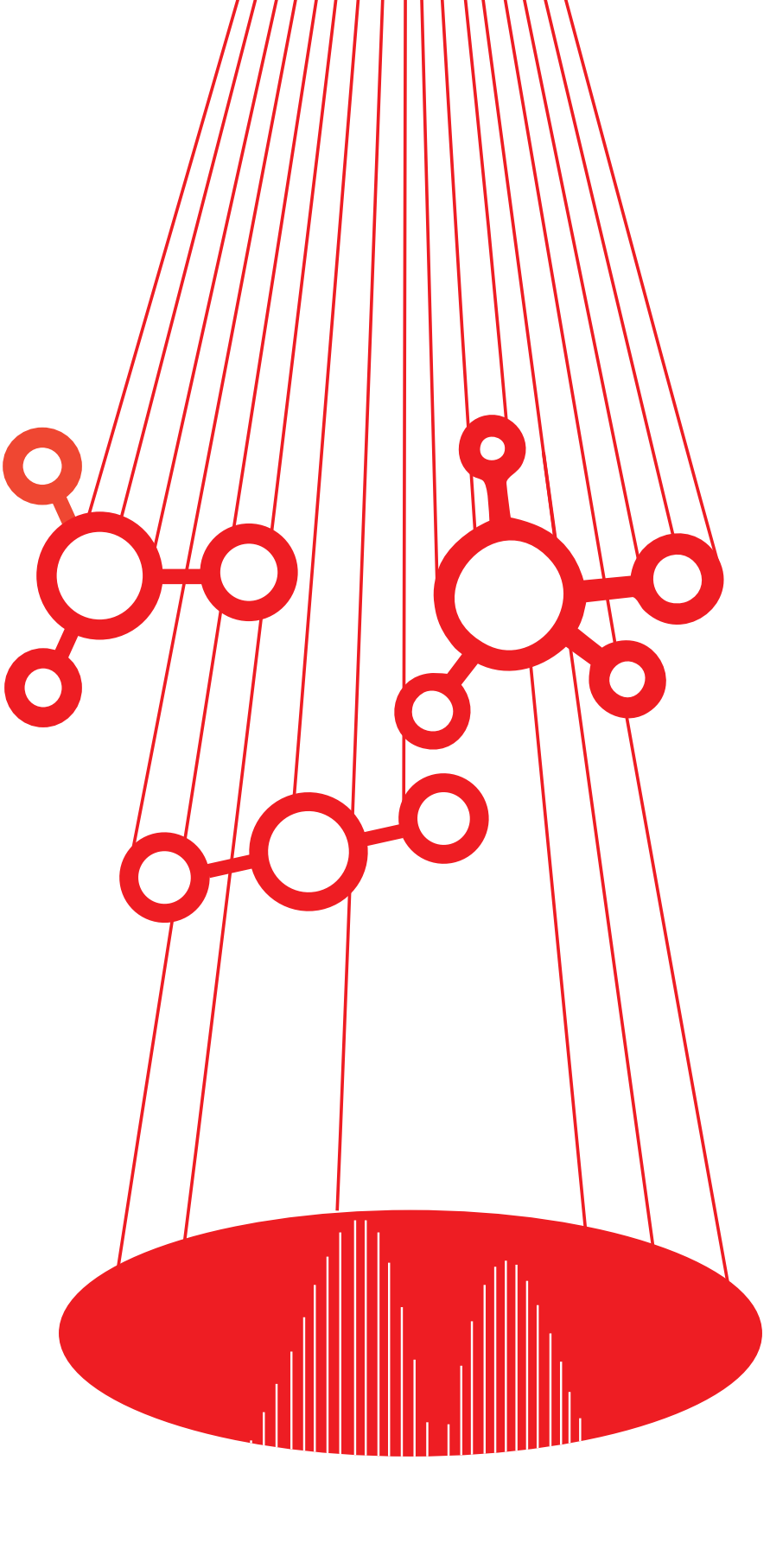
Examples of chemical processes targeted to be replaced by plasma-driven processes are (catalytic) pyrolysis and dry reforming methods for producing syngas (CO and H_2) and oxygenates from CO_2 and CH_4. Furthermore, CH_4 steam reforming can be replaced by plasma-catalytic conversion of CH_4 to produce hydrogen as a feedstock for ammonia or methanol. Sustainable plasma chemistry research is focused on using non-thermal plasmas to drive these (highly-)endothermic chemical gas conversions. The high CO_2 footprint of the current processes results from the required high temperatures and pressures supplied by burning fossil fuels, and are to be replaced by plasma-driven processes driven by sustainable green energy.

Scientific research into plasma-driven processes and conversions aims to enhance the fundamental understanding of the processes, improve the efficiency and reaction-specificity of the process, and make the processes viable for the chemical industry. Largely, the analytical methods for investigating plasma-driven processes can be divided into in situ methods and efflux analysis. While efflux analysis is largely dominated by mass spectrometric methods, in-situ measurements can be done using various techniques, such as Raman, CARS, emission spectroscopy, and infrared absorption spectroscopy.

In the subsequent chapters, supercontinuum-based spectroscopy is used for both in-situ and reaction product studies. Focusing on conversions of methane and carbon dioxide, the formation of C2-reaction products (i.e., ethane, ethylene, or acetylene) was monitored inside the plasma, and low-concentration byproducts such as formaldehyde, acetone, and acetaldehyde quantified in the outflow.

References

1. Hollas JM. Modern spectroscopy. 2nd ed. Chichester; New York: Wiley; 1992. 407 p.
2. Woyessa G, Kwarkye K, Dasa MK, Petersen CR, Sidharthan R, Chen S, et al. Power stable 1.5–10.5 μm cascaded mid-infrared supercontinuum laser without thulium amplifier. Opt Lett. 2021;46(5):1129–1132.
3. Brauer CS, Johnson TJ, Blake TA, Sharpe SW, Sams RL, Tonkyn RG. The Northwest Infrared (NWIR) gas-phase spectral database of industrial and environmental chemicals: recent updates. In: Vo-Dinh T, Lieberman RA, Gauglitz GG, editors. Advanced Environmental, Chemical, and Biological Sensing Technologies XI. SPIE; 2014. p. 910604.
4. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf. 2022;277:107949.
5. Khodabakhsh A, Ramaiah-Badarla V, Rutkowski L, Johansson AC, Lee KF, Jiang J, et al. Fourier transform and Vernier spectroscopy using an optical frequency comb at 3–54 μm. Opt Lett. 2016;41(11):2541.
6. Muraviev AV, Smolski VO, Loparo ZE, Vodopyanov KL. Massively parallel sensing of trace molecules and their isotopologues with broadband subharmonic mid-infrared frequency combs. Nat Photon. 2018;12(4):209–14.
7. Ycas G, Giorgetta FR, Baumann E, Coddington I, Herman D, Diddams SA, et al. High-coherence mid-infrared dual-comb spectroscopy spanning 2.6 to 5.2 μm. Nat Photon. 2018;12(4):202–8.
8. Ycas G, Giorgetta FR, Friedlein JT, Herman D, Cossel KC, Baumann E, et al. Compact mid-infrared dual-comb spectrometer for outdoor spectroscopy. Opt Express. 2020;28(10):14740–14752.
9. Gianella M, Nataraj A, Tuzson B, Jouy P, Kapsalidis F, Beck M, et al. High-resolution and gapless dual comb spectroscopy with current-tuned quantum cascade lasers. Opt Express. 2020;28(5):6197.
10. Griffith DWT, Deutscher NM, Caldow C, Kettlewell G, Riggenbach M, Hammer S. A Fourier transform infrared trace gas and isotope analyser for atmospheric applications. Atmos Meas Tech. 2012;5(10):2481–98.
11. Snoeckx R, Bogaerts A. Plasma technology: A novel solution for CO_2 conversion? Chem Soc Rev. 2017;46(19):5805–5863.
12. Bogaerts A, De Bie C, Snoeckx R, Kozák T. Plasma based CO_2 and CH_4 conversion: A modeling perspective. Plasma Processes and Polymers. 2017;14(6):1–21.
13. Sadiek I, Fleisher AJ, Hayden J, Huang X, Hugi A, Engeln R, et al. Dual-comb spectroscopy of ammonia formation in non-thermal plasmas. Commun Chem. 2024;7(1):110.
14. Klarenaar BLM, Engeln R, Van Den Bekerom DCM, Van De Sanden MCM, Morillo-Candas AS, Guaitella O. Time evolution of vibrational temperatures in a CO_2 glow discharge measured with infrared absorption spectroscopy. Plasma Sources Science and Technology. 2017;26(11):115008.
15. Fridman AA. Plasma chemistry. Cambridge University Press; 2008. 978 p.



Chapter 4

Mid-infrared Supercontinuum-based Fourier-Transform Spectroscopy for plasma analysis[1]

[1.] This chapter is based on the following article:
Krebbers R, Liu N, Jahromi KE, Nematollahi M, Bang O, Woyessa G, Petersen CR, van Rooij G, Harren FJM, Khodabakhsh, A, Cristescu, SM. Mid-infrared supercontinuum-based Fourier transform spectroscopy for plasma analysis. *Sci Rep. 2022;12(1):9642.*

Abstract

Broadband mid-infrared (MIR) spectroscopy is a well-established and valuable diagnostic technique for reactive plasmas. Plasmas are complex systems and consist of numerous (reactive) types of molecules; it is challenging to measure and control reaction specificity with a good sensitivity. Here, we demonstrate the first use of a novel MIR supercontinuum (SC) source for quantitative plasma spectroscopy. The SC source has a wide spectral coverage of 1300-2700 cm^{-1} (wavelength range 3.7-7.7 μm), thus enabling broadband multispecies detection. The high spatial coherence of the MIR SC source provides long interaction path lengths, thereby increasing the sensitivity for molecular species. The combination of such a SC source with a custom-built FTIR spectrometer (0.1 cm^{-1} spectral resolution) allows detection of various gases with high spectral resolution. We demonstrate its potential in plasma applications by accurate identification and quantification of a variety of reaction products (e.g. nitrogen oxides and carbon oxides) under low-pressure conditions, including the molecular species with overlapping absorbance features (e.g. acetone, acetaldehyde, formaldehyde, etc).

4.1 Introduction

One of the most promising emerging methods for reducing carbon dioxide (CO_2) emissions in chemical industrial processes is plasma-based gas conversion (1). This is particularly of interest for the conversion of two common greenhouse gasses, CO_2 and methane (CH_4) into carbon monoxide (CO) and hydrogen (H_2), also known as syngas (2,3). Although traditional (catalytic) pyrolysis and dry reforming methods are considered efficient ways to produce syngas and oxygenates from CO_2 and CH_4, high temperatures and pressures are usually required (4–6). To meet these requirements, the energy is largely supplied by burning fossil fuels, which is therefore accompanied by high CO_2 emissions. Compared to thermal methods, discharges can be powered by fully-sustainable green energy, avoiding CO_2 release (1,7–12). However, the plasmas formed in electric discharges are complex systems, consisting of numerous (reactive) species, and it is a challenge to control its reaction specificity. Therefore, it is of great importance to develop a detection system with good sensitivity and specificity to analyze the reaction products from the discharge for different plasma conditions (e.g., different feed gases, electric discharge power and mixing ratios).

The preferred analytical method for characterizing reaction products of an electric discharge is absorption spectroscopy (13–15), since it can combine high sensitivity and selectivity with excellent time resolution. The mid-infrared (MIR) wavelength region (2.5–25 μm) is of particular interest to the spectroscopy community, as many molecular gas species exhibit strong and unique absorption features in this region. Therefore, several different types of MIR sources are used for plasma diagnostics, such as MIR thermal sources (lamps) (15,16), tunable diode lasers (14), Quantum Cascade Lasers (QCLs) (17–19), Interband Cascade Lasers (ICLs) (20) and optical frequency combs (21–22). While narrowband lasers, i.e., tunable diode lasers, ICLs, and QCLs, are well suited for sensitive detection of a specific molecule, they lack the ability to easily detect different types of molecules simultaneously.

Broadband absorption techniques, on the other hand, can detect and identify multiple species simultaneously, as is the case for the classical Fourier-Transform InfraRed (FTIR) spectrometers, based on IR lamps (23). However, the omnidirectionality and divergence of light from these sources limits the optical interaction path length with the gas sample species. Moreover, the low spectral brightness of these sources requires the spectrum to be averaged over longer periods to achieve an acceptable signal-to-noise ratio (SNR), especially for high spectral resolution measurements.

Optical frequency combs overcome these limitations but have (with some notable exceptions (24)) a relatively narrow spectral coverage. Furthermore, their high price and technical complexity limit their widespread use by end-users in real-life applications.

MIR supercontinuum (SC) sources are very suitable for broadband absorption spectroscopy, as they provide high spatial coherence, high spectral brightness, and ultra-broadband spectral coverage, outperforming conventional thermal sources and even some synchrotrons (25–28). SC sources have demonstrated powerful capabilities in emitting in the visible and near-infrared range using silica fibers. Sources emitting up to 4 μm using fluoride fibers have become commercially available in recent years (29). Several applications for these types of SC sources have been demonstrated, such as absorption spectroscopy with high sensitivity and selectivity for simultaneous measurement of multiple compounds (30), spectroscopic standoff detection (31), or optical coherence tomography in the MIR range providing real-time and high-resolution images (32,33). Current developments in MIR SC sources extend the spectral range of the source beyond 4 μm to cover a larger part of the MIR fingerprint region (34). Applications in absorption spectroscopy (35) and hyperspectral imaging (36,37) have been demonstrated for these highly experimental, new sources. Moreover, as the technology matures, this type of MIR SC source is expected to be used for compact and cost-effective air quality sensor networks (38).

SC sources in the visible and near-infrared wavelength region have been previously used for plasma diagnostics (39), but to our knowledge, this report is the first demonstration of plasma diagnostics using a MIR SC source. Here, we demonstrate the advantages of MIR SC sources as broad spectral coverage (3.7–7.7 μm) and high sensitivity, achieved by the long optical path length of the spatially coherent beam.

4.2 Results

4.2.1 Experimental setup

The developed spectroscopic system consists of the MIR SC source, whose beam transmits through a multipass cell (MPC) containing the products of the plasma reaction at a low pressure. The transmitted beam is sent towards a custom-built Fourier Transform Spectrometer (FTS) with a spectral resolution of 0.1 cm^{-1} (3 GHz), allowing detection of narrow molecular absorption lines of gas species at a low pressure. Outside the MPC, the SC beam path was shielded and purged continuously with N_2 gas. The experimental setup is presented in Figure 4.1. The light source is a

newly developed, fiber-coupled MIR SC source (DTU Fotonik, average power ~ 86 mW, pulse duration ~ 0.5 ns, repetition frequency 3 MHz (40)). The MPC has a 31.2 m effective interaction length (Thorlabs, HC30L/M-M02) and is connected to the discharge cell, such that the reaction products of the plasma are sent to the MPC for interaction with the SC beam. The custom-built FTS has been demonstrated and discussed in detail in our previous work (35). A brief description is presented in the Methods section, as well.

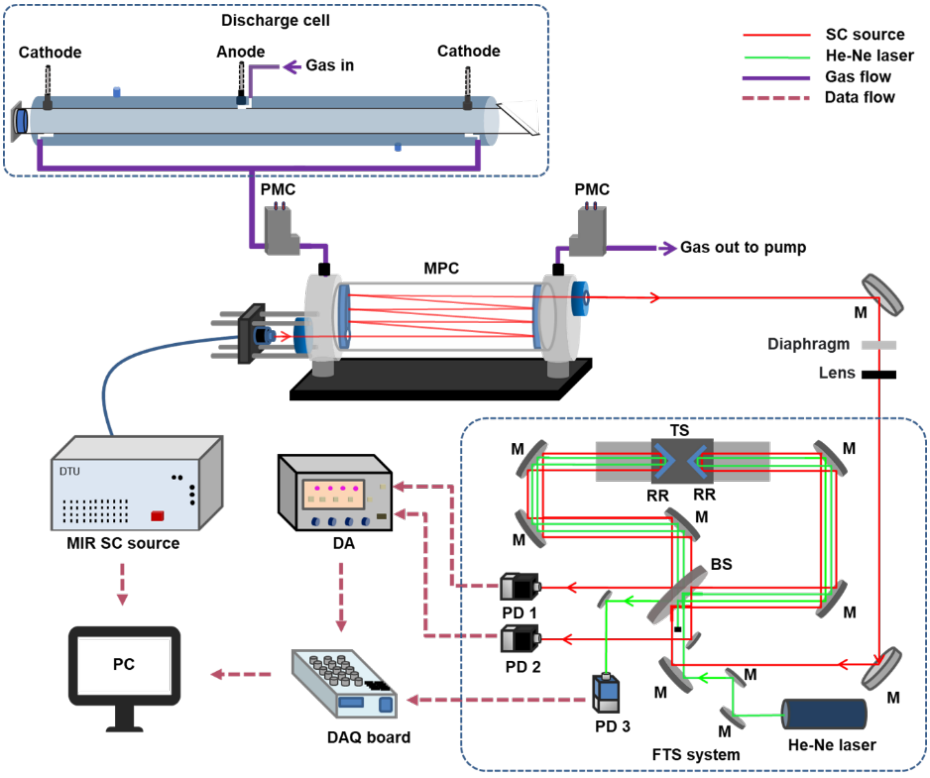


Figure 4.1: Experimental setup of the SC source and FTS-based multi-species detection system. MPC: multipass cell PMC: pressure meter & controller, M: mirror, TS: translation stage, RR: retroreflector mirror, BS: beamsplitter, He-Ne laser: Helium-Neon laser, PD: photodetector, DA: differential amplifier, DAQ board: Data acquisition board, PC: computer.

4.2.2 Source characterization

The spectral coverage of the MIR SC source was characterized with the FTS system by sending the SC light through the MPC under vacuum ($< 10^{-3}$ bar). The resulting power spectral density is shown in Figure 4.2a and covers the spectral region between 1300 and 2700 cm^{-1} (3.7–7.7 µm). While most of the spectral power

is between 1700 and 2100 cm^{-1}, spectroscopy in the other wavelength ranges is still feasible with high SNR, due to the high average power of the MIR SC source and sensitive photodetectors. Highly absorbing water (H_2O) and CO_2 lines can be observed in the spectrum, despite the measurement being performed under vacuum conditions. In Figure 4.2a, the etalon fringes are also visible in the spectrum due to the partial overlap of the SC spots of the consecutive reflections on the MPC mirrors. However, the measured absorbance spectrum is not affected as the etalon fringes are stable and are canceled out by normalizing the sample spectra by the background spectra (35). The simulated spectral line intensities of the relevant species from the HITRAN2020 database (41) between 1300 and 2500 cm^{-1} are shown in Figure 4.2b. Due to the overlap of the absorption lines, a low pressure (16.5 mbar) and a high spectral resolution (3 GHz/0.1 cm^{-1}) are required to distinguish between the absorption lines.

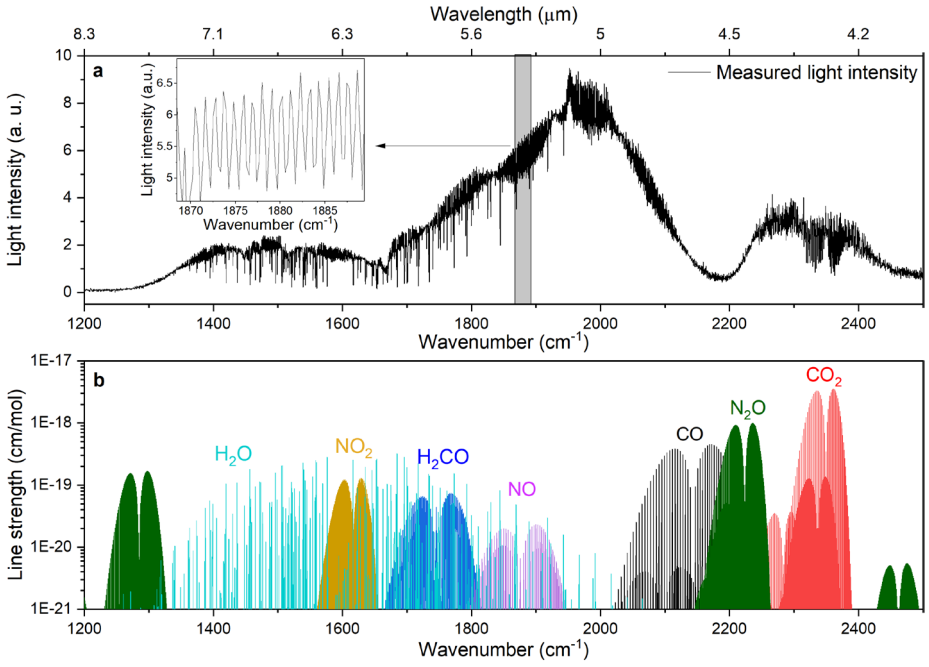


Figure 4.2: Spectralcoverage of the MIR SC source and absorbing molecular species. **a** The spectrum of the MIR SC source with a spectral resolution of ~0.27 cm^{-1} (8 GHz) measured under vacuum conditions, insert: an enlargement of the spectrum, showing the etalon fringes in the measured spectrum. **b** Simulated spectral line intensity of different species from HITRAN2020 between 1300 cm^{-1} and 2500 cm^{-1}.

4.2.3 System validation

To initially evaluate the performance of the MIR SC FTS-based system, we measured the spectrum of a gas mixture of 495 ppm CO_2 in N_2 at 16.5 mbar pressure. Here, we dilute a calibrated mixture of 5% CO_2 in N_2 (Linde Gas) with pure N_2, down to 495 ppm, using two flow controllers. The measured spectrum is shown in Figure 4.3 (in black, 500 averaged scans in ~16 min, 1.9 s per scan) alongside a fitted, modelled CO_2 spectrum (in red, inverted for clarity). The model spectrum is calculated using the HITRAN database parameters and a Voigt profile, convolved with a sinc function. The retrieved concentration from the fit is 485 ± 12 ppm. The uncertainty is calculated from the standard deviation of the noise in the residual of the fit. In Figure 4.3a, the full rotational-vibrational band of CO_2 is shown. To demonstrate the agreement between the measurement and the fitting routine, an enlargement of the spectral features between 2357 cm^{-1} and 2363 cm^{-1} is displayed in Figure 4.3b. The residuals are shown in the bottom panels (Figure 4.3c,d). The rather featureless residuals demonstrate the high precision of the frequency calibration, as well as the good quality of the fitting routine. The small peaks which are still visible in the residual can be contributed to the influence of carbon dioxide in the beam path outside the MPC. The light intensity at these specific peaks is reduced in both the background and measurement spectrum to an extent that it degrades the fitting quality.

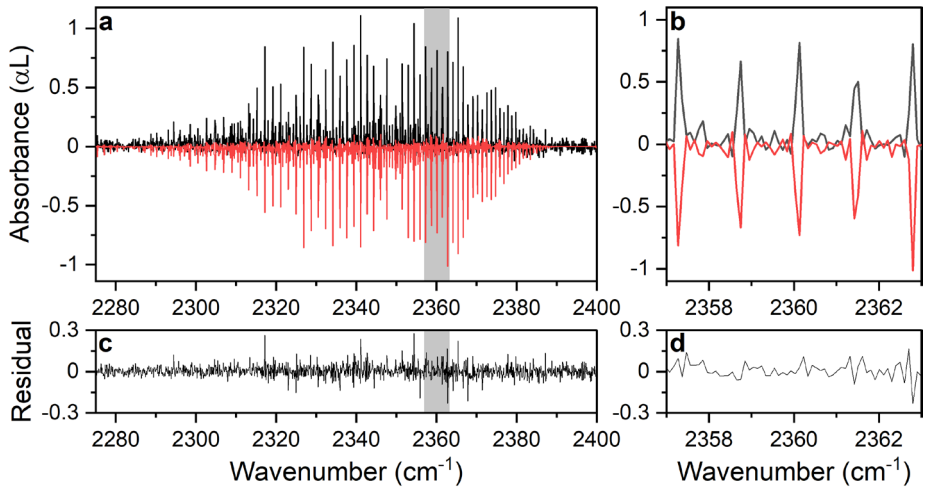


Figure 4.3: Experimental spectrum of CO_2. **a** Measured spectrum (in black, 0.1 cm^{-1} spectral resolution, pressure 16.5 mbar, 500 averages in ~16 minutes) of 495 ppm CO_2 diluted in N_2 and a fitted modelled spectrum (in red, inverted) from the HITRAN database using a Voigt profile and convolved with a sinc-function. **b** Details of a part of the spectrum indicated in a as a grey rectangle. **c** and **d** Residual of the fit.

The linear response of the system to different applied CO_2 concentrations was evaluated in a dynamic range between 0.05% and 2.5% of CO_2 in N_2, by diluting the 5% CO_2 in N_2 mixture further using pure N_2 gas. Each measurement consists of 500 averaged spectra, measured in ~ 16 min. The reference spectra for the linear fitting routine were simulated from the HITRAN database as described before. The retrieved concentrations from the fit versus the applied concentrations are shown in Figure 4.4, together with the corresponding errors, exhibited in the lower panel. The linear fit shows a Pearson correlation coefficient square value of 0.9995, demonstrating a very good agreement between the measured concentration values and the applied concentrations. The corresponding relative errors are within a ± 4% margin for the entire range, with an average error of 2%.

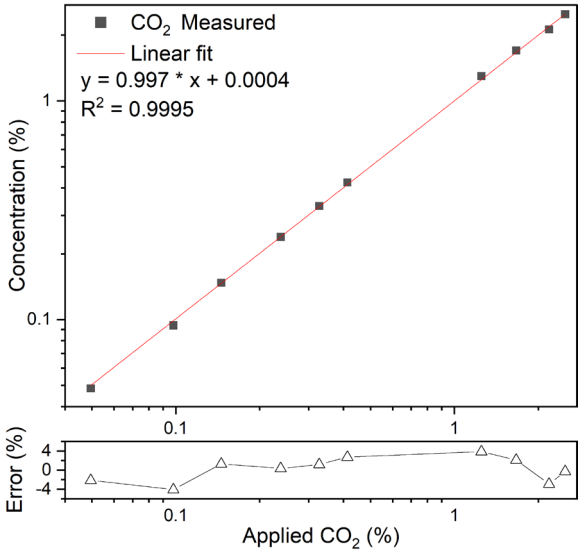


Figure 4.4: Retrieved concentrations of CO_2 for various dilutions of CO_2 in N_2, along with a linear fit to demonstrate the linear response of the system. The error provided is the relative difference between the applied and retrieved concentration.

4.2.4 Product analysis of CO_2/N_2 plasmas

To further assess the performance of the system and to demonstrate its ability to quantify complex mixtures of reaction products, the outflow of a discharge was used. The discharge was generated in a flowing mixture of 50% CO_2 in N_2 (flow of 2 l_n/h) in a discharge cell at a pressure of 25 mbar. The applied voltage was 17.5 kV with a current of 10 mA that results in a specific energy input of 7.1 MJ/mol. The outflow of the discharge cell was guided to the MPC, which had a controlled

pressure of 16.5 mbar. In Figure 4.5, the measured absorbance spectra of nitrogen dioxide (NO_2), nitric oxide (NO), nitrous oxide (N_2O), CO, CO_2 and H_2O are shown (in black) together with corresponding simulated spectra (inverted, in color, using HITRAN parameters). The concentrations of the species retrieved from fitting the simulations to the absorption spectra are shown in Table 4.1. The practically featureless residual of the fit, shown in the bottom panel of Figure 4.5, indicates a very good fit for all detected species. The few spikes that remain visible in the residual can be attributed to the effect of highly absorbing water vapor lines in the atmospheric air outside the MPC, degrading the fitting quality.

A wide variety of different species are detected over a range of concentration levels from hundreds of ppm to percentage-level. The broad spectral coverage does not only allow for detection of absorption lines of various compounds, but also to select absorption lines of a certain species with an appropriate line strength for the given concentration, preventing limitations arising from absorption lines which absorb almost 100% of the light.

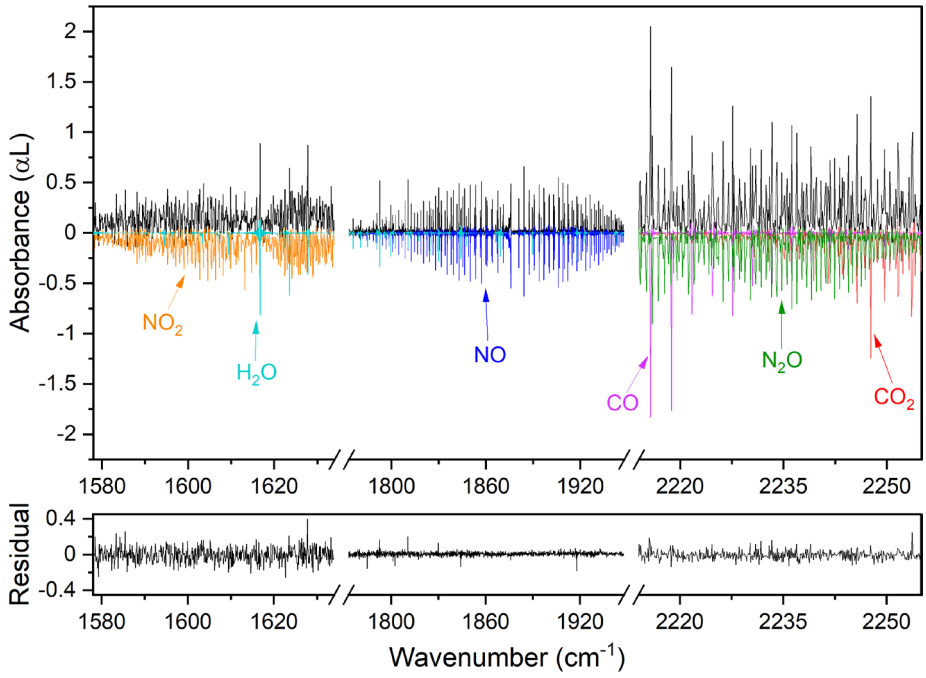


Figure 4.5: Measured spectrum (in black, 0.1 cm^{-1} spectral resolution, 16.5 mbar pressure, ~16 minutes) of reaction products of a 50% CO_2 / 50% N_2 discharge and the fitted modelled spectra (in colors, inverted) based on the HITRAN database parameters and Voigt profiles convolved with a sinc function. The residual of the fit is shown in the lower panel.

Table 4.1: Retrieved concentrations of electric discharge products

	Retrieved concentration (% or ppm)	
	50% CO_2 / 50% N_2	**30% CO_2 / 70% CH_4**
Carbon dioxide (CO_2)	13.0 ± 0.6 %	13.9 ± 0.7 %
Carbon monoxide (CO)	14.6 ± 0.7 %	32.1 ± 1.7 %
Nitrous oxide (N_2O)	570 ± 30 ppm	--
Nitrogen dioxide (NO_2)	154 ± 13 ppm	--
Nitric oxide (NO)	0.33 ± 0.01 %	--
Ethylene (C_2H_4)	--	0.73 ± 0.03 %
Formaldehyde (H_2CO)	--	0.13 ± 0.01 %
Acetone (C_3H_6O)	--	0.3 ± 0.1 %
Acetaldehyde (C_2H_4O)	--	0.3 ± 0.1 %

4.2.5 Product analysis of CO_2/CH_4 plasmas

A bigger challenge for the spectroscopic analysis of a complex mixture of reaction products is the dry reforming of methane, as more complex molecules are formed, which cannot be found in HITRAN. Therefore, to extend the evaluation of the system, a CO_2/CH_4 discharge is generated. A discharge voltage of 18 kV with a current of 15 mA provided a specific energy input of 11 MJ/mol for a mixture of 70% CH_4 and 30% CO_2 (flow 2 l_n/h, 19 mbar pressure). The absorbance features of the detected products are presented in Figure 4.6. Here, a broad absorbance feature is visible in the 1700–1800 cm^{-1} wavenumber region. In general, this indicates the presence of molecular species with a large number of closely spaced rotational transitions in the vibrational band, which cannot be resolved spectroscopically at this pressure and temperature. Using the PNNL database, this specific absorbance profile was found to be likely belonging to acetone (C_3H_6O) and acetaldehyde (C_2H_4O), both with a concentration of 0.3 ± 0.1%. The concentration could however not be determined exactly, as PNNL is constructed with experimental data at 1 atmosphere pressure, which is significantly different from our experiment. We confirmed the presence of these two molecular species in our mixture using proton-transfer-reaction mass spectrometry (PTR-MS)(42) and gas chromatography–mass spectrometry (GC–MS)(43). Moreover, the residual shows the ability of the system to detect overlapping absorbance features of multiple species, as both the acetone, acetaldehyde, formaldehyde (H_2CO) and most H_2O spectral features are fitted well with their reference spectra. Furthermore, the absorbance lines of ethylene (C_2H_4) around 1880 cm^{-1} are not included in the HITRAN database. Therefore, the GEISA database (44) was used to create a simulated reference spectrum for C_2H_4, indicating a concentration of 0.73 ± 0.03%.

In summary, this experiment demonstrates the system's ability to accurately detect numerous molecular species created in electric discharges of CO_2/CH_4, even for the ones with overlapping spectral features.

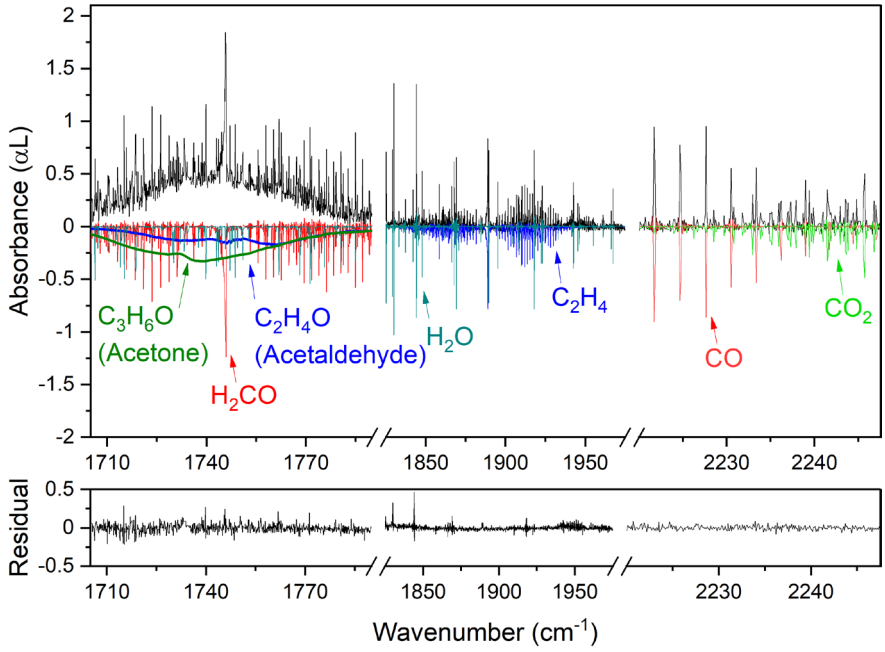


Figure 4.6: Measured spectrum of reaction products of a discharge of 70% CH_4 and 30% CO_2 (in black, 0.1 cm^{-1} spectral resolution, 16.5 mbar pressure, ~16 minutes). Reference spectra of formaldehyde, water, carbon monoxide, carbon dioxide (from HITRAN), acetone and acetaldehyde (from PNNL) and ethylene (from GEISA) shown inverted.

4.2.6 Plasmas with varying ratio of CO_2 and CH_4

To demonstrate the possibilities of the system for plasma analysis and study, a quantitative analysis of the products formed in the electric discharge is made, using a series of measurements with a varying CO_2/CH_4 ratio (gas flow 2 l_n/h, 18 kV, 15 mA, specific energy input of 11 MJ/mol). In Figure 4.7a, the measured CO concentration is displayed for varying mixtures of CO_2/CH_4, along with the assumed limit in CO production, calculated from the number of carbon (in red) and oxygen atoms (in blue) available in the system to react to CO. These limits are calculated from the retrieved CO_2 concentrations, which are shown in Figure 4.8. From the difference between the CO_2 concentrations with the discharge on or off, the maximum number of carbon and oxygen atoms available for conversion is determined. The highest CO production is found for a 50/50-mixture of CO_2/CH_4. The CO values are within the

expected maximum available carbon and oxygen atoms generated in the plasma, indicating a correct reaction balance. As the concentration of CO (~ 53%) is higher than the concentration of converted CO_2, this indicates that CO is not only formed by removal of an oxygen atom from CO_2, but also by recombination of the removed oxygen atom from CO_2 with a dehydrogenated carbon atom from CH_4. Furthermore, there is additional production of C_2H_4 and H_2CO (Figure 4.6), of which the retrieved concentrations are presented in Figure 4.7b.

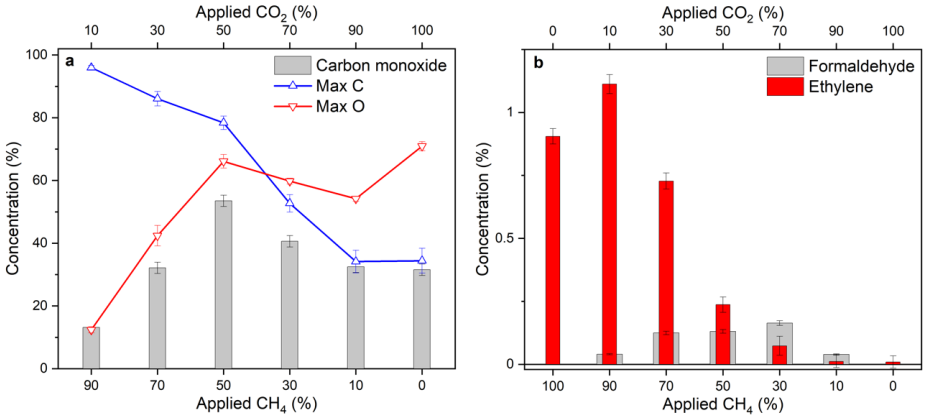


Figure 4.7: Retrieved concentrations of reaction products. **a** Measured CO concentrations for varying CH_4/CO_2 ratio generated by the plasma (in grey). Maximum amount of carbon atoms (Max C, blue) and oxygen atoms (Max O, red), available from the CO_2 and CH_4 depletion in the discharge, indicate the maximum CO concentration that can be generated. The measured CO-levels are below these values. **b** Measured concentrations of ethylene (in red) and formaldehyde (in grey) generated by the plasma for varying ratios of CH_4/CO_2.

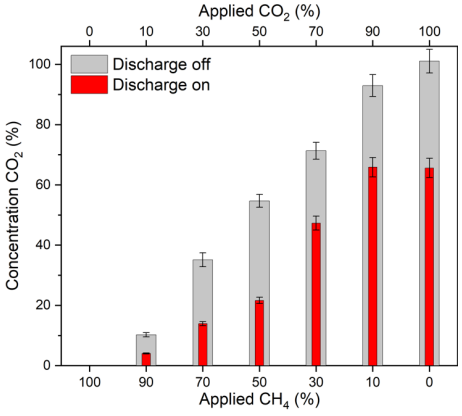


Figure 4.8: Measured CO_2 concentrations with (in grey) and without (in red) the discharge applied for a varying CH_4/CO_2 ratio. The difference between the two indicates the amount of CO_2 that was consumed in the reaction within the discharge.

4.3 Discussion

A MIR SC FTS system is presented for the first time for plasma analysis. We showed that various molecular species of interest can be detected and quantified across the spectral range of the supercontinuum source, while the high spectral resolution from the FTS allowed for identification of molecular species with overlapping absorption features. Therefore, the MIR SC FTS system provides a unique set of characteristics that are especially useful for plasma diagnosis, combining the broad spectral coverage, high spectral power, and high spectral resolution with a high spatial coherence, stability and ease of use. The high detection sensitivity makes it possible to monitor the molecular reaction products in a low-pressure plasma, minimizing spectral interference between different species and accurately determining their concentrations.

A rather detailed comparison of the MIR SC FTS with other spectroscopy systems in the MIR range is presented in our previous work (35). The noise equivalent absorption sensitivity of the MIR SC FTS is one to two orders of magnitude lower than MIR optical frequency comb-based spectrometers and it also provides coarser spectral resolution compared to comb-based systems. However, the broader spectral coverage of the MIR SC FTS allows simultaneous detection of a large number of molecular species. In addition, the ease-of-use, compactness, robustness, and lower cost of MIR SC sources make them better suited for real-life plasma applications.

Classical FTIRs can match the spectral coverage of our system. As mentioned elsewhere (35), they currently provide a better Noise Equivalent Absorption Sensitivity compared to the MIR SC FTS presented here. However, classical FTIRs tend to operate at a lower spectral resolution, as the combination of high resolution with the relatively low spectral power density of a thermal source would increase the need for spectral averaging and consequently lengthen the measurement time. Most importantly, the thermal sources lack spatial coherence, making it challenging to create long interaction path lengths (e.g. by using an MPC with a long effective path length) to maximize the sensitivity. Furthermore, a well-collimated beam with a small beam diameter is required for *in-situ* monitoring in long (> 1 m) plasma cells, which is also difficult to obtain using thermal sources.

Our system's performance can also be compared specifically to previous studies on plasma analysis using QCLs (20) and tunable diode lasers (14), which are currently the most frequently used laser-based systems for absorption spectroscopy on plasmas. Their narrow spectral range, which generally is in the order of a few wavenumbers (with some exceptions for external cavity QCLs), limits the number of detectable molecular species as well as the detection and identification of unknown compounds and introduces difficulties for overlapping/

broadband spectra. For instance, detection, identification, and quantification of the overlapping spectra of acetaldehyde, acetone, and formaldehyde as in Figure 4.6 would have been significantly more difficult with such a system.

In conclusion, the reaction products were quantified for a plasma with a varying ratio of CO_2 and CH_4. The CO yield was found to increase when CH_4 was added to the CO_2 plasma, which could be explained by the recombination of the removed oxygen atom from CO_2 with a dehydrogenated carbon atom from CH_4. The observed trends in the generation of the other observed products, H_2CO and C_2H_4, demonstrated a good agreement with previous studies. H_2CO should be composed from H-atoms of CH_4 and O-atoms of CO_2 and is indeed only found in mixtures containing both CH_4 and CO_2 (45). The formation of C_2H_4 increases towards higher concentrations of CH_4 in the mixture and can be easily explained as C_2H_4 is a product from the dehydrogenation of CH_4 (46).

Currently, the relative intensity noise of the MIR SC source limits the sensitivity of the spectrometer, although this is not essential for the demonstrated application, in identification and quantification of the products of a plasma reaction. Moreover, with the current rapid developments in MIR SC sources, it is anticipated that the intensity noise of these sources will be reduced further in the near future (29,34). Finally, the spatial coherence of the MIR SC beam could also be exploited for other applications, such as for accurate *in-situ* probing in plasmas, to monitor the plasma dynamics, as well as to detect intermediate products in the plasma.

Finally, besides the application of this system for electrolytic chemical conversions, it can potentially be applied as a spectroscopic tool for many more plasma applications. For example, within the synthesis of carbon nanotubes using plasma-enhanced chemical vapor deposition (PECVD), monitoring a correct flow of reactants and the formation of gas phase intermediates and products is currently done with mass spectrometry (47). Our system could replace mass spectrometry, providing on-line measurements and being able to distinguish species with identical masses. Furthermore, in plasma medicine cold atmospheric plasma jets are used to treat human skin or wounds (48). Optical Emission Spectroscopy (OES) and FTIR are used to investigate the species formed in these plasmas. As OES lacks quantification (49) and FTIR lacks sensitivity due to limited interaction lengths and spectral resolution (50), a MIR SC FTS system might be preferred to quantify the species formed in plasma jets for these applications. Similarly, the system could be of interest for other plasma-based techniques, such as wastewater treatment (51), or surface and material modification and sterilization (52,53).

4.4 Methods

4.4.1 Fourier-transform spectrometer

The overview of the full experimental system is presented in Figure 4.1. The output fiber of the MIR SC source is connected to a reflective collimator that is mounted on the MPC to limit the optical path outside the MPC. After passing through the MPC, the MIR SC beam is aligned to the Fourier-transform spectrometer (FTS) using two flat mirrors. A diaphragm is used to filter stray light and a long-focal-point lens was used to match the beam waist to the center of the translation stage in the FTS. The setup of the FTS system has been described in detail in our previous work (35). Briefly, the FTS system is based on a Michelson interferometer; the incoming light beam is split by a beamsplitter into two arms, each leading to a hollow retroreflector mirror. The retroreflector mirrors reflect the beams with slight horizontal and vertical shifts, so that the reflected beams recombine on the beamsplitter and the two resulting superpositions of the beams (interfering pair of beams) can both propagate towards separate photovoltaic detectors (PVI-4TE-10.6, Vigo Systems). The two interference patterns at the output of the detectors are subtracted from each other by a differential amplifier (SR560, Stanford Research system) in a balanced detection scheme. Since the two interference patterns are out-of-phase, but the intensity noise is in-phase, the balanced detection scheme highly reduces intensity noise, which leads to an effective improvement of the signal-to-noise ratio (SNR). The two retroreflectors are mounted back-to-back on a linear-motor translator stage. Scanning the translator stage over a distance of 2.5 cm results in an optical path difference of 10 cm between the two beams, which provides a spectral resolution of 0.1 cm^{-1} (3 GHz). A He–Ne laser beam is sent along the SC beam path in the FTS to create an interferogram for a known and stable wavelength. The interferogram of the He–Ne laser is recorded by a separate photovoltaic detector. To prevent the interference of absorption by H_2O and CO_2 in the ambient air with the measurement, the FTS is confined in a closed box and continuously purged with N_2 gas.

4.4.2 Data analysis

For real-time analysis of the detected interferograms, a LabVIEW-based data processing system is developed. A flowchart of the algorithms used for this analysis is shown in Figure 4.9. The wavenumber calibration of the MIR SC source is performed through a resampling process using a He–Ne laser. In this process, explained in detail in Ref. (35), the algorithm finds the zero-crossings of the He–Ne laser interferogram (whose beam is propagating in parallel to the MIR SC beam) by linear interpolation. Then, it linearly interpolates the MIR SC interferogram

at the zero-crossing positions of the He–Ne laser interferogram. Therefore, the interpolated MIR SC interferogram in the zero-crossing domain is calibrated with an optical path difference step size equal to the zero-crossing intervals. After the calibration, a Fast Fourier Transform (FFT) of the resampled interferograms yields the spectra. Note that we used a natural boxcar apodization function for the interferograms. The absorbance spectrum is obtained by dividing the transmission spectrum of the sample to the background (the latter is taken after purging the MPC with pure N_2 gas) and taking the natural logarithm. The absorbance spectra of the different species are simulated, using parameters from the HITRAN2020 (41) or GEISA (44) database and convoluted with a sinc instrument line shape function, to fit the natural boxcar apodization of the interferograms. GEISA was used to simulate the absorbance of ethylene (C_2H_4) around 1900 cm^{-1}, since this ethylene band is not completely included in HITRAN. The PNNL (54) database was used to identify absorbance from acetone (C_3H_6O) and acetaldehyde (C_2H_4O). However, PNNL contains actual FTIR measurements of absorbance spectra at atmospheric pressure, which cannot correctly be recalculated for low pressures. Finally, the concentrations of different species are calculated using a multiple linear regression algorithm based on the least squares method (35). Given the limitations of the PNNL data, the concentration of acetone and acetaldehyde can only be approximated.

4.4.3 Plasma conditions

For the generation of the plasma, a water-cooled discharge tube (length 180 cm, internal diameter of 1.5 cm) is used to create a uniform electric discharge. The anode and gas inlet are located in the center, while two hollow cathodes (4 cm long each) and the gas outlets are located at either end of the discharge tube. A current-stabilized high-voltage (HV) power supply (Haefely Hipotronics, US, providing up to 25 kV, 40 mA) is used for generating a direct current (DC) electric discharge in the tube (22). The gas inflow is regulated with flow controllers and the pressure inside the discharge tube is regulated with a back-pressure controller (25 mbar). The outflow of the gas from the discharge tube is directed to a MPC (effective path length of 31.2 m), which is regulated by another back-pressure controller to a pressure lower than in the discharge tube (16.5 mbar) to ensure a proper gas flow.

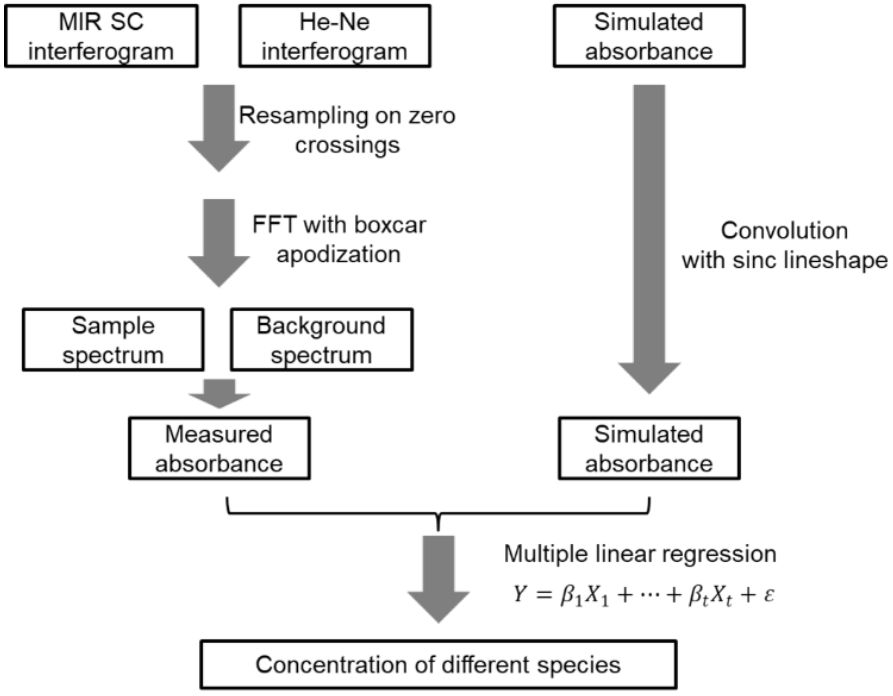


Figure 4.9: Flowchart of the developed analysis algorithms.

References

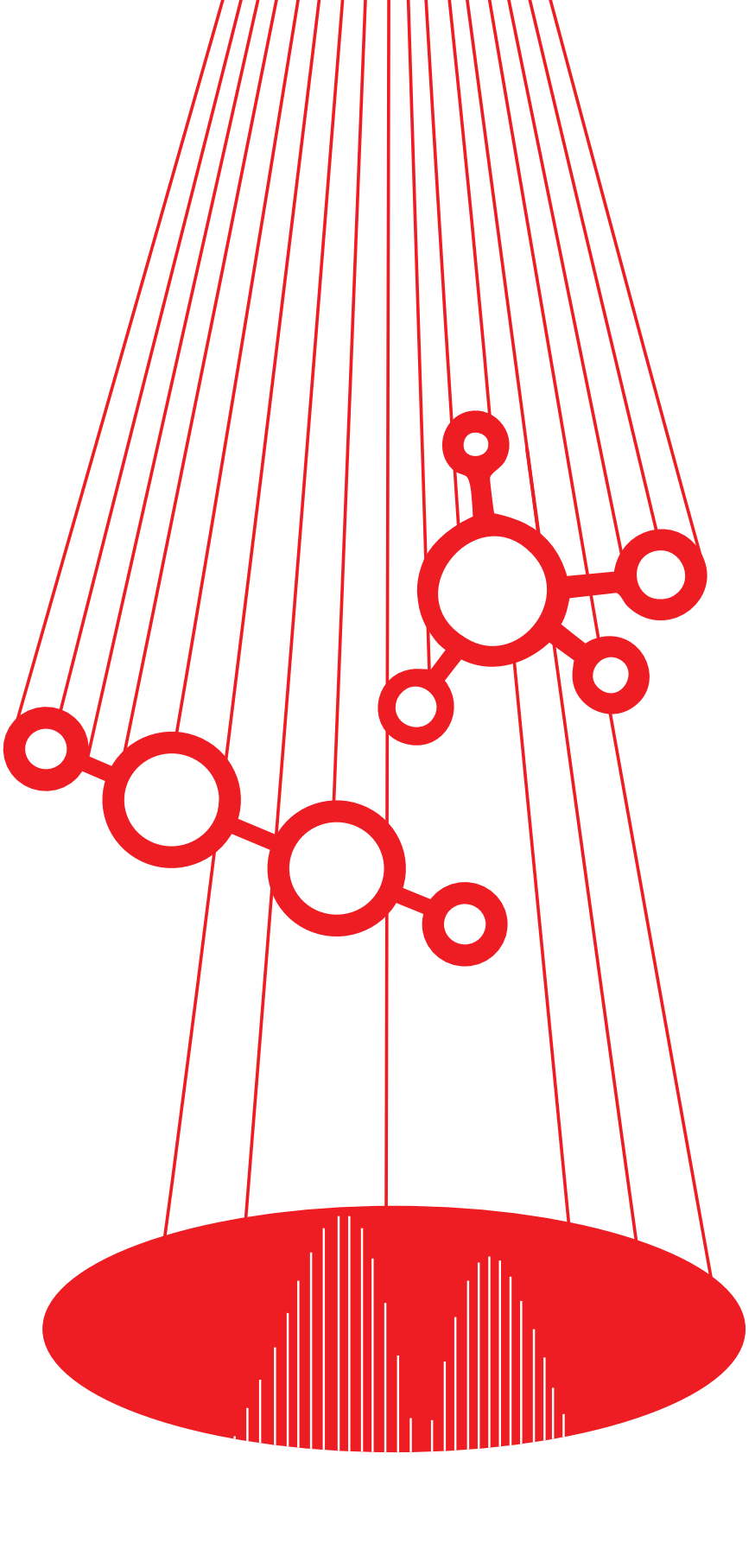
1. Snoeckx R, Bogaerts A. Plasma technology: A novel solution for CO_2 conversion? Chem Soc Rev. 2017;46(19):5805–5863.
2. Lebouvier A, Iwarere SA, D'Argenlieu P, Ramjugernath D, Fulcheri L. Assessment of carbon dioxide dissociation as a new route for syngas production: A comparative review and potential of plasma-based technologies. Energy Fuels. 2013;27(5):2712–2722.
3. Xia Y, Lu N, Li J, Jiang N, Shang K, Wu Y. Combined steam and CO_2 reforming of CH_4 for syngas production in a gliding arc discharge plasma. J. CO2 Util. 2020;37:248–259.
4. Westermann A, Azambre B, Bacariza MC, Graça I, Ribeiro MF, Lopes JM, et al. Insight into CO_2 methanation mechanism over NiUSY zeolites: An operando IR study. Appl Catal B. 2015;174–175:120–125.
5. Fan MS, Abdullah AZ, Bhatia S. Catalytic technology for carbon dioxide reforming of methane to synthesis gas. ChemCatChem. 2009;1(2):192–208.
6. Pakhare D, Spivey J. A review of dry (CO_2) reforming of methane over noble metal catalysts. Chem Soc Rev. 2014;43(22):7813–7837.
7. Liu S, Winter LR, Chen JG. Review of plasma-assisted catalysis for selective generation of oxygenates from CO_2 and CH_4. ACS Catal. 2020;10(4):2855–2871.
8. Li D, Li X, Bai M, Tao X, Shang S, Dai X, et al. CO_2 reforming of CH_4 by atmospheric pressure glow discharge plasma: A high conversion ability. Int J Hydrog Energy. 2009;34(1):308–313.
9. Paulmier T, Fulcheri L. Use of non-thermal plasma for hydrocarbon reforming. Chem Eng J. 2005;106(1):59–71.
10. Bogaerts A, Neyts EC. Plasma technology: An emerging technology for energy storage. ACS Energy Lett. 2018;3(4):1013–1027.
11. Tao X, Bai M, Li X, Long H, Shang S, Yin Y, et al. CH_4–CO_2 reforming by plasma: Challenges and opportunities. Prog Energy Combust Sci. 2011;37(2):113–124.
12. Puliyalil H, Lašič Jurković D, Dasireddy VDBC, Likozar B. A review of plasma-assisted catalytic conversion of gaseous carbon dioxide and methane into value-added platform chemicals and fuels. RSC Adv. 2018;8(48):27481–27508.
13. Xu S, Chansai S, Xu S, Stere CE, Jiao Y, Yang S, et al. CO poisoning of Ru catalysts in CO_2 hydrogenation under thermal and plasma conditions: A combined kinetic and diffuse reflectance infrared Fourier transform spectroscopy-mass spectrometry study. ACS Catal. 2020;10(21):12828–12840.
14. Röpcke J, Lombardi G, Rousseau A, Davies PB. Application of mid-infrared tuneable diode laser absorption spectroscopy to plasma diagnostics: A review. Plasma Sources Sci Technol. 2006;15(4):S148–S168.
15. Azzolina-Jury F, Thibault-Starzyk F. Mechanism of low pressure plasma-assisted CO_2 hydrogenation over Ni-USY by microsecond time-resolved FTIR spectroscopy. Top Catal. 2017;60(19-20): 1709–1721.
16. Horvth G, Skaln JD, Mason NJ. FTIR study of decomposition of carbon dioxide in dc corona discharges. J Phys D: Appl Phys. 2008;41(22):225207.
17. Damen MA, Martini LM, Engeln R. Temperature evolution in a pulsed CO_2-N_2 glow discharge measured using quantum cascade laser absorption spectroscopy. Plasma Sources Sci Technol. 2020;29(6):065016.

18. Marinov D, Lopatik D, Guaitella O, Hübner M, Ionikh Y, Röpcke J, et al. Surface vibrational relaxation of N_2 studied by CO_2 titration with time-resolved quantum cascade laser absorption spectroscopy. J Phys D: Appl Phys. 2012;45(17):175201.
19. Welzel S, Hempel F, Hübner M, Lang N, Davies PB, Röpcke J. Quantum cascade laser absorption spectroscopy as a plasma diagnostic tool: An overview. Sensors. 2010;10(7):6861–6900.
20. Röpcke J, Davies PB, Lang N, Rousseau A, Welzel S. Applications of quantum cascade lasers in plasma diagnostics: A review. J Phys D: Appl Phys. 2012;45(42):423001.
21. Abbas MA, van Dijk L, Jahromi KE, Nematollahi M, Harren FJM, Khodabakhsh A. Broadband time-resolved absorption and dispersion spectroscopy of methane and ethane in a plasma using a mid-infrared dual-comb spectrometer. Sensors. 2020;20(23):6831.
22. Abbas MA, Pan Q, Mandon J, Cristescu SM, Harren FJM, Khodabakhsh A. Time-resolved mid-infrared dual-comb spectroscopy. Sci Rep. 2019;9(1):17247.
23. Dębek R, Azzolina-Jury F, Travert A, Mauge F, Thibault-Starzyk F. Low-pressure glow discharge plasma-assisted catalytic CO_2 hydrogenation: The effect of metal oxide support on the performance of the Ni-based catalyst. Catal Today. 2019;337:182–194.
24. Vasilyev S, Moskalev IS, Smolski VO, Peppers JM, Mirov M, Muraviev AV, et al. Super-octave longwave mid-infrared coherent transients produced by optical rectification of few-cycle 25-µm pulses. Optica. 2019;6(1):111–114.
25. Petersen CR, Møller U, Kubat I, Zhou B, Dupont S, Ramsay J, et al. Mid-infrared supercontinuum covering the 1.4-13.3 µm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre. Nat Photon. 2014;8(11):830–834.
26. Petersen CR, Moselund PM, Huot L, Hooper L, Bang O. Towards a table-top synchrotron based on supercontinuum generation. Infrared Phys Technol. 2018;91:182–186.
27. Wang Y, et al. Mid-infrared supercontinuum generation spanning from 1.9 to 5.7 µm in a chalcogenide fiber taper with ultra-high NA. Infrared Phys Technol. 2018;88:102–105.
28. Kwarkye K, Jensen M, Engelsholm RD, Dasa MK, Jain D, Bowen P, et al. In-amplifier and cascaded mid-infrared supercontinuum sources with low noise through gain-induced soliton spectral alignment. Sci Rep. 2020;10(1):8230.
29. Sylvestre T, Genier E, Ghosh AN, Bowen P, Genty G, Troles J, et al. Recent advances in supercontinuum generation in specialty optical fibers [Invited]. J Opt Soc Am B. 2021;38(12):F90.
30. Eslami Jahromi K, Nematollahi M, Pan Q, Abbas MA, Cristescu SM, Harren FJM, et al. Sensitive multi-species trace gas sensor based on a high repetition rate mid-infrared supercontinuum source. Opt Express. 2020;28(18):26091–26101.
31. Kilgus J, Duswald K, Langer G, Brandstetter M. Mid-infrared standoff spectroscopy using a supercontinuum laser with compact Fabry-Perot filter spectrometers. Appl Spectrosc. 2018;72(4):634–642.
32. Israelsen NM, Petersen CR, Barh A, Jain D, Jensen M, Hannesschläger G, et al. Real-time high-resolution mid-infrared optical coherence tomography. Light. 2019;8(1):11.
33. Zorin I, Su R, Prylepa A, Kilgus J, Brandstetter M, Heise B. Mid-infrared Fourier-domain optical coherence tomography with a pyroelectric linear array. Opt Express. 2018;26(25):33428–33439.
34. Zorin I, Gattinger P, Ebner A, Brandstetter M. Advances in mid-infrared spectroscopy enabled by supercontinuum laser sources. Opt Express. 2022;30(4):5222–5254.

35. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
36. Petersen CR, Prtljaga N, Farries M, Ward J, Napier B, Lloyd, GR, et al. Mid-infrared multispectral tissue imaging using a chalcogenide fiber supercontinuum source. Opt Lett. 2018;43(5):999–1002.
37. Kilgus J, Langer G, Duswald K, Zimmerleiter R, Zorin I, Berer T, et al. Diffraction limited mid-infrared reflectance microspectroscopy with a supercontinuum laser. Opt Express. 2018;26(23):30644–30654.
38. Napier B, Bang O, Markos C, Moselund P, Huot L, Harren FJM, et al. Ultra-broadband infrared gas sensor for pollution detection: The TRIAGE project. J Phys Photonics. 2021;3(3):031003.
39. Durocher-Jean A, Jean-Ruel H, Dion-Bertrand LI, Blais-Ouellette S, Stafford L. Ultra-high-resolution optical absorption spectroscopy of DC plasmas at low pressure using a supercontinuum laser combined with a laser line tunable filter and a HyperFine spectrometer. J Phys D: Appl Phys. 2021;54(8):085204.
40. Woyessa G, Kwarkye K, Dasa MK, Petersen CR, Sidharthan R, Chen S, et al. Power stable 1.5–10.5 μm cascaded mid-infrared supercontinuum laser without thulium amplifier. Opt Lett. 2021;46(5):1129–1132.
41. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transfer. 2022;277:107949.
42. Danner H, Samudrala D, Cristescu SM, van Dam NM. Tracing Hidden Herbivores: Time-Resolved Non-Invasive Analysis of Belowground Volatiles by Proton-Transfer-Reaction Mass Spectrometry (PTR-MS). J Chem Ecol. 2012;38(6):785–794.
43. Sola Martinez RA, et al. Data preprocessing workflow for exhaled breath analysis by GC/MS using open sources. Sci Rep. 2020;10(1):22008.
44. Delahaye T, Armante R, Scott NA, Jacquinet-Husson N, Chédin A, Crépeau L, et al. The 2020 edition of the GEISA spectroscopic database. J Mol Spectrosc. 2021;380:111510.
45. Kozlov KV, Michel P, Wagner HE. Synthesis of organic compounds from mixtures of methane with carbon dioxide in dielectric-barrier discharges at atmospheric pressure. Plasmas Polym. 2000;5(3-4):129–150.
46. Delikonstantis E, Scapinello M, Stefanidis GD. Low energy cost conversion of methane to ethylene in a hybrid plasma-catalytic reactor system. Fuel Process Technol. 2018;176:33–42.
47. Nozaki T, Okazaki K. Carbon nanotube synthesis in atmospheric pressure glow discharge: A review. Plasma Process Polym. 2008;5(4):300–321.
48. Fridman G, Friedman G, Gutsol A, Shekhter AB, Vasilets VN, Fridman A. Applied plasma medicine. Plasma Process Polym. 2008;5(6):503–533.
49. Winter J, Wende K, Masur K, Iseni S, Dünnbier M, Hammer MU, et al. Feed gas humidity: A vital parameter affecting a cold atmospheric-pressure plasma jet and plasma-treated human skin cells. J Phys D: Appl Phys. 2013;46(29):295401.
50. Schmidt-Bleker A, Winter J, Iseni S, Dünnbier M, Weltmann KD, Reuter S. Reactive species output of a plasma jet with a shielding gas device—combination of FTIR absorption spectroscopy and gas phase modelling. J Phys D: Appl Phys. 2014;47(14):145201.
51. Wang X, Zhou M, Jin X. Application of glow discharge plasma for wastewater treatment. Electrochim Acta. 2012;83:501–512.
52. Ren J, Yao M, Yang W, Li Y, Gao J. Recent progress in the application of glow-discharge electrolysis plasma. Cent Eur J Chem. 2014;12(12):1213–1221.

53. Samanta K, Jassal M, Agrawal AK. Atmospheric pressure glow discharge plasma and its applications in textile. Indian J Fibre Text Res. 2006;31(1):83–98.

54. Sharpe SW, Johnson TJ, Sams RL, Chu PM, Rhoderick GC, Johnson PA. Gas-phase databases for quantitative infrared spectroscopy. Appl Spectrosc. 2004;58(12):1452–1461.



Chapter 5

In-situ analysis of methane plasmas with mid-infrared supercontinuum-based Fourier-transform spectroscopy[1]

1. This chapter is based on the following article:
Krebbers R, Liu N, Averink W, Harren FJM, Butterworth T, van Rooij G, Khodabakhsh, A, Cristescu, SM. *In-situ* analysis of methane plasmas with mid-infrared supercontinuum-based Fourier-transform spectroscopy. *Appl Phys B. 2025;131(1):5.*
Selected as ***Editor's Choice Article Highlight*** *and* ***Issue cover image***

Abstract

The conversion of methane (CH_4) into the hydrocarbons ethane (C_2H_6) and acetylene (C_2H_2) in plasma is investigated in the first *in-situ* study using a mid-infrared supercontinuum source and a Fourier transform spectrometer. Due to the spatial coherence of the supercontinuum source, it is possible to probe the plasma over a path length of 3.5 meters and observe CH_4, C_2H_6, and C_2H_2 with concentrations below percent level at a few mbar pressure. The broad spectral range of the supercontinuum source (1.4 – 4.1 µm) covers entire rovibrational bands of molecules, allowing simultaneous detection of the rotational and vibrational temperature of CH_4 and C_2H_2. The ability to study these experimental parameters *in situ* opens new possibilities for understanding and optimizing plasma-based chemical conversions.

5.1 Introduction

Most of the methane (CH_4) used in industry is for the production of grey hydrogen as a feedstock for ammonia, methanol, and refining (1). This grey hydrogen causes ~2% of global carbon dioxide (CO_2) emissions (2). Most carbon emissions originate from the combustion of methane used to drive the endothermic reactions, together with the intrinsic CO_2 emissions from steam-methane reforming and the water-gas shift reaction.

To minimize the carbon footprint, plasma-driven conversions are an attractive alternative (3–7), as the energy input can be provided by electricity from sustainable energy with zero CO_2 emissions (e.g. wind or solar energy). However, plasma-driven conversions demand further improvements to match (or possibly even surpass) both the efficiency and throughput of the established industrial processes, and to make the green transition more viable. Therefore, there is always a need for novel tools and methods to provide a better understanding of plasma-driven chemistry.

Chemical analysis of the plasma reaction is usually done by analysis of the reaction products upon efflux (7,8). To investigate in more detail the properties of the plasma and the behavior of chemical reactions in the plasma, *in-situ* measurements are necessary. Absorption-based spectroscopy is preferred because it allows non-invasive, *in-situ* measurements of temperature and species concentrations. It directly determines the population of the energy states and hence the rotational and vibrational temperature. In contrast, Optical Emission Spectroscopy (OES) requires modeling to derive the ratio between ground-state and excited-state populations, and consequently the temperature of the gas (9).

In a plasma, there can be a non-equilibrium between electronic, vibrational, and translational energy distributions, which can result in chemical reaction rates that are significantly different from traditional thermal chemical reactions. The constant, high-voltage induced, electron current creates radicals, ions, and converted neutrals, all in electronic, vibrational, and rotational excited states. The relaxation (de-excitation) rate from their excited level depends strongly on the molecule and pressure (collision rate) and temperature of the gas. Within an electronic state, there is a large difference between the de-excitation rate of rotational states and vibrational states. Vibrational levels require thousands of collisions to be de-excited to a lower vibrational state. In contrast, the de-excitation rate of a rotational level requires only a few collisions, and its energy is largely in equilibrium with the translational energy (10). Therefore, the Maxwell Boltzmann distribution for the rotational states within a vibrational band can be attributed to a rotational temperature (T_{rot}) and between the vibrational bands to a vibrational temperature (T_{vib}). For this reason, broad spectral coverage is needed to obtain the

entire molecular vibrational bands to simultaneously determine the T_{rot} and T_{vib} of the molecular species in the plasma.

Previously, *in-situ* plasma analysis was performed using absorption spectroscopy with an infrared Quantum Cascade Laser (QCL) to measure the concentration and temperature of CO_2 (11). The use of such a single-frequency laser is promising for rapid detection of species in the plasma, but it is difficult to achieve multi-species detection due to the generally limited spectral coverage of these types of lasers. As a result, only a few molecular absorption lines can be observed, limiting the ability to distinguish molecular rotational and vibrational temperatures in the plasma. Laser-based systems with limited spectral coverage typically determine only the rotational (12) or vibrational (11) temperature of the molecules. To overcome this, multiple laser systems are used to achieve multi-species detection. For example, Damen and coworkers used three collinear QCLs to determine the time evolution of the gas temperature, including rotational and vibrational temperatures, in a pulsed CO_2-N_2 glow discharge (13). Although time-resolved measurements can be improved through an intermittent scanning scheme (14), the use of multiple lasers and time-division multiplexing technology makes the system more expensive and complicated.

Another approach is to use Fourier Transform Infrared (FTIR) absorption spectroscopy, which has been applied within CO_2-glow-discharge-plasma research (15–18). However, classical FTIR systems have limited spectral resolution and/or long measurement times due to their incoherent light sources, which is undesirable for simultaneous discrimination of multiple species and rapid analysis. Furthermore, using an incoherent light source limits the optical path length and thus the detection sensitivity for weak absorption lines and/or for low gas concentrations.

Broadband spectroscopy has recently demonstrated great potential for the study of methane discharge plasmas via Dual-Comb Spectroscopy (DCS). The method allows a detailed plasma analysis of the recombination products and the kinetic processes (19–21). However, the ongoing challenge is to obtain broad wavelength coverage in the mid-infrared wavelength region, with high spectral resolution and sufficient detection sensitivity, to cover entire rovibrational molecular absorption bands in a simple experimental setup.

The emergence of mid-infrared supercontinuum (SC) sources offers an attractive new avenue for plasma-based gas spectroscopy. The unique combination of the broad spectral coverage and spatial coherence of SC sources has proven capabilities for the sensitive, simultaneous detection of multiple gas concentrations (22–24). Compared to classical thermal sources, the SC sources have the advantage of possessing spatial coherence, which not only enables longer optical interaction path lengths to enhance the detection sensitivity but also allows the plasma to be probed for *in-situ* analysis. Compared to DCS, the reduced

experimental complexity of an SC source-based setup generally allows a broader spectral coverage as temporal coherence no longer is a requirement (25). Here, we present for the first time *in-situ* CH_4 plasma analysis using mid-infrared SC sources. Our SC source enables the simultaneous detection of methane and its main reaction products acetylene and ethane. Based on the large number of simultaneously observed rovibrational transitions, the molecular excitation mechanisms in the plasma can be determined.

5.2 Experimental set-up

For *in-situ* detection of methane discharge products, a water-cooled glow-discharge plasma reactor is used, as illustrated in Figure 5.1. The discharge cell has a length of 175 cm and an inner diameter of 1.5 cm. A constant flow of water is used to cool the plasma cell. A ZnSe window is mounted on the beam inlet port of the cell at Brewster angle, and a mirror is placed at the other end of the discharge cell to double the light interaction path length to 3.5 m. The plasma is generated by applying DC power with an HV power supply (Hippotronics Inc., 40 kV, 40 mA) and regulating the current using a tetrode tube (Eimac 8188, screen voltage 1.5 kV) yielding a maximum potential input power of 1.5 kW. The anode and gas inlet are in the middle of the cell, and two hollow cathodes and gas outlets are located at each end of the cell, creating a symmetrical discharge.

Two mass flow controllers (EL-FLOW select, Bronkhorst Instruments) are used to control the flow rate of the incoming gas mixture, ensuring a total flow rate of ~2 ln/h. The discharge pressure was controlled by a pressure controller (EL-PRESS, Bronkhorst Instruments) in the gas outlet towards a vacuum pump (Adixen ACP15). Excessive discharge currents can result in soot production (e.g. carbon) accumulating on the discharge wall. To remove these deposits, a discharge in a mixture of O_2 and Helium is required.

The SC source consists of a fiber-coupled broadband mid-infrared source (SuperK MIR, NKT Photonics, power 500 mW, spectral coverage 2 – 4.1 µm, repetition rate 2.5 MHz). The fiber-coupled output of the SC source is connected to an off-axis parabolic mirror collimator (RC02FC-P01, Thorlabs) that sends the light beam to the discharge cell. Returning from the discharge cell it is directed towards a custom-built Fourier transform spectrometer (FTS). The set-up of the FTS is based on a Michelson interferometer and has been described in our previous work (24). In short, the incoming light beam is split into two arms by a beamsplitter (BSW711, Thorlabs), each leading to a cubic retroreflector mirror (HRR201-P01, Thorlabs). Upon reflection, the beams are recombined on the beamsplitter, and the resulting

recombined beams propagate toward two photovoltaic detectors (PVI-4TE-4, Vigo Systems). The two signals are subtracted using a differential amplifier (SR560, Stanford Research Systems) in a balanced detection scheme, and the final signal is collected by a data acquisition card (DAQ, BNC-2110, National Instruments) and recorded by a LabVIEW-based program. The two retroreflectors are mounted on a linear-motor translation stage (DDSM100/M, Thorlabs). Scanning the translation stage over a distance of 2.5 cm results in an optical path difference of 10 cm between the two beams, which provides a spectral resolution of 1/10 = 0.1 cm^{-1} or 3 GHz.

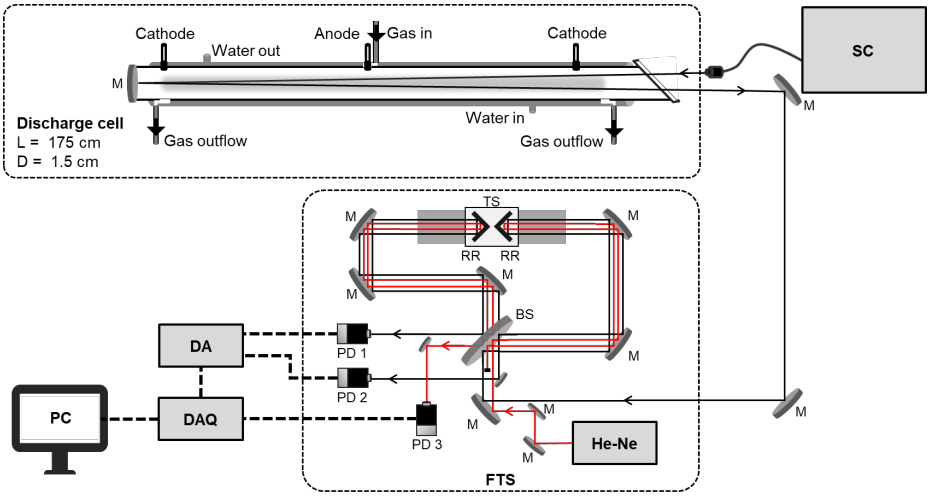


Figure 5.1: Set-up with discharge cell and SC-based Fourier transform spectrometer. M: mirror, TS: translation stage, RR: retroreflector, BS: beam splitter, PD: photodetector, He-Ne: helium-neon laser, DA: differential amplifier, DAQ: data acquisition card.

5.2.1 Retrieving gas concentrations and vibrational/rotational temperatures

To identify the compounds inside the plasma and to find their concentrations and associated temperatures, the absorbance spectra measured with the FTS are compared to simulated spectra using the parameters in the HITRAN2020 database (26). The simulated spectra are calculated with fixed input parameters for optical path length L, pressure P, and a Voigt profile convolved with a sinc-function (to account for the instrument line shape function of the FTS). It is then iteratively fitted to the measured spectra with varying parameters for concentration c and temperature T, ultimately yielding an optimum value when the squared sum of the residuals is minimal. To take into account the effect of the temperature inside the plasma on the observed line intensity S, the line strength as provided by HITRAN for a reference temperature of 298 K, needs to be recalculated for a given temperature T as

$$S_{ij}(T) = S_{ij}(T_{\mathrm{ref}}) \frac{Q(T_{\mathrm{ref}})}{Q(T)} \frac{\exp\left(-\frac{c_2 E''}{T}\right)\left[1-\exp\left(-\frac{c_2 \nu_{ij}}{T}\right)\right]}{\exp\left(-\frac{c_2 E''}{T_{\mathrm{ref}}}\right)\left[1-\exp\left(-\frac{c_2 \nu_{ij}}{T_{\mathrm{ref}}}\right)\right]}. \tag{5.1}$$

In this equation, Q is the total internal partition function, E'' the lower state energy, ν_{ij} the energy difference (in wavenumber) between lower state i and higher state j (all provided by HITRAN), and c_2 the second radiation constant $c_2 = hc/k_B$, in which h is the Planck constant, c the speed of light and k_B the Boltzmann constant (27). As the temperature is determined both vibrationally and rotationally, equation (5.1) needs to be rewritten as (28):

$$S_{ij}(T) = S_{ij}(T_{\mathrm{ref}}) \frac{Q_{\mathrm{vib}}(T_{\mathrm{ref}})}{Q_{\mathrm{vib}}(T_{\mathrm{vib}})} \frac{Q_{\mathrm{rot}}(T_{\mathrm{ref}})}{Q_{\mathrm{rot}}(T_{\mathrm{rot}})} \frac{\exp\left(-\frac{c_2 E''_{\mathrm{vib}}}{T_{\mathrm{vib}}}\right)\exp\left(-\frac{c_2 E''_{\mathrm{rot}}}{T_{\mathrm{rot}}}\right)\left[1-\exp\left(-\frac{c_2 \nu_{ij}}{T}\right)\right]}{\exp\left(-\frac{c_2 E''_{\mathrm{vib}}}{T_{\mathrm{ref}}}\right)\exp\left(-\frac{c_2 E''_{\mathrm{rot}}}{T_{\mathrm{ref}}}\right)\left[1-\exp\left(-\frac{c_2 \nu_{ij}}{T_{\mathrm{ref}}}\right)\right]} \tag{5.2}$$

To separate the partition functions into a vibrational and rotational part, the Total Internal Partition Sum (TIPS) from HITRAN is used. Generally, the TIPS can be written as $Q(T) = Q_{\mathrm{vib}} \cdot Q_{\mathrm{rot}}$. In which Q_{vib} can then be approximated by the harmonic oscillator approximation:

$$\mathrm{Q_{vib}\,(T_{vib})} = \prod_{\substack{\text{fundamental}\\ \text{vibrations}}} \left[\left(1-\exp\left(\frac{c_2 G_v}{T}\right)\right)\right]^{-1} \tag{5.3}$$

where G_v is the term value of the vibrational state (29). Q_{rot} is consequently calculated by dividing the TIPS by Q_{vib}. As G_ν and consequently E''_{vib} are not

5

provided in HITRAN, this information needed to be added for the fundamental vibrational levels of methane and acetylene and their major isotopes, of which CH_4 (30), $^{13}CH_4$ (31), C_2H_2 (30), $H^{13}C^{12}CH$ (32,33), and C_2HD (30) were used.

The rotational and vibrational temperatures are determined separately. The rotational temperature results from the ratio of the absorbance of rovibrational lines belonging to different rotational energy levels (J states) within the same vibrational transition. For the vibrational band used, the molecules are transitioning from the vibrational ground state, and consequently the vibrational lower state energy (E''_{vib}) is zero, canceling out this term. The vibrational temperature of methane was determined by making a fit to the hot bands of methane and retrieving the temperature for which the simulated absorbance of the hot bands matches the measured absorbance, at the concentration previously retrieved for the fit on the ν_3 transitions of methane. To complete the fit a baseline correction function was used, consisting of a 12th-order polynomial needed to compensate for distortions in the signal caused by refractive index changes of the plasma, and a sum of a few low-frequency sine waves (simultaneously fit to the measured spectrum) to remove the baseline and the etalon fringes, respectively.

5.3 Results

5.3.1 The vibrational and rotational temperature of methane inside a plasma

Previously, we demonstrated multi-species trace gas detection and reaction product analysis at varying concentrations and pressures, with a supercontinuum source and a custom-built FTS. To retrieve a molecular concentration, we used the spectroscopic database HITRAN2020 (8,24). For the current plasma experiments, we aim to determine the temperatures in the plasma from the spectra. In our first experiment, a mixture of 10% methane and 90% CO_2 was fed to the discharge tube with a flow of 5 ln/h. The pressure inside the tube was 10 mbar, and a discharge was created by applying a voltage of 8.5 kV at a current of 15 mA to the tube, yielding a Specific Energy Input (SEI) of 2.1 MJ/mol. The corresponding absorbance spectrum, continuously averaged for 16 minutes, is shown in Figure 5.2 (in red). The difference with a methane spectrum at room temperature (5% methane in 95% CO_2, flow 5 ln/h) without discharge is also shown in Figure 5.2 (inverted in blue). To highlight the differences in the respective spectra, magnifications are shown in Figures 5.2.b and 5.2.c. The increased absorbance of the hot bands (indicated with asterisks) implies vibrational excitation. The stronger absorbance of the lines around 3200 cm^{-1}

indicates a rotational energy distribution towards higher temperatures, as these lines relate to higher rotational levels in the R-branch (also visible in the P-branch).

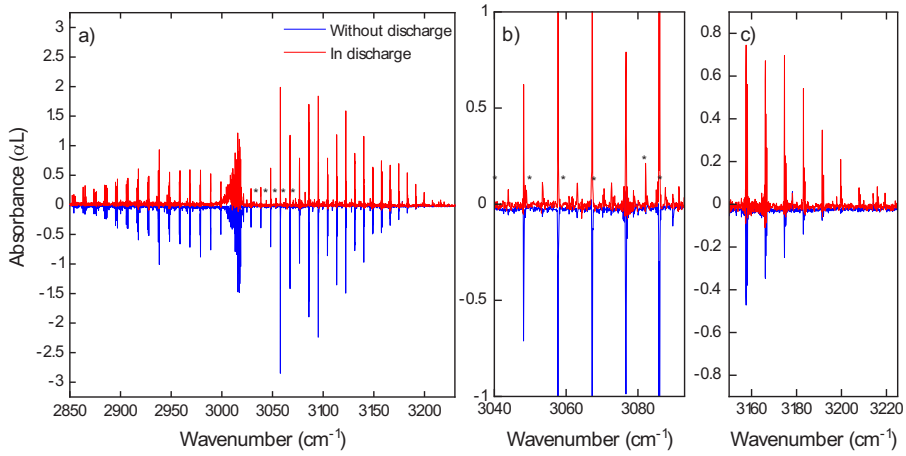


Figure 5.2: a) Absorbance spectra of the ν_3 band of 10% methane in 90% CO_2 in a discharge (SEI of 2.1 MJ/mol, 10 mbar), exhibiting features of an increased rotational and vibrational temperature (in red), and 5% methane in 95% CO_2 under room temperature conditions (in blue, inverted), both measured in 16 minutes. The asterisks indicate the location of hot bands. **b)** and **c)** enlargements of the spectrum, showing the features related to, respectively, vibrational and rotational excitation in more detail

To demonstrate the high accuracy of the retrieved parameters in discharge, Figure 5.3 contains a zoomed-in measured spectrum of the R-branch of the ν_3-band of CH_4 during discharge, together with the HITRAN-simulated, fitted CH_4 spectra. For the retrieved parameters (T_{rot} of 567 ± 2 K, T_{vib} of 604 ± 5 K, and concentration of 0.9 %), this results in a featureless residual, indicating the good quality of the fit.

The difference that $T_{\mathrm{vib}} > T_{\mathrm{rot}}$ was systematically found in our other experiments where methane was mixed with other gases (see Figure 5.4); this will be discussed later. Furthermore, for identical discharge parameters (SEI of 2.1 MJ/mol, pressure 10 mbar), both T_{vib} and T_{rot} in methane are higher in mixtures with CO_2 and N_2 than in mixtures with a noble gas, such as helium or argon.

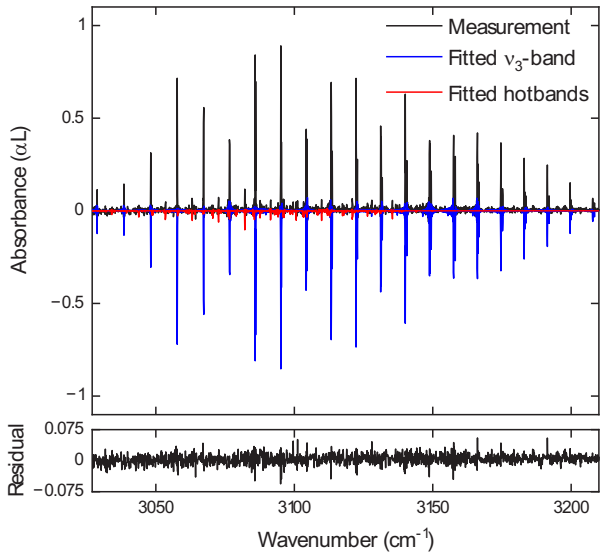


Figure 5.3: Measured absorbance spectrum of the R-branch of the ν_3-band CH_4 in a discharge (SEI of 2.1 MJ/mol, 10 mbar, measured and averaged for 16 minutes, in black), with fitted, simulated spectra of CH_4 (in blue and red). The featureless residual indicates a good agreement between the fit and measured spectra.

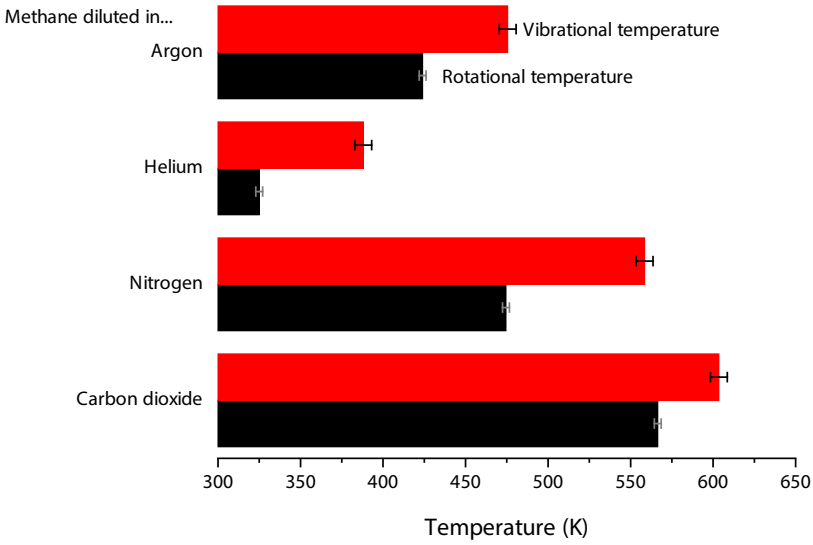


Figure 5.4: Retrieved rotational and vibrational temperatures of 10% methane mixed with various gases in a discharge (SEI 2.1 MJ/mol, pressure 10 mbar). The uncertainty indicated by the error bar is the average standard deviation for the 55 data samples used in this study.

5.3.2 Methane/helium plasma with varying Specific Energy Input

The absorbance spectrum of methane under varying plasma conditions was measured (10% methane in 90% helium, 5.9 mbar). The SEI was varied by controlling the current at a constant voltage (3.5 kV, flow 2 ln/h). Figure 5.5 shows the absorbance spectrum with an SEI of 0.8 MJ/mol, containing absorption features of methane and its reaction products ethane (C_2H_6) and acetylene (C_2H_2). The simulated spectra of methane, ethane, and acetylene are shown inverted. These spectra were fitted to the measured spectrum to determine the concentration of these species. The wide spectral range of the source enables us to determine T_{rot} of methane and acetylene. For methane, this was done specifically for the R-branch of the ν_3 band, as these transitions are largely free of interferences from ethane lines. For acetylene, all visible transitions, located in the observed window of 3200 – 3280 cm^{-1} were considered. The retrieved concentrations and rotational temperatures are shown in Figures 5.6.a and 5.6.b, respectively. Each point in these figures is the average over five consecutive measurements (six minutes each), with its standard deviation shown as error bars. In the case of ethane, limitations in the HITRAN2020 database prevent the ability to retrieve a rotational temperature for this molecule. Compared to previous editions of HITRAN, the number of available ethane transitions around 2.9 µm has improved significantly in the HITRAN2020 edition. However, the available line list is still incomplete as there are no hot band transitions listed. In addition, the spectrum is complex due to interfering combinational bands and overtones (26). This effect is particularly pronounced for ethane spectra at elevated temperatures. Therefore, the fit of the simulated ethane spectrum to the measured spectra is not perfect, leaving some features visible in the residual. This is also evident from the increased error bars in the concentration of ethane, compared to those of methane and acetylene.

With increasing SEI, a decrease in the methane concentration can be observed, indicating an increasing conversion of methane (Figure 5.6.a). The concentrations of acetylene and ethane increase up to 1 MJ/mol and then decrease, indicating that these reaction products are in turn also dissociated at higher discharge powers or other reaction pathways begin to prevail. Probably most of the molecules are converted into soot depositions, which clearly form on the walls of the plasma cell at higher SEIs. The relative decrease is larger for ethane than for acetylene, which can be explained by a reaction pathway where acetylene is formed from the dehydrogenation of ethane. The rotational temperatures of methane and acetylene increase as more power is applied to the discharge. Without plasma the temperature of methane is, as expected, close to room temperature. As acetylene is not present without plasma, a sample of acetylene was measured in a gas cell and the fitting

routing was applied. A temperature of 292 K was found, which indicates that the fitting routine also yields reliable results for the rotational temperature of acetylene.

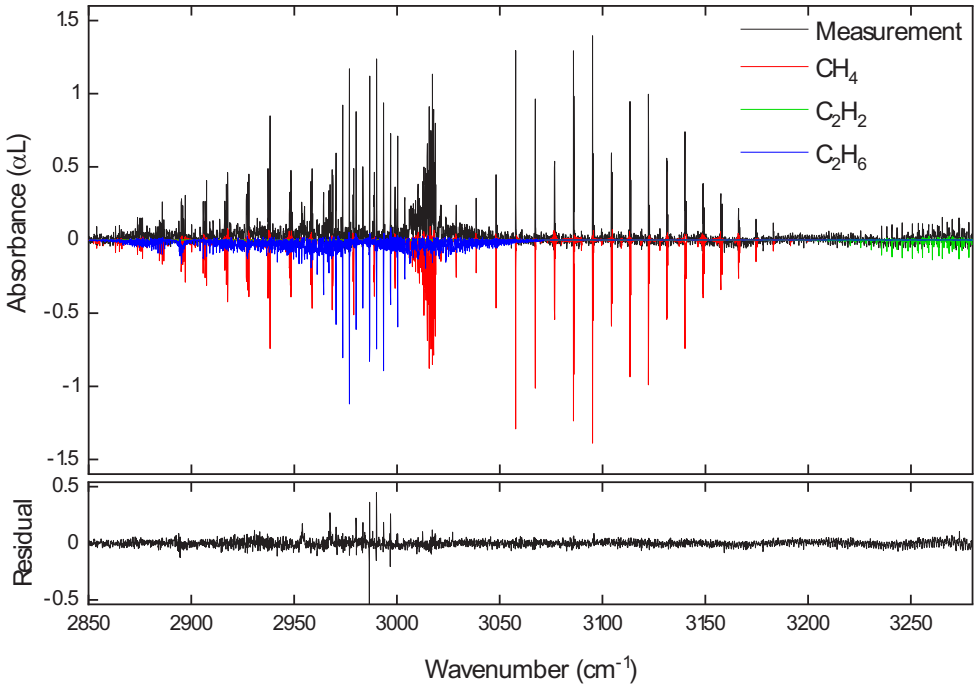


Figure 5.5: Absorbance spectrum of 10% methane in helium in a plasma (SEI 0.8 MJ/mol, measured in 6 minutes). Fitted simulated spectra for methane, acetylene, and ethane are shown inverted. The residual of the fit is shown in the bottom panel.

Figure 5.7 shows the vibrational temperature of methane at various SEI values. A significant discrepancy can be observed between the rotational and vibrational temperature of methane, which will be discussed in the next section. The higher uncertainty in the vibrational temperature can be explained by the low absorbance and limited number of hot band transitions in comparison with the regular fundamental ν_3 transitions used for the rotational fit, which can be observed in Figure 5.2.

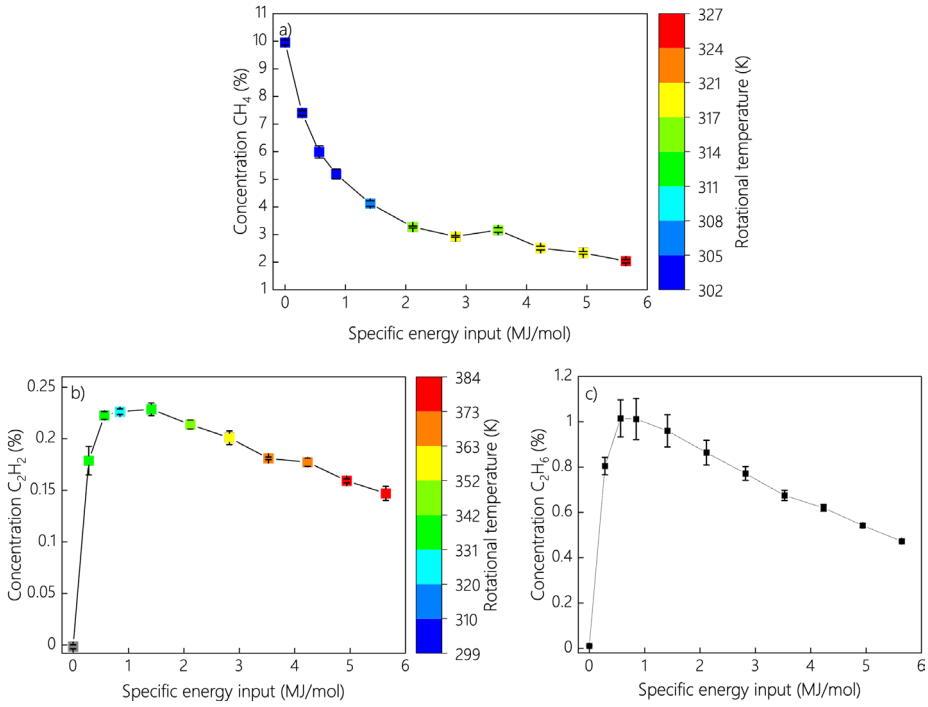


Figure 5.6: Retrieved concentration of **a)** methane, **b)** acetylene, and **c)** ethane, and associated rotational temperatures for methane and acetylene inside the methane/helium plasma over varying energy input. The error bar is the standard deviation of 5 consecutive measurements.

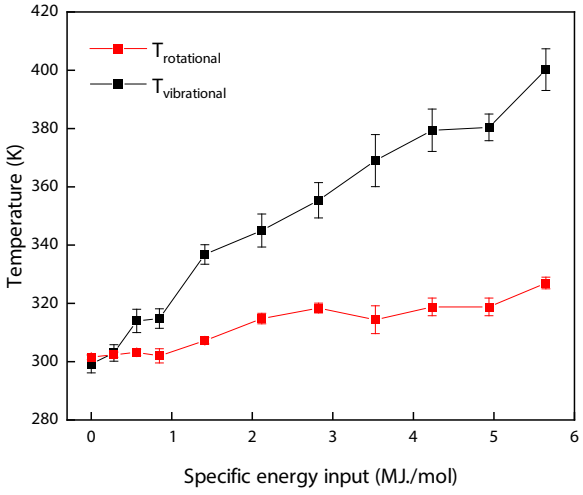


Figure 5.7: Retrieved (rotational and vibrational) temperatures of methane inside the methane/helium plasma over varying SEI, the error bar is the standard deviation of 5 consecutive measurements.

5.4 Discussion

The use of MIR SC sources for spectroscopy in plasmas allows obtaining a broader spectral coverage than conventional laser sources, while maintaining long interaction pathlengths with high brightness, outperforming incoherent light sources of classical FTIRs. The results in this work serve as proof of principle of these features for this particular application. However, the results also provide openings for further discussion, especially on the reported temperatures of the species present in the plasma.

Generally, the reported temperatures indicate that the chemistry happening inside the plasma is primarily driven by electron impact dissociation, as they are too low for thermally driven chemistry. Interestingly, there is a difference between the observed rotational temperatures of methane and acetylene. Methane, even under highly applied currents, shows only a moderate temperature increase, while acetylene appears to reach higher temperatures. The most probable explanation for this is the location of these molecules in the discharge. As methane is a starting compound of the plasma reaction, its density is expected to be the highest around the inlet of the plasma cell. On the other hand, acetylene is expected to be mostly located further towards the outflow of the plasma cell, as it is a reaction product of methane plasmas. Therefore, acetylene is in the region of the plasma that is expected to be the hottest, resulting in a higher temperature than the methane molecules. This way, it seems reasonable to assume that the rotational temperature of a molecule might serve as a parameter to indicate the density distribution of the molecule inside the plasma cell, given there is a temperature gradient inside the discharge.

Moreover, the results show a difference between T_{vib} and T_{rot} of methane. We consider the explanation for this to be either an artifact from the measurement or to be a genuine characteristic of the plasma. The former, which seems most probable, can be explained by the line of sight of the measurement. As the SC beam travels through the length of the tube, it comes across regions of the plasma with different temperatures, which was already indicated by the temperature discrepancies between methane and acetylene. It seems reasonable to assume that the line-averaged spectrum of methane is a composite of methane from regions of the plasma with different temperatures. Simulations indicate that such spatial temperature heterogeneity could indeed appear in a composite spectrum as a difference between T_{vib} and T_{rot}. Future work could verify this hypothesis by measuring the plasma perpendicular to the length of the discharge tube and obtaining T_{vib} and T_{rot} for different regions of the plasma. As this would reduce the interaction path length with the discharge, and consequently the SNR of the measurement,

measurements in this geometry would benefit from a multi-pass arrangement, for example with a set of corner-cube mirrors or two concave mirrors in a Herriott-type multi-pass set-up. Alternatively, a different plasma source with a wider diameter or operating at a higher pressure could compensate the loss in SNR. Secondly, we hypothesize that the inhomogeneity of T_{vib} and T_{rot} is a genuine characteristic of the plasma. Differences between T_{vib} and T_{rot} have also been reported by studies on glow discharges using coherent anti-Stokes Raman scattering (34,35). This would be explained by the difference in relaxation times of excited rotational and vibrational energy levels. Given that the rotational de-excitation is faster (only a few collisions needed) than vibration collisional de-excitation, T_{rot} will be smaller than T_{vib} (10,36). Another explanation could be that the energies of rotational and vibrational energy levels behave independently, which has been shown for other polyatomic molecules (37). This would imply that inside the discharge the excitation of the rotational and vibrational levels occurs independently, rather than that the differences follow from different relaxation rates. Apart from the exact mechanism behind the differences in T_{vib} and T_{rot}, its implication would then be important to consider. As we have observed that T_{rot} can be up to 15% lower than T_{vib}, it is important to distinguish rotational and vibrational temperatures when using rovibrational absorption spectroscopy on non-equilibrium media such as plasmas. In the literature there are examples where these parameters are provided (such as (13,17,21)), but also where these parameters are not mentioned at all (11,38,39).

Interestingly, the conversion of methane did not yield an observable concentration of ethylene (C_2H_4). From the noise in the residual of the spectrum, the limit of detection (3σ) for ethylene was determined to be 500 ppm. Therefore, the line-averaged concentration will have an upper limit of 500 ppm, roughly 3 to 4 times lower than the observed concentrations of acetylene. Theoretically, ethylene could have formed through two main pathways (40–42). For instance, ethylene forms from the dehydrogenation of the produced ethane. The absence of ethylene could then be explained by further dehydrogenation of ethylene towards acetylene, assuming the reaction rate of ethylene-acetylene is much higher than ethane-ethylene. In this case, the ethylene produced is short-lived and therefore yields a low concentration averaged through the tube. The other possibility is that ethylene forms from two CH_2 radicals, and the reaction rate of both the dehydrogenation of methane to CH_3 radicals and the CH_2 to CH radicals is higher than the dehydrogenation of CH_3 to CH_2 radicals. Further research aimed specifically at detecting the potential presence of these radicals (CH_3, CH_2, and CH) could reveal the pathway at play.

5.5 Conclusion

For the first time, we have demonstrated the possibility of *in-situ* plasma spectroscopy using MIR SC sources. The broad spectral coverage of the MIR SC source compared to narrowband lasers (such as QCLs) enables simultaneous detection of many more absorbance lines, which is a great advantage for multi-species detection. Using the MIR SC source in combination with our custom-built FTS, the conversion of CH_4 into the larger hydrocarbons C_2H_6 and C_2H_2 could be observed, and the concentration of all three molecules could be determined simultaneously. Moreover, the spatially coherent beam of the MIR SC proves to be an advantage over classical broadband light sources, previously used in FTIRs, as the path length could easily be improved over an order of magnitude compared to previous FTIR *in-situ* plasma studies (17). Therefore, the broad spectrum allows observing enough absorbance peaks to determine the rotational temperature, while the spatial coherence ensures the sensitivity necessary to observe sufficient absorbance peaks to determine the vibrational temperature of CH_4 in the discharge. The combination of the temperature parameters together with the observed concentrations of the gas species provides valuable insights into the conditions in the plasma, opening up a new perspective on and tools for improving plasma-based gas conversions. In its current configuration, the beam from the SC is traveling throughout the length of the discharge tube, yielding line-averaged concentrations and associated temperatures of the compounds inside the plasma. In future work, the beam could be directed perpendicular to the plasma to investigate distinct regions of the plasma separately. This would enable a further examination of the apparent spatial and/or rotational-vibrational temperature inhomogeneity and density distributions of the various molecular species in the plasma.

References

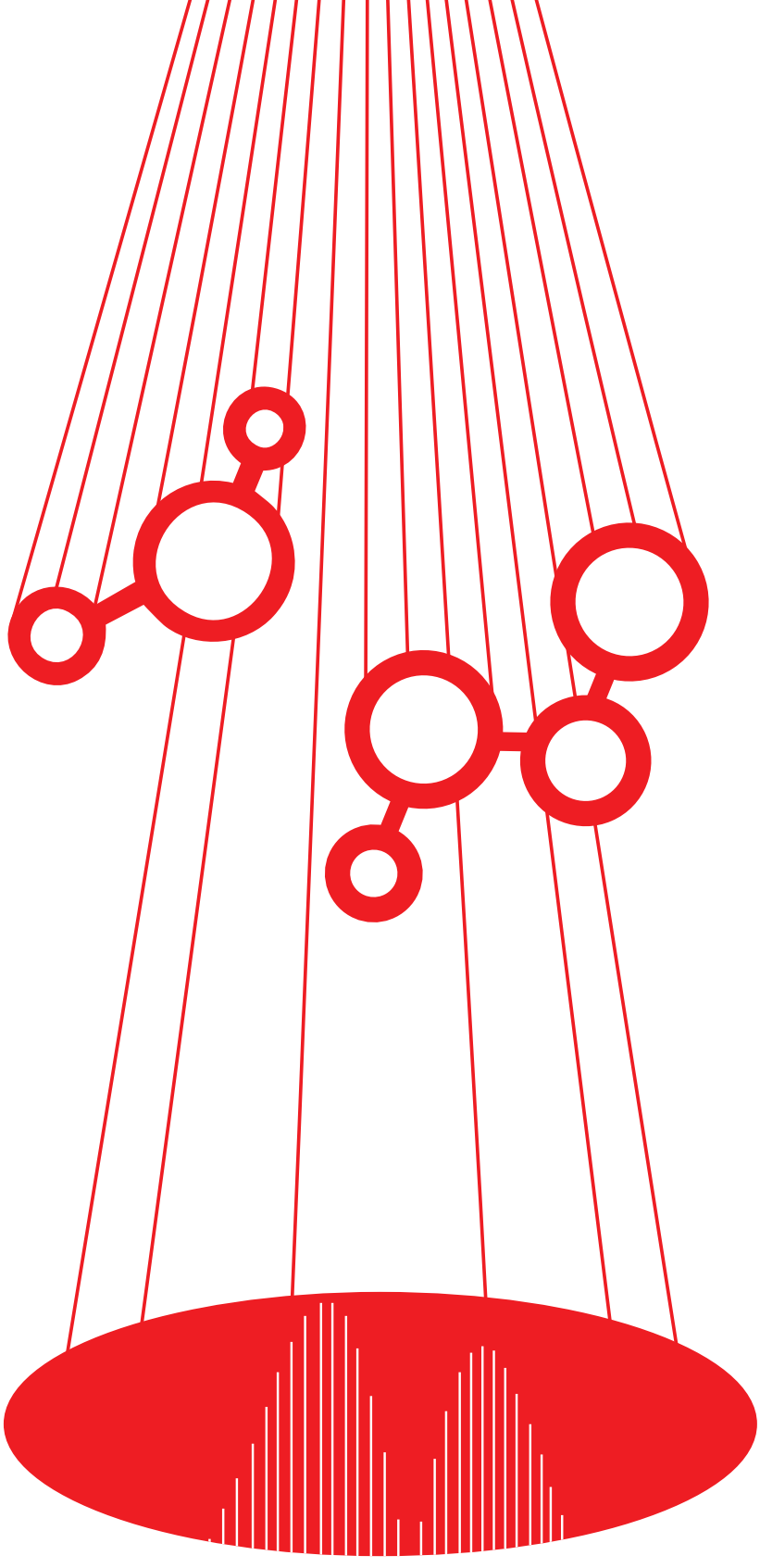
1. Iulianelli A, Liguori S, Wilcox J, Basile A. Advances on methane steam reforming to produce hydrogen through membrane reactors technology: A review. Catal Reviews – Sci Eng. 2016;58(1):1-35.
2. Gül T, Turk D, Bennett S, Remme U. The Future of Hydrogen. Paris: IEA; 2019.
3. Indarto A, Choi JW, Lee H, Song HK. Decomposition of greenhouse gases by plasma. Environ Chem Lett. 2008;6(4):215.
4. Snoeckx R, Bogaerts A. Plasma technology: A novel solution for CO_2 conversion? Chem Soc Rev. 2017;46(19):5805–5863.
5. Tao X, Bai M, Li X, Long H, Shang S, Yin Y, et al. CH_4–CO_2 reforming by plasma: Challenges and opportunities. Prog Energy Combust Sci. 2011;37(2):113–124.
6. Puliyalil H, Lašič Jurković D, Dasireddy VDBC, Likozar B. A review of plasma-assisted catalytic conversion of gaseous carbon dioxide and methane into value-added platform chemicals and fuels. RSC Adv. 2018;8(48):27481.
7. Wanten B, Maerivoet S, Vantomme C, Slaets J, Trenchev G, Bogaerts A. Dry reforming of methane in an atmospheric pressure glow discharge: Confining the plasma to expand the performance. J CO_2 Utilization. 2022;56:101869.
8. Krebbers R, Liu N, Jahromi KE, Nematollahi M, Bang O, Woyessa G, et al. Mid-infrared supercontinuum-based Fourier transform spectroscopy for plasma analysis. Sci Rep. 2022;12(1):9642.
9. Reuter S, Sousa JS, Stancu GD, Van Helden JPH. Review on VUV to MIR absorption spectroscopy of atmospheric pressure plasma jets. Plasma Sources Sci Technol. 2015;24(5):054001.
10. Lambert JD. Vibrational and Rotational Relaxation in Gases. Oxford: Clarendon Press; 1977.
11. Herklotz F, Rubin T, Sinnreich M, Helmke A, von Haimberger T, Heyne K. Fast Simultaneous CO_2 Gas Temperature and Concentration Measurements by Quantum Cascade Laser Absorption Spectroscopy. Appl Sci. 2022;12(10):5057.
12. Spearrin RM, Ren W, Jeffries JB, Hanson RK. Multi-band infrared CO_2 absorption sensor for sensitive temperature and species measurements in high-temperature gases. Appl Phys B. 2014;116(4):855.
13. Damen MA, Martini LM, Engeln R. Temperature evolution in a pulsed CO_2 –N_2 glow discharge measured using quantum cascade laser absorption spectroscopy. Plasma Sources Sci Technol. 2020;29(6):065016.
14. Fischer M, Tuzson B, Hugi A, Brönnimann R, Kunz A, Blaser S, et al. Intermittent operation of QC-lasers for mid-IR spectroscopy with low heat dissipation: tuning characteristics and driving electronics. Opt Express. 2014;22(6):7014.
15. Dębek R, Wierzbicki D, Motak M, Galvez ME, Da Costa P, Azzolina-Jury F. Operando FT-IR study on basicity improvement of Ni(Mg, Al)O hydrotalcite-derived catalysts promoted by glow plasma discharge. Plasma Sci Technol. 2019;21(4):045503.
16. Azzolina-Jury F, Thibault-Starzyk F. Mechanism of Low Pressure Plasma-Assisted CO_2 Hydrogenation Over Ni-USY by Microsecond Time-resolved FTIR Spectroscopy. Top Catal. 2017;60(19-20):1709.
17. Klarenaar BLM, Engeln R, Van Den Bekerom DCM, Van De Sanden MCM, Morillo-Candas AS, Guaitella O. Time evolution of vibrational temperatures in a CO_2 glow discharge measured with infrared absorption spectroscopy. Plasma Sources Sci Technol. 2017;26(11):115008.
18. Silva T, Grofulović M, Klarenaar BLM, Morillo-Candas AS, Guaitella O, Engeln R, et al. Kinetic study of low-temperature CO_2 plasmas under non-equilibrium conditions. I. Relaxation of vibrational energy. Plasma Sources Sci Technol. 2018;27(1):015019.

19. Abbas MA, van Dijk L, Jahromi KE, Nematollahi M, Harren FJM, Khodabakhsh A. Broadband Time-Resolved Absorption and Dispersion Spectroscopy of Methane and Ethane in a Plasma Using a Mid-Infrared Dual-Comb Spectrometer. Sensors. 2020;20(23):6831.
20. Hoghooghi N, Chang P, Egbert S, Burch M, Shaik R, Diddams SA, et al. GHz repetition rate mid-infrared frequency comb spectroscopy of fast chemical reactions. Optica. 2024;11(6):876.
21. Sadiek I, Fleisher AJ, Hayden J, Huang X, Hugi A, Engeln R, et al. Dual-comb spectroscopy of ammonia formation in non-thermal plasmas. Commun Chem. 2024;7(1):110.
22. Jahromi KE, Nematollahi M, Krebbers R, Abbas MA, Khodabakhsh A, Harren FJM. Fourier transform and grating-based spectroscopy with a mid-infrared supercontinuum source for trace gas detection in fruit quality monitoring. Opt Express. 2021;29(8):12381–12397.
23. Zorin I, Gattinger P, Ebner A, Brandstetter M. Advances in mid-infrared spectroscopy enabled by supercontinuum laser sources. Opt Express. 2022;30(4):5222–5254.
24. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
25. Krebbers R, Van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, et al. Ultra-broadband spectroscopy using a 2–11.5 μm IDFG-based supercontinuum source. Opt Express. 2024;32(8):14506–14520.
26. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf. 2022;277:107949.
27. Šimečková M, Jacquemart D, Rothman LS, Gamache RR, Goldman A. Einstein A-coefficients and statistical weights for molecular absorption transitions in the HITRAN database. J Quant Spectrosc Radiat Transf. 2006;98(1):130–155.
28. Butterworth TD, Amyay B, v d Bekerom D, v d Steeg A, Minea T, Gatti N, et al. Quantifying methane vibrational and rotational temperature with Raman scattering. J Quant Spectrosc Radiat Transf. 2019;236:106562.
29. Gamache RR, Roller C, Lopes E, Gordon IE, Rothman LS, Polyansky OL, et al. Total internal partition sums for 166 isotopologues of 51 molecules important in planetary atmospheres: Application to HITRAN2016 and beyond. J Quant Spectrosc Radiat Transf. 2017;203:70.
30. Shimanouchi T. Tables of Molecular Vibrational Frequencies, Consolidated Volume I. Washington: National Bureau of Standards; 1972.
31. Dang-Nhu M, Pine AS, Robiette AG. Spectral intensities in the v_3 bands of $^{12}CH_4$ and $^{13}CH_4$. J Mol Spectrosc. 1979;77(1):57-68.
32. Alboni F, Dilonardo G, Ferracuti P, Fusina L, Venuti E, Mohamed KA. Vibration-Rotation Spectra of ^{13}C-Containing Acetylene: The Stretching Fundamentals. J Mol Spectrosc. 1995;169(1):148-153.
33. Robert S, Fayt A, Di Lonardo G, Fusina L, Tamassia F, Herman M. The vibrational energy pattern in acetylene VII: $^{12}C^{13}CH_2$. J Chem Phys. 2005;123(17):174302.
34. Dedic CE, Meyer TR, Michael JB. Single-shot ultrafast coherent anti-Stokes Raman scattering of vibrational/rotational nonequilibrium. Optica. 2017;4(5):563.
35. Dakroub L, Scherman M, McGuire S, Pivard C, Dorval N, Packan D, et al. 1 kHz rate measurement of vibrational/rotational nonequilibrium using hybrid femto/picosecond coherent anti-Stokes Raman scattering in plasma. In: Proceedings Optica Sensing Congress 2024 (AIS, LACSEA, Sensors, QSM). Optica Publishing Group; 2024.
36. Menard-Bourcin F, Boursier C, Doyennette L, Menard J. Rotational and Vibrational Relaxation of Methane Excited to $2v_3$ in CH_4/H_2 and CH_4/He Mixtures at 296 and 193 K from Double-Resonance Measurements. J Phys Chem A. 2005;109(14):3111-3119.

37. Klarenaar BLM, Morillo-Candas AS, Grofulović M, van de Sanden MCM, Engeln R, Guaitella O. Excitation and relaxation of the asymmetric stretch mode of CO_2 in a pulsed glow discharge. Plasma Sources Sci Technol. 2019;28(3):035011.
38. Van Hest MFAM, De Graaf A, Van De Sanden MCM, Schram DC. Use of in situ FTIR spectroscopy and mass spectrometry in an expanding hydrocarbon plasma. Plasma Sources Sci Technol. 2000;9(4):615.
39. Stepanov S, Welzel S, Röpcke J, Meichsner J. Time resolved QCLAS measurements in pulsed cc-rf CF_4/H_2 plasmas. J Phys Conf Ser. 2009;157:1.
40. Gao Y, Neal L, Ding D, Wu W, Baroi C, Gaffney AM, et al. Recent Advances in Intensified Ethylene Production—A Review. ACS Catal. 2019;9(9):8592.
41. Fridman AA. Plasma Chemistry. Cambridge: Cambridge University Press; 2008.
42. Scapinello M, Delikonstantis E, Stefanidis GD. The panorama of plasma-assisted non-oxidative methane reforming. Chem Eng Process Intensif. 2017;117:120.

Part B

Spectroscopy applications with an IDFG-based SC source



Chapter 6

Ultra-broadband spectroscopy using a 2–11.5 µm IDFG-based supercontinuum source[1]

[1] This chapter is based on the following article:
Krebbers R, van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, Khodabakhsh, A, Cristescu, SM. Ultra-broadband spectroscopy using a 2–11.5 µm IDFG-based supercontinuum source. *Opt Express. 2024;32(8):14506–14520.*
Selected as ***Editors' pick***

Abstract

Supercontinuum sources based on intrapulse difference frequency generation (IDFG) from mode-locked lasers open new opportunities in mid-infrared gas spectroscopy. These sources provide high power and ultra-broadband spectral coverage in the molecular fingerprint region with very low relative intensity noise. Here, we demonstrate the performance of such a light source in combination with a multipass cell and a custom-built Fourier-transform spectrometer (FTS) for multispecies trace gas detection. The light source provides a low-noise, ultra-broad spectrum from 2–11.5 µm with ~3 W output power, outperforming existing mid-infrared supercontinuum sources in terms of noise, spectral coverage, and output power. This translates to an excellent match for spectroscopic applications, establishing (sub-)ppb sensitivity for molecular hydrocarbons (e.g., CH_4, C_2H_4), oxides (e.g., SO_2, NO_x), and small organic molecules (e.g., acetone, ethyl acetate) over the spectral range of the supercontinuum source with a measurement time varying from seconds to minutes. We demonstrate a practical application by measuring the off-gas composition of a bioreactor containing an acidic ammonia-oxidizing culture with the simultaneous detection of multiple nitrogen oxides (NO, NO_2, N_2O, etc.). As the different species absorb various parts of the spectrum, these results highlight the functionality of this spectroscopic system for biological and environmental applications.

6.1 Introduction

Most molecules have their strongest vibrational absorption bands with unique spectral features in the mid-infrared wavelength range (2–12 µm). This region covers both parts of the molecular fingerprint region and the functional group region. Broadband, high-resolution absorption spectroscopy in this wavelength region provides the opportunity to detect a large number of species in the gas phase simultaneously and with high detection sensitivity. Traditionally, Fourier-transform infrared spectroscopy (FTIR) setups based on thermal light sources have been used for this purpose. However, thermal sources have a low spectral brightness and an omnidirectional light beam. These properties make it difficult to produce a powerful and well-collimated beam to increase the interaction path length in the gas sample, e.g., in a multipass cell (MPC). In addition, the low spectral brightness of the thermal sources usually requires cryogenically cooled infrared detectors to reduce detector noise. Furthermore, achieving a spectrum with a high signal-to-noise ratio (SNR) and a high spectral resolution – for example 0.1 cm^{-1}, which is needed for the detection of certain atmospheric gas species – usually translates into a long averaging time on the order of tens of minutes.

Recently, broadband and ultra-broadband supercontinuum (SC) sources based on highly nonlinear fibers have been developed whose spectra are extending deep into the mid-infrared region (1,2). Their performance and usability, in combination with different types of spectrometers, have been demonstrated in infrared molecular spectroscopy on mixtures of gaseous species (3–5) with various applications (6,7). These SC sources operate by pumping highly nonlinear fibers with long ps to ns light pulses, usually generated by semiconductor lasers. Typically, a cascade of fibers made of different soft glasses is used to extend deeper into the mid-infrared wavelength range. Although these SC sources are rather inexpensive and easy to use, they demonstrate high relative intensity noise (RIN) that arises during the noise amplification in the nonlinear processes (8,9). More specifically, ultra-broad SC sources with an even greater number of cascaded nonlinear fibers demonstrate higher RIN levels that limit the SNR of the obtained spectra in standard spectroscopy systems. A major part of the RIN can be reduced by using alternative detection schemes, such as lock-in amplification (4), boxcar averaging (10), and balanced detection (5). Note that short fs light pulses, e.g. from mode-locked lasers, can also be incorporated in these types of SC sources to generate a very low noise SC by coherent spectral broadening (11,12); however, the low noise operation imposes several restrictions on the pulse peak power, pulse duration, and fiber lengths. These restrictions effectively limit the achievable bandwidth and power of the SC light generated by coherent spectral broadening from nonlinear fibers (13), making them less attractive for ultra-broadband spectroscopy.

In addition to fiber-based SC sources, high-power ultra-broadband or widely tunable broadband spectra can be generated in the mid-infrared region using mode-locked lasers and nonlinear conversion in a crystal by means of optical parametric oscillation (OPO) (14) and difference frequency generation (DFG) (15). These sources demonstrate much less RIN compared to SC sources based on highly nonlinear fibers pumped by semiconductor lasers; however, they are bulky and more expensive. For spectroscopic applications, one can use these sources in combination with a broadband spectrometer (16–20) or use two of them in a dual-comb configuration providing very short single-shot measurement times without moving parts (21–23). Traditionally, OPO-based sources demonstrate higher power than DFG-based sources, but they are more complex and need a resonance cavity.

Recently, high-power DFG- and intrapulse DFG-based (IDFG-based) systems have been developed that extend deep into the mid-infrared wavelength range (24,25) and have shown an excellent performance for molecular absorption spectroscopy. Both in traditional (linear) dual-comb spectroscopy (DCS) (26–28) and electric field sampling DCS (29,30), their potential in high resolution and precision ultra-broadband spectroscopy was demonstrated. Here, we use an IDFG-based SC source with a particularly interesting ultra-broad spectrum from 2–11.5 µm (31) in combination with our custom-built FTS (5) and a multipass cell that allows us to perform spectroscopic measurements over the full range of the source. We demonstrate the capabilities of the system for simultaneous multispecies trace gas detection and compare the performances of the IDFG-based source with the established fiber-based SC sources. We also demonstrate the possibilities for real-life applications by analyzing the gas composition in the off-gas of a bioreactor containing an acidic ammonia-oxidizing bacterial culture.

6.2 Methods

The IDFG-based SC source is shown in detail in Figure 6.1 and as part of the entire setup in Figure 6.2. It consists of a mode-locked master oscillator (MO) and a single-pass power amplifier (PA), both based on polycrystalline Cr:ZnS gain elements and optically pumped by off-the-shelf Er-doped fiber lasers (28,30,31). The output of the Cr:ZnS MOPA (~19 fs, 50 nJ pulses at 2.4 µm, repetition rate $f_r = 80$ MHz) is coupled to the IDFG stage based on a $ZnGeP_2$ (ZGP) crystal. The resulting output of the source is a ~3 W continuum spanning a wavelength range between 2–11.5 µm (see Figure 6.3), with ~300 mW in the 7–11 µm wavelength range. The main distinctive feature of the source in comparison to the established SC sources is the use of bulk polycrystalline Cr:ZnS laser medium providing the direct access to ~2-cycle pulses

with >2 MW peak power at the wavelength 2.4 µm. According to recent studies (32), these pulses are very well suited for IDFG in non-oxide nonlinear materials with broad IR transparency and high figures of merit, e.g. ZGP and GaSe. Another important feature of IDFG with the few-cycle 2.4 µm pulses is their significant augmentation and red-shift during the propagation through the nonlinear crystal. Therefore, the output SC spectrum consists of a long-wave infrared IDFG component that is merged with the augmented fundamental component. These two superimposed components have slightly different spatial distributions, beam divergences, and amplitude noises. Moreover, in their overlapping region between 4 and 7 µm, they beat at the offset frequency of the master oscillator (f_0). The frequency f_0 can in turn be set to any value $0 \leq f_0 \leq f_r/2$ via the control of the oscillator's pump power (31).

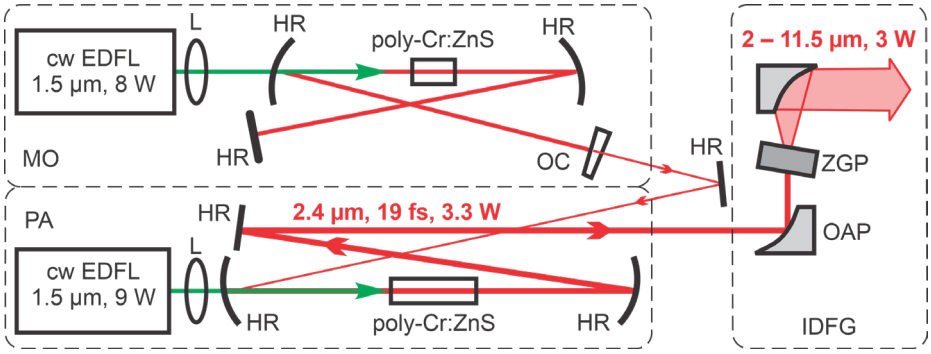


Figure 6.1: Internal structure of the SC source, EDFL: Erbium-doped fiber laser, MO: master oscillator, PA: power amplifier, IDFG: intrapulse difference frequency generation, L: lens, HR: high reflectors, Cr:ZnS: chromium-doped zinc-sulfide crystal, OC: output coupler, OAP: off-axis parabolic mirror, ZGP: zinc-germanium-phosphide crystal.

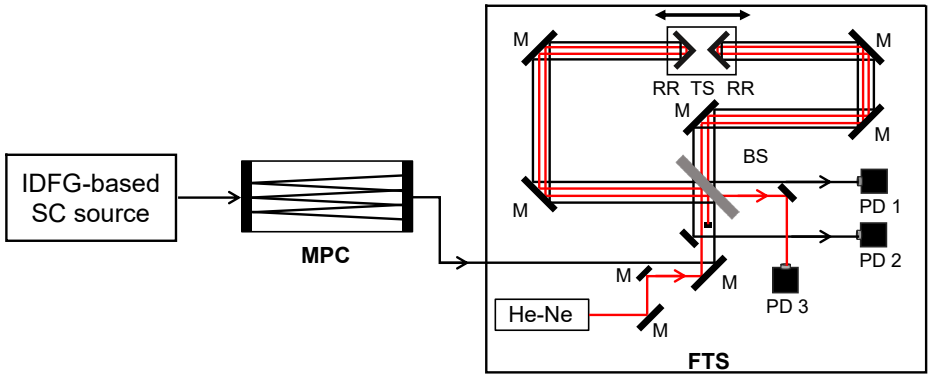


Figure 6.2: Experimental setup, consisting of the SC source, MPC, and FTS. He-Ne: helium-neon reference laser, M: mirror, BS: beam splitter, RR: retroreflector, TS: motorized linear translation stage, PD: photodetector.

The free-space output of the IDFG-based SC source is directed to a Herriott multipass cell (MPC, Thorlabs HC30 L/M-M02) with an optical path length of 31.2 m and then to a custom-built Fourier-transform spectrometer (FTS), as illustrated in Figure 6.2. The FTS design has been described in our previous work (5). The CaF_2 windows of the MPC were replaced by uncoated ZnSe windows to be able to transmit light over the whole spectral range of the source. The FTS is based on a Michelson interferometer, in which the incoming light is split into two arms by a beamsplitter (BS, Thorlabs BSW711). The beam in each arm is reflected by a cubic retroreflector that is mounted back-to-back on the same motorized linear translation stage (TS, Thorlabs DDSM100/M) to double the optical path difference (OPD). In this way, moving the stage over 2.5 cm results in an OPD of 10 cm between the two beams, which corresponds to a spectral resolution of 0.1 cm^{-1} (or 3 GHz). The reflected beams recombine on the beamsplitter and propagate to two photovoltaic detectors (PD 1 and PD 2 in Figure 6.2, Vigo Photonics PVI-4TE-10.6), which have a spectral response between 3 and 11.5 µm. Two pinholes in front of the detectors were used to prevent saturation. Using balanced detection, the two signals are subtracted by a differential amplifier to decrease the RIN of the SC source. The obtained stream of data is then digitized with 16-bit resolution using a data acquisition card (National Instruments PCIe-6361). To track the OPD during the measurement, the output beam from a stable reference He-Ne laser is sent along the beam path of the SC source through the FTS and is detected by a separate photodetector (PD 3 in Figure 6.2, Thorlabs PDA8A2). The interferogram of the reference He-Ne laser is used to calibrate the OPD. In general, the similarity of the spectrometer used in this work with the one from our previous work (5) allows an objective comparison between the performances of different SC sources and their potential for spectroscopic applications.

During the measurements, the gas flow towards the MPC was controlled using mass flow controllers (Bronkhorst EL-FLOW FG-201CV and F-201CV) and the pressure was regulated to 1 bar using a back-pressure regulator (Bronkhorst EL-PRESS P-702CV). The FTS was set to a spectral resolution of 0.1 cm^{-1} (3 GHz), which is sufficient for resolving spectral features of the small gaseous molecules at atmospheric pressure (1 bar) (33). Background spectra to determine the baseline laser intensity (as presented in Figure 6.3) were measured by flushing the MPC with nitrogen gas. Reference spectra are fitted to the acquired absorbance spectra using a least-square fit with simulated spectra from the HITRAN (34) or PNNL (35) databases. For the HITRAN spectra, we used a Voigt profile to account for Doppler and pressure broadening, and a sinc function convolution to account for the instrumental line shape due to the finite OPD window of the FTS (boxcar apodization). In the fit, baseline offsets were compensated using a constant offset (DC-offset), or up to a 3^{rd}-order polynomial offset in the case of severe baseline drifts.

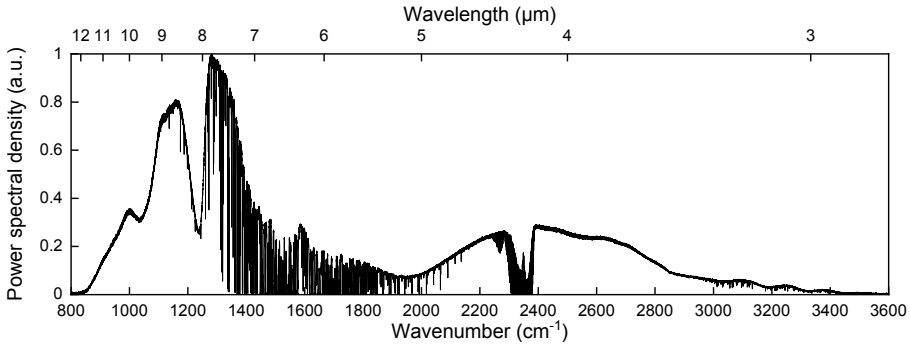


Figure 6.3: Power spectral density of the IDFG-based SC source as measured by the FTS at 3 GHz spectral resolution. The dips shown in the spectrum are absorption lines from water vapor (~5 –8 µm, and ~3.3 µm) or carbon dioxide (~4.3 µm) in the surrounding air.

We demonstrated the capability of the system with the IDFG-based SC source by detecting various nitrogen-containing compounds including NO_x species, as well as CO_2, in a biological application. Gas samples were taken from the gas outflow of a bioreactor containing an enrichment culture of an acid-tolerant ammonia-oxidizing bacterium (36). The bioreactor was supplied with a mineral medium at a constant flow rate of 1 L day^{-1}, consisting of 70 mM NH_4Cl, 1 mM NaCl, 0.8 mM $MgSO_4 \cdot 7\ H_2O$, 1 mM K_2HPO_4, 0.3 mM $CaCl_2 \cdot 2\ H_2O$, and 1 mL each of trace element solutions I and II (36). The bioreactor was operated at room temperature and stirred at 850 rpm; pH 4 was maintained by dosing of 1 M $KHCO_3$ solution via an automatically controlled pump. The working volume of the bioreactor was maintained at 2.2 L using a level sensor controller, with biomass retention via a membrane filter. In addition, biomass was removed continuously at a rate of 0.2 L day^{-1}. The bioreactor was flushed with a mixture of 2% CO_2 in air at a flow rate of 300 mL min^{-1}. For gas analysis, the off-gas was captured in a 3 L Tedlar bag and transferred to the MPC.

6.3 Results

The unique combination of a high power, ultra-broadband IDFG-based SC source with low RIN makes it very well suited for sensitive multispecies trace gas detection. Firstly, we measured individual gas species, such as CO_2 around 4.8–5.0 µm and 10.1–11.0 µm, and SO_2 around 7–7.5 µm, next to a complex gas mixture of acetone, methanol, ethyl acetate, and ethylene between 7–11.5 µm, to demonstrate the ability of the system to retrieve the concentrations of species with both narrow and broadband absorption features in different spectral regions. Secondly, the linear response curve and detection limits have been determined for ethylene at 10–11 µm.

Thirdly, to characterize the setup and quantify its performance compared to other SC-based FTS setups, we have assessed the sensitivity of the setup for different gases in various parts of the spectrum using Allan-Werle deviation analysis. The noise equivalent absorption sensitivity (NEAS) of the system was characterized. Finally, to demonstrate the applicability of the system, a gas sample from a bioreactor containing highly enriched ammonia-oxidizing bacteria was analyzed.

6.3.1 Single and multispecies gas detection

To demonstrate the performance of the system, we measured the spectrum of CO_2, consisting of individual absorption lines, and that of SO_2, which has a broadband absorption spectrum with many overlapping lines. The absorbance spectrum of a certified mixture of 4.9 ± 0.1% CO_2 in synthetic air (Linde Gas) was measured simultaneously in two wavelength regions, from 4.8–5.0 µm and 10.1–11.0 µm, and averaged for 100 seconds. The result is shown in Figure 6.4 along with the HITRAN-based fitted spectrum of CO_2. The retrieved concentration of the fit is 5.04 ± 0.07%, where the uncertainty is determined by the standard deviation of 5 consecutive measurements. Points with saturated absorption were not included in the fit. Apart from some remnants of very strong lines of atmospheric water vapor removed during the fitting procedure, no residual CO_2 features can be detected in the residual of the fit, indicating an accurate fit. It should be noted that the absorbance strength of CO_2-lines at 10.4 µm is an order of magnitude weaker compared to those around 4.8 µm. This poses no challenge to the system, which can detect both CO_2 bands simultaneously.

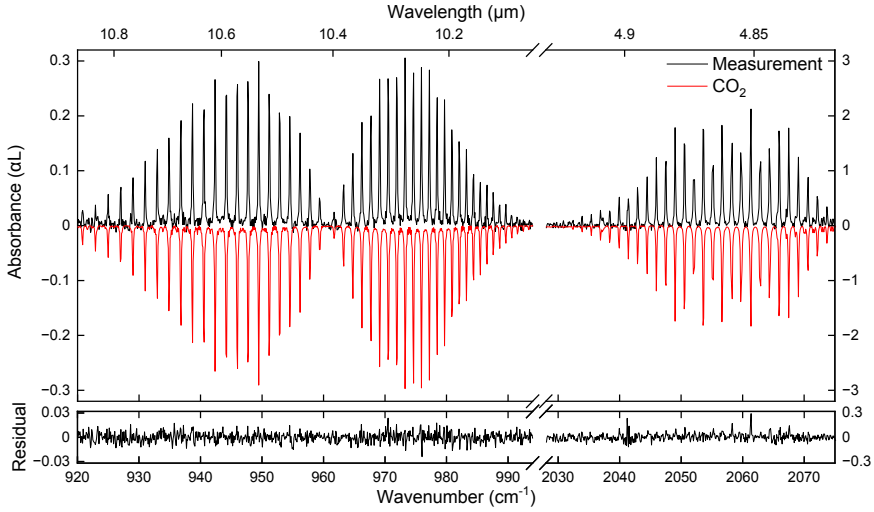


Figure 6.4: Absorbance spectrum of 4.9 ± 0.1 % CO_2 in N_2 (in black) measured in 100 seconds, with the simulated spectrum using HITRAN of CO_2 (red, shown inverted). The residual of the fit is shown in the bottom.

The spectrum of a mixture of 32.6 ± 0.7 ppm SO_2 in N_2 (Linde Gas) was measured at 7–7.5 µm in 28 seconds. To optimize the spectral power in these wavelengths, a 6.7 µm optical longpass filter was inserted in the optical beam path. The obtained absorbance spectrum is shown in Figure 6.5 along with the HITRAN-simulated SO_2 spectrum. A concentration of 32.1 ± 0.12 ppm was obtained, with the uncertainty based on the standard deviation of 10 consecutive measurements. The small features remaining in the residual can be attributed to atmospheric water absorbance lines which are close to 100% absorption in both the measurement and background spectrum (i.e. the spectrum without the SO_2 mixture). Saturated absorption line profiles are difficult to model and fit properly in the spectrum. Since the remainder of the residual is largely flat, it follows that SO_2 can be accurately fitted to the simulated spectrum. The spectrum of SO_2 in this spectral region has been investigated previously by a state-of-the-art chalcogenide-based SC source (5).

The sensitivity of a system was assessed by considering the standard deviation (σ) of a noise-dominated (thus flat) part of the residual (marked grey in Figure 6.5) and normalizing it to the measurement time (T): Sensitivity = $\sigma\sqrt{2T}$.

It results that the sensitivity of the system using the IDFG-based SC source for SO_2 detection is 0.032 $Hz^{-1/2}$. As the sensitivity of the chalcogenide-based SC source in an identical setup with MPC and FTS was 7.8 $Hz^{-1/2}$ (5), a comparison between the two systems indicates an improvement in sensitivity for this spectral region of over 240x in favor of the IDFG-based SC source.

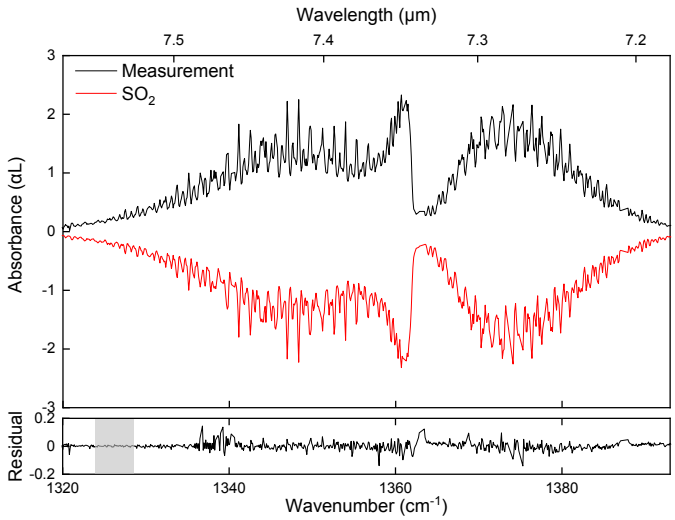


Figure 6.5: Measured absorbance spectrum of 32.6 ± 0.7 ppm SO_2 in N_2 (in black) with the simulated spectrum containing SO_2 (in red, inverted). Water lines have been removed for clarity. Marked in grey is the noise dominated part of the residual used for sensitivity calculations.

The strength of covering such a broad wavelength range between 7–11 µm is the ability to detect multiple gases simultaneously, especially when these gases exhibit broad and overlapping absorbance features. To demonstrate this, we used mass flow controllers (Bronkhorst EL-FLOW FG-201CV and F-201CV) to generate a dynamic mixture of acetone (20.0 ± 1.0 ppm), ethyl acetate (19.0 ± 0.4 ppm), ethylene (19.6 ± 0.4 ppm), and methanol (19.6 ± 0.4 ppm) in N_2. To retrieve the concentration from the spectrum, reference spectra were taken from the HITRAN (for ethylene) and PNNL database (for the other gases). The measured and calculated spectra are shown in Figure 6.6. The retrieved concentrations from the fit are 20.1 ± 0.11 ppm ethyl acetate, 20.2 ± 0.6 ppm ethylene, 20.1 ± 0.8 ppm methanol, and 17.2 ± 0.3 ppm acetone, in which the uncertainty is given by the standard deviation of 10 consecutive measurements (each spectrum averaged for 190 seconds). In Figure 6.6, it can be clearly noticed that the very broad absorption bands of ethyl acetate overlap significantly with the absorption band of methanol and those of acetone, but thanks to the wide spectral range it is still possible to determine the concentration of all gas species, even though the measured acetone concentration is lower than expected. Since acetone is known to have strong wall adsorption properties (being a so-called "sticky gas molecule") and the fit of acetone leaves no visible remaining features in the residual, we are confident that the retrieved concentration of acetone is accurate and the lower concentration measured is due to the partial condensation of acetone on the tubing or walls of the MPC.

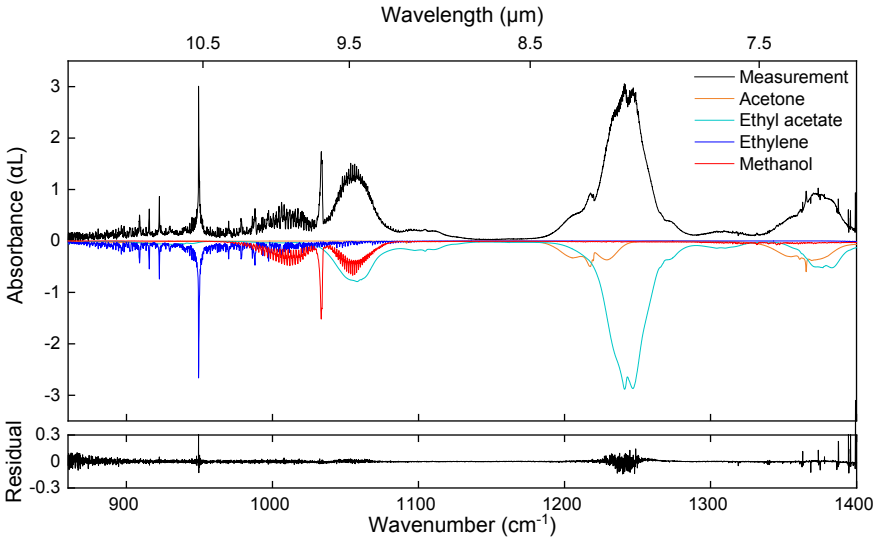


Figure 6.6: Measured absorbance spectrum of a mixture of acetone, ethyl acetate, ethylene, and methanol (each around 20 ppm) in N_2 (in black) and corresponding reference spectra from HITRAN (ethylene) or PNNL (others) (in color, shown inverted).

Furthermore, the small visible peaks in the residual between 1350 and 1400 cm^{-1} can be attributed to water vapor, whose concentration seemingly varied in the atmosphere during the measurements. The increased noise visible in the residual at the peak intensity of ethyl acetate (1230–1250 cm^{-1}) and the Q-branch of ethylene (950 cm^{-1}) is caused by saturation of these strong absorbances. However, this does not affect the ability to retrieve accurate concentrations of the species, thanks to the broadband nature of the source that allows to measure the weaker absorption bands of these species as well.

6.3.2 System linearity, detection sensitivity and long-term stability

The dynamic range of the system can be demonstrated by characterizing the linear response of the system with dilution series of a gas. Ethylene (C_2H_4) is chosen because of its specific absorption spectrum at the end of the spectrum of the SC source at 10 µm. The concentration was varied between 25 ppb and 1 ppm using dilutions from a certified and calibrated mixture of 1.04 ± 0.05 ppm ethylene in N_2 (Linde Gas). The measured concentrations are shown for each of the applied concentrations in Figure 6.7(a), with the uncertainty in dilution resulting from the MFCs (which have ±0.5% reading and ±0.1% full scale uncertainty), and the measurement uncertainty resulting from the standard deviation of 8 consecutive measurements (each averaged for 190 seconds). In Figure 6.7(b), the absorbance spectrum of 1 ppm C_2H_4 is shown, with the corresponding HITRAN spectrum of C_2H_4 in red (inverted). The excellent match between measurement and simulation is shown by the flat and low noise in the residual. A linear fit yields an R^2-value of 0.9994. To demonstrate that the linearity and dynamic response of the system extends even further to almost 5 orders of magnitude, two dilutions of 49 and 98 ppm ethylene were measured and averaged for 190 seconds from a certified, calibrated mixture of 98 ± 2 ppm ethylene in N_2 (Linde Gas) and shown in Figure 6.7(a).

To determine the long-term stability of the system, the optimal averaging time, and the detection sensitivity, we have assessed the Allan-Werle deviation in two different spectral regions of the source. The first region of interest is around 3–3.5 µm, where most hydrocarbons have strong absorbance bands related to their C-H stretching vibrational modes. Moreover, it is a wavelength range covered by commercially available ZBLAN fiber-based SC sources (10). Therefore, we made a comparison between the detection sensitivity of the system equipped with the IDFG-based SC source, and with a commercial (ZBLAN) fiber-based SC source (37). The second region of interest is at 10–11 µm, as it is an interesting part of the molecular fingerprint region with strong distinctive absorption bands of many molecules. Moreover, to this date, this spectral region has been unexplored for spectroscopic research with SC sources.

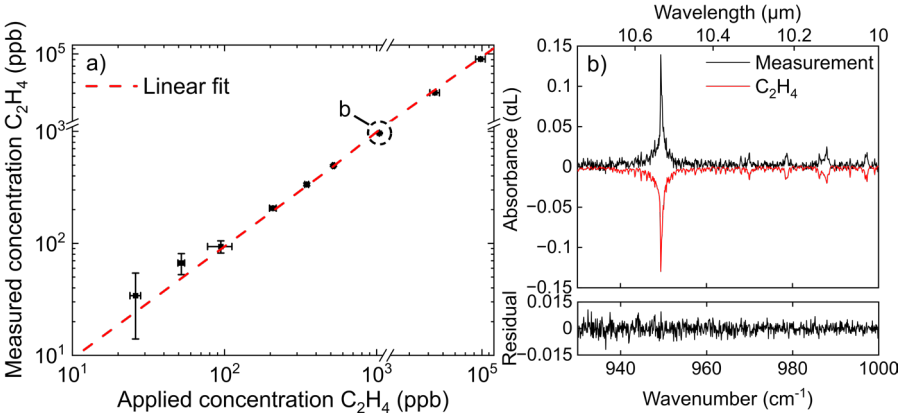


Figure 6.7: a) Linear response of the system to a dilution series of C_2H_4 (ethylene). The retrieved linear fit on the data points between 25 and 1000 ppb (shown in red) yields an R^2-value of 0.9994, **b)** The measured absorbance spectrum of 1 ppm C_2H_4 (in black) with corresponding reference spectrum from HITRAN (shown in red, inverted) and residual of the fit.

The Allan-Werle plots in Figure 6.8 were obtained by measuring spectra of pure nitrogen gas for several hours. Noise equivalent absorbance spectra were obtained by using the first measured spectrum of nitrogen as background for the rest of the measurements. In the case of the obtained Allan-Werle deviation curve near 3 μm (Figure 6.8(a)), a 3–3.5 μm bandpass filter (Thorlabs FB3250-500) was placed in the beam path to enhance the spectral brightness for this wavelength region (similar to (37)), and methane was fitted to these spectra in order to acquire noise equivalent concentrations (NECs) for methane in this spectral window. For the long-wavelength range of the IDFG-based SC source, we repeated a similar experiment with an 8–12 μm bandpass filter (Edmund Optics 11-999) and fitted ammonia in the spectral window 900–1250 cm^{-1} (8–11 μm) to determine the NEC of ammonia, resulting in the Allan-Werle plot shown in Figure 6.8(b).

The white noise can be reduced over 10^3 seconds as it follows a $\tau^{-1/2}$-dependency, yielding an optimum averaging time of ~1000 seconds with a detection sensitivity of ~200 ppt for methane. This is a 60-fold improvement over commercial ZBLAN fiber-based SC sources (37), which is a remarkable result, as that the main strength of the IDFG-based SC source is its much broader spectrum extending further in the mid-infrared wavelength region. The Allan-Werle deviation curve of ammonia starts deviating from the $\tau^{-1/2}$-dependency after about 100 seconds, indicating that the stability of the source at the extreme end of its long-wavelength emission is suffering slightly more from long-term drifts. However, a detection limit of 6.2 ppb ammonia in 500 seconds can be achieved.

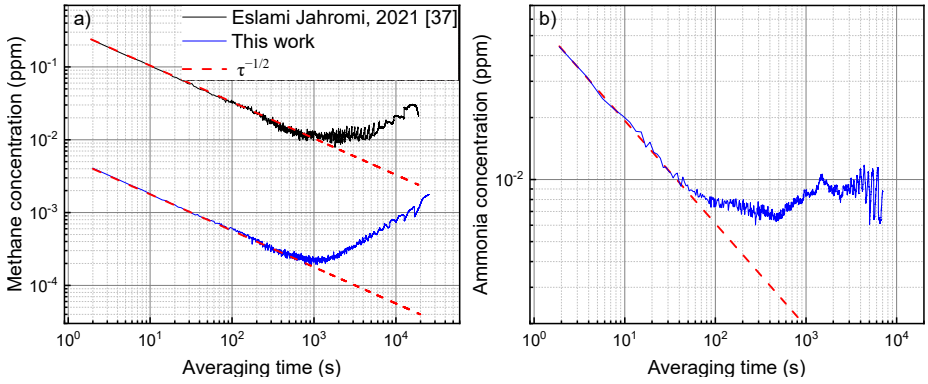


Figure 6.8: Allan-Werle plot of the NEC of **a)** methane around 3 – 3.5 µm for this work (in blue) and previous work (in black) (37), and **b)** ammonia around 8 – 11 µm, with the $\tau^{-1/2}$-dependency of white noise (dashed line).

6.3.3 Noise-equivalent absorption sensitivity (NEAS)

An elegant way to compare the performance of the system with the IDFG-based SC source to systems with other SC sources is the noise-equivalent absorption sensitivity (NEAS). The NEAS is defined as

$$\mathrm{NEAS} = \frac{\sigma\sqrt{2T}}{L\sqrt{M}}, \tag{6.3}$$

in which σ is the standard deviation of the noise in a normalized absorption spectrum, T the measurement time of this spectrum (e.g., 2 seconds for the sample measurement plus 2 seconds for the background measurement), L the effective interaction path length (or optical path length, 31.2 m) and M the number of spectral elements (given by the width of the obtained spectrum divided by the resolution of the spectrometer). The noise behavior of the IDFG-based SC source is not identical over its full spectral range. As mentioned in the Methods section, the full spectral coverage of the source is a combination of two co-propagating, superimposed beams. The first beam covers a spectral range between 2–7 µm, while the second beam covers a range between 4–11.5 µm, with the two beams having a spectral overlap in the 4–7 µm region. As the two beams exhibit different RIN levels, two different figures for the NEAS can be calculated: one covering the entire spectrum with the higher associated RIN, and another one in which only the spectrum in the longer wavelengths is considered with the lower associated RIN. For the entire spectrum, which has 27500 spectral elements (2750 cm^{-1} / 0.1 cm^{-1}) and a σ of $5.1 \cdot 10^{-2}$, the NEAS was found to be $2.7 \cdot 10^{-7}$ cm^{-1} $Hz^{-1/2}$ per spectral

element. For the long wavelength range (using a 6.7 µm longpass filter, providing 6500 spectral elements and a σ of $9.1 \cdot 10^{-3}$), the NEAS was calculated to be $9.8 \cdot 10^{-8}$ cm^{-1} $Hz^{-1/2}$ per spectral element.

In comparison to other systems (5), the current system with IDFG-based SC source outperforms the other SC-based FTS systems (NEAS of $9.9 \cdot 10^{-7}$ and $5.3 \cdot 10^{-6}$ cm^{-1} $Hz^{-1/2}$ per spectral element for short and long wave mid-infrared wavelength range, respectively) by one to two orders of magnitudes. Moreover, the performance aligns with other fs-laser-based systems, with the clear advantage that the IDFG-based system presented in this study can perform over a much wider spectral range.

6.3.4 Application: bioreactor off-gas samples

To demonstrate the applicability of the system, we analyzed the off-gas from a bioreactor containing an enrichment culture of acid-tolerant ammonia-oxidizing bacteria (36). The reactor content was constantly stirred and purged with a mixture of 2% CO_2 in air at a flow rate of 300 mL min^{-1}. In this study, the off-gas of the bioreactor was collected and measured to test for nitrogen oxides.

During canonical ammonia oxidation under neutral to slightly alkaline conditions, NH_3 is oxidized to NO_2^- via the intermediates NH_2OH and NO (38). The exact physiology of ammonia-oxidizing microorganisms at acidic pH remains uncertain, partially due to the chemical instability of NO_2^- and NO, especially at low pH. Consequently, biotic and abiotic reactions occur simultaneously in the bioreactor, and a wide variety of nitrogen species are expected to be present. This bioreactor setup is therefore of particular interest to test the IDFG-based system. At acidic conditions, NO_2^- protonates partially to HNO_2, which then chemically degrades mostly into volatile compounds such as NO, NO_2, and N_2O_3 (39). In addition, N_2O can be produced both biotically and abiotically in this system (40). Simultaneous observations of multiple nitrogen oxides such as NO_x and N_2O are challenging, given their complexity to be measured by mass spectrometric methods, as well as that their absorption spectra are in different parts of the mid-infrared spectrum. For example, the main absorption band of NO_2 is between 6–6.4 µm, NO between 5.1–5.6 µm, N_2O between 4.4–4.6 µm, and HNO_2 between 7.7–8.3 µm and 5.7–6.3 µm (both bands being about equally strong). Given the strong water absorption between 6–7 µm, the NO_x bands in this spectral window are not useful, and consequently, the weaker NO_2 absorption band around 3.4–3.5 µm had to be used. The detection of these nitrogen oxides (including NO_x, N_2O, and HNO_2) would thus either require several laser sources or a source with at least a spectral bandwidth of 3.4–8.3 µm. Therefore, the broad spectral coverage of the IDFG-based SC source opens new possibilities to detect and quantify the major NO_x species simultaneously.

As shown in Figure 6.9, in a single measurement (averaged for 190 seconds), absorption bands of NO, NO_2, N_2O, HNO_2, and CO_2, could be detected. The top panel contains the full coverage of the power spectral density obtained from the measurement, of which only the narrow parts with a white background were used to retrieve the concentration of the six species of interest. Even though the full absorbance bands of the species of interest can be observed, the high spectral resolution allows us to zoom into specific parts of the spectrum with maximum absorbance strength and minimal interference with, e.g., strongly absorbing water lines. Consequently, the regions shown in the absorbance spectrum (middle panel) were used for retrieving the concentrations of the species of interest (Figure 6.9). Fitting these spectra to simulations using the HITRAN database (or PNNL in the case of HNO_2) yields concentrations of 89 ± 9 ppm NO, 270 ± 40 ppm NO_2, 11.2 ± 0.10 ppm N_2O, 357 ± 2 ppm HNO_2, and 1.34 ± 0.02% CO_2, in which the uncertainty is given by the standard deviation of 4 consecutive measurements.

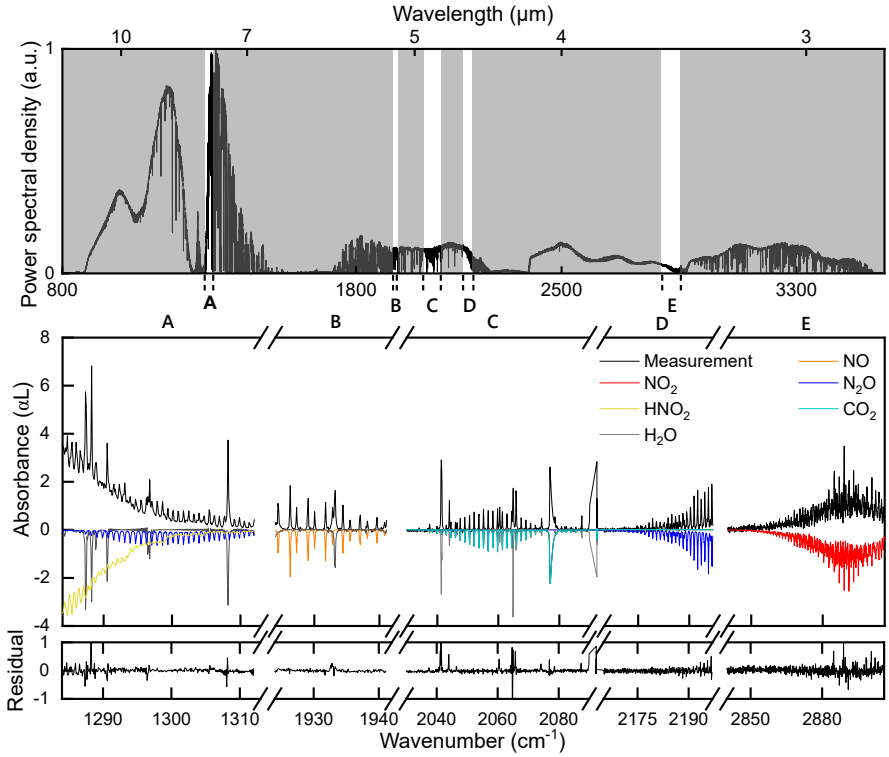


Figure 6.9: Measured power spectral density (top) and absorbance spectrum (middle) of a bioreactor off-gas sample from an acidic ammonia-oxidizing enrichment culture (in black). The absorbance spectrum is zoomed to the regions used for fitting gas species concentrations and fitted reference spectra from PNNL (HNO_2) and HITRAN (others) are shown in color and inverted. The residual of the fit is shown in the bottom.

The ability to simultaneously observe and quantify all these nitrogen-containing species offers the possibility of constructing nitrogen mass balances of the pathways converting ammonia under acidic conditions. In contrast to previous methods (41–43) targeting only a few of these molecules mainly by colorimetric assays, gas and ion chromatography, the method described here is a major step towards establishing a completely closed mass balance.

6.4 Discussion

The IDFG-based SC source in combination with FTS shows great possibilities for sensitive, multispecies gas detection. However, there are some technical considerations which need further discussion. As mentioned before, the output of the source consists of a combination of two co-propagating, superimposed beams that each cover a different part of the spectrum with an overlap in the middle (4–7 µm). Previous works (31,32) describe how those two beams are resulting from the interaction within the ZGP crystal, where the broadened and red-shifted fundamental component stretches up to 7 µm and the IDFG-generated component spans from 4–11.5 µm. Performing dual-comb spectroscopy with two IDFG-based SC sources over the entire spectral coverage of the source is rather complicated since the carrier-envelope offset frequencies of these two components of the beam are different. Therefore, the demonstrations of dual-comb spectroscopy using two of these sources have been limited to a spectral coverage between ~6.6 and 11.5 µm, which is the spectral range that is solely generated by the IDFG process (28,30), limiting the capability of the detection system. Here, we have shown that it is possible to use both components in the beam simultaneously for spectroscopic applications using an FTS, yielding – to the best of our knowledge – an unprecedented spectral coverage from 3–11.5 µm for sensitive and simultaneous multi-species trace gas detection based on SC sources.

However, as explained in the Methods section, these two superimposed components of the beam have slightly different spatial distributions and beam divergences although the two beams are co-propagating. As a result, only one of these two components can be properly mode matched with the correct entrance angle to the MPC and achieve the nominal effective path length, while the other component will be lost (cut off in the cell). During the experiments, where we used a bandpass or a longpass filter and only transmitted one of the two beam components, it was possible to achieve the maximum effective path length of the MPC (32.1 m). However, by removing the filter the maximum achievable path length for both beam components, without one of them being cut off, was shorter. We measured and calibrated this path length by using an independently certified

CO_2/N_2 mixture in the MPC that absorbs in both shorter and longer wavelengths, where the two beam components cover separately. This calibration was performed prior to the reported experiments in this manuscript. Despite being shorter than the nominal interaction length, the path lengths for both beams agreed over the entire spectrum to be 25.6 m. This is the interaction length that we used to retrieve concentration of CO_2 in the two spectral ranges shown in Figure 6.4. The featureless residual of the fit and the unique retrieved concentration shows that the interaction length was properly calibrated.

Replacing the MPC with a resonant cavity would be an option to further enhance the sensitivity of the system, considering the recent developments in the ultralow loss mid-infrared crystalline mirrors (44). However, it is not possible to use enhancement cavities for the entire wavelength range simultaneously, because the HR-coatings of the cavity mirrors have a limited bandwidth. Moreover, the IDFG process removes the carrier-envelope offset frequency, while this is a necessary tuning parameter when optimizing the coupling to the cavity (in the tight-locking scheme), given that the dispersion in the cavity leads to changes in the regularity of the free spectral range in the cavity over the spectrum. In addition, the high spectral power of the IDFG-based SC source makes it possible to use MPCs with a higher number of reflections and still have enough power on the detectors to maintain a good SNR. This can increase the detection sensitivity without sacrificing the spectral bandwidth of the source.

6.5 Conclusion

The development of IDFG-based SC sources opens new opportunities in mid-infrared spectroscopy. In combination with an MPC and a custom-built FTS, the IDFG-based SC source presented here proves to be an excellent tool for multispecies gas detection. In comparison to existing mid-infrared SC sources, the current system provides one to two orders of magnitude better NEAS. Moreover, the IDFG-based SC source extends much further into the IR spectrum, with a major part of its spectral power in the 7–11.5 µm wavelength region, thereby accessing the highly interesting molecular fingerprint region. Although there are existing DFG- and OPO-based systems that also access (a part of) this spectral window (37), the immense wavelength range of 2–11.5 µm proves to be an interesting advantage, which to the best of our knowledge, has not been demonstrated yet for SC-based spectroscopy.

Characterizing the performance of the source for various gaseous molecules, (sub-) ppb levels of sensitivity are achieved for hydrocarbons (e.g., CH_4, C_2H_4), oxides (e.g., SO_2, NO_x), and small organic molecules (e.g., acetone, ethyl acetate) in measurement times of the order of seconds to a few minutes. In comparison to the fiber-based SC sources, the system equipped with an IDFG-based SC source proved to be 60 times more sensitive than using a ZBLAN-based SC source (for CH_4 around 3 μm) and 240 times more sensitive than using a chalcogenide-based SC source (for SO_2 around 7 μm).

In a real-life application, various nitrogen species including nitrogen oxides were measured simultaneously in the off-gas of a bioreactor demonstrating the advantage of a wide spectral range of the system. Further applications of such a system include the analysis of notoriously complex gas mixtures, such as gas phase chemical conversions in electrical discharges or plasmas (7), or human breath, which contains different compounds in ppb- to ppm-level concentrations while the spectrum is significantly obscured by highly absorbing spectral lines of CO_2 and H_2O (45–48). Finally, an interesting application for this source is open-path spectroscopy for environmental monitoring, where the combination of the broad spectral coverage and spatial coherence of the source could be an advantage over the current state-of-the-art dual-comb systems used to monitor multiple greenhouse gases or pollutants in larger outdoor areas (49–55).

References

1. Petersen CR, Moselund PM, Huot L, Hooper L, Bang O. Toward a table-top synchrotron based on supercontinuum generation. Infrared Phys Technol. 2018;91:182–186.
2. Petersen CR, Møller U, Kubat I, Zhou B, Dupont S, Ramsay J, et al. Mid-infrared supercontinuum covering the 1.4–13.3 µm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre. Nat Photonics. 2014;8(11):830–834.
3. Jahromi KE, Pan Q, Høgstedt L, Friis SMM, Khodabakhsh A, Moselund PM, et al. Mid-infrared supercontinuum-based upconversion detection for trace gas sensing. Opt Express. 2019;27(17):24469.
4. Jahromi KE, Nematollahi M, Pan Q, Abbas MA, Cristescu SM, Harren FJM,et al. Sensitive multi-species trace gas sensor based on a high repetition rate mid-infrared supercontinuum source. Opt Express. 2020;28(18):26091.
5. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
6. Jahromi KE, Pan Q, Khodabakhsh A, Sikkens C, Assman P, Cristescu SM, et al. A broadband mid-infrared trace gas sensor using supercontinuum light source: applications for real-time quality control for fruit storage. Sensors. 2019;19(10):2334.
7. Krebbers R, Liu N, Jahromi KE, Nematollahi M, Bang O, Woyessa G, et al. Mid-infrared supercontinuum-based Fourier transform spectroscopy for plasma analysis. Sci Rep. 2022;12(1):9642.
8. Newbury NR, Washburn BR, Corwin KL, Windeler RS. Noise amplification during supercontinuum generation in microstructure fiber. Opt Lett. 2003(11);28:944.
9. Møller U, Sørensen ST, Jakobsen C, Johansen J, Moselund PM, Thomsen CL, et al. Power dependence of supercontinuum noise in uniform and tapered PCFs. Opt Express. 2012;20(3):2851.
10. Zorin I, Gattinger P, Ebner A, Brandstetter M. Advances in mid-infrared spectroscopy enabled by supercontinuum laser sources. Opt Express. 2022;30(4):5222–5254.
11. Heidt AM, Hartung A, Bosman GW, Krok P, Rohwer EG, Schwoerer H, et al. Coherent octave spanning near-infrared and visible supercontinuum generation in all-normal dispersion photonic crystal fibers. Opt Express. 2011;19(4):3775.
12. Klimczak M, Soboń G, Kasztelanic R, Abramski KM, Buczyński R. Direct comparison of shot-to-shot noise performance of all normal dispersion and anomalous dispersion supercontinuum pumped with sub-picosecond pulse fiber-based laser. Sci Rep. 2016;6(1):19284.
13. Rao DS, Jensen M, Grüner-Nielsen L, Olsen JT, Heiduschka P, Kemper B, et al. Shot-noise limited, supercontinuum-based optical coherence tomography. Light Sci Appl. 2021;10(1):133.
14. Vodopyanov KL. Ultra-broadband mid-infrared frequency combs produced by optical subharmonic generation. Quantum Electron. 2022;52(4):307–312.
15. Krzempek K, Tomaszewska D, Głuszek A, Martynkien T, Mergo P, Sotor J, et al. Stabilized all-fiber source for generation of tunable broadband fCEO-free mid-IR frequency comb in the 7–9 µm range. Opt Express. 2019;27(26):37435.
16. Adler F, Masłowski P, Foltynowicz A, Cossel KC, Briles TC, Hartl I, et al. Mid-infrared Fourier transform spectroscopy with a broadband frequency comb. Opt Express. 2010;18(21):21861.
17. Khodabakhsh A, Ramaiah-Badarla V, Rutkowski L, Johansson AC, Lee KF, Jiang J, et al. Fourier transform and Vernier spectroscopy using an optical frequency comb at 3–54 µm. Opt Lett. 2016;41(11):2541.
18. Fleisher AJ, Bjork BJ, Bui TQ, Cossel KC, Okumura M, Ye J. Mid-infrared time-resolved frequency comb spectroscopy of transient free radicals. J Phys Chem Lett. 2014;5(13):2241–2246.

19. Hjältén A, Germann M, Krzempek K, Hudzikowski A, Głuszek A, Tomaszewska D, et al. Optical frequency comb Fourier transform spectroscopy of $^{14}N_2{}^{16}O$ at 7.8 µm. J Quant Spectrosc Radiat Transf. 2021;271:107734.
20. Iwakuni K, Bui TQ, Niedermeyer JF, Sukegawa T, Ye J. Comb-resolved spectroscopy with immersion grating in long-wave infrared. Opt Express. 2019;27(3):1911.
21. Muraviev AV, Smolski VO, Loparo ZE, Vodopyanov KL. Massively parallel sensing of trace molecules and their isotopologues with broadband subharmonic mid-infrared frequency combs. Nat Photonics. 2018;12(4):209–214.
22. Ycas G, Giorgetta FR, Baumann E, Coddington I, Herman D, Diddams SA, et al. High-coherence mid-infrared dual-comb spectroscopy spanning 2.6 to 5.2 µm. Nat Photonics. 2018;12(4):202–208.
23. Abbas MA, Pan Q, Mandon J, Cristescu SM, Harren FJM, Khodabakhsh A. Time-resolved mid-infrared dual-comb spectroscopy. Sci Rep. 2019;9(1):17247.
24. Gaida C, Gebhardt M, Heuermann T, Stutzki F, Jauregui C, Antonio-Lopez J, et al. Watt-scale super-octave mid-infrared intrapulse difference frequency generation. Light Sci Appl. 2018;7(1):94.
25. Vasilyev S, Moskalev I, Smolski V, Peppers J, Mirov M, Muraviev A, et al. Multi-octave visible to long-wave IR femtosecond continuum generated in Cr:ZnS-GaSe tandem. Opt Express. 2019;27(11):16405.
26. Timmers H, Kowligy A, Lind A, Cruz FC, Nader N, Silfies M, et al. Molecular fingerprinting with bright, broadband infrared frequency combs. Optica. 2018;5(6):727.
27. Lind AJ, Kowligy A, Timmers H, Cruz FC, Nader N, Silfies M, et al. Mid-infrared frequency comb generation and spectroscopy with few-cycle pulses and $\chi^{(2)}$ nonlinear optics. Phys Rev Lett. 2020;124(13):133904.
28. Vasilyev S, Muraviev A, Konnov D, Mirov M, Smolski V, Moskalev I, et al. Longwave infrared (6.6–11.4 µm) dual-comb spectroscopy with 240,000 comb-mode-resolved data points at video rate. Opt Lett. 2023;48(9):2273.
29. Kowligy AS, Timmers H, Lind AJ, Elu U, Cruz FC, Schunemann PG, et al. Infrared electric field sampled frequency comb spectroscopy. Sci Adv. 2019;5(6):eaaw8794.
30. Konnov D, Muraviev A, Vasilyev S, Vodopyanov K. High-resolution frequency-comb spectroscopy with electro-optic sampling and instantaneous octave-wide coverage across mid-IR to THz at a video rate. APL Photonics. 2023;8(11):110801.
31. Vasilyev S, Smolski V, Peppers J, Moskalev I, Mirov M, Barnakov Y, et al. Middle-IR frequency comb based on Cr:ZnS laser. Opt Express. 2019;27(24):35079.
32. Vasilyev S, Smolski V, Moskalev I, Peppers J, Mirov M, Barnakov Y, et al. Multi-octave infrared femtosecond continuum generation in Cr:ZnS-GaSe and Cr:ZnS-ZGP tandems. In: Schunemann PG, Schepler KL, editors. Nonlinear Frequency Generation and Conversion: Materials and Devices XIX. San Francisco, United States: SPIE; 2020. p. 6.
33. Kim J, Lee S, Lee W, Lee J. Design optimization and implementation of a Fourier transform spectrometer with rotating motion for 0.1 cm^{-1} resolution spectroscopy. Opt Express. 2023;31(20):33041.
34. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transfer. 2022;277:107949.
35. Sharpe SW, Johnson TJ, Sams RL, Chu PM, Rhoderick GC, Johnson PA. Gas-phase databases for quantitative infrared spectroscopy. Appl Spectrosc. 2004;58(12):1452–1461.
36. Picone N, Pol A, Mesman R, van Kessel MAHJ, Cremers G, van Gelder AH, et al. Ammonia oxidation at pH 2.5 by a new gammaproteobacterial ammonia-oxidizing bacterium. ISME J. 2021;15(4):1150–1164.

37. Jahromi KE, Nematollahi M, Krebbers R, Abbas MA, Khodabakhsh A, Harren FJM. Fourier transform and grating-based spectroscopy with a mid-infrared supercontinuum source for trace gas detection in fruit quality monitoring. Opt Express. 2021;29(8):12381–12397.

38. Kuypers MMM, Marchant HK, Kartal B. The microbial nitrogen-cycling network. Nat Rev Microbiol. 2018;16(5):263–276.

39. Udert KM, Larsen TA, Gujer W. Chemical nitrite oxidation in acid solutions as a consequence of microbial ammonium oxidation. Environ Sci Technol. 2005;39(11):4066–4075.

40. Quick AM, Reeder WJ, Farrell TB, Tonina D, Feris KP, Benner SG. Nitrous oxide from streams and rivers: A review of primary biogeochemical pathways and environmental variables. Earth-Sci Rev. 2019;191:224–262.

41. Roth M. Fluorescence reaction for amino acids. Anal Chem. 1971;43(7):880–882.

42. Gordon SA, Fleck A, Bell J. Optimal conditions for the estimation of ammonium by the Berthelot reaction. Ann Clin Biochem. 1978;15(1–6):270–275.

43. Wang S, Lin K, Chen N, Yuan D, Ma J. Automated determination of nitrate plus nitrite in aqueous samples with flow injection analysis using vanadium (III) chloride as reductant. Talanta. 2016;146:744–748.

44. Truong GW, Perner LW, Bailey DM, Winkler G, Cataño-Lopez SB, Wittwer VJ, et al. Mid-infrared supermirrors with finesse exceeding 400 000. Nat Commun. 2023;14(1):7846.

45. Henderson B, Khodabakhsh A, Metsälä M, Ventrillard I, Schmidt FM, Romanini D, et al. Laser spectroscopy for breath analysis: towards clinical implementation. Appl Phys B. 2018;124(8):161.

46. Fenske JD, Paulson SE. Human breath emissions of VOCs. J Air Waste Manage Assoc. 1999;49(5):594–598.

47. Mathew T, Pownraj P, Abdulla S, Pullithadathil B. Technologies for clinical diagnosis using expired human breath analysis. Diagnostics. 2015;5(1):27–60.

48. Wang C, Sahay P. Breath analysis using laser spectroscopic techniques: breath biomarkers, spectral fingerprints, and detection limits. Sensors. 2009;9(10):8230–8262.

49. Herman DI, Weerasekara C, Hutcherson LC, Giorgetta FR, Cossel KC, Waxman EM, et al. Precise multispecies agricultural gas flux determined using broadband open-path dual-comb spectroscopy. Sci Adv. 2021;7(14):eabe9765.

50. Mead GJ, Waxman EM, Bon D, Herman DI, Baumann E, Giorgetta FR, et al. Open-path dual-comb spectroscopy of methane and VOC emissions from an unconventional oil well development in Northern Colorado. Front Chem. 2023;11:1202255.

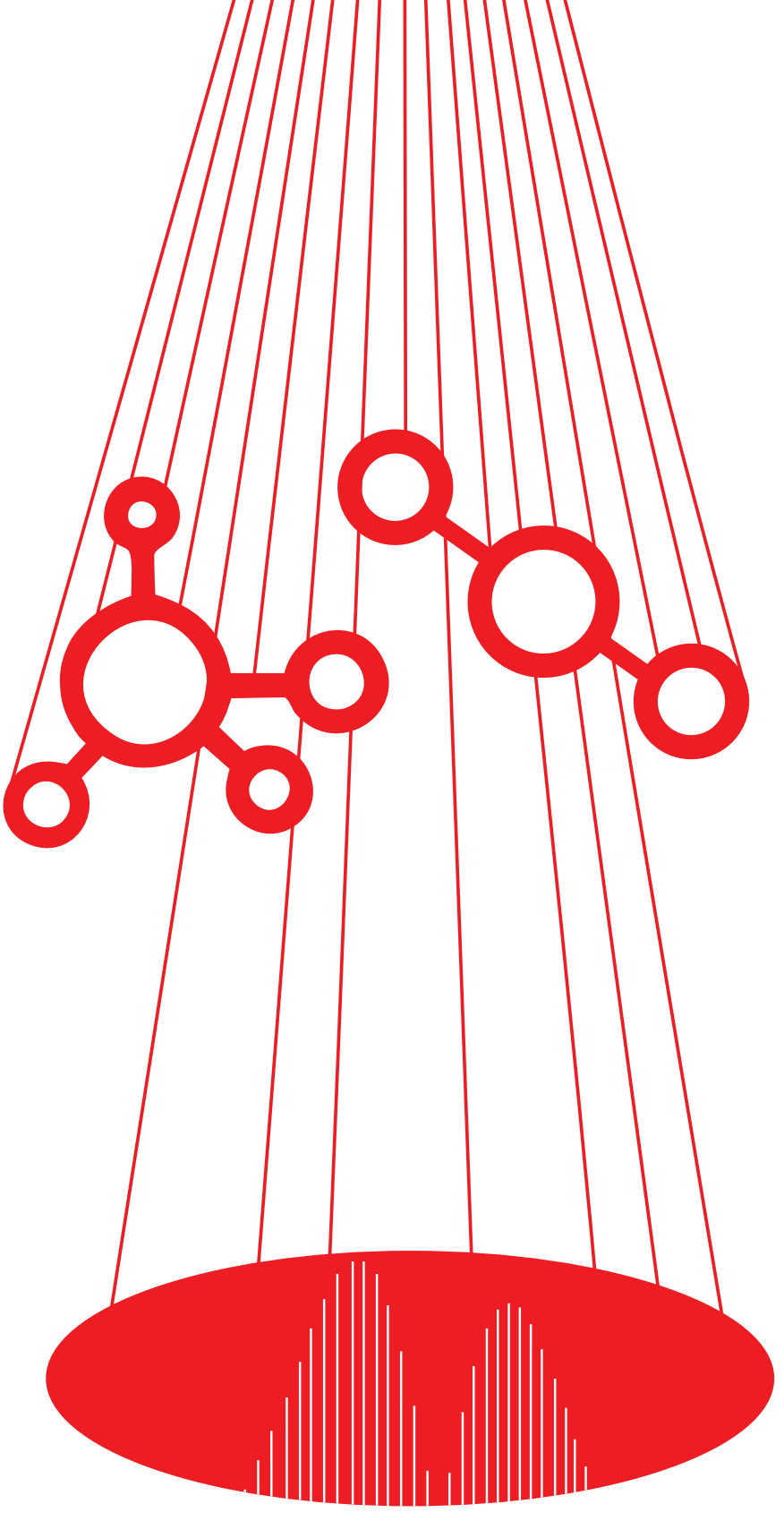
51. Giorgetta FR, Peischl J, Herman DI, Ycas G, Coddington I, Newbury NR, et al. Open-path dual-comb spectroscopy for multispecies trace gas detection in the 4.5–5 µm spectral region. Laser Photonics Rev. 2021;15(9):2000583.

52. Cossel KC, Waxman EM, Giorgetta FR, Cermak M, Coddington IR, Hesselius D, et al. Open-path dual-comb spectroscopy to an airborne retroreflector. Optica. 2017;4(7):724.

53. Rieker GB, Giorgetta FR, Swann WC, Kofler J, Zolot AM, Sinclair LC, et al. Frequency-comb-based remote sensing of greenhouse gases over kilometer air paths. Optica. 2014;1(5):290.

54. Alden CB, Coburn SC, Wright RJ, Baumann E, Cossel K, Perez E, et al. Single-blind quantification of natural gas leaks from 1 km distance using frequency combs. Environ Sci Technol. 2019;53(5):2908–2917.

55. Coburn S, Alden CB, Wright R, Cossel K, Baumann E, Truong GW, et al. Regional trace-gas source attribution using a field-deployed dual frequency comb spectrometer. Optica. 2018;5(4):320.



Chapter 7

Ultra-Broadband Coherent Open-Path Spectroscopy for Multi-Gas Monitoring in Wastewater Treatment[1]

[1.] This chapter is based on the following article:
Krebbers R, van Kempen K, Lin Y, Meurs J, Hendriks L, Aben R, Paranaiba, JR, Fritz, C, Veraart, AJ, Khodabakhsh, A, Cristescu, SM. Ultra-Broadband Coherent Open-Path Spectroscopy for Multi-Gas Monitoring in Wastewater Treatment. *Environ Sci Ecotechnology. 2025;25:100554.*

Abstract

Wastewater treatment plants significantly contribute to greenhouse gas emissions, including nitrous oxide (N_2O), carbon dioxide (CO_2), and methane (CH_4). Current methods to measure these emissions typically target specific molecular compounds, providing limited scope and potentially incomplete emissions profiles. Here, we show an innovative ultra-broadband coherent open-path spectroscopy (COPS) system capable of simultaneously monitoring multiple greenhouse gases. This novel approach combines Fourier-transform spectroscopy with a coherent, ultra-broadband mid-infrared light source spanning 2–11.5 µm at approximately 3 W power. Positioned above an aeration tank, the COPS system selectively detected absorption signatures for CH_4, CO_2, N_2O, ammonia (NH_3), carbon monoxide (CO), and water vapor (H_2O), enabling real-time, path-integrated concentration measurements with a temporal resolution of 40 seconds. Elevated concentrations of CH_4 and CO_2 were clearly identified within emission plumes traversing the beam path above the aeration tank. Additionally, CH_4 emission patterns closely tracked variations in ammonium loading from incoming wastewater, whereas CO_2 emissions correlated strongly with oxygen concentrations introduced during aeration. Measurements of N_2O, NH_3, and CO were stable and aligned closely with traditional point-based measurements from commercial gas analyzers. Our findings demonstrate that COPS offers a robust, comprehensive solution for the simultaneous real-time monitoring of multiple gases in complex and heterogeneous emission environments. This capability significantly enhances atmospheric and industrial emission assessments, potentially transforming the approach to emissions quantification and environmental management.

7.1 Introduction

Wastewater treatment plants (WWTPs) are known to contribute to greenhouse gas (GHG) emissions of carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) (1,2). However, only a few GHG emission datasets with sufficient spatial and temporal coverage are available from WWTPs. This limits the feasibility of modeling approaches, especially with the high heterogeneity that the current datasets show (3,4). The best currently available practical methods for estimating GHG emissions revolve around two approaches: point sampling and open-path (also known as free-path) spectroscopy (5,6).

Point sampling measurements for GHG emission estimates at WWTPs can be taken via discrete samples from surface waters and the atmosphere combined with (often highly uncertain) estimates of air-water gas exchange (7,8), or by collecting off-gas from covered WWTP components (9). Discrete sampling methods are limited by their temporal and spatial resolution: their sampling frequency is low compared to continuous monitoring techniques, and the spatial coverage is limited to specific points within the study area. Moreover, the potential contribution of microbubbles is excluded, which may be particularly important in WWTPs where aeration takes place (10,11).

Another point sampling approach is to use open and closed floating flux chambers, usually combined with online GHG analyzers (e.g., off-axis integrated cavity output spectrometers or cavity ring-down spectrometers) (12).

Flux chamber measurements offer snapshots of flux estimates and can be ideal for analyzing small-scale variations in GHG fluxes within aquatic ecosystems. However, flux chamber measurements face operational difficulties in the fast-flowing waters of, e.g., WWTPs, and the presence of the chamber itself can challenge the assumption that measured fluxes represent those outside of the chamber, for example, due to artificially enhanced within-chamber turbulence (13–15). Moreover, both discrete water or gas sampling and flux chamber methods are currently labor-intensive and cover only a small surface area, which is a problem in systems with significant spatial heterogeneity.

Continuous spatiotemporal measurements with open-path spectroscopic techniques overcome the limitations of point sampling approaches. These techniques send a light beam over a free path through the area of interest to interact directly with the gas. This contrasts with point measurement techniques, where gas samples are captured in a gas cell for interaction with the light. Open-path techniques offer non-invasive, *in situ*, stand-off, and real-time detection of gas species, providing faster measurements and avoiding sampling complications with gases that are difficult to sample in gas cells due to wall-adhesion effects (such as ammonia (NH_3)).

Differential optical absorption spectroscopy (DOAS) is a classical technique in open-path spectroscopy (16–18) using sunlight or a thermal source (usually in ultraviolet/visible spectrum) for gas detection (e.g. of GHGs and nitrogen oxides) typically in combination with a Fourier-transform spectrometer (FTS) or grating-based spectrometer. As the brightness of the thermal source is limited, DOAS needs a long measurement time, provides a low spectral resolution, and thus results in limited applicability (19).

Coherent open-path spectroscopy (COPS) replaces thermal incoherent sources in DOAS systems with spatially and/or temporally coherent sources. These sources can be optimized for the near- (1–2.5 μm) and especially the mid-infrared (MIR) spectral range (2.5–20 μm) where most of the interesting gases present the strongest molecular absorption features. Dual-comb spectroscopy (DCS) systems are frequently used for COPS-based gas sensing, operating mostly in the near-infrared (NIR) spectral range (20–26). Even though the technology for this spectral range is mature, gas sensing in NIR mostly relies on (weaker) overtone vibrational bands of molecular species, resulting in a limited number of detectable species and low detection sensitivities. More recently, DCS COPS systems have been employed in the MIR, however, either not extending further than 5 μm (27–31) or exclusively covering the region around 10 μm (32).

High-power supercontinuum (SC) sources are a compelling choice for COPS systems, given their high brightness, broad spectral coverage, and spatial coherence facilitating transmission over long distances and shorter measurement times (33). In recent years, there have been significant developments in their availability for the MIR spectral region (34,35). Current state-of-the-art SC sources achieve broad spectral coverage in MIR via nonlinear effects in special fibers (36–39) or intrapulse difference frequency generation (IDFG) in nonlinear crystals (40–42). With the exceptional spectral coverage of state-of-the-art SC sources, it is possible to cover the strongest, fundamental rovibrational bands of different molecular species (e.g., hydrocarbons, oxides, and small organic molecules) in the MIR spectral range from 2–11.5 μm (42), enabling multispecies detection of these molecules with high sensitivity.

Here, we present the results of the first field deployment of a COPS system with unprecedented spectral coverage (2–11.5 μm) in the MIR based on IDFG, yielding high detection sensitivities for a broad range of molecules. The system was deployed at a WWTP to characterize the gases present above the aeration tank. The results of the COPS system were compared to state-of-the-art commercial GHG analyzers, which measured the concentrations of GHGs in the emission plume from the aeration tank.

7.2 Materials and methods

7.2.1 The COPS system: a platform for simultaneous detection of greenhouse and polluting gases

Coherent open-path spectroscopy is a unique technique that combines laser-based gas spectroscopy with ultra-broadband spectral coverages, enabling sensitive measurements with a high dynamic range for a great variety of different gases simultaneously. It also provides a high resolving power for identification and interference-free detection of these gas species. The COPS system used in this study comprises an in-house developed and built Fourier- transform spectrometer (36) combined with a mid-infrared supercontinuum source (SC-FTS). The supercontinuum (SC) source is a novel laser source with a high-power and low-noise ultrabroad spectrum in the mid-infrared spectral region (~3 W, 2–11.5 µm) (42,43). The SC source is water-cooled, and its laser head is purged with dry air. It consists of a mode-locked master oscillator and a single-pass power amplifier, both based on polycrystalline Cr:ZnS gain elements and optically pumped by Er-doped fiber lasers. The broad spectral coverage is produced using IDFG in a non-linear $ZnGeP_2$ crystal. The source's unique, broad spectral range covers fundamental rovibrational bands of most small molecular gases (e.g., GHGs and polluting gases, as well as many other hydrocarbons, oxides, and small organic molecules). To the best of our knowledge, this is the first field deployment of such an ultra-broadband IDFG-based SC source out of a laboratory environment. The free-space output from the SC source is reflected by two off-axis parabolic mirrors (MPD129-P01 and MPD169-P01, Thorlabs) that expand the beam and optimize the collimation. The expanded beam is then sent over an open path for interaction with the targeted sample area. A 4.6-cm cubic retroreflector (HRR203-P01, Thorlabs) is used to reflect the beam and return it to the SC-FTS, where a second pair of off-axis parabolic mirrors (MPD249-P01 and MPD019-P01, Thorlabs) reduces the beam size and directs it to the FTS. The spectrum is measured every two seconds by the FTS with a 3 GHz (= 0.1 cm^{-1}) spectral resolution. The high spectral resolution of FTS resolves the absorption features at the natural atmospheric broadening, enabling optimal discrimination between gases with overlapping absorption features. The FTS is a Michelson-type interferometer that uses a balanced detection scheme to improve the signal-to-noise ratio (SNR) and enhance the detection sensitivity (36). The entire system is integrated into a transportable 1.3 m × 1.3 m × 0.8 m cart.

7.2.2 Data collection

The FTS records an interferogram every two seconds, which is converted in post-processing to a power spectrum using fast Fourier transformation. Every power

spectrum contains the entire spectral coverage of the source, meaning that all gas species absorbing the SC light are detected simultaneously. The power spectra are averaged for 40 seconds to further enhance the SNR and increase the sensitivity of the system. The averaged power spectra are transformed into absorbance spectra, so that the obtained absorption features scale linearly with the concentration of the molecules, following the Beer-Lambert law. This transformation is done by normalizing the power spectrum to a global baseline and a summation of sinusoidal functions to account for the general shape of the power spectrum and etalons introduced by the optics, most notably the beam splitter inside the FTS. Despite the beam splitter being wedged (Thorlabs BSW711), it is very common for FTS systems to have some residual etalon in the detected spectra that needs to be removed numerically. The HITRAN database (44) is used to construct reference spectra of the expected molecules (e.g., NH_3, carbon monoxide (CO), N_2O, CO_2, CH_4, and water (H_2O)) in the environmental conditions during the measurement campaign (temperature: 17 ± 1 °C, atmospheric pressure 1013 ± 1 hPa). A fit of these reference spectra to the measurement spectra yields the concentration of each molecule integrated over the entire line-of-sight of the laser. Sinusoidal and polynomial reference spectra account for power spectrum fluctuations and etalons. The residual (the measured absorbance spectrum minus the concentration-scaled reference spectra) is minimized through this fitting; leaving no remaining absorption features indicates all gas species have been accounted for and their concentrations have been appropriately retrieved. A schematic overview of this data-processing procedure is provided in Figure 7.1.

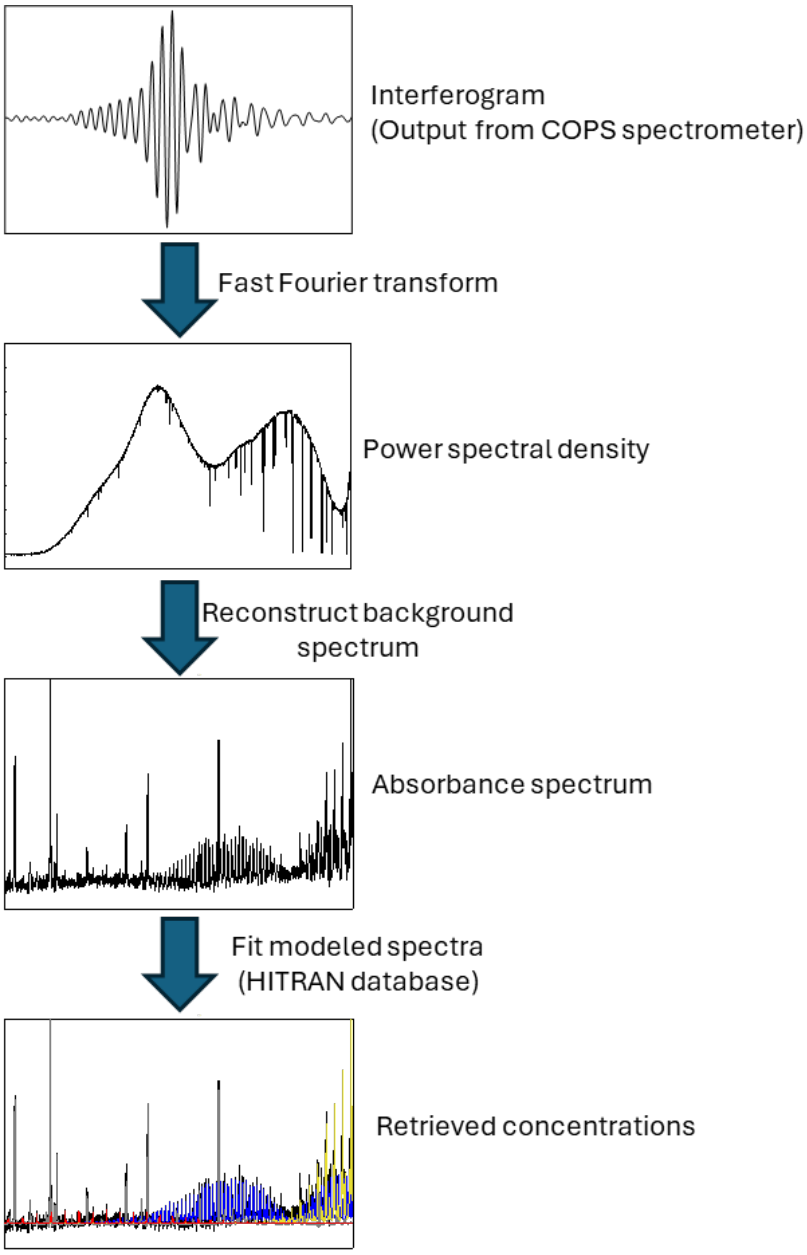


Figure 7.1: Schematic overview of the data processing procedure.

7.2.3 The WWTP site and experimental protocol

The measurements were performed above the aeration tank (AT) of the WWTP in Remmerden, the Netherlands (Figure 7.2), in October 2023. The COPS system (A) was placed next to the AT, and the laser beam traveled towards a retroreflector (B_1) over the AT and back to the spectrometer (red lines). To identify the background concentration of the gases, a second beam path was constructed perpendicular to the beam path over the AT and facing upwind (blue lines to B_2). With a range finder (Coolshot Pro II, Nikon), both beam paths were verified to be 60.0 ± 0.8 m long. The COPS system was continuously monitoring the gases above the AT for approximately four hours, after which the background concentrations were recorded for approximately 30 minutes.

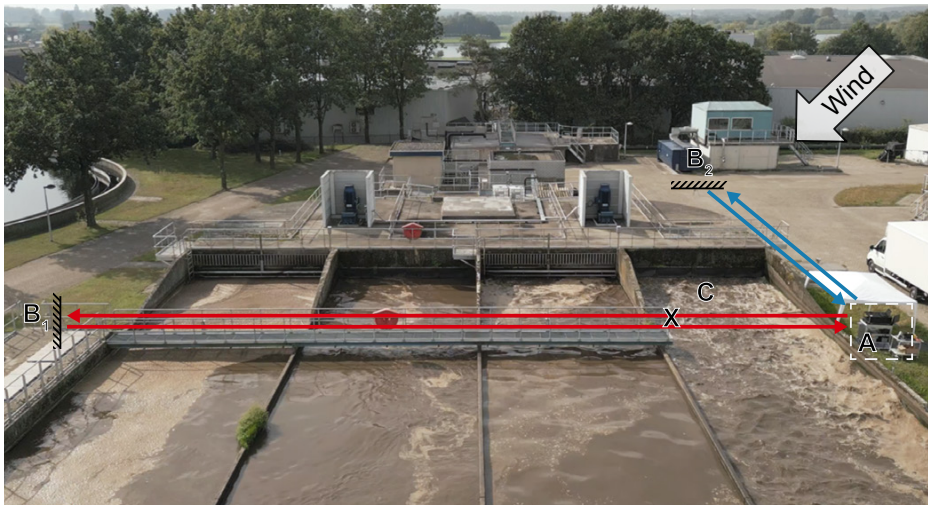


Figure 7.2: View of the aeration tank (AT) at the wastewater treatment plant with the coherent open-path spectroscopy system (A), retroreflectors (B_1 and B_2), the AT beam path (red), and upwind beam path (blue), both 60 meters. The point measurement sampling location (X) is above the lane with the main aerator of the AT (C).

During the COPS measurements, we employed several state-of-the-art commercial point measurement systems to measure the gases simultaneously retrieved by the COPS system. All these commercial systems use an internal gas cell where the gas is sampled from a location (marked X in Figure 7.2) above the main aerator of the AT and along the beam path of the COPS system. CH_4 and CO_2 were measured using an off-axis integrated cavity output spectroscopy (OA-ICOS-based) system (Ultra-portable Greenhouse Gas Analyzer, ABB, Inc.). CO and N_2O were measured using a mid-infrared laser-based system with a multipass cell (MIRA Ultra analyzer, Aeris

Technologies, Inc.). NH_3 was measured using a photoacoustic system (LSE NH_3 Monitor 1700, LSE Monitors) equipped with a specially perfluoroalkoxy alkane (PFA)-coated cell and heated (~40 °C) PFA tubing to minimize the adhesion of NH_3.

The AT of the WWTP is a four-lane carrousel system equipped with two surface aerators to introduce oxygen to the water. The main aerator is located at the top of the right lane (C, Figure 7.2). The AT is equipped with a probe to measure the dissolved oxygen (O_2) concentration (LDO, Hach) and dissolved ammonium (NH_4^+) concentration (ANI-SE, Hach). Aeration is increased when the NH_4^+ concentration exceeds a threshold of ~3.5 mg L^{-1}.

7.3 Results and discussion

7.3.1 Determining detected molecular species and path-averaged concentrations

The 40-second averaged power spectrum of the measurement of a 60-meter open path over the AT covering the entire wavelength region is shown in Figure 7.3a. This is a significant part of the mid-IR fingerprint region related to the C–H stretching and C=O vibrational modes and, therefore, opens possibilities for the identification and sensitive detection of a great variety of gases (41). The power spectrum predominantly contains transmission dips related to highly absorbing H_2O and CO_2 features. Relevant areas containing features of other gas species are marked white. The absorbance spectrum of these spectral ranges is shown in Figure 7.3b, together with fitted, modeled reference spectra of the detected molecules (NH_3, CO, N_2O, CO_2, CH_4, and H_2O). These spectra are shown in color and inverted for clarity. The fit of these reference spectra to the measured spectra yields the concentration of each molecule. The residual of this fit (Figure 7.3c) demonstrates the high quality of the fitting since no visible absorption features are left, apart from a few remaining ones linked to highly absorbing (saturating) water lines. Therefore, these specific saturated water features were not considered when retrieving the water concentration, and they do not interfere with detecting other molecules. Thanks to the broadband spectrum, it is possible to select spectral regions with little water features or resolve features at a spectral resolution sufficient to discriminate the water features from the other molecules. Therefore, this demonstrates that the method is interference-free and able to resolve molecular concentrations without needing to calibrate or compensate for the water concentration.

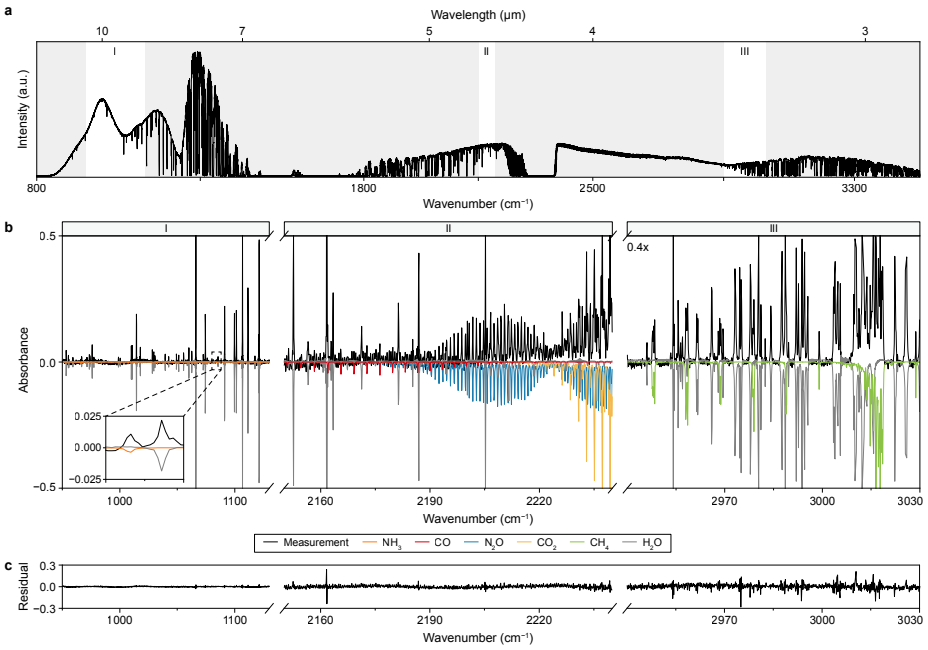


Figure 7.3: a–b, Power spectrum measured in 40 seconds over the aeration tank **(a)** with absorbance spectrum **(b)** zoomed to spectral ranges I, II, and III of the power spectrum for simultaneous detection of NH_3, CO, N_2O, CO_2, CH_4, and H_2O. The absorbance spectrum is fitted to reference spectra (in color and inverted), and the right window is magnified 0.4× to visualize relevant absorbance lines. **c**, The residual of the fit of reference spectra to the measured absorbance spectrum.

For each spectrum (i.e., every 40 seconds), the concentrations of the absorbing molecules detected by the COPS system are retrieved; their dynamics are displayed in Figure 7.4.

The reported concentration is an average for the entire 60-meter beam path, even though the concentration throughout the area might be unevenly distributed. The concentration might be higher in a plume from an emission source; this will be considered in section 7.3.2. The concentration can also be reported as the product of volume fraction and path length, which gives insights into the detection sensitivity of the system independent of the optical path length used. These values, which are called path-integrated concentrations, can be found in Supplementary Materials Figure 7.1. From 11:00 to 15:15, the COPS system measured over the AT (black line). For a shorter period, from 15:30 onwards, the system recorded the background concentrations upwind of the AT (blue line). In this period, the species detected and monitored over time were CH_4, CO_2, N_2O, CO, H_2O, and NH_3. CH_4 and CO_2 have higher concentrations above the AT than the background concentrations and show a clear dynamic pattern. The concentrations of N_2O, CO, NH_3, and H_2O are

similar over the AT and the upwind beam path, retrieving typical concentrations of 329 ± 5 ppb, 157 ± 19 ppb, 11.9 ± 1.5 ppb, and 1.53 ± 0.01 %, respectively, that remain stable over time, confirming that there are no or only minor (detectable) emissions from the AT of these compounds over the considered measuring period. The increase in H_2O relates to an increasing ambient temperature during the day.

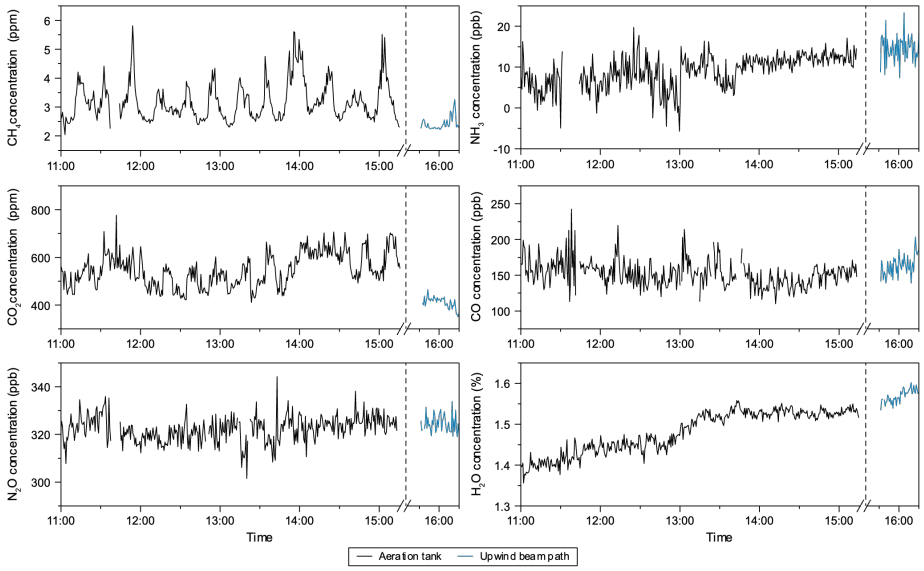


Figure 7.4: Path-averaged concentrations (40-second averages) of detected species over the aeration tank (black lines) and path-averaged background concentration over the upwind beam path (blue lines).

The stable concentrations of N_2O, CO, and NH_3 measured in the upwind and the AT measurement path by the COPS system agree with the average concentrations detected by the point measurement devices (Table 7.1).

Table 7.1: Time-averaged concentrations of CO, NH_3, and N_2O measured above the aeration tank with the coherent open-path spectroscopy system (COPS) and the point measurement devices. The concentration was averaged over one hour, with the standard deviation as uncertainty and a 40-second time interval per measurement point.

Molecule	Average concentration (COPS)	Average concentration (Point measurements)
CO	157 ± 19 ppb	165 ± 4 ppb
NH_3	11.9 ± 1.5 ppb	11.6 ± 1.7 ppb
N_2O	329 ± 5 ppb	333.2 ± 0.6 ppb

NH_3 is detected at a concentration close to the detection limit of the COPS system, as indicated by the high noise in the retrieved concentrations (Figure 7.4); however, it matches well with the concentration from the point-measurement instrument. Furthermore, these concentrations are in line with the literature on expected concentrations of these molecules in (semi-)urban environments (45–48).

7.3.2 Estimating emission plume concentrations

The temporal dynamics of the CH_4 and CO_2 concentrations observed with the COPS system (Figure 7.4) indicate clear emission moments of these gases from the AT. Even though the COPS system monitors the integrated concentration over the entire line of sight of the beam, the concentration within the emission plume could be estimated. For comparison, the point measurement devices were placed specifically above the aeration point of the WWTP to sample from the emission plume and detect the concentration of these gases inside the plume. For the point measurements, it is crucial to sample at the right position in the plume, as the dynamic and elevated concentrations could not be observed at a different sampling point above the other lanes of the AT. The COPS system, however, does not depend on these strict limitations on the sampling location. To estimate the plume concentrations with the COPS system, the path-integrated concentration was considered for a path length of twice the width of the aeration channel (10 m), and atmospheric background concentrations (CH_4: 2.3 ppm, CO_2: 420 ppm) were subtracted for the remaining path length. The resulting concentrations show the same trend as those found by the point measurements (Figure 7.5). The CH_4 concentrations show a high level of correlation, while the CO_2 concentrations demonstrate the same dynamics albeit at a slight offset between 12:00 and 12:45, possibly caused by diurnal fluctuations in the background CO_2 level. Extended measurements of the background concentration in the upwind beam path would improve the accuracy of these plume concentration estimates. The correlations demonstrate that the COPS system can detect the concentrations within the emission plume with similar precision as a point measurement instrument held specifically at the emission point, even though the COPS system is observing its entire beam path at once (i.e. stand-off detection). Thus, the COPS system offers more flexibility in regions that can be monitored (as not all spots at the WWTP are easily accessible for point sampling), has a higher chance of detecting unforeseen emission points, and can detect multiple points simultaneously. Furthermore, retrieving a spatially-integrated concentration reduces the chance of errors due to spatial heterogeneity. In addition, the COPS system can simultaneously monitor many more molecular species than all the point measurement devices used in this study, combined.

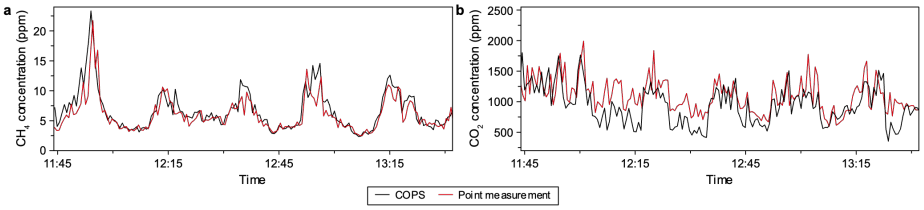


Figure 7.5: Plume concentrations of CH_4 **(a)** and CO_2 **(b)** at the aerator, measured and estimated with the coherent open-path spectroscopy system (COPS, black lines) and sampled from the plume with the OA-ICOS-based ABB point measurement system (red lines).

7.3.3 Relation to WWTP operation

In the AT of this specific WWTP, air is pushed through the wastewater during aeration using a surface aerator. The turbulence at the water-air interface enhances the exchange of dissolved CH_4 and CO_2 to the atmosphere. The operation of the aeration system is regulated by the measured dissolved ammonium (NH_4^+) concentration in the AT. NH_4^+ is an indicator of inflowing wastewater to the AT, as it is naturally present in domestic sewage and gets oxidized to nitrate (NO_3^-) by nitrifying microorganisms in the aerobic treatment process (49). The detected gases above the AT can be related to the operation of the WWTP. In Figure 7.6, the concentrations of CH_4 and CO_2 above the AT retrieved by the COPS system are shown with the periods in which the aeration is maximized (grey background) and the measured concentration of oxygen and NH_4^+ dissolved in the water. The CH_4 and CO_2 concentrations are the same as reported in Figure 7.5 but smoothened using adjacent averaging to improve the clarity of the trends.

Both the emission of CO_2 and CH_4 can be linked to maximal system aeration (Figure 7.6). The concentrations of CH_4 and NH_4^+ (black and green lines) follow the same trend, indicating that the CH_4 emissions peak when wastewater flows into the AT. CH_4 is produced by methanogenic archaea as a product of the anaerobic conversion of the fermentation products acetate and hydrogen within the sewage and the pipelines of the wastewater treatment process before entering the AT (9,50,51). Aeration enhances the oxygen concentration in the wastewater for the aerobic nitrification of NH_4^+ to NO_3^-, which can then be converted to nitrogen gas (N_2) under oxygen-limited conditions. Similarly, aerobic processes convert organic carbon-rich compounds into CO_2, while anaerobic processes can convert these compounds to CH_4 in addition to (some) CO_2 (52). Therefore, the similar dynamics between gaseous CO_2 and dissolved O_2 (blue and red lines) can be well explained. During every aeration period, O_2 is brought into the wastewater. The dissolved O_2 rapidly decreases when the aeration is tuned down. The remaining O_2 is converted to CO_2 that dissolves well in water. When the aeration is turned on again, CO_2 emission peaks.

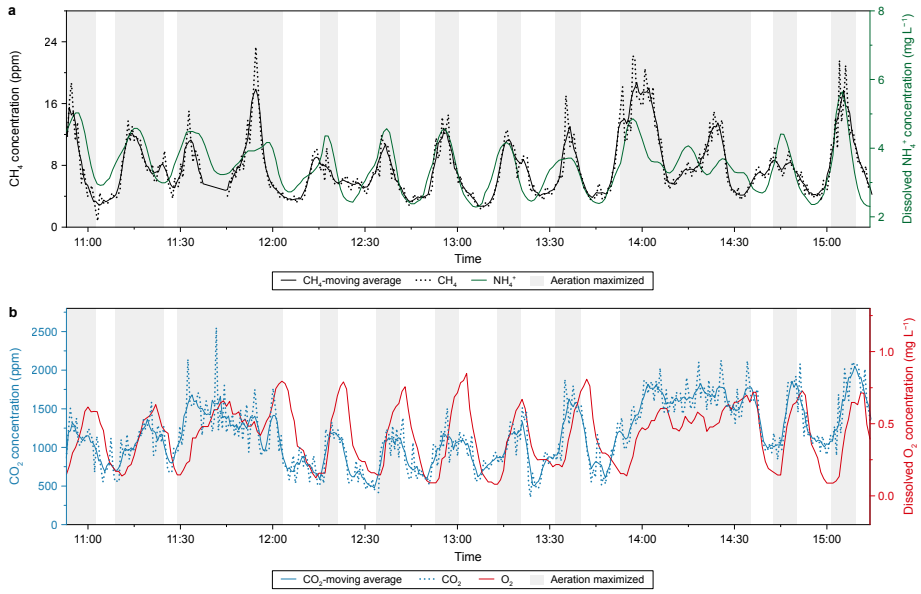


Figure 7.6: Plume concentrations of CH_4 **(a)** and CO_2 **(b)** above the aeration tank, measured by the coherent open-path spectroscopy system, with the periods of maximized aeration in grey shades and the NH_4^+ **(a)** and O_2 **(b)** concentration.

7.3.4 Implications

The broadband coverage and the high spectral resolution of the COPS system enable a high resolving power for identification and interference-free detection of an extensive number of gas species. While CH_4, CO_2, N_2O, NH_3, CO, and H_2O were monitored at the WWTP, these gases do not represent an exhaustive list of detectable species. In previous work, it has been demonstrated that the COPS system can detect, among other molecules, hydrocarbons, oxides, and small organic molecules in a laboratory setting (42). Simultaneous measurements of more (trace) gases and pollutants enable a deep understanding of biogeochemical processes and validation of multi-scale models.

Although N_2O was monitored during this study, no sustained emissions have been found. The stable, matching concentrations between the AT and upwind beam path indicate that the system successfully detects this greenhouse gas and the AT did not emit considerable amounts of this gas at the time of measurement. N_2O emissions from WWTPs are of high interest due to the limited available information on the total emission of this strong GHG from the WWTP, as well as its variability over time. The emissions are known to be related to incomplete (de)nitrification

and have a strong temporal dynamic, showing to be spiking at specific moments, such as during rapidly changing processes or environmental conditions (1). For this reason, having measurement solutions that can continuously monitor areas is instrumental and a major advantage over labor-intensive discrete sampling approaches. The COPS system is a well-suited candidate for such continuous measurement campaigns. The system can operate autonomously, resulting in an efficient, cost-effective solution for emission monitoring and characterization over prolonged periods, including day-night cycles and seasonal variations.

A clear benefit of the COPS system over point-measurement systems is the ability to detect these emissions without carefully searching and moving the system to different positions to find the emission location/area. Incorrect positioning of point measurement systems can cause over- or underestimation of gas concentrations emitted from an unexpected or inhomogeneous source. By scanning over an open path, a bigger area can be monitored simultaneously, overcoming the spatial limitations of point measurement techniques in heterogeneous environments. Future refinements of the technique will allow higher sampling rates that would facilitate combined analysis of gas concentration changes and atmospheric mass flow (53). Furthermore, as the system does not need to be moved, it is also easier to keep track of temporal changes in the emission rates. Finally, another advantage of COPS is its calibration-free methodology, as well as the fact that no gas sample treatment is needed (i.e., preventing wall-adhesions effects in sampling NH_3).

7.4 Conclusions and outlook

The uniquely broad spectral coverage of the SC source in the COPS system (2–11.5 µm) in combination with ~3 W power enables the simultaneous detection of a wide variety of molecular gases with a high sensitivity and a high dynamic range. For molecules with stable atmospheric concentrations (CO, NH_3, and N_2O), estimated sensitivities were reached of 19, 1.5, and 5 ppb in 40 seconds, respectively. This novel COPS system has significantly broader spectral coverage than current dual-comb systems (28,32), increasing the scope of detectable molecular species and the dynamic range due to the high number of detectable absorption lines (42).

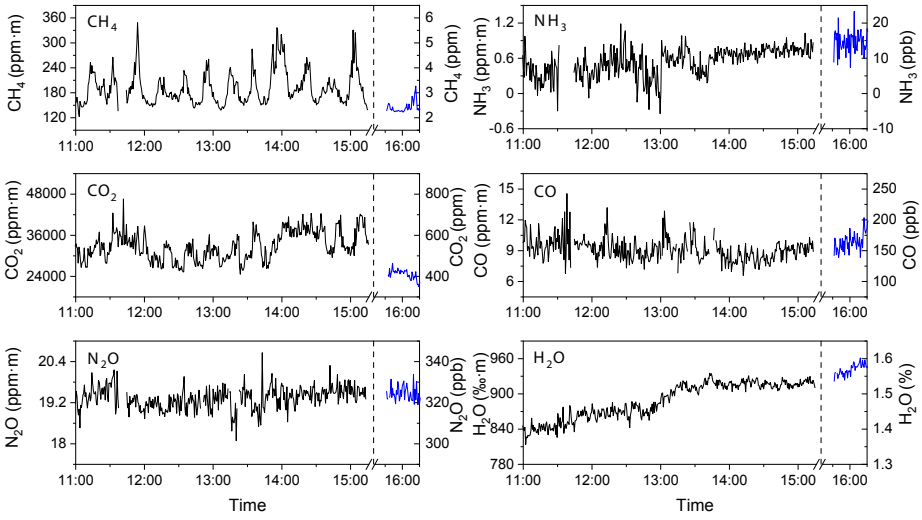
We demonstrated the first field deployment of such a system. Over an open path, gas concentrations were monitored above and upwind of the AT of a WWTP. Several species, such as CH_4, CO_2, N_2O, NH_3, CO, and H_2O, were present at detectable concentrations at the WWTP. The temporal dynamics and elevated concentrations of CH_4 and CO_2 indicated clear emissions of these compounds from the AT. Similar dynamics were observed using commercial laser-based point-sampling

spectroscopy instruments. These temporal dynamics link to the operation of the WWTP, showing the relation of CO_2 and CH_4 emissions to the inflow of wastewater to the AT, aerobic processes in the AT, and the aerator speed.

The system successfully detected N_2O, CO, and NH_3 in the measurements of the beam path over the AT and the upwind beam path. We conclude that the AT did not emit considerable amounts of these gases at the time of measurement, as stable, matching concentrations were found between both paths.

The results presented here demonstrate the COPS technology in a field application at a WWTP. This represents the first step towards real-time emission monitoring and quantification of GHG and other gases relevant to complementing the nitrogen and carbon balances of the WWTPs. In particular, quantifying these emissions over extended periods, ranging from day–night cycles to varying weather conditions or seasonal variations, will provide valuable data to WWTP operators for more accurate annual GHG budgets and better decision-making on mitigation strategies to ensure WWTP sustainability.

Supplementary materials



Supplementary Figure 7.1: Path-integrated concentrations (40-second averages) of detected species over the AT (in black) and path-integrated background concentration over the upwind beam path (in blue)

The reported concentration in Figure 7.4 is an average for the entire 60-meter beam path, even though the concentration throughout the area might be unevenly distributed. As the absorbance strength of the gas species scales linearly with both the path length and concentration, it is also possible to report path-integrated concentrations as the product of volume fraction and path length instead of the volume fraction only. This gives insights into the detection sensitivity of the system, independent of the optical path length used. Therefore, Supplementary Figure 7.1 reports path-integrated concentrations (in units of the product of volume fraction and path length, left axis) as well as the path-averaged concentrations (in units of volume fraction, right axis).

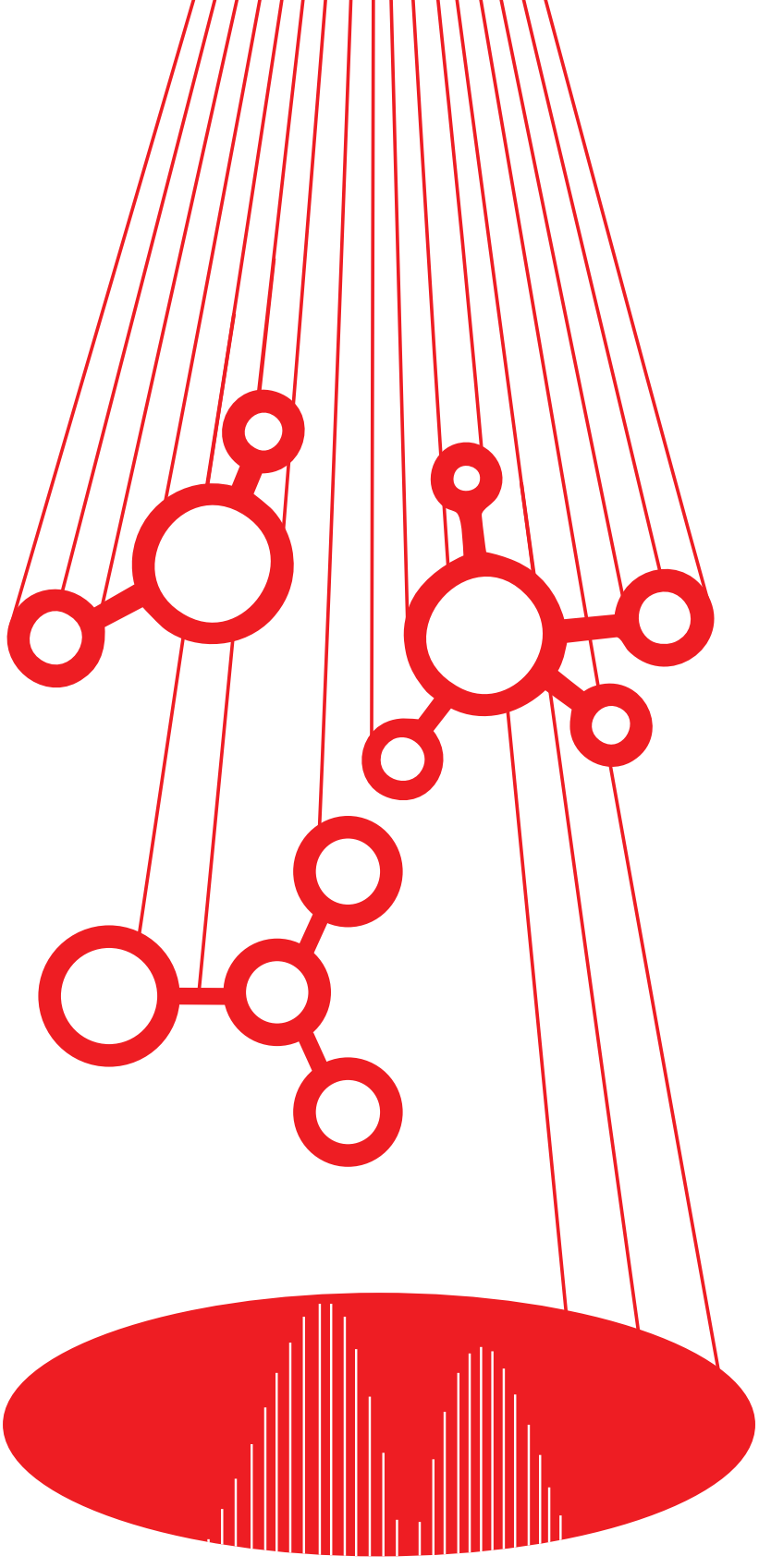
References

1. Kampschreur MJ, Temmink H, Kleerebezem R, Jetten MSM, Van Loosdrecht MCM. Nitrous oxide emission during wastewater treatment. Water Res. 2009;43(17):4093–4103.
2. Campos JL, Valenzuela-Heredia D, Pedrouso A, Val Del Río A, Belmonte M, Mosquera-Corral A. Greenhouse gases emissions from wastewater treatment plants: minimization, treatment, and prevention. Journal of Chemistry. 2016;2016:1–12.
3. Gruber W, Villez K, Kipf M, Wunderlin P, Siegrist H, Vogt L, et al. N_2O emission in full-scale wastewater treatment: proposing a refined monitoring strategy. Sci Total Environ. 2020;699:134157.
4. Masuda S, Suzuki S, Sano I, Li Y-Y, Nishimura O. The seasonal variation of emission of greenhouse gases from a full-scale sewage treatment plant. Chemosphere. 2015;140:167–173.
5. Bai M, Wang Z, Lloyd J, Seneviratne D, Flesch T, Yuan Z, et al. Long-term onsite monitoring of a sewage sludge drying pan finds methane emissions consistent with IPCC default emission factor. Water Res X. 2023;19:100184.
6. Bastviken D, Wilk J, Duc NT, Gålfalk M, Karlson M, Neset T-S, et al. Critical method needs in measuring greenhouse gas fluxes. Environ Res Lett. 2022;17(10):104009.
7. Raymond PA, Caraco NF, Cole JJ. Carbon dioxide concentration and atmospheric flux in the Hudson River. Estuaries. 1997;20(2):381.
8. Cole JJ, Caraco NF. Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF6. Limnol Oceanogr. 1998;43(4):647–656.
9. Daelman MRJ, Van Voorthuizen EM, Van Dongen UGJ, Volcke EIP, Van Loosdrecht MCM. Methane emission during municipal wastewater treatment. Water Res. 2012;46(11):3657–3670.
10. McGinnis DF, Kirillin G, Tang KW, Flury S, Bodmer P, Engelhardt C, et al. Enhancing surface methane fluxes from an oligotrophic lake: exploring the microbubble hypothesis. Environ Sci Technol. 2015;49(2):873–880.
11. Rosentreter JA, Maher DT, Ho DT, Call M, Barr JG, Eyre BD. Spatial and temporal variability of CO_2 and CH_4 gas transfer velocities and quantification of the CH_4 microbubble flux in mangrove dominated estuaries. Limnol Oceanogr. 2017;62(2):561–578.
12. Gerardo-Nieto O, Vega-Peñaranda A, Gonzalez-Valencia R, Alfano-Ojeda Y, Thalasso F. Continuous measurement of diffusive and ebullitive fluxes of methane in aquatic ecosystems by an open dynamic chamber method. Environ Sci Technol. 2019;53(9):5159–5167.
13. Lorke A, Bodmer P, Noss C, Alshboul Z, Koschorreck M, Somlai-Haase C, et al. Technical note: drifting versus anchored flux chambers for measuring greenhouse gas emissions from running waters. Biogeosciences. 2015;12(23):7013–7024.
14. Mannich M, Fernandes CVS, Bleninger TB. Uncertainty analysis of gas flux measurements at air–water interface using floating chambers. Ecohydrol Hydrobiol. 2019;19(4):475–486.
15. Vachon D, Prairie YT, Cole JJ. The relationship between near-surface turbulence and gas transfer velocity in freshwater systems and its implications for floating chamber measurements of gas exchange. Limnol Oceanogr. 2010;55(4):1723–1732.
16. Griffith DW, Pöhler D, Schmitt S, Hammer S, Vardag SN, Platt U. Long open-path measurements of greenhouse gases in air using near-infrared Fourier transform spectroscopy. Atmos Meas Tech. 2018;11(3):1549–1563.
17. Nasse J-M, Eger PG, Pöhler D, Schmitt S, Frieß U, Platt U. Recent improvements of long-path DOAS measurements: impact on accuracy and stability of short-term and automated long-term observations. Atmos Meas Tech. 2019;12(8):4149–4169.

18. Volten H, Bergwerff JB, Haaima M, Lolkema DE, Berkhout AJC, Van Der Hoff GR, et al. Atmospheric measurement techniques two instruments based on differential optical absorption spectroscopy (DOAS) to measure accurate ammonia concentrations in the atmosphere. Atmos Meas Tech. 2012;5(2):413–427.
19. Schmitt TD, Kuhn J, Kleinschek R, Löw BA, Schmitt S, Cranton W, et al. An open-path observatory for greenhouse gases based on near-infrared Fourier transform spectroscopy. Atmos Meas Tech. 2023;16(24):6097–6110.
20. Rieker GB, Giorgetta FR, Swann WC, Kofler J, Zolot AM, Sinclair LC, et al. Frequency-comb-based remote sensing of greenhouse gases over kilometer air paths. Optica. 2014;1(5):290.
21. Coburn S, Alden CB, Wright R, Cossel K, Baumann E, Truong G-W, et al. Regional trace-gas source attribution using a field-deployed dual frequency comb spectrometer. Optica. 2018;5(4):320.
22. Alden CB, Coburn SC, Wright RJ, Baumann E, Cossel K, Perez E, et al. Single-blind quantification of natural gas leaks from 1 km distance using frequency combs. Environ Sci Technol. 2019;53(5):2908–2917.
23. Alden CB, Wright RJ, Coburn SC, Caputi D, Wendland G, Rybchuk A, et al. Temporal variability of emissions revealed by continuous, long-term monitoring of an underground natural gas storage facility. Environ Sci Technol. 2020;54(22):14589–14597.
24. Cossel KC, Waxman EM, Giorgetta FR, Cermak M, Coddington IR, Hesselius D, et al. Open-path dual-comb spectroscopy to an airborne retroreflector. Optica. 2017;4(7):724–728.
25. Cossel KC, Waxman EM, Hoenig E, Hesselius D, Chaote C, Coddington I, et al. Ground-to-UAV, laser-based emissions quantification of methane and acetylene at long standoff distances. Atmos Meas Tech. 2023;16(22):5697–5707.
26. Herman DI, Weerasekara C, Hutcherson LC, Giorgetta FR, Cossel KC, Waxman EM, et al. Precise multispecies agricultural gas flux determined using broadband open-path dual-comb spectroscopy. Sci Adv. 2021;7(14):eabe9765.
27. Ycas G, Giorgetta FR, Baumann E, Coddington I, Herman D, Diddams SA, et al. High-coherence mid-infrared dual-comb spectroscopy spanning 2.6 to 5.2 μm. Nat Photon. 2018;12(4):202–208.
28. Ycas G, Giorgetta FR, Friedlein JT, Herman D, Cossel KC, Baumann E, et al. Compact mid-infrared dual-comb spectrometer for outdoor spectroscopy. Opt Express. 2020;28(10):14740–14752.
29. Mead GJ, Waxman EM, Bon D, Herman DI, Baumann E, Giorgetta FR, et al. Open-path dual-comb spectroscopy of methane and VOC emissions from an unconventional oil well development in Northern Colorado. Front Chem. 2023;11:1202255.
30. Giorgetta FR, Peischl J, Herman DI, Ycas G, Coddington I, Newbury NR, et al. Open-path dual-comb spectroscopy for multispecies trace gas detection in the 4.5–5 μm spectral region. Laser Photonics Rev. 2021;15(9):2000583.
31. Kara O, Sweeney F, Rutkauskas M, Farrell C, Leburn CG, Reid DT. Open-path multi-species remote sensing with a broadband optical parametric oscillator. Opt Express. 2019;27(15):21358–21366.
32. Westberg J, Teng CC, Chen Y, Liu J, Patrick L, Shen L, et al. Urban open-air chemical sensing using a mobile quantum cascade laser dual-comb spectrometer. APL Photonics. 2023;8(12):120803.
33. Sylvestre T, Genier E, Ghosh AN, Bowen P, Genty G, Troles J, et al. Recent advances in supercontinuum generation in specialty optical fibers [Invited]. J Opt Soc Am B. 2021;38(12):F90.

34. Petersen CR, Møller U, Kubat I, Zhou B, Dupont S, Ramsay J, et al. Mid-infrared supercontinuum covering the 1.4-13.3 μm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre. Nat Photon. 2014;8(11):830–834.
35. Woyessa G, Kwarkye K, Dasa MK, Petersen CR, Sidharthan R, Chen S, et al. Power stable 1.5–10.5 μm cascaded mid-infrared supercontinuum laser without thulium amplifier. Opt Lett. 2021;46(5):1129–1132.
36. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
37. Jahromi KE, Nematollahi M, Krebbers R, Abbas MA, Khodabakhsh A, Harren FJM. Fourier transform and grating-based spectroscopy with a mid-infrared supercontinuum source for trace gas detection in fruit quality monitoring. Opt Express. 2021;29(8):12381–12397.
38. Zorin I, Gattinger P, Ebner A, Brandstetter M. Advances in mid-infrared spectroscopy enabled by supercontinuum laser sources. Opt Express. 2022;30(4):5222–5254.
39. Bizot R, Tiliouine I, Désévédavy F, Gadret G, Strutynski C, Serrano E, et al. All-fiber supercontinuum absorption spectroscopy for mid-infrared gas sensing. APL Photonics. 2024;9:111303.
40. Vasilyev S, Moskalev IS, Smolski VO, Peppers JM, Mirov M, Muraviev AV, et al. Super-octave longwave mid-infrared coherent transients produced by optical rectification of few-cycle 25-μm pulses. Optica. 2019;6(1):111–114.
41. Vasilyev S, Moskalev I, Smolski V, Peppers J, Mirov M, Muraviev A, et al. Multi-octave visible to long-wave IR femtosecond continuum generated in Cr:ZnS-GaSe tandem. Opt Express. 2019;27(11):16405–16412.
42. Krebbers R, Van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, et al. Ultra-broadband spectroscopy using a 2–11.5 μm IDFG-based supercontinuum source. Opt Express. 2024;32(8):14506–14520.
43. Vasilyev S, Muraviev A, Konnov D, Mirov M, Smolski V, Moskalev I, et al. Longwave infrared (6.6–11.4 μm) dual-comb spectroscopy with 240,000 comb-mode-resolved data points at video rate. Opt Lett. 2023;48(9):2273–2276.
44. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf. 2022;277:107949.
45. Sgobba F, Sampaolo A, Patimisco P, Giglio M, Menduni G, Ranieri AC, et al. Compact and portable quartz-enhanced photoacoustic spectroscopy sensor for carbon monoxide environmental monitoring in urban areas. Photoacoustics. 2022;25:100318.
46. Nazaroff WW, Weschler CJ. Indoor acids and bases. Indoor Air. 2020;30(4):559–644.
47. Galán Madruga D, Fernández Patier R, Sintes Puertas MA, Romero García MD, Cristóbal López A. Characterization and local emission sources for ammonia in an urban environment. Bull Environ Contam Toxicol. 2018;100(4):593–599.
48. Thoning K, Dlugokencky E, Lan X, NOAA Global Monitoring Laboratory. Trends in globally-averaged CH_4, N_2O, and SF_6; 2022.
49. Okabe S, Aoi Y, Satoh H, Suwa Y. Nitrification in wastewater treatment. In: Ward BB, Arp DJ, Klotz MG, editors. Nitrification. Washington, DC, USA: ASM Press; 2014:405–433.
50. Song C, Zhu J-J, Willis JL, Moore DP, Zondlo MA, Ren ZJ. Methane emissions from municipal wastewater collection and treatment systems. Environ Sci Technol. 2023;57(6):2248–2261.
51. El-Fadel M, Massoud M. Methane emissions from wastewater management. Environ Pollut. 2001;114(2):177–185.
52. Bastviken D. Methane. In: Encyclopedia of Inland Waters. Elsevier; 2009:783–805.

53. Rodrigo PJ, Larsen HE, Ashik AS, Vechi NT, Kissas K, Fredenslund AM, et al. Fast horizontal radial plume mapping of N_2O using open-path absorption spectroscopy with a quantum-cascade laser. Atmos Environ. 2024;328:120510.



Chapter 8

Optimizing Data Analysis for Broadband Mid-Infrared Absorption Spectroscopy: A Hybrid Dataset Approach[1]

1. This chapter is based on the following article:
Krebbers R*, Sluyterman LAÆ*, Meurs J, Khodabakhsh A, Cator EA, Cristescu SM. Optimizing Data Analysis for Broadband Mid-Infrared Absorption Spectroscopy: A Hybrid Dataset Approach. *Anal Chim Acta. 2025;344303.*.
* These authors contributed equally to this work.

Abstract

Broadband mid-infrared spectroscopy for gas sensing is a rapidly progressing field with a main frontier in novel mid-infrared ultra-broadband laser sources. These sources provide increasingly complex and high-resolution spectra that are useful for highly sensitive multispecies (trace) gas detection. However, laser-based sources also add a high level of instrument-specific noise and baseline drifts. To efficiently handle the complexity of data and simultaneously overcome the noise and baseline drift, a demand arises to assess and enhance the processing of the acquired spectra. We propose a workflow for optimizing the detection of gas compounds from broadband mid-infrared spectra by training/optimizing models on spectra closely resembling measurement conditions. This is done by combining simulated absorbance spectra with measured (featureless) background intensity spectra to create a hybrid dataset, scalable in size and complexity. Instrument-specific effects, such as baseline drifts, are incorporated and the effect of interfering species in complex mixtures is considered. This approach benefits both commonly used processing techniques, such as classical least squares (CLS) fitting, which can be fine-tuned, and statistical models such as partial least squares (PLS), which demand a (relatively) large realistic training set. In addition, the hybrid dataset can be used to estimate an instrument- and application-specific detection limit *a priori*. The workflow was applied to measurements in the 8 – 11 μm wavelength region to detect trace levels of acetone in CO_2- and water vapor-rich exhaled human breath samples. The workflow presented here provides a means to optimize, train, and assess processing techniques for broadband mid-infrared gas spectra. It can be implemented for any gas absorption spectroscopic instrument with broadband coverage, as long as it is possible to determine the instrument-specific characteristics. Furthermore, other types of processing models, both physical and statistical, can be applied to the workflow and their performance can be evaluated for any instrument-specific spectra.

8.1 Introduction

The mid-infrared spectral region is of high interest for gas spectroscopy due to strong and unique molecular absorption features. For sensitive detection of (trace) gases, spatially coherent sources are preferable as they enable long interaction path lengths with gas samples. The advances in the development of mid-infrared (temporally and/or spatially) coherent broadband sources, such as frequency combs (1) and supercontinuum sources (2), have been the driving force in the field of mid-infrared absorption spectroscopy for multispecies gas sensing (3). The recent breakthroughs in ultra-broadband and ultra-low noise mid-infrared supercontinuum sources (4,5) open new opportunities in trace gas sensing for real-life applications (6). With such sources, it is possible to acquire high-resolution spectra with high detection sensitivity over broad spectral windows. Recent examples demonstrate applications in the field of trace gas sensing (6,7), such as fruit quality monitoring (8), plasma and combustion science (9–11), and open-path spectroscopy (12–14).

Other applications, e.g., breath analysis, have been demonstrated with near- and mid-infrared broadband laser technologies in proof-of-principle experiments (15,16). However, a wide adaptation of these technologies has not happened yet, due to the limited spectral bandwidth and poor sensitivity of early-generation systems.

Following the development of ultra-broadband sources, the field of mid-infrared spectroscopy-based (trace) gas detection is shifting towards more complex data, containing hundreds to thousands of absorption features. An enhancement in the processing of the acquired spectra compared to using classical methods, i.e. least-squares fitting of reference spectra calculated from physical parameters, is thus needed to take full advantage of the broadband coverage and sensitivity of the system.

Generally, different data-processing approaches can be applied to detect the presence and retrieve the concentration of the gas species from the absorption features in a spectrum. Griffiths and De Haseth categorized the approaches used in the broad field of Fourier-transform infrared (FTIR) spectroscopy as classical least squares (CLS), regression techniques (such as partial least squares (PLS)), multivariate curve resolution, and neural networks (17).

Traditionally, the most popular approach is to use CLS models (examples in (6–14), and methodological backgrounds in (18,19)). However, these models face challenges with the increased complexity of the data. Although different models such as neural networks and PLS models may be able to handle this increased

complexity (20), they require large amounts of spectra with known concentrations for training.

To acquire sufficient training data, an intuitive approach would be to train a model on simulated absorbance spectra with known concentrations (21–23). The resulting models, however, might not be able to handle instrument-specific noise effects and baseline drifts when applied to real, measured spectra of the novel ultra-broadband laser-based systems. Instead, a large dataset of measured spectra with precisely known concentrations would be required. This approach was already explored in the 90s and early 00s for classical incoherent FTIRs, but creating such a dataset is highly time-consuming and expensive as it requires many different calibrated gas mixtures with high precision (24–27).

In this paper, we present an alternative approach by constructing a hybrid dataset that is made by combining measured background spectra and simulated absorption coefficients. These datasets can be used to improve the data processing that is currently limiting the full potential of mid-infrared spectroscopy. This hybrid dataset is instrument-specific and contains the relevant noise sources and baseline drifts.

Using simulated absorption coefficients, a large set of spectra with precise concentrations for each molecular species in a complex mixture is provided, while measured blank spectra account for the unique noise patterns and baseline drifts of each specific instrument, resulting in an easily scalable configurable dataset.

The ability to create scalable instrument-specific hybrid datasets offers multiple key advantages:

- Existing data analysis procedures can be fine-tuned and optimized for the specific application due to the availability of a large set of realistic and labeled data.
- New types of models, such as PLS or machine learning techniques such as neural networks can be developed with and trained on the hybrid dataset.
- The dataset can be applied in experimental design. Before conducting measurements of compounds of interest, the hybrid dataset workflow can estimate the detection limit of a certain compound within specific experimental conditions. With this information, the experiment can be tailored to, for example, reach detection limits suitable for the targeted application. Additionally, the simulations provide an estimate of the accuracy of the data analysis.

8.2 Methods

To improve gas concentration retrieval methods from molecular absorbance spectra, we propose creating a dataset that contains all spectral features of the measurement instrument. This dataset serves as a training and testing environment for various data analysis procedures. As the dataset contains realistic spectra for which the concentrations are known precisely, it can be used to optimize our modeling choices, parameters and boundary conditions. Moreover, it facilitates the training of statistical models, such as a PLS model, which requires substantial amounts of accurate training data with precisely known concentrations. The workflow of this approach is depicted in Figure 8.1 and consists of the following steps:

Step 1 -- Measure: The spectroscopic instrument is used to acquire blank spectra with minimal absorption features. Using actual measurements, we incorporate the baseline drift and noise features specific to the instrument used.

Step 2 -- Simulate: The absorption features are simulated for various gas compositions at the measurement conditions of the targeted applications. This is done using physical parameters from databases such as HITRAN (28) or, alternatively, using measurement data from databases such as PNNL (29).

Step 3 -- Combine: The simulated absorption coefficients from Step 2 are combined with pairs of measured blank spectra from Step 1 to create realistic, hybrid absorbance spectra. Combining real measurements with accurately simulated absorption coefficients results in a dataset of absorbance spectra that strongly resembles actual measurements of the specific instrument. In contrast to using measured spectra of calibrated gas mixtures, this hybrid synthetic-authentic dataset can easily be scaled up in quantity, has high precision, and can be readily adjusted to different gas mixture compositions.

Step 4 -- Utilize: To show the possibilities offered by the new dataset, this is utilized to train a PLS model and fine-tune a CLS model. Moreover, the model performance is evaluated. The obtained performance parameters give an *a priori* estimate of the detection limit specific to the spectroscopic setup, gas mixture, and analysis method, which can be used for experimental design. Finally, the trained and optimized models are used on actual measurements of absorbance spectra.

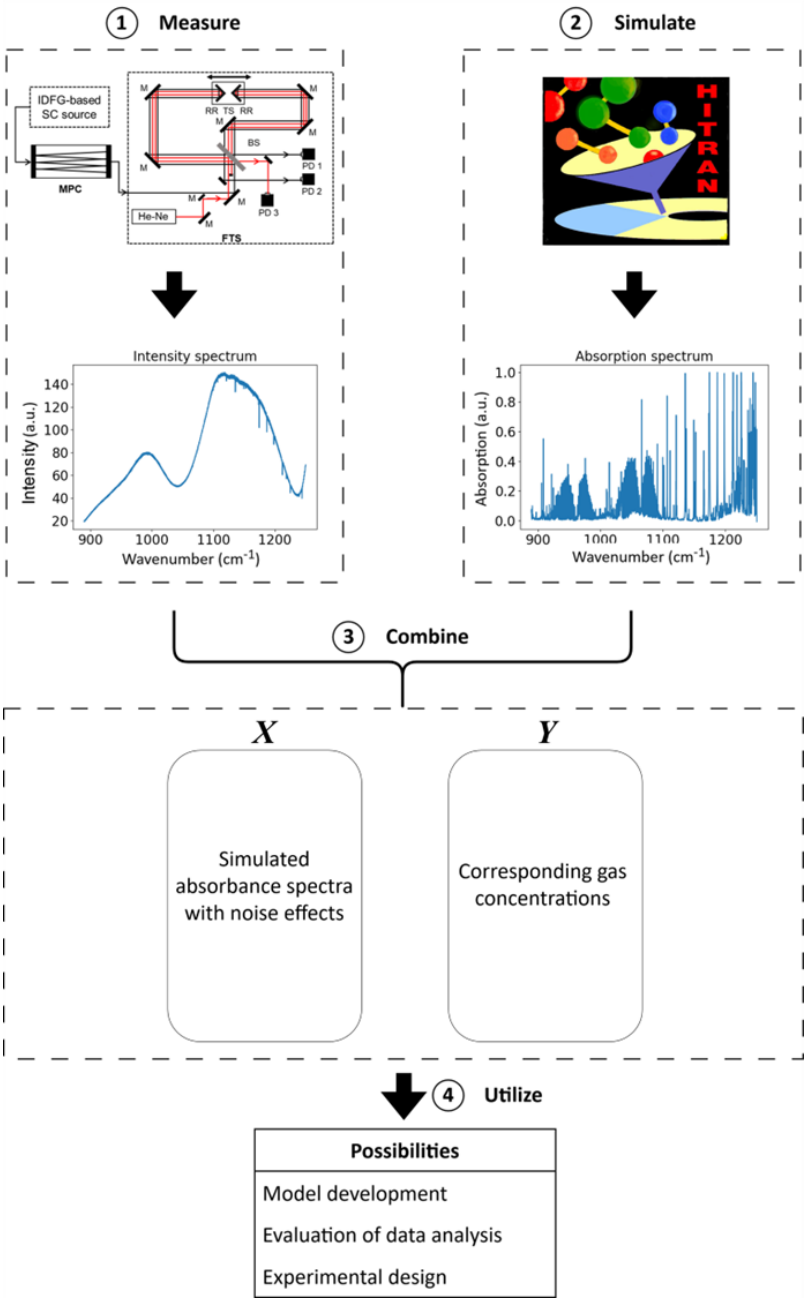


Figure 8.1: A visualization of the four steps of our approach: 1 – Blank power spectra are measured. 2 – Absorption coefficients are simulated for various gas mixtures. 3 – The measurements and simulations are combined to construct realistic absorbance spectra. 4 – The constructed dataset of absorbance spectra (X) with associated gas concentrations (Y) can be utilized for various tasks, such as model development or experimental design.

8.2.1 Step 1: Measuring blank spectra with instrument-specific features

To create a realistic hybrid dataset of absorbance spectra to train, improve, and evaluate our models, it is important to incorporate features such as baseline drifts and noise that are realistic to those found in the measured spectra to which the model will be applied.

The instrument used for the acquisition of our experimental measurement data has been described in more detail in a previous work (6). It consists of a mid-infrared supercontinuum light source, a multipass gas cell, and a Fourier-transform spectrometer (FTS). The light source is an intrapulse difference frequency generation (IDFG)-based supercontinuum source, which provides ~3 W output power as a low-noise and ultra-broadband spectrum in the spectral range of 2 – 11.5 µm. The beam from this source is directed to a Herriot-type multipass gas cell (Thorlabs HC30L/M-M02) with an interaction path length of 31.2 m. The spectrometer is set to a spectral resolution of 3 GHz (0.1 cm^{-1}), at which one measurement scan takes 1.9 seconds. The range between 8 and 11.2 µm (890 – 1250 cm^{-1}) is considered for spectral analysis; this is a water-transparent part of the molecular fingerprint region, even for the long interaction path length and high water vapor content of the considered measurements (Table 8.1), and within the spectral range of the light source used.

For the dataset, power spectra with minimal absorbance features are required, as the absorbance features will be added using simulations (Steps 2 and 3). Therefore, the multipass cell is flushed with a constant flow of nitrogen gas, and measurements (so-called *background* or blank measurements for their absence of absorbance features) are performed several times throughout the day, divided over two days. Each background measurement consists of 310 scans (1.9 seconds per scan, ~ 10 minutes averaging). By measuring at different times of the day, we incorporate power spectrum fluctuations in the measurements, which can be used to create baseline fluctuations in the dataset.

8.2.2 Step 2: Simulating absorption features

The second step, which is described schematically as line 2 in *Algorithm 8.1*, is to simulate a set of absorbance features corresponding to different concentrations of the compounds of interest. The absorbance features in the simulated dataset are mostly built up from physical parameters. The absorbance spectrum as a function of the wavenumber, $\alpha(\nu)$, is defined here as

$$\alpha(\nu) = \kappa(\nu)L = SNX(\nu)L. \tag{8.1}$$

The absorbance is the product of the absorption coefficient (κ) and the interaction path length (L). The absorption coefficient is the product of the integrated line

strength (S), the number density of the particular gas species (N), and the line shape (X). When possible, the absorbance is calculated for a specified concentration, temperature, and pressure using parameters from the HITRAN2020 database (28,30). The line shape is calculated using a Voigt profile. Finally, a sinc convolution is applied to account for the FTS instrumental broadening, due to boxcar apodization.

Using predefined ranges for experimental parameters such as the expected concentrations of the relevant gas species, the interaction path length, gas cell pressure, gas cell temperature, and the spectral resolution, 1000 spectra are simulated, each consisting of the absorption features of a unique mixture of the defined gases. The HITRAN database was used as the primary source for modeling. For molecules not included in the HITRAN database, the PNNL database was used (29). The considered gases and corresponding concentrations are shown in Table 8.1. The choice of gases and concentrations is related to values typically found in human breath (31,32). The high concentrations of CO_2 and H_2O cause the absorbance coefficients of these gases to be highly dominant, making the detection of gases with lower concentrations significantly more difficult. The interesting mixture of molecules poses different challenges to the model. For example, at atmospheric pressure, the narrow absorbance features of CO_2 or CH_4 can be more easily resolved than the broadband absorbance structure of a molecule such as acetone, whose structures are more susceptible to being hidden in noise or baseline drifts. Furthermore, the effect of (highly-absorbing) overlapping features from other molecules might pose challenges.

Table 8.1: Molecules considered in the simulated absorbance spectra, with the range of concentrations used.

Molecule	Simulated concentrations	Database
Carbon dioxide (CO_2)	3 – 6 %	HITRAN
Methane (CH_4)	2 – 200 ppm	HITRAN
Ammonia (NH_3)	50 – 1000 ppb	HITRAN
Methanol (CH_3OH)	0.1 – 2 ppm	HITRAN
Nitrous oxide (N_2O)	0.3 – 1.6 ppm	HITRAN
Ethylene (C_2H_4)	1 – 250 ppb	HITRAN
Isoprene (C_5H_8)	10 – 600 ppb	PNNL
Acetaldehyde (CH_3CHO)	3 – 5000 ppb	PNNL
Acetone ($(CH_3)_2CO$)	1 – 3000 ppb	PNNL
Ethanol (C_2H_5OH)	10 – 1000 ppb	PNNL
Water (H_2O)	1 – 8 %	HITRAN

8.2.3 Step 3: Combining measured spectra with simulated absorption features

To create absorbance spectra, the measured blank background power spectra, $B = \{\boldsymbol{I}^{(1)}, \ldots \boldsymbol{I}^{(n_b)}\}$, are combined with the simulated absorbance coefficients, α, with corresponding concentrations, $\boldsymbol{y}$, $A = \{(\boldsymbol{\alpha_1}, \boldsymbol{y_1}), \ldots, (\boldsymbol{\alpha}_{\boldsymbol{n_c}}, \boldsymbol{y}_{\boldsymbol{n_c}})\}$. First, two spectra, $(\boldsymbol{I}^{(j)}, \boldsymbol{I}^{(k)})$, are randomly selected from the list of blank background power spectra (line 4). The first spectrum of the pair will be used as the background spectrum and the second as the measurement spectrum containing absorption features. Different combinations of power spectra will result in a variety of baseline drifts typical to this spectroscopic instrument. Next, other random noise effects are incorporated via bootstrapping (line 5). This is possible because each blank background spectrum consists of 310 measurement scans (measured in ~10 minutes). These individual scans can be averaged to reduce the noise. A bootstrapped sample, with replacement, of these 310 scans is taken to get a different realization of these measured blank background spectra. These 310 bootstrapped scans of the blank background and measurement spectra, denoted with $(\widehat{\boldsymbol{I}^{(j)}}, \widehat{\boldsymbol{I}^{(k)}})$, are then averaged, as would be done in practice (line 6). Finally, a random pair of absorbance coefficients and compound concentrations, $(\boldsymbol{\alpha}_l, \boldsymbol{y}_l)$, is selected and applied to the averaged and bootstrapped measurement spectrum, $\widehat{\boldsymbol{I}^{(k)}}$ to get an absorbance spectrum, $\boldsymbol{x_i}$, with known compound concentrations, $\boldsymbol{y_i}$ (lines 7 and 8).

Algorithm 8.1: This pseudo-code explains the general simulation approach

1: **Inputs** n_s, the desired size of the simulated set. $B = \{\boldsymbol{I}^{(1)}, \ldots \boldsymbol{I}^{(n_b)}\}$ A set of background measurements. Each background measurement consists of n_2 scans, where $\boldsymbol{I}^{(i)} \in \mathbb{R}^{n_2 \times n_{grid}}$. In our case, n_2=310;

2: Simulate a set $A = \{(\boldsymbol{\alpha_1}, \boldsymbol{y_1}), \ldots, (\boldsymbol{\alpha}_{\boldsymbol{n_c}}, \boldsymbol{y}_{\boldsymbol{n_c}})\}$, where $\boldsymbol{y}_i$ is a vector containing the concentrations of the compounds of interest and $\boldsymbol{\alpha}_i$ are the corresponding absorbance coefficients;

3: For i in range (*1*, n_s):

4: Select a pair, ($\widehat{I^{(j)}}$, $\widehat{I^{(k)}}$) of background spectra from *B* at random;

5: Bootstrap both background spectra by resampling the rows (the different scans) with replacement, denote the resulting pair with ($\widehat{I^{(j)}}$, $\widehat{I^{(k)}}$);

6: Average ($\widehat{I^{(j)}}$, $\widehat{I^{(k)}}$) column-wise, i.e., over the 310 bootstrapped scans. Denote these averages with $\overline{\widehat{\boldsymbol{I}^{(j)}}}$ and $\overline{\widehat{\boldsymbol{I}^{(k)}}}$;

7: Select a random pair $(\boldsymbol{\alpha}_l, \boldsymbol{y}_l)$ from A;

8: $\boldsymbol{x_i} \coloneqq -\log\left(\frac{\overline{\widehat{\boldsymbol{I}^{(k)}}} e^{-\boldsymbol{\alpha}_l}}{\overline{\widehat{\boldsymbol{I}^{(j)}}}}\right)$, and $\boldsymbol{z}_i = \boldsymbol{y}_l$;

9: **Return** $\mathcal{D} \coloneqq \{(\boldsymbol{x}_1, \boldsymbol{z}_1), \ldots, (\boldsymbol{x}_{n_s}, \boldsymbol{z}_{n_s})\}$, a set of simulated absorbance spectra with corresponding concentration vectors.

Since most random effects, such as baseline drifts, are highly instrument- and condition-specific, it is not feasible to simulate these as is done for the absorbance coefficients. However, blank background measurements of an instrument are typically abundantly available and are otherwise not difficult to obtain. A modestly sized set of 20 background spectra results in $\binom{20}{2} = 190$ distinct baseline drifts.

8.2.4 Step 4: Utilizing the dataset

Before utilizing the dataset for developing a new model or improving and evaluating an existing procedure, it is necessary to have well-separated training, validation, and test sets. This is done by separating the blank power spectra and the simulated absorption spectra of known concentrations into three parts (for training, validating, and testing, respectively) and applying Algorithm 8.1 on each of them. A model can be trained on the training set while using the validation set to find the optimal hyperparameters, e.g., the number of latent variables for a PLS model or the type of baseline correction. This is done by creating a script that goes through all the various combinations of parameters and baseline corrections and then selects the model with the best performance on the validation set. The resulting model is evaluated on the previously unseen test set to get an indication of the generalization error.

Crucially, the three different datasets should not share any information. Consequently, the same blank background spectrum cannot be used in the different sets. Similarly, we cannot use the same simulated concentrations with corresponding absorbance spectra. For the experiments presented here, 1000 different concentration combinations with corresponding absorbance coefficients were simulated. These were subsequently split into a training, validation, and test set of sizes 600, 200, and 200, respectively. For each set, we used 15, 7, and 7 distinct blank background spectra, respectively, each spectrum consisting of 310 scans, to create the final training, validation, and test sets using Algorithm 8.1. These final sets consisted of 1500, 500, and 200 absorbance spectra.

8.2.4.1 Application to data processing models: Classical Least Squares and Partial Least Squares

The generated training, validation, and test sets can be applied to various data processing models; in this study these are CLS and PLS. The choice for these models was not made to claim that these are best suited for this type of data, and it is not the goal to provide a comparison between these two approaches, but rather to use them to demonstrate the benefits of our hybrid dataset.

The CLS model fits the measured spectrum to reference spectra that have been constructed either from physical data, such as line strength and broadening parameters (from HITRAN), or experimental data (from PNNL). Thus, this model can

be applied to the test dataset directly without prior training. For identification and quantification of the concentration of the molecular gases in a sample, experiment-specific information on parameters such as the gas pressure and the optical interaction path length inside the gas cell is required to simulate realistic reference data.

The CLS model is optimized by tuning the fit conditions of the reference data to the measurement data. Furthermore, a common practice among users of CLS models is the introduction of polynomial functions in the reference spectra dataset to account for the baseline fluctuations (7,18,29). However, to optimize the model in a statistically sound way, the split between test and training data is relevant. Using the test datasets, the performance of the CLS model has been assessed with and without the fit conditions optimized on the training set applied.

For broadband spectra, which contain hundreds to thousands of absorption features, other methods, such as PLS or neural networks, are attractive alternatives to CLS, as these models are purely statistical and rely on calibration measurements (20), removing the constraint of having to fully reconstruct these complex spectra from physical parameters.

A PLS model reduces the dimensionality of the covariates such that the resulting latent variables and targets have maximum covariance. A regression is then performed on the latent variables. Maximizing the covariance is essential in gas spectroscopy applications since the covariates with the most variance are not necessarily predictive of the concentrations. For example, in a spectrum with (sub-) ppm-levels of acetone and a high percentage-level water vapor content, the absorbance features corresponding to the water features will have a far greater variance but will not be informative for predicting the acetone concentration.

Using the train/validation/test split, a PLS model was trained on the training set while choosing the optimal hyperparameters, such as the number of latent variables, on the validation set. The performance of the model was then evaluated on the unseen test set.

The key difference between CLS and PLS is that the latter is trained to directly identify the relation between the absorbance spectra and the concentrations. For this approach, many spectra of known corresponding concentrations are required. While some examples make use of simulated spectra to build up a training dataset (30–32), we hypothesize that such a noise-free dataset does not suffice for applications to broadband laser-based instruments such as the one used here, which possesses baseline fluctuations and non-uniform noise.

8.2.4.2 Experimental methods

The experiments performed to evaluate the value of the hybrid dataset workflow are separated into three parts. The first section (Section 8.3.1) validates the added

value of creating the hybrid synthetic-authentic dataset over a noise-free simulated dataset, exploring the performance of PLS models trained on either dataset. Note that, unlike the PLS models, the CLS models do not require any training and can be applied directly to the test data. Section 8.3.2 evaluates the performance of PLS and CLS models on the hybrid dataset and explores how the hybrid dataset could aid the optimization and evaluation of the models. The performance is assessed using the train/validation/test split to estimate the sensitivity and detection limit of the data analysis for this specific instrument and target application. Finally, Section 8.3.3 reviews measured spectra of human breath samples to assess whether the results obtained on the hybrid test dataset translate to real, measured data.

The first experiment tests whether it is possible to train a model on absorbance spectra, free from any random effects, instrument-specific noise sources, or baseline drifts. A simpler version of the presented workflow was used to develop a dataset that does not incorporate the instrument-specific noise sources and baseline drifts. This was achieved by only applying line 2 of Algorithm 8.1 and simulating a set of concentrations with corresponding absorbance coefficients, which yielded a dataset of 1000 combinations of 11 compounds (see Table 8.1) with corresponding absorbance coefficients without instrument-specific noise and baseline drifts. This dataset was randomly split into a training and test set of sizes 800 and 200. A PLS model with 30 latent variables (to have roughly three latent variables per compound) was trained on this training set. The targets, i.e., the concentrations, were standardized (by dividing with the standard deviation of the concentrations in the training set) prior to training. The covariates, i.e., the absorbance coefficients, were not pre-processed. The PLS model trained on the simplified dataset was compared with a model trained on the hybrid dataset by evaluating test sets of the simplified and hybrid dataset.

Additionally, the entire workflow is demonstrated for a PLS model using the training, validation, and test sets described before. To optimize the PLS model for realistic hybrid spectra, the model was trained using the training and validation sets of the hybrid dataset, which contains the instrument-specific noise and baseline features. This model predicts the concentrations of all 11 compounds in Table 8.1 simultaneously. The model optimization consisted of the following steps:

- The concentrations were standardized using the training set. Ensuring the concentrations of the different compounds have the same order of magnitude, prevents the model from prioritizing concentration retrieval of one compound over another.
- A Savitzky–Golay filter (SG filter) was used as a baseline correction (33,34). This filter has three parameters: the order of the polynomial, the order of the derivative, and the width of the filter. A second-order polynomial and derivative were used.

- The width of the filter and the number of latent variables were tuned during an exhaustive hyperparameter search. The options for the number of latent variables were [10, 15, 25, 50, 75, 100] and for the width of the window of the SG filter the options were [50, 75, 100, 125, 150, 175, 200]. A width of 1 corresponds to ~0.1 cm^{-1}. A model was fitted for every combination and the combination was selected with the lowest average root mean squared error (RMSE) as

$$RMSE = \sqrt{\frac{1}{n_{val} \cdot n_{\text{compounds}}} \sum_{i=1}^{n_{val}} \sum_{j=1}^{n_{\text{compounds}}} \left(y_{ij} - \widehat{y_{ij}}\right)^2},$$

  where y_{ij} and $\widehat{y_{ij}}$ are the true and predicted concentrations of the *j*-th compound for the *i*-th data point in the validation set on the standardized scale.

Although a CLS model does not require a training set, it is still useful to have realistic spectra of known concentrations. We can leverage the same training set to optimize the fitting procedure of the CLS model and the test set to evaluate its performance. The spectral window, absorbance threshold for removing highly absorbing water lines, and order of the polynomial baseline correction are optimized on the training dataset and evaluated on the unseen test dataset.

Finally, experimental data was gathered to test the optimized models. The measurements consist of breath samples from a healthy participant who provided written consent (the studies involving human participants were reviewed and approved by Research Ethics Committee of the Faculty of Science (Radboud University) under ethical approval number REC22013). The participant exhaled through a mouthpiece into an open-end heated buffer pipe equipped with a CO_2 sensor. As soon as the CO_2 concentration exceeded 5%, an automated valve switched to start sampling the end-tidal part of the exhalation to a multipass gas cell. The gas cell (with a volume of 0.85 L) was filled with two exhalations. The breath composition in the gas cell was measured using the supercontinuum-based FTS system described in section 8.2.1 at a spectral resolution of 3 GHz (0.1 cm^{-1}). After the measurement, a small pump (KNF Laboport N86) was used to transfer the content of the multipass cell to a 3-L Tedlar® bag (MediSense). This bag was incubated at 40 °C for 10 minutes in an oven (Thermo Scientific Heratherm™) before directly connecting it to the inlet line of a proton transfer reaction time-of-flight mass spectrometer (PTR-ToF-MS, Ionicon Analytik PTR-TOF 8000). The PTR-ToF-MS serves as a reference instrument to the results obtained from the PLS and CLS models. To introduce the content of the Tedlar bag in the drift tube of the PTR-ToF-MS, the sampling rate was set to 60 mL/min. The temperature, pressure, and voltage in the drift tube were 80 °C, 2.0 mbar and 480 V, respectively, resulting in a reduced

electric field of approximately 130 Td. Mass spectra were acquired in the range *m/z* 0-234 for 2 minutes at a rate of 2 Hz. After acquisition, the *m/z* scale was calibrated using *m/z* 21.0221 ($H_3{}^{18}O^+$) and *m/z* 203.943 (Ionicon Analytik PerMasCal) and the peak areas were normalized by the number of reagent ions ($H_3{}^{18}O^+ \times 488$). In this study we focused on acetone that was measured at m/z 59.05. For calculation of the concentration of acetone, a calibration series was prepared in the range 234.0 – 936.0 ppm through successive dilution of a calibrated gas standard (Linde Gas).

To prevent condensation and compounds stickiness, all the gas sampling lines were of generally inert material (polytetrafluoroethylene (PTFE) and perfluoroalkoxy alkane (PFA) and together with the gas cell were heated to 40 °C.

8.3 Results

To highlight the improvements enabled by our hybrid data workflow, we evaluated its impact on PLS and CLS model performance by systematically comparing models developed with and without key workflow components. For PLS, we examined the importance of incorporating instrument-specific noise and baseline drifts (step 1 of the workflow) by showing that a model trained without these factors fails to generalize to a realistic setting. Furthermore, for a CLS model, we show that finetuning the fitting procedure on the hybrid data set, for instance what wavelengths to use and how to deal with water saturation, leads to significant improvements gains on previously unseen data. Finally, we demonstrate for both a PLS and a CLS model that leveraging the hybrid dataset results in considerable performance gains on the real-life application of breath measurements.

8.3.1 Training a PLS model on raw absorbance spectra does not work

The hypothesized necessity of incorporating instrument-specific noise sources and baseline drifts in the hybrid dataset was assessed by comparing the PLS model trained on absorbance spectra without any sources of noise with the PLS model trained on data from the hybrid data workflow. An example absorbance spectrum is shown for both the spectra without and with noise in Figure 8.2. In Figure 8.2.c, the baseline drift is visible, while the enlarged section in Figure 8.2.d demonstrates the noise and a small etalon. Figure 3 illustrates that the PLS model trained on absorbance spectra without any random effects makes effectively perfect predictions on an unseen test set that also does not contain any random effects, i.e., yielding a goodness-of-fit (Q^2-value) very close to 1. The figures for the other seven compounds were similar and can be found in Appendix A. If, however, this model is applied to the test data from the hybrid dataset, which was

generated according to Algorithm 8.1 and contains all relevant instrument-specific noise factors and baseline drifts, the model fails to generalize (Figure 8.4). Even for molecules with strong and distinct absorption features, such as CO_2, the predictions are significantly off. Thus, models trained on a simplified training set without instrument-specific noise and baseline drift cannot be used for measured spectra with realistic levels of noise, highlighting the need for a model to be optimized for the relevant instrument-specific sources of noise and baseline drifts.

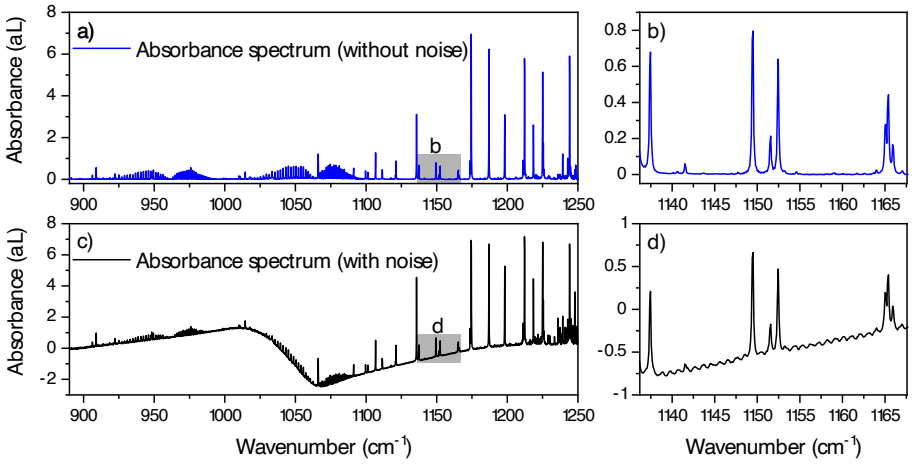


Figure 8.2: Absorbance spectrum of a typical spectrum **a)** without and **c)** with instrument-specific noise and baseline drift. **b)** and **d)** are enlarged segments extracted from spectra a) and **c)**, respectively.

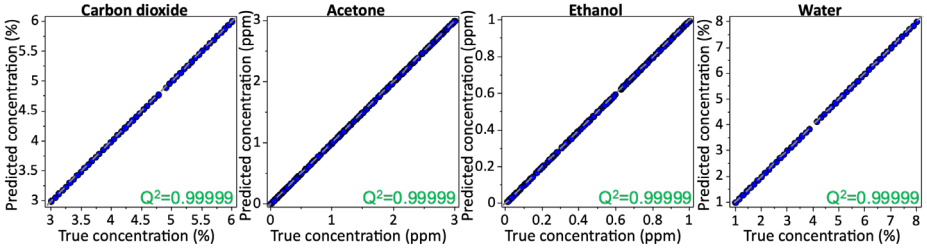


Figure 8.3: Predictions of a PLS model trained on training data without instrument-specific noise and baselines, on test data without instrument-specific noise and baselines. In the absence of noise and baselines, the model can perfectly predict the individual concentrations of a complex mixture on a previously unseen test set.

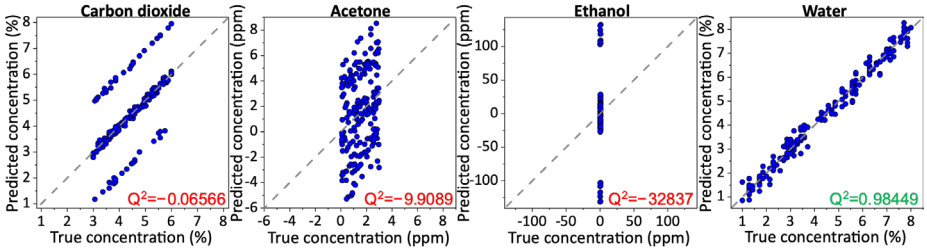


Figure 8.4: Predictions of a PLS model trained on training data without instrument-specific noise and baselines, on test data with instrument-specific noise and baselines. Except for water, the performance is dramatically worse. For ethanol, the model makes predictions that range from roughly -100 to +100 ppm even though the true concentrations range between 0.01 and 1 ppm.

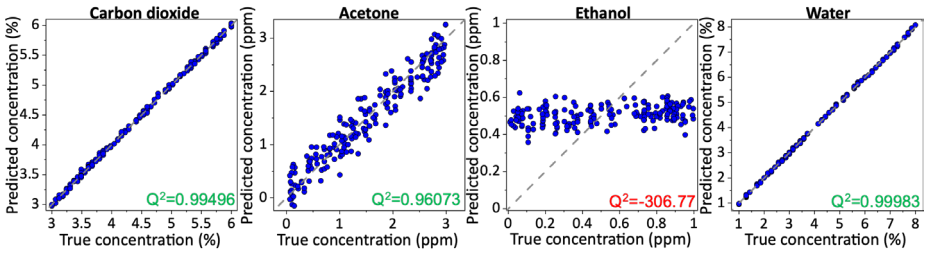


Figure 8.5: Predictions of a PLS model trained and optimized on a training set with instrument-specific noise and baselines, evaluated on a test set with instrument-specific noise and baselines. Carbon dioxide and water are predicted very well. Acetone is also predicted well but has a noticeably higher error. Ethanol is not retrieved well and only an average value is fitted.

8.3.2 Demonstration of the entire workflow

8.3.2.1 PLS models

A PLS model was trained, following the tuning explained in Subsection 8.2.4, on the training set of the hybrid dataset. The optimal hyperparameters were 15 latent variables and a window width of 75. The resulting model was used on the test set of the hybrid dataset, which is the same test set used for the results in Figure 8.4. With this new model, five of the 11 compounds are predicted very well (Q^2-values larger than 0.9). Three compounds are predicted reasonably well ($0 < Q^2 < 0.9$) but suffer from random effects, and three are predicted very poorly ($Q^2 < 0$). For the latter, this is to be expected, as for these species, at the given experimental conditions, their absorbance features are too small to be distinguishable from noise. The results of four of the eleven compounds are given in Figure 8.5. The figures for the remaining compounds can be found in Appendix C. Comparing Figure 8.4 and Figure 8.5, the model trained on the hybrid dataset performs significantly better than a model trained on synthetic data without instrument-specific noise and baseline drifts.

Notably, this combination of hyperparameters is not necessarily optimal for every compound, but gives the best average results. The hyperparameter optimization could also be repeated using the RMSE of one specific compound, e.g., acetone, as the criterion. This is exemplified by optimizing the model for acetone but is in principle possible for any of the gases included in the dataset. This resulted in 75 latent variables and a window length of 75. A model fine-tuned on acetone performance achieved an RMSE of 179 ppb compared to 242 ppb when selecting hyperparameters for good average performance. This highlights that hyperparameters can be tested and optimized for different compounds with the hybrid dataset workflow and that we yield an indication of the model's performance and the detection limit.

8.3.2.2 CLS models

The workflow and the resulting hybrid dataset were also used to optimize and assess the performance of the CLS model similar to the PLS model. First, the CLS model can be applied to the simplified test dataset without instrument-specific noise or baseline drifts. Secondly, the baseline drifts and noise can be introduced to the dataset to see their effect on the fit results when the CLS models have not been optimized. Finally, the CLS model can be optimized on the hybrid training dataset by adding reference spectra of polynomial functions to account for baseline drifts and setting conditions to the fit of the reference spectra to the dataset.

The benefit of a CLS model is that, in contrast to the PLS model, it does not need to be trained on labeled data: it only relies on the set of reference spectra used. However, it is still possible to optimize the parameters of the model to obtain better results. To demonstrate the results from the CLS model and be able to compare it to the PLS model, the same two distinct datasets described in Section 8.3.1 (the simplified and hybrid dataset) were used.

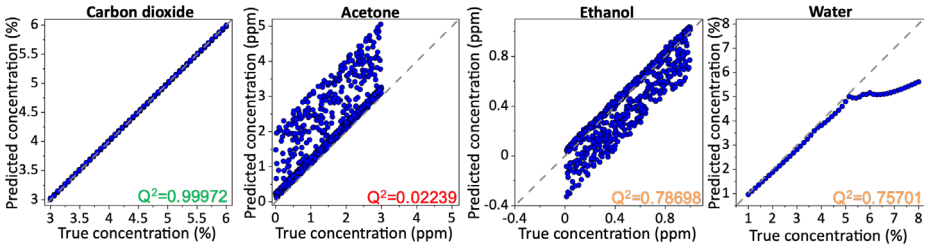


Figure 8.6: Predictions of the CLS model on the test dataset without instrument-specific noise and baseline drift, without further optimizations. For carbon dioxide, acetone, and ethanol, the model gives a linear response, even though in certain spectra the acetone and ethanol are systematically over- and underestimated, respectively. The linear response for water breaks around 5%, where a saturation threshold is reached.

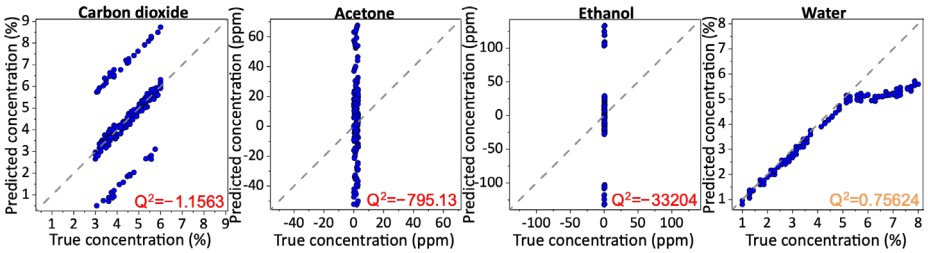


Figure 8.7: Predictions of the CLS model on the test dataset with instrument-specific noise and baseline drifts, without further optimizations. The performance of the model for acetone and ethanol is dramatically poor ($Q^2<0$), with predicted concentrations between -60 to +60 ppm, and -120 to +120 ppm, while the actual concentration of these compounds was between 0.001 to 3 ppm, and 0.01 to 1 ppm, respectively.

The results of the CLS model on the simplified test dataset without instrument noise and baseline drift (Figure 8.6), show a significant discrepancy between the true and predicted concentrations of molecules such as acetone, ethanol, and water. This is not the effect of noise or a baseline drift (as these are still absent in this dataset), but can be linked to saturated water features in the spectrum. Consequently, the predicted water concentration is underestimated for water concentration higher

than 5 percent, and the acetone and ethanol concentrations are, respectively, over- and underestimated when the water concentration exceeds this threshold. This highlights the importance of optimizing the CLS with some constraints to highly absorbing (water) features. In contrast, molecules that do not have overlapping absorption features with water, such as carbon dioxide, can be retrieved well, yielding a Q^2 larger than 0.999.

The CLS model performs significantly worse as soon as noise effects and baseline drifts are introduced, yielding goodness-of-fits (Q^2) lower than 0. Contrary to the PLS model, the CLS model is not trained on this data, and therefore, the model does not have any optimization to handle baseline fluctuations or noise. The fit results and predicted concentrations for each molecule can be found in Figure 8.7 (and for the remaining molecules in Appendix E). Especially for molecules with small absorption coefficients and broadband structures, such as acetone and ethanol, the predicted concentrations can vary by orders of magnitude from the actual concentration, as these molecules are therefore very susceptible to both the introduction of noise and baseline fluctuations. The concentrations of molecules with sharp, distinct, and highly absorbing features, such as CO_2 and water, can still be partially retrieved, as their absorbance features are significantly stronger than the noise and baseline effects.

In Figure 8.8, the results are shown for the optimized CLS model (remaining molecules in Appendix F). Using the training dataset, an optimum was found for the order of the polynomial baseline functions, the spectral range used, and the intensity threshold to ignore saturated water features. The ability to tune the parameters of these optimizations to a training dataset and evaluating it on a test dataset helps to improve the reliability of the model and its overall outcomes and provides a way to visualize the limitations and challenges of the model.

The predicted concentrations of acetone improve significantly from these optimizations, as the saturated water features overlap with the broadband structure of acetone and the broadband structure of acetone also makes the fit more susceptible to baseline drifts. In Figure 8.9, the error in the predicted concentration of acetone is presented as a function of the water concentration in the spectrum for the simplified test dataset without instrument noise and baseline drift, and without optimization of the fit (red) and for the hybrid dataset with instrument-specific noise and baseline drift and with optimizations to the fit (blue). The error in acetone concentration is related to the saturated water features, which is visible in the red data points. The error increases strongly for higher water concentrations.

Using the optimized model, which ignores saturated water features, the dependency of the error on the water concentration vanishes. Thus, in Figure 8.8, the predicted concentrations of water improved for the same reason. The

predicted concentrations of ethanol did not improve sufficiently to discriminate the simulated concentration range of ethanol. This is a useful result to indicate that the simulated concentration range is below the detection limit for the current combination of model and experimental setup, which can be used and considered in the experimental design for an application of this system.

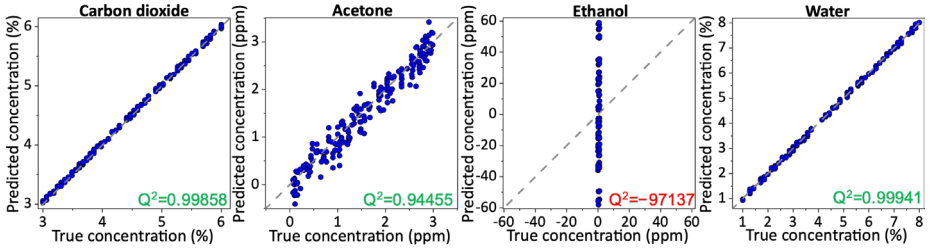


Figure 8.8: Predictions of the CLS model on the test dataset with instrument-specific noise and baseline drift, with optimizations to the fit. Carbon dioxide, acetone, and water are predicted accurately ($Q^2>0$), while ethanol cannot be retrieved well: the predicted concentrations of ethanol range between -60 and +60 ppm, while the true concentrations of ethanol are between 0.01 and 1 ppm.

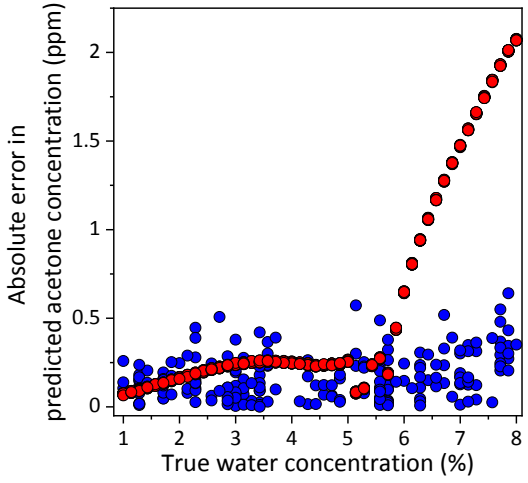


Figure 8.9: Absolute error in the predicted acetone concentration for the CLS model as a function of the applied water concentration on simplified test data without instrument-specific noise and baseline drift, and without optimizations (red), and on hybrid test data with instrument-specific noise and baseline drift, with optimizations to the model (blue). The error and its relation to the water content of the spectrum are effectively reduced by optimizing the parameters of the CLS model.

8.3.3 Retrieving acetone concentrations from measured breath spectra

To put the optimized models to a final test, the hybrid dataset does not suffice. Instead, actual measurements of gas samples should be used to compare the performance of both the optimized PLS and CLS models and test the value of training and optimizing models using the workflow with hybrid data. For this, the measurements of exhaled breath samples of a healthy participant were used.

Using the PLS and CLS models, and the results from the PTR-ToF-MS (which served as an external reference for validation of the models' results), the acetone concentrations of 8 breath samples (measured over one working day) were analyzed. Acetone was chosen for its detectability with PTR-ToF-MS, its known presence in exhaled breath (31,32), and most importantly, for the great difficulty of obtaining reliable concentration readings in the supercontinuum-based FTS system. This difficulty arises from two main elements: 1) The overlap of the weak absorption features of less than a few ppm acetone with the strong, abundant absorption features of percentage-level concentration of water; 2) The baseline drift of the SC-based FTS system interferes with the broad structure of the absorption band of acetone, which has a significant effect given the weak absorption intensity for the expected acetone concentration compared to the amplitude of the baseline drift. This effect could be observed before in the retrieved concentration of acetone in the hybrid test dataset in Figure 8.4 and Figure 8.7 (for the PLS and CLS model, respectively) in comparison to the baseline drift-free dataset in Figure 8.3 and Figure 8.6.

The optimized PLS and CLS models were calibrated using the results from the test dataset (Figure 8.5 and Figure 8.8), which yielded a systematic bias in the PLS and CLS model for acetone of +105 ppb and -60 ppb in the range 0 – 1 ppm, respectively. In Figure 8.10, the performances of the optimized and unoptimized CLS (a) and PLS (b) models are assessed by comparing the retrieved concentrations of acetone from these models with the PTR-ToF-MS (blue dashed line, open markers). The unoptimized CLS and PLS models are the models assessed in Figure 8.4 and Figure 8.7, which have not been trained or optimized using the hybrid dataset. The unoptimized models are off by a large margin. Both unoptimized models demonstrate more than 1 ppm of inaccuracy; particularly the CLS model has a high inaccuracy with unrealistic (negative) values. The CLS and PLS models trained and optimized on the hybrid dataset (see Figure 8.5 and Figure 8.8) yield results that align with the PTR-ToF-MS findings, with the concentrations mostly within the uncertainty margin of the models.

Thus, both CLS and PLS clearly benefit from the hybrid dataset approach, demonstrating a good performance when applied to measurement data. The optimized CLS and PLS models mostly match the PTR-ToF-MS, despite methodological

challenges in measuring the same breath sample with both the SC-based FTS and the PTR-ToF-MS, i.e. transferring the sample from the multipass cell of the SC-based FTS to the PTR-ToF-MS might result in losses due to (bag) wall adsorption effects (35), and dilution with ambient air during the pumping process.

In conclusion, the results indicate that the hybrid dataset workflow enhances the performance of both types of models, yielding good results for the challenging retrieval of acetone concentrations in human breath samples. This shows that the hybrid dataset workflow hits the target of improving the concentration retrieval; the unoptimized models demonstrated systematic and random errors of >1 ppm acetone, while the optimized CLS and PLS models improved the detection limit (RMSE) for the hybrid test data to 210 and 180 ppb. Finally, the performance on real breath samples demonstrates that the optimized models can be applied to real data and that the acetone concentration in human breath samples can be monitored over time for concentrations from high ppb to low ppm.

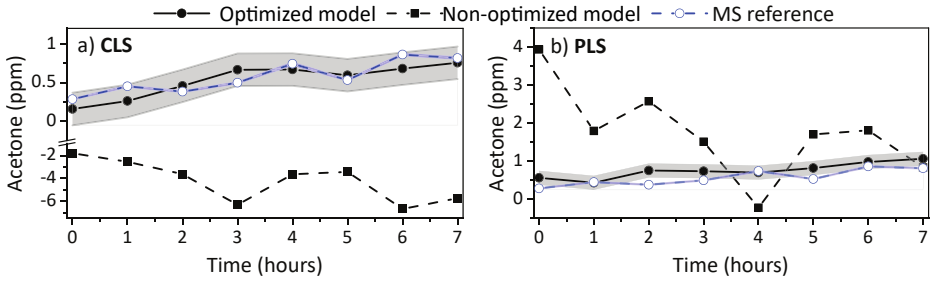


Figure 8.10: Retrieved acetone concentrations in breath samples from one individual throughout the day, using the CLS **(a)** and PLS model **(b)** with an optimized (solid line, round markers) and unoptimized model (dashed line, square markers). The shaded area of the optimized models indicates the uncertainty, given by the RMSE on the test dataset. Mass spectrometry measurements from the PTR-ToF-MS (MS, blue dashed lines, uncertainty as blue shaded area) of the same samples are a reference of the expected concentrations.

Discussion

The workflow presented here leans heavily on the assembly of a hybrid dataset of simulated spectra for precise absorption features and measured spectra for instrument-specific noise and baseline drift. The hybrid dataset serves as a realistic set of spectra that can easily be adapted to different conditions (gas mixtures, measurement parameters, etc.) and scaled to appropriate sample sizes. This approach has three major benefits. 1) The realistic dataset with precisely known concentrations can be used to fine-tune and optimize fitting models, such as CLS. 2) The training dataset provides a way to train statistical models, such as the PLS model, as they require a substantial number of spectra with precisely known concentrations of compounds, which is difficult to acquire merely using measurement data. 3) Using the test dataset, limits of detection and potential limitations for the detection of numerous compounds can be estimated for the specific instrument and application before doing actual measurements.

The results shown here serve as a demonstration of these benefits. We proved that a PLS model can be constructed and optimized using the hybrid dataset, which was used to retrieve acetone from real measurements of breath samples. However, this does not mean that this model is the ultimate choice for these types of measurements, nor that the PLS model could not have been optimized differently. For example, using the workflow presented here, other statistical models, such as neural networks, can also be trained and optimized. For the optimization, other choices could also be easily applied, such as an entirely different type of baseline correction or using a different region of the spectrum.

The hybrid dataset serves to assess the quality of data processing approaches and models on real data. As shown, the quality of retrieving acetone concentrations can be affected by interfering species. In breath samples, the strength of the water absorption features (present in concentrations of 1-8% percent) compared to the acetone absorption features (present in concentrations of up to a few ppm) highly affects accurate retrieval of the acetone concentration in CLS models, as was demonstrated in Figure 8.9. A CLS model tries to fit all features in the spectrum by minimizing the residual spectrum when subtracting the reference spectra from the measured spectrum, and consequently is heavily impacted by unresolvable saturated water vapor peaks. However, PLS should inherently be able to deal with the effect of these high-intensity water peaks better, as the model is trained to these concentrations and adapted to only consider wavenumbers with relevant covariance between the measured spectrum and the molecules' concentrations. In other words, PLS automatically ignores unresolvable features, and thus, in contrast to the CLS models, which try to minimize the residual of the saturated absorption peaks in vain, interfering saturated water lines did not impact the retrieved acetone

concentration in any PLS model. Moreover, it was shown that in an optimized CLS model, the effect of the water absorption peaks could also be effectively reduced by manually introducing thresholds (as shown in Figure 8.8).

A crucial element for the success of this workflow is the quality of the simulations. If the simulated spectra do not resemble the actual measurements that the instrument will perform, the resulting model is likely to give unreliable results. For example, if the sample's temperature or pressure differs from the simulated conditions, or the simulated mixture is missing compounds or contains drastically different mixing ratios, the model outcomes will likely be off. For some of these parameters, such as the effective optical interaction path length, it is possible to perform calibrations instead of constructing a new model from scratch. Since the Beer-Lambert law dictates that the absorbance (A) scales linearly with the absorption strength of a specific gas species (ε), its concentration (c), and the optical interaction path length (L) via $A = \varepsilon l c$, an unaccounted systematic error in the optical interaction path length will cause a systematic error in the retrieved concentration. Using some measured spectra with known concentrations, it will be possible to verify or calibrate the model for any (minor) deviations from the measured spectra and correct such systematic errors.

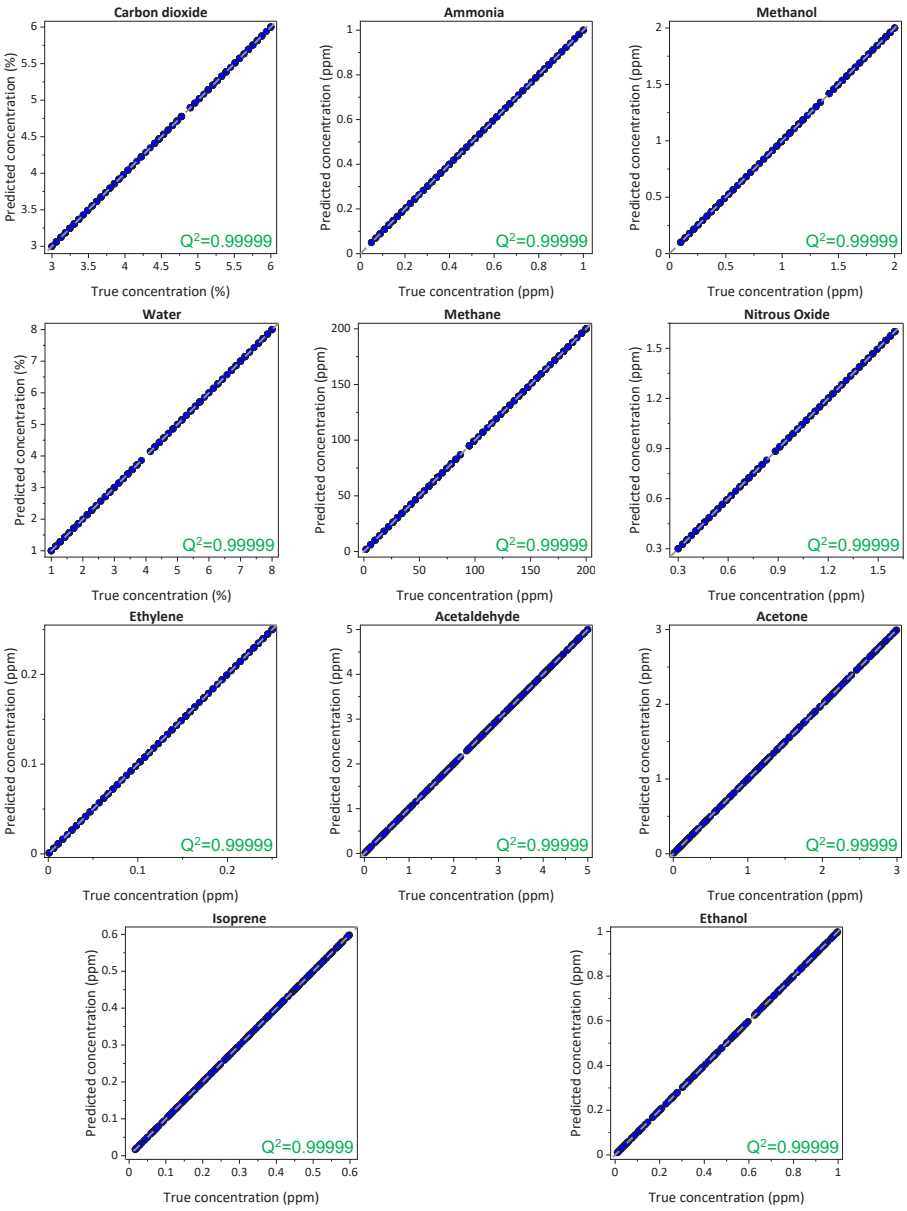
Another important element of the workflow is the incorporation of measured instrument-specific noise and baseline drift, which is a contrast to some previous studies using PLS on spectra of gaseous compounds acquired with a classical FTIR (30). While the advantages of broadband laser-based systems, i.e., the enhanced sensitivity due to the source's coherence and brightness, over classical FTIRs equipped with incoherent lamps are well-known (3), it introduces challenges with non-uniform noise distributions and baseline drifts, compared to the typically white-noise-dominated spectra with low-order baseline drifts in classical FTIR spectra. Consequently, incorporating the measured instrument-specific noise in the hybrid dataset demonstrated a significant improvement in the performance of the PLS model on the spectra of the laser-based instrument used here.

8.4 Conclusion

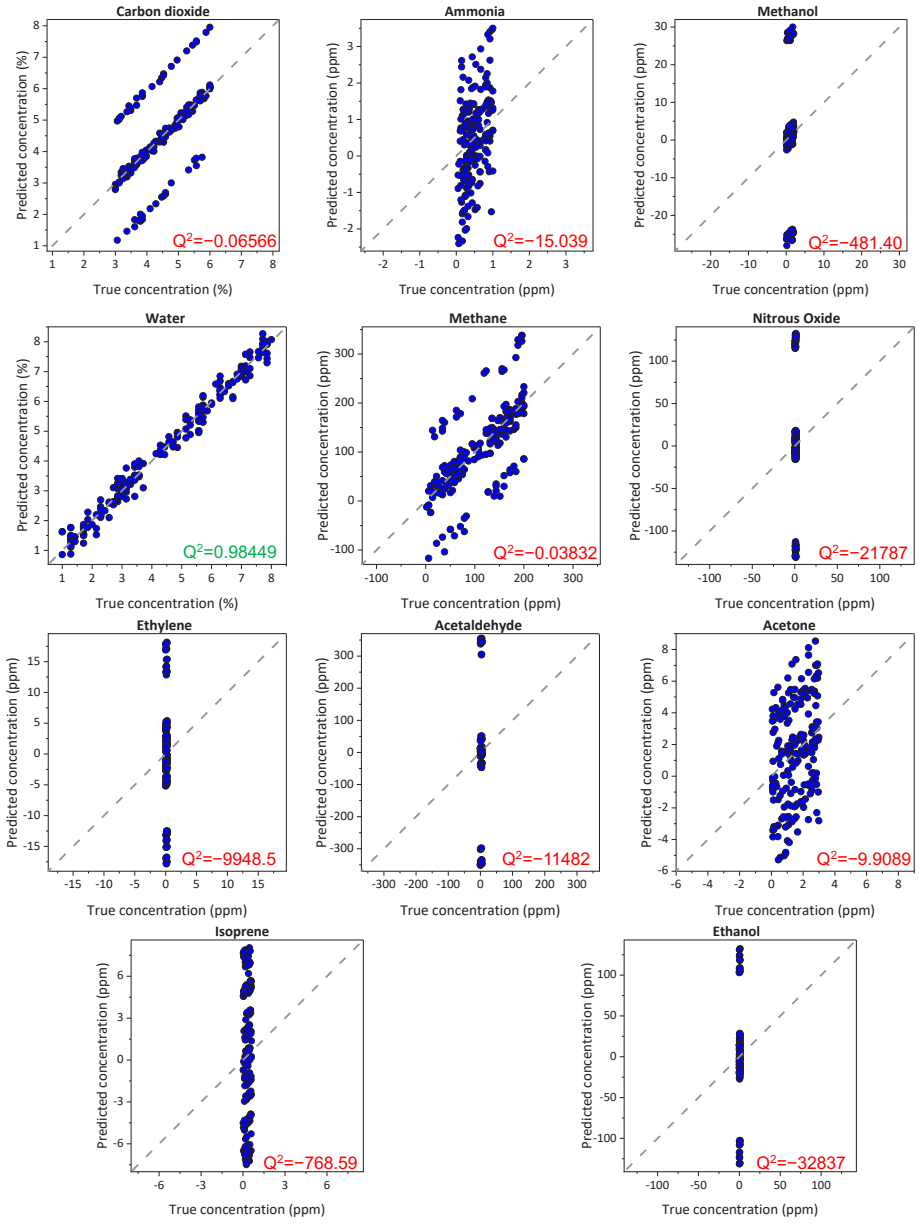
In this paper, we presented a workflow to optimize trace-level gas detection using broadband mid-infrared spectra from laser-based instruments. We created a hybrid synthetic-authentic dataset by combining simulated absorbance features of known gases with blank instrument-specific power spectra. This dataset can be used to improve classical processing techniques like CLS, facilitate the training of statistical models like PLS, or improve experimental designs by estimating system detection limits. Applying this workflow to human breath samples demonstrated that significant performance gains can be achieved both for PLS and CLS models. This approach is broadly applicable to any broadband gas absorption spectrometer and could serve as a vital tool to provide sufficient training data to machine learning models.

Supplementary materials

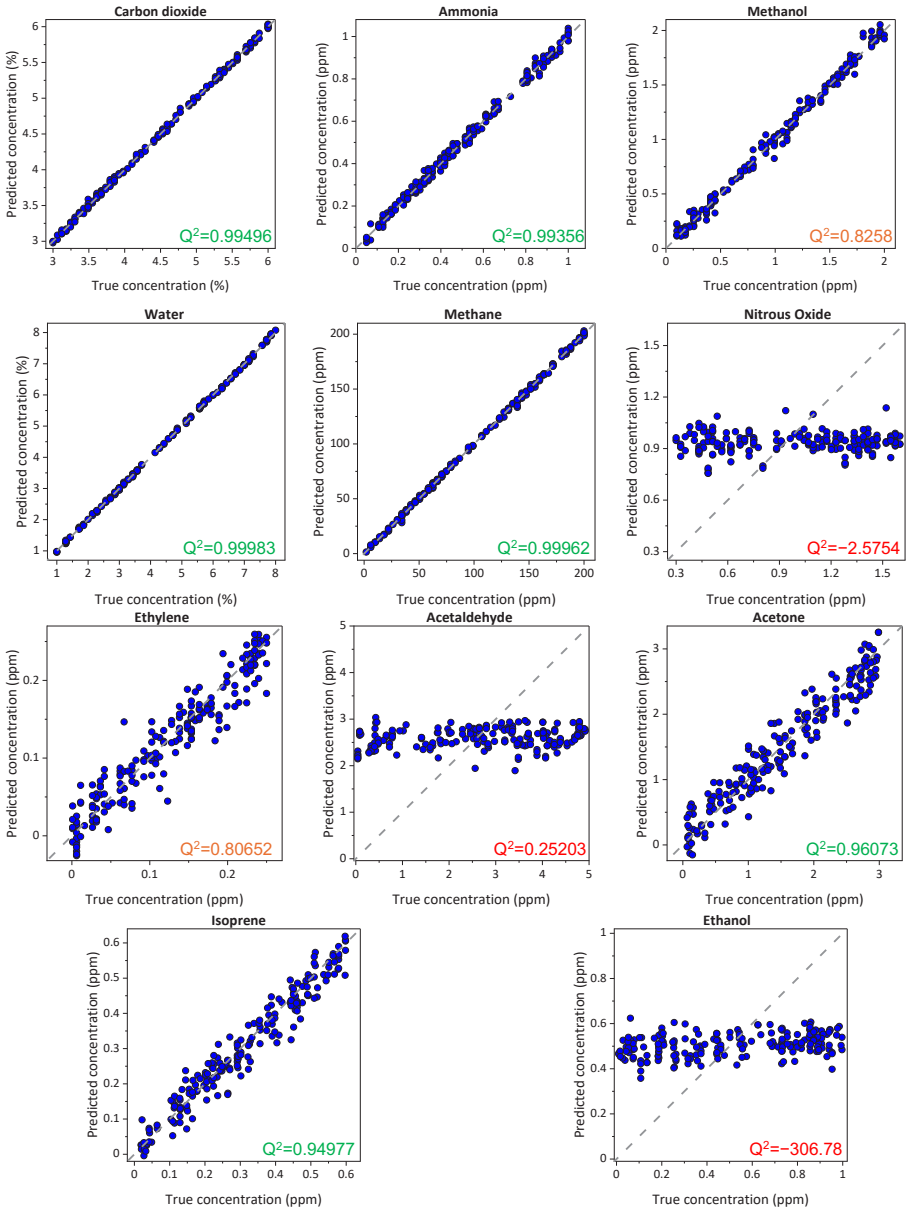
Appendix A: PLS model trained and tested on simulated data without instrument-specific noise



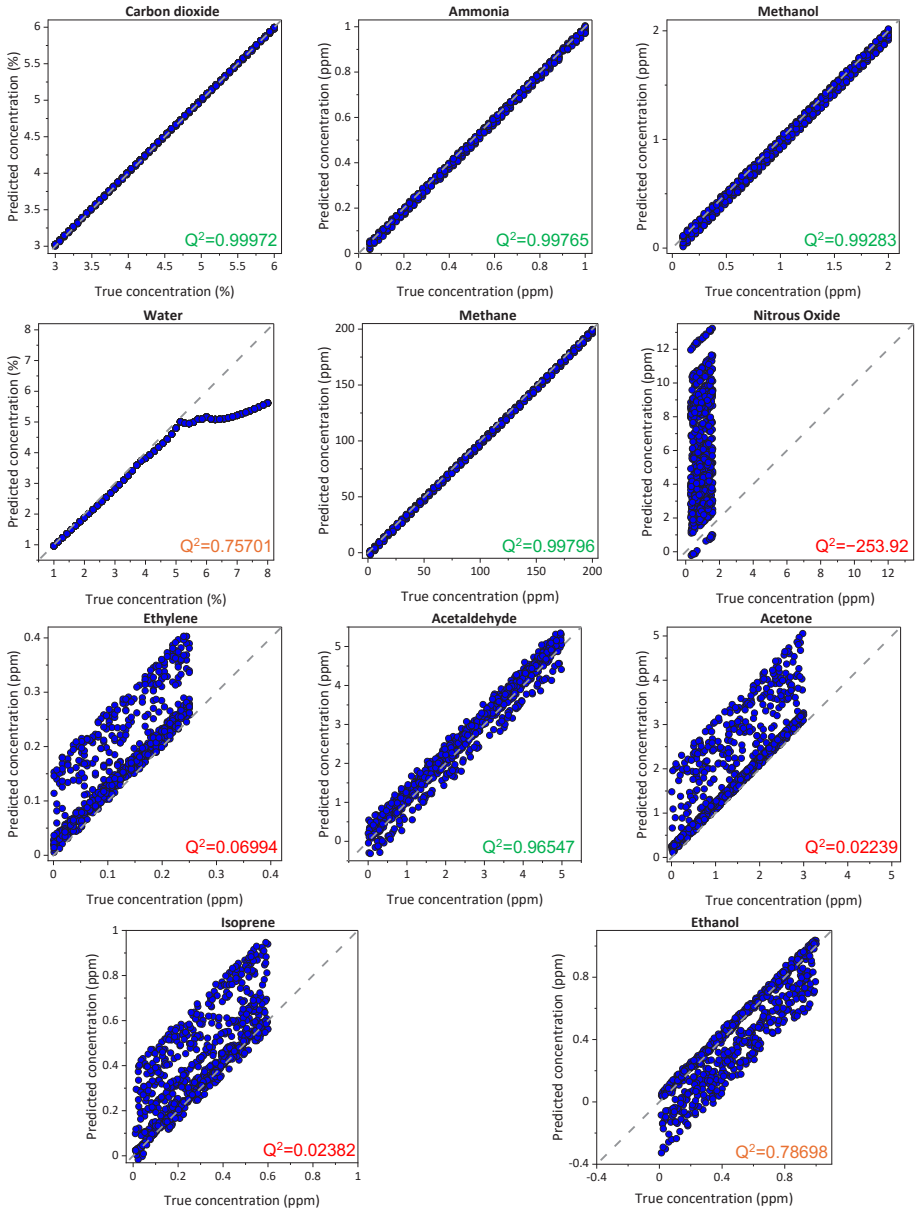
Appendix B: PLS model trained on simulated data without instrument-specific noise, tested on simulated data with instrument-specific noise



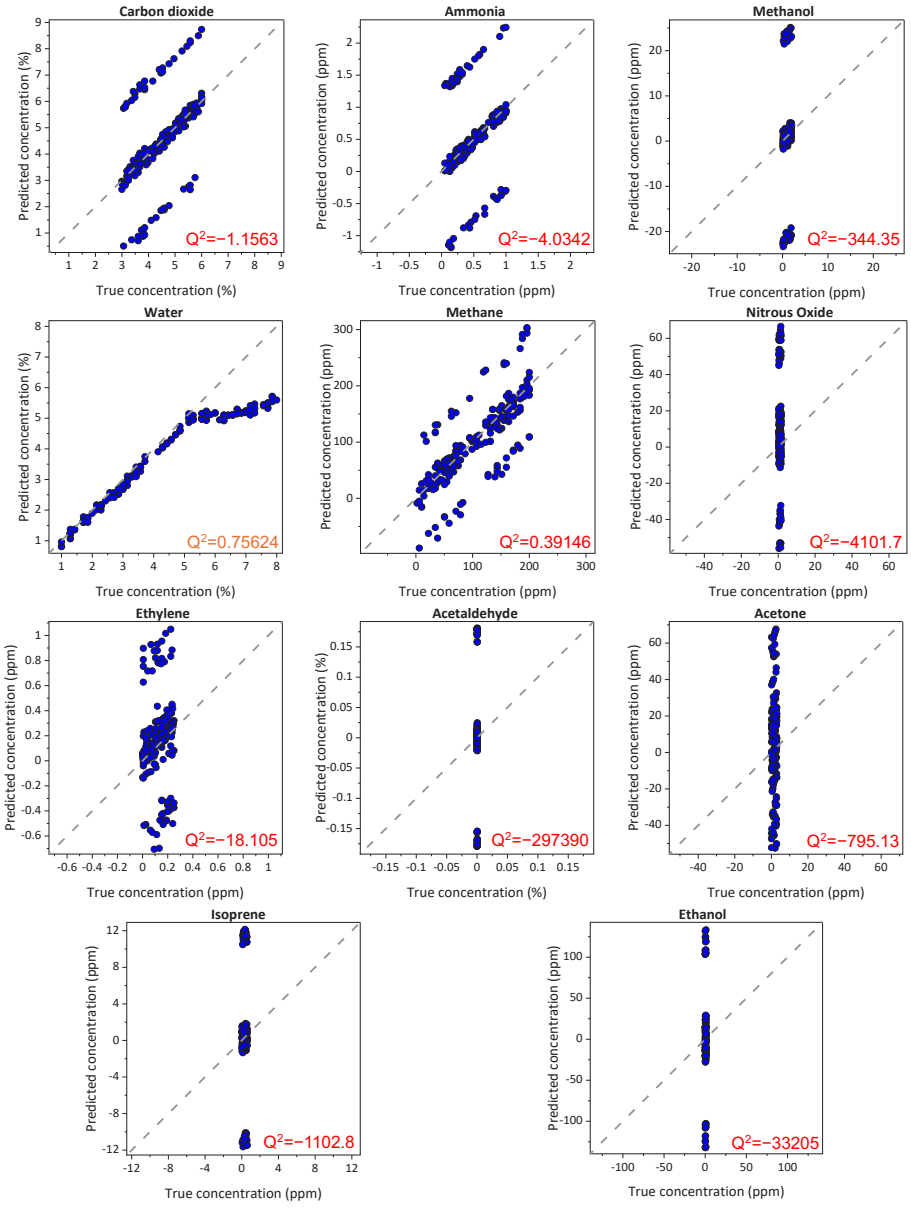
Appendix C: PLS model trained and tested on simulated data with instrument-specific noise



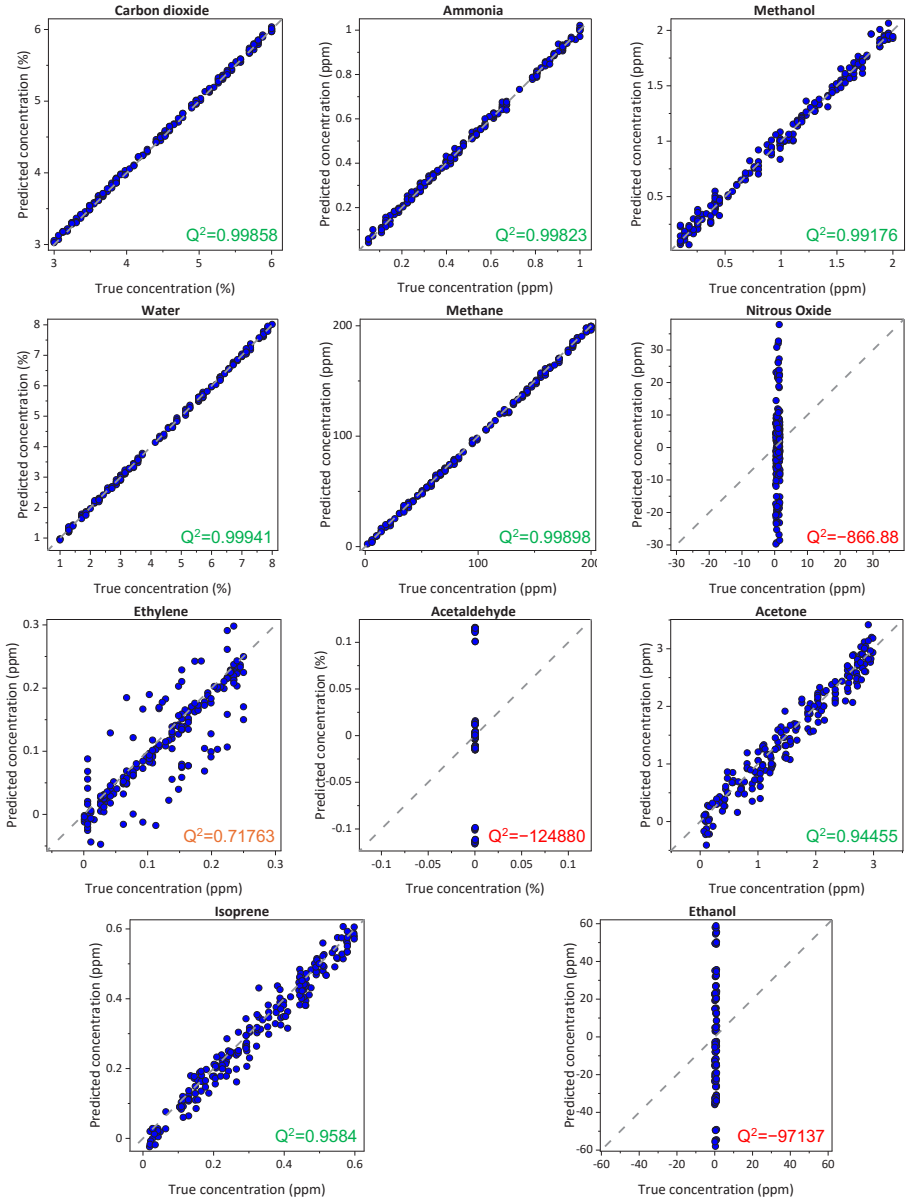
Appendix D: CLS model without optimizations and tested on simulated data without instrument-specific noise



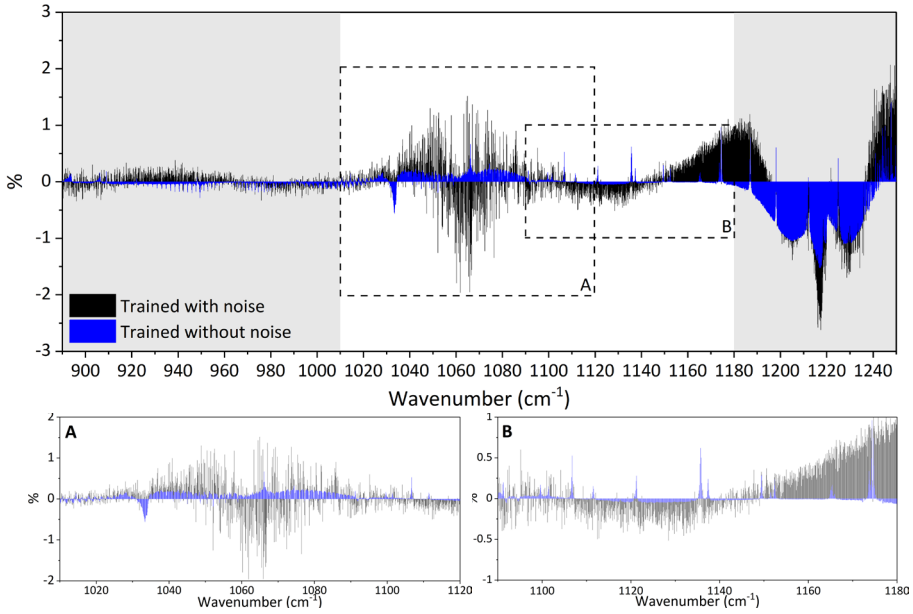
Appendix E: CLS model without optimizations tested on simulated data with instrument-specific noise



Appendix F: Optimized CLS model tested on simulated data with instrument-specific noise



Appendix G: Sensitivity analysis of PLS model



We aim to understand how training a PLS model on absorbance spectra with realistic noise sources differs from training it on raw absorbance spectra. To this end, we trained identical PLS models, each with 15 latent variables and no baseline correction, to predict acetone concentration using the two datasets described in Section 8.3.1 of the main text—one with noise and baseline drifts and the other without. The differences between the two models were evaluated through the following sensitivity analysis.

We first calculated the prediction of each model when given the average of their respective training sets. For each wavenumber, we then evaluated the percentage change in the predicted acetone concentration when adding 0.1 to the absorbance of that wavenumber. This provides an insight not only into the importance of the wavenumbers but also into the direction of the effect of that wavenumber.

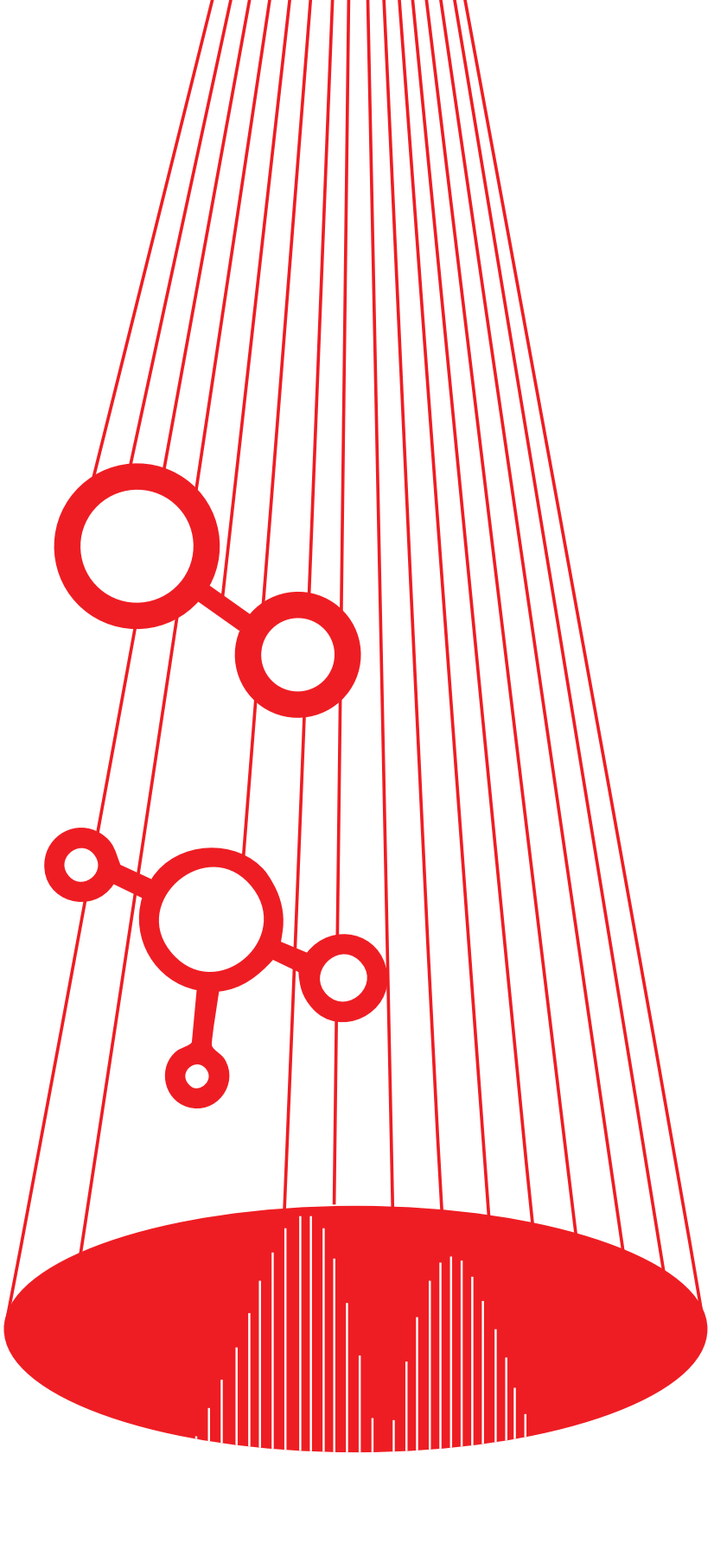
We see two stark differences. Firstly, in the region from 1000 to 1100 cm^{-1}, the model trained on a noisy dataset uses alternating directions, meaning that increasing the absorbance at neighboring wavenumbers has an opposite effect on the predicted concentration. This reduces the sensitivity to a baseline drift. In case of a drift, all neighboring wavenumbers will increase or decrease and using alternating weights mitigates the effect.

The second difference is that we see that the model trained on a noisy data set uses substantially more different wavenumbers as opposed to relying solely on a small set. This is particularly evident in the region 1100 to 1150 cm^{-1}. While in theory, the exact concentration can be determined with only and handful of wavenumbers, this is highly sensitive to noise. By making a prediction based on more different wavenumbers, the model becomes less sensitive to noise.

References

1. Picqué N, Hänsch TW. Frequency comb spectroscopy. Nat Photon. 2019;13(3):146–157.
2. Sylvestre T, Genier E, Ghosh AN, Bowen P, Genty G, Troles J, et al. Recent advances in supercontinuum generation in specialty optical fibers [Invited]. J Opt Soc Am B. 2021;38(12):F90.
3. Shumakova V, Heckl OH. A short guide to recent developments in laser-based gas phase spectroscopy, applications, and tools. APL Photonics. 2024;9(1):010803.
4. Vasilyev S, Moskalev IS, Smolski VO, Peppers JM, Mirov M, Muraviev AV, et al. Super-octave longwave mid-infrared coherent transients produced by optical rectification of few-cycle 25-μm pulses. Optica. 2019;6(1):111–114.
5. Vasilyev S, Moskalev I, Smolski V, Peppers J, Mirov M, Muraviev A, et al. Multi-octave visible to long-wave IR femtosecond continuum generated in Cr:ZnS-GaSe tandem. Opt Express. 2019;27(11):16405–16412.
6. Krebbers R, Van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, et al. Ultra-broadband spectroscopy using a 2–11.5 μm IDFG-based supercontinuum source. Opt Express. 2024;32(8):14506–14520.
7. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
8. Jahromi KE, Nematollahi M, Krebbers R, Abbas MA, Khodabakhsh A, Harren FJM. Fourier transform and grating-based spectroscopy with a mid-infrared supercontinuum source for trace gas detection in fruit quality monitoring. Opt Express. 2021;29(8):12381–12397.
9. Sadiek I, Fleisher AJ, Hayden J, Huang X, Hugi A, Engeln R, et al. Dual-comb spectroscopy of ammonia formation in non-thermal plasmas. Commun Chem. 2024;7(1):110.
10. Krebbers R, Liu N, Jahromi KE, Nematollahi M, Bang O, Woyessa G, et al. Mid-infrared supercontinuum-based Fourier transform spectroscopy for plasma analysis. Sci Rep. 2022;12(1):9642.
11. Makowiecki AS, Herman DI, Hoghooghi N, Strong EF, Cole RK, Ycas G, et al. Mid-infrared dual frequency comb spectroscopy for combustion analysis from 2.8 to 5 μm. Proc Combust Inst. 2021;38(1):1627–1635.
12. Giorgetta FR, Peischl J, Herman DI, Ycas G, Coddington I, Newbury NR, et al. Open-Path Dual-Comb Spectroscopy for Multispecies Trace Gas Detection in the 4.5–5 μm Spectral Region. Laser Photonics Rev. 2021;15(9):2000583.
13. Ycas G, Giorgetta FR, Friedlein JT, Herman D, Cossel KC, Baumann E, et al. Compact mid-infrared dual-comb spectrometer for outdoor spectroscopy. Opt Express. 2020;28(10):14740–14752.
14. Kara O, Sweeney F, Rutkauskas M, Farrell C, Leburn CG, Reid DT. Open-path multi-species remote sensing with a broadband optical parametric oscillator. Opt Express. 2019;27(15):21358–21366.
15. Foltynowicz A, Masłowski P, Fleisher AJ, Bjork BJ, Ye J. Cavity-enhanced optical frequency comb spectroscopy in the mid-infrared application to trace detection of hydrogen peroxide. Appl Phys B. 2013;110(2):163–175.
16. Liang Q, Chan Y-C, Changala PB, Nesbitt DJ, Ye J, Toscano J. Ultrasensitive multispecies spectroscopic breath analysis for real-time health monitoring and diagnostics. Proc Natl Acad Sci U S A. 2021;118(40):e2105063118.
17. Griffiths PR, de Haseth JA. Fourier Transform Infrared Spectrometry. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2007.

18. Adler F, Masłowski P, Foltynowicz A, Cossel KC, Briles TC, Hartl I, et al. Mid-infrared Fourier transform spectroscopy with a broadband frequency comb. Opt Express. 2010;18(21):21861–21872.
19. Kranendonk LA, Caswell AW, Sanders ST. Robust method for calculating temperature, pressure, and absorber mole fraction from broadband spectra. Appl Opt. 2007;46(19):4117-4124.
20. Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, et al. A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. Anal Chim Acta. 2015;879:10–23.
21. Esler MB, Griffith DW, Wilson SR, Steele LP. Precision trace gas analysis by FT-IR spectroscopy. 1. Simultaneous analysis of CO_2, CH_4, N_2O, and CO in air. Anal Chem. 2000;72(1):206–215.
22. Ouyang T, Wang C, Yu Z, Stach R, Mizaikoff B, Liedberg B, et al. Quantitative analysis of gas phase IR spectra based on extreme learning machine regression model. Sensors. 2019;19(24):5535.
23. Ponte S, Andrade JM, Vázquez C, Ferreiro B, Cobas C, Pérez A, et al. Modeling the natural gas knocking behaviour using gas-phase infrared spectra and multivariate calibration. J Nat Gas Sci Eng. 2021;90:103944.
24. Johansen I-R, Lines GT, Honne A, Midtgaard T. Calibration of an FT-IR spectrometer for ambient air monitoring using PLS. Appl Spectrosc. 1997;51(10):1540–1546.
25. Yost MG, Rose MA, Morgan MS. An evaluation of Fourier transform infrared (FTIR) spectroscopy for detecting organic solvents in expired breath. Appl Occup Environ Hyg. 2003;18(3):160–169.
26. Lin C-H, Grant RH, Heber AJ, Johnston CT. Application of open-path Fourier transform infrared spectroscopy (OP-FTIR) to measure greenhouse gas concentrations from agricultural fields. Atmos Meas Tech. 2019;12(6):3403–3415.
27. Makhoukhi N, Péré E, Creff R, Pouchan C. Determination of the composition of a mixture of gases by infrared analysis and chemometric methods. J Mol Struct. 2005;744–747:855–859.
28. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf. 2022;277:107949.
29. Sharpe SW, Johnson TJ, Sams RL, Chu PM, Rhoderick GC, Johnson PA. Gas-phase databases for quantitative infrared spectroscopy. Appl Spectrosc. 2004;58(12):1452–1461.
30. Šimečková M, Jacquemart D, Rothman LS, Gamache RR, Goldman A. Einstein A-coefficients and statistical weights for molecular absorption transitions in the HITRAN database. J Quant Spectrosc Radiat Transf. 2006;98(1):130–155.
31. Mathew T, Pownraj P, Abdulla S, Pullithadathil B. Technologies for clinical diagnosis using expired human breath analysis. Diagnostics. 2015;5(1):27–60.
32. Fenske JD, Paulson SE. Human breath emissions of VOCs. J Air Waste Manage Assoc. 1999;49(5):594–598.
33. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Anal Chem. 1964;36(8):1627–1639.
34. Zimmermann B, Kohler A. Optimizing Savitzky–Golay parameters for improving spectral resolution and quantification in infrared spectroscopy. Appl Spectrosc. 2013;67(8):892–902.
35. Teshima N, Li J, Toda K, Dasgupta PK. Determination of acetone in breath. Anal Chim Acta. 2005;535(1-2):189–199.



Chapter 9

Broadband spectroscopic analyzer with unprecedented profile of biomarkers in exhaled breath[1]

1. This chapter is based on the following article:
Krebbers R, Huisman M, van Kempen K, Meurs J, Khodabakhsh A, Cristescu SM. Broadband spectroscopic analyzer with unprecedented profile of biomarkers in exhaled breath. *In preparation*.

Abstract

Human breath contains hundreds of different compounds, often at trace levels (ppmv or below), that are indicative of metabolic processes in the body. Breathomics analyses these compounds non-invasively, providing insights into the actual state of the biochemical and pathological condition. With infrared spectroscopy, several classes of small molecular compounds can be detected with great selectivity and sensitivity. The narrowband spectrum of most laser-based instruments highly specializes it for one specific compound, while infrared techniques based on thermal light sources (e.g. FTIR) broaden the scope of detectable compounds at the cost of limited sensitivity. Here we introduce a novel spectroscopic approach in breathomics based on broadband infrared laser spectroscopy. This is a versatile analyzer that combines high selectivity and sensitivity with a broad range of detectable species. For this, we built an optical system using a broadband supercontinuum laser with an unprecedented, instantaneous spectral bandwidth of 2.9-11.5 µm (870 - 3450 cm^{-1}) or 2580 cm^{-1} at a spectral resolution of 0.1 cm^{-1}, equipped with a dedicated breath sampling setup for on-line measurements. Through a series of case studies, we demonstrate how this system serves as a platform for breath analysis for the detection of multiple biomarkers, ranging from inorganic compounds (e.g., ammonia), hydrocarbons (e.g., methane, isoprene), and ketones (e.g., acetone), to oxides (e.g., carbon monoxide, nitrous oxide). With sensitivities of tens of ppbv in 3-minute measurements, we were able to trace the changes in the metabolic processes due to fasting, protein intake, and smoking, as reflected in the participants' breath profiles. This analyzer is the first demonstration of a versatile laser-based breath analysis platform, which may further accelerate the proliferation of breath analysis in research and clinical practice.

9.1 Introduction

Human breath contains hundreds of different compounds in concentrations ranging from trace levels (ppbv) to ppmv and percentage levels. Some of these may serve as biomarkers for various health conditions or diseases. Breathomics - a branch of metabolomics, is an emerging field studying exhaled breath, primarily focused on volatile compounds and their biological relevance for human health. Many studies have reported on various methods and analyses on specific compounds or untargeted breath profiles (1,2). However, for a certain metabolite or set of metabolites to move into clinical practice, understanding their biochemical pathways is paramount.

Laser-based optical spectroscopic methods have been a preferred method for breath analysis when a single molecular species is targeted (3,4). Such methods provide fast measurements with high selectivity and sensitivity. Examining the existing FDA-approved breath tests highlights that most of them rely on the detection of a single compound as an indicator for a certain health condition: e.g., nitric oxide as an indicator of symptomatic asthma (5). Moreover, these compounds have known biochemical pathways and, with a few exceptions, are typically well-detectable with optical methods. Whilst the majority of the laser-based systems are highly specialized to a specific compound, the clinical practice might demand a broader characterization of health conditions or diseases. Therefore, measuring multiple breath metabolites simultaneously will provide a broader view of the metabolic processes involved and thus provide the clinician with tools to aid decision-making on diagnosis, prognosis, and treatment.

On the other hand, mass spectrometric (MS) methods such as SIFT-MS (6), PTR-ToF-MS (6,7), and GC-MS (8) are versatile and flexible systems widely used for breath analysis. They are often not focused on a single molecular compound, but rather specialized in classes or groups of molecules. Currently, mass spectrometry-based methods stand as the current golden standard in (un)targeted breath analysis in clinical studies (4,9). Methods such as GC-MS cannot be used in real-time and require long processing times. On the other hand, SIFT-MS and PTR-ToF-MS are used in on-line settings and allow real-time analysis of breath samples (6,7). Furthermore, MS techniques face challenges in identifying isomeric compounds and often rely on reference and calibration measurements to quantify the concentration of a compound. Finally, detecting small molecular compounds, e.g., oxides such as NO_x or carbon oxides, which have high relevance, also for studies on exogenous compounds, poses an additional challenge to these techniques.

The introduction of broadband infrared laser-based spectroscopic methods offers a practical solution that combines the advantages of MS systems and laser-

based spectroscopy. The broadband spectral coverage of the emerging laser sources enables the versatility and flexibility of detecting many molecular compounds simultaneously in a single device, similar to MS systems. Moreover, laser-based spectroscopic methods offer high sensitivity, unambiguous identification, and quantification of molecular compounds. Breath analysis with broadband infrared spectroscopy has been demonstrated using thermal lamps to achieve the broadest spectral coverage (3500 cm^{-1}), at the cost of limited sensitivity (10). On the other hand, mid-infrared (frequency comb) laser-based spectroscopy was used to achieve the highest sensitivity, with the broadest spectral coverage reported being 1010 cm^{-1}(11). It should be noted that this coverage was achieved by combining separate measurements of two different laser systems, covering the range from 1850 to 2230 cm^{-1} and from 2700 to 3330 cm^{-1}. While these studies demonstrated spectroscopic measurements of breath compounds, the setups used are still optical laboratory-bound and lack correct breath sampling and/or on-line measurement capabilities.

Standardization of exhaled breath analysis across different (MS-based) measurement platforms remains challenging. An attempt to move towards a protocol has been proposed and investigated in the *Peppermint Experiment* for different MS instruments (7,12,13). Here, it was found that standardization of the breath sampling is needed as well as data processing to correct for instrumental differences. Drawing lessons from this study, a broadband infrared laser-based system requires a standardized, reproducible breath-sampling strategy. Moreover, absorption spectroscopic methods benefit from inherently reproducible analyses, as it is a direct measurement technique of the compounds present. Therefore, it does not require calibration measurements. Most MS-based techniques, such as SIFT-MS and PTR-ToF-MS, are indirect techniques requiring calibration measurements, as the detectable signal depends on the ionization of the target molecules.

The novel, comprehensive breath analyzer presented here is based on broadband infrared laser spectroscopy and comprises a breath sampling system to directly sample targeted phases of the exhalation. Using a novel, broadband infrared laser source, an unprecedented spectral coverage in the infrared of nearly 2580 cm^{-1} (from 870 to 3450 cm^{-1}) in a single scan is achieved. The spectrum is acquired with a home-built Fourier-transform spectrometer at a resolution of 0.1 cm^{-1} (14). The system is compact, transportable, and functional outside laboratory conditions (as shown in Chapter 7). Sophisticated spectroscopic analyses of the acquired data are used to fully utilize the advantages of the sensitive, broad spectral data (see Chapter 8). This results in the detection and selective identification of many metabolites with known biochemical pathways over a large dynamic range from percentage level to tens of ppb. We demonstrate that simultaneous detection of biologically relevant compounds in breath such as acetone, isoprene, ammonia,

carbon monoxide, nitrous oxide, and methane is achievable with this instrument. Moreover, the effect of several interventions known to trigger changes in specific breath compounds is investigated in three case studies, highlighting the feasibility of the breath analyzer for potential use in clinical settings.

Table 9.1: Typical compounds in breath samples with associated physiological basis. Their typical concentration range in end-tidal breath for healthy persons is given, together with their origin.

Compound	Concentration range	Physiological basis	Reference
Acetaldehyde	0-5 ppm	Ethanol metabolism	(18)
Acetic acid	0-50 ppb	Fermentation of non-digestible carbohydrates by the gut bacteria	(19–21)
Acetone	100-2000 ppb	Decarboxylation of acetoacetate and the dehydrogenation of isopropanol	(22–25)
Ammonia	50-1000 ppb	Catabolism of proteins and amino acids	(25,26)
Butyric acid	0-10 ppb	Fermentation of non-digestible carbohydrates by the gut bacteria	(20,21)
Carbon dioxide	4-6 %	Cellular respiration	(27)
Carbon monoxide	Non-smoker < 2 ppm Smoker 5-12 ppm	Degradation of free hemoglobin and cellular hemoproteins by heme oxygenase in blood	(27–30)
Ethane	0-10 ppb	Lipid peroxidation	(31,32)
Ethylene	0-100 ppb	Lipid peroxidation of poly- and mono-unsaturated fatty acids	(33,34)
Ethanol	10-1000 ppb	Fermentation of carbohydrates by gut microbes	(35,36)
Isoprene	10-600 ppb	Lipolytic cholesterol metabolism in the skeletal muscle	(37–39)
Methane	2-20 ppm	Methanogenesis of gut microbiota	(23,24,40)
Methanol	100-400 ppb	Methionine amino acid metabolism via S-adenosyl methionine or fermentation of methylated compounds by gut bacteria	(23,24,41)
Nitric oxide	0-50 ppb	Oxidation of L-arginine by nitric oxide synthase	(42–44)
Nitrous oxide	300-1600 ppb	Reduction of nitrates compounds by denitrifying bacteria in the gut and oral cavity	(45,46)
Propionic acid	0-50 ppb	Fermentation of non-digestible carbohydrates by the gut bacteria	(20,21)
Water	2-4 %	Cellular respiration	(27)

9.1.1 Metabolites in exhaled breath

The human breath is composed of compounds with both endogenous and exogenous origins. Endogenous compounds are metabolically produced inside the body and thus require methods such as exhaled breath analysis or more invasive studies (e.g., blood tests) to be detected. Exogenous compounds are compounds that, depending on their source, either enter the body through inhalation or other external sources (e.g., food, drinks, medication, cosmetics, etc.). Gaseous or air-borne exogenous compounds are inhaled and enter the bloodstream by alveolar exchange. Measuring both exogenous and endogenous compounds, the effect of exposure to exogenous compounds could be monitored through changes in metabolic processes that can be observed through exhaled endogenous compounds.

Many endogenous compounds well-suited to be measured using infrared spectroscopy have known metabolic pathways, as shown in Table 9.1. They play an important role in several health conditions. For example, acetone serves as a biomarker for diabetic ketoacidosis (15), isoprene is a biomarker for lipid metabolic disorders (16), and ethylene is an early indicator of lipid peroxidation (17).

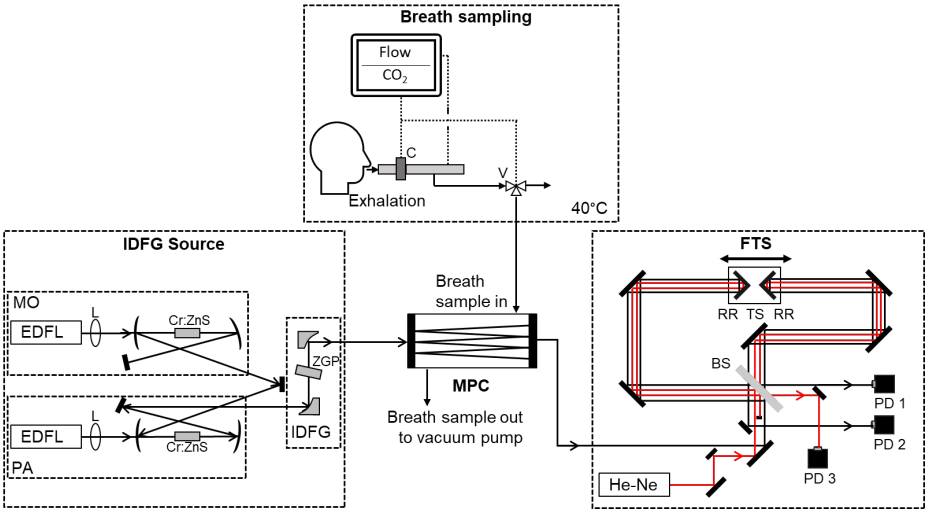


Figure 9.1: Breath analyzer setup with breath-sampling system, laser source, and spectrometer. EDFL: erbium-doped fiber laser, L: lens, MO: master oscillator, PA: power amplifier, IDFG: intrapulse difference-frequency generation, Cr:ZnS: chromium-doped zinc selenide crystal, ZGP: zinc germanium phosphide crystal, C: Capnograph, V: three-way valve, MPC: multi-pass cell, FTS: Fourier-transform spectrometer, He-Ne: helium-neon laser, BS: beam splitter, TS: translation stage, RR: retroreflector, PD: photodetector.

9.2 Methods

In this work, we developed a novel, compact, comprehensive breath analyzer using ultra-broadband infrared laser-based spectroscopy. It consists of three sub-systems: a breath sampling system, a novel ultra-broadband infrared laser, and a Fourier-transform spectrometer (Figure 9.1).

9.2.1 IDFG-based Fourier-transform spectroscopy

An intrapulse difference-frequency generation (IDFG)-based infrared super-continuum source and a home-built Fourier-transform spectrometer (FTS) with turn-key operation were used for spectroscopic investigations in the spectral region 3 – 11.5 µm. The spectroscopic setup has been described in detail in previous work (14,47). Therefore, we restrict our description to the most essential elements.

The light source used is a novel intra-pulse difference-frequency generation (IDFG)-based supercontinuum source built on a Cr:ZnS ultrafast mid-infrared laser platform (CLPF-SC, IPG Photonics) (48,49). It produces a low-noise, ultrabroad 3 W continuum from 2 – 11.5 µm, of which ~300 mW in the 7 – 11.5 µm wavelength range with an 80 MHz repetition rate. The source is a three-stage system, consisting of a mode-locked master oscillator and a single-pass power amplifier, both pumped by an erbium-doped fiber laser, and an IDFG stage based on a zinc germanium phosphide crystal. The output of this source is directed to a multi-pass gas cell (MPC, HC30 L/M-M02, Thorlabs) with an optical interaction path length between the light beam and (breath) sample of 31.2 m. The MPC is equipped with uncoated ZnSe windows (WW70530, Thorlabs) to ensure transmission over the full wavelength range. After interaction with the sample, the light is directed to the Fourier-transform spectrometer (FTS). The FTS is custom-built, using a zinc selenide beam splitter (BSW711, Thorlabs) to split the intensity of the beam into two equal parts propagating to a separate retroreflector (HRR201-P01, Thorlabs) mounted back-to-back on a single translation stage (DDSM100/M, Thorlabs). The beams are recombined on the beam splitter and propagate to two photovoltaic detectors (PVI-4TE-10.6, Vigo Photonics) with optimal detectivity in the spectral region 3 – 11.5 µm. As the interferograms on both detectors have opposite phases, a balanced detection scheme was used, which effectively doubled the interferogram signal while removing common-mode noise (47). The light beam of a He-Ne laser is sent in parallel to the IDFG beam and its interferogram is detected by a separate photodetector (PDA8A2, Thorlabs) to calibrate the optical path difference. A measurement is performed by scanning the translation stage over 2.5 cm, resulting in an optical path difference of 10 cm, which corresponds to a spectral resolution of 0.1 cm^{-1} (3 GHz). The acquisition

of a single spectrum takes roughly 1.9 seconds. The spectra are averaged for 100 scans to reduce the white noise in the spectrum and improve the sensitivity.

9.2.2 Breath sampling and case study design

The breath sampling system is designed to allow breath phase selection in order to directly sample a targeted portion of the exhalation. On-line sampling enables direct measurements of the breath samples. Participants exhale through a mouthpiece into an open-end buffer pipe at a self-regulated flow of 50 mL/s. An interface (GSI, Loccioni) for the participant provides feedback to self-regulate the exhalation flow rate for measurement standardization and indicates the exhalation phase using the CO_2 concentration measured by capnography. The capnograph (Capnostat 5, Philips) located between the mouthpiece and buffer pipe is used to trigger a three-way valve to open towards the MPC as soon as the CO_2 concentration in the exhalation reaches 5 %, ensuring only the end-tidal part of the exhalation to be sampled into the MPC. Two consecutive exhalations are sufficient to fill the MPC (volume: 0.85 L). The entire system in contact with breath samples is heated to 40 °C to prevent condensation of relevant compounds and all gas sampling lines were made of the inert materials polytetrafluoroethylene (PTFE) and perfluoroalkoxy alkane (PFA).

Three different case studies were designed to demonstrate changes in the exhaled breath composition of participants over time. The samples are taken from healthy participants who provided written consent; the studies involving human participants were reviewed and approved by Research Ethics Committee of the Faculty of Science (Radboud University) under ethical approval number REC22013. The case studies consisted of interventions, e.g., fasting, consuming a protein-rich meal, or smoking. In the case of the protein-rich meal and smoking, the t_0 measurement was taken before the intervention, while for the fasting case study, reference measurements were taken of the participants on a day without dietary restrictions. The participants of the fasting studies were asked to refrain from eating and consuming beverages (except for drinking water) after dinner on the evening before the study. During the measurement day, the participants were only allowed to drink water.

For the smoking study, regular smokers were asked not to consume a cigarette (or equivalent) on the morning before the study. After a t_0 measurement, the smokers consumed one Marlboro Gold cigarette. The first sample after smoking was taken within 10 minutes of smoking the cigarette.

For the protei n study, the participants did not consume any food in the 12 hours prior to the experiment. Following a t_0 measurement, the participants consumed a protein-rich shake (HiPRO Protein drink) with milk-based proteins yielding an equivalent of 0.83 g/kg of protein per body weight ratio. Subsequently,

the participants provided a breath sample hourly. During the entire experiment, the participants did not consume any food and were asked to only drink water.

9.2.3 Data processing

The interferograms of the He-Ne and IDFG source were digitized using an analog-to-digital converter (NI 6251, National Instruments). Initial processing of the IDFG interferogram, such as resampling the optical path difference, obtaining the (power) spectrum through FFT, and averaging multiple scans, are performed in a LabVIEW environment. In postprocessing, the spectra were normalized to absorbance spectra and a least-squares fit was used to fit reference spectra simulated from the HITRAN2020 database (50) or retrieved from the PNNL database (51). The fit was optimized using baseline correction with polynomial functions, setting the spectral window used for fitting, and introducing a threshold discarding highly absorbing absorbance features from the fit.

9.2.4 Statistical Analysis

The precision of the measurement instrument is compound-specific, due to the specific absorption strength of the compound, its absorption profile, its spectral window (as the IDFG source has a non-uniform spectral density), and potential spectral overlap with competing compounds. Therefore, the precision is estimated for each sample by measuring each sample five times. As the composition of the sample in the MPC is expected to be stable, deviations within the five measurements of the same sample are attributed to the precision of the instrument and shown as their error margin in Figures 9.4-9.7.

9.3 Results

Eighty-four breath samples from eight healthy volunteers participating in three control-case studies, were measured. The breath samples contained the end-tidal portion of exhaled breath, where blood-gas exchange occurs. The transmission spectra acquired from the samples have a spectral coverage of 3 – 11.5 μm with a spectral resolution of 0.1 cm^{-1} (3 GHz). One measurement of a breath sample is acquired in ~3 minutes. An example of a typical transmission spectrum of a breath sample is shown in Figure 9.2.

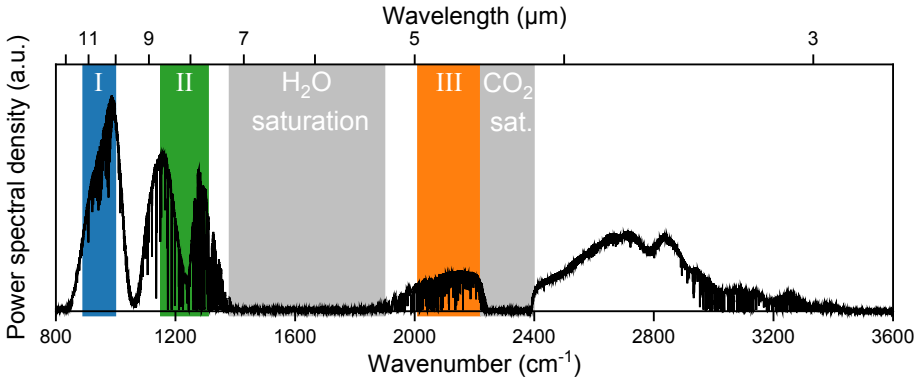


Figure 9.2: Power spectral density of the ultra-broadband source. The spectrum of the IDFG-based source was measured by the FTS at 3 GHz spectral resolution after passing through a breath sample. Strong absorption features from water and carbon dioxide are observed. Relevant compounds in breath can be found absorbing throughout different parts of the spectrum, of which regions I, II, and III are shown in Figure 9.3.

Transformation of the intensity spectra to absorbance spectra is required to evaluate the absorption by various molecules in the sample. This conversion is done in two ways: to either retrieve the absolute concentrations of the molecules in one sample or to retrieve a trend or relative change in concentrations within a measurement series of one participant. To extract the absolute concentrations of the molecules present in the breath sample, the measured intensity spectra are normalized to a blank measurement of dry nitrogen gas. Absolute concentrations are of interest when comparing to base levels of the concentrations of relevant molecular compounds, or when performing a comparative study between participants. As shown in Figure 9.2, significant parts of the intensity spectrum are completely absorbed by water and carbon dioxide (grey areas), some of the most abundant compounds in exhaled breath. Therefore, the absorbance spectrum is zoomed to specific parts of the spectrum (regions I, II, and III), with maximal absorption of relevant compounds and minimal absorption of these interfering species. However, although the absolute-concentration absorbance spectrum is still dominated by absorption features related predominantly to these two compounds, the high spectral resolution facilitates the discrimination of the presence of the other compounds and the retrieval of their concentrations (Figure 9.3, left).

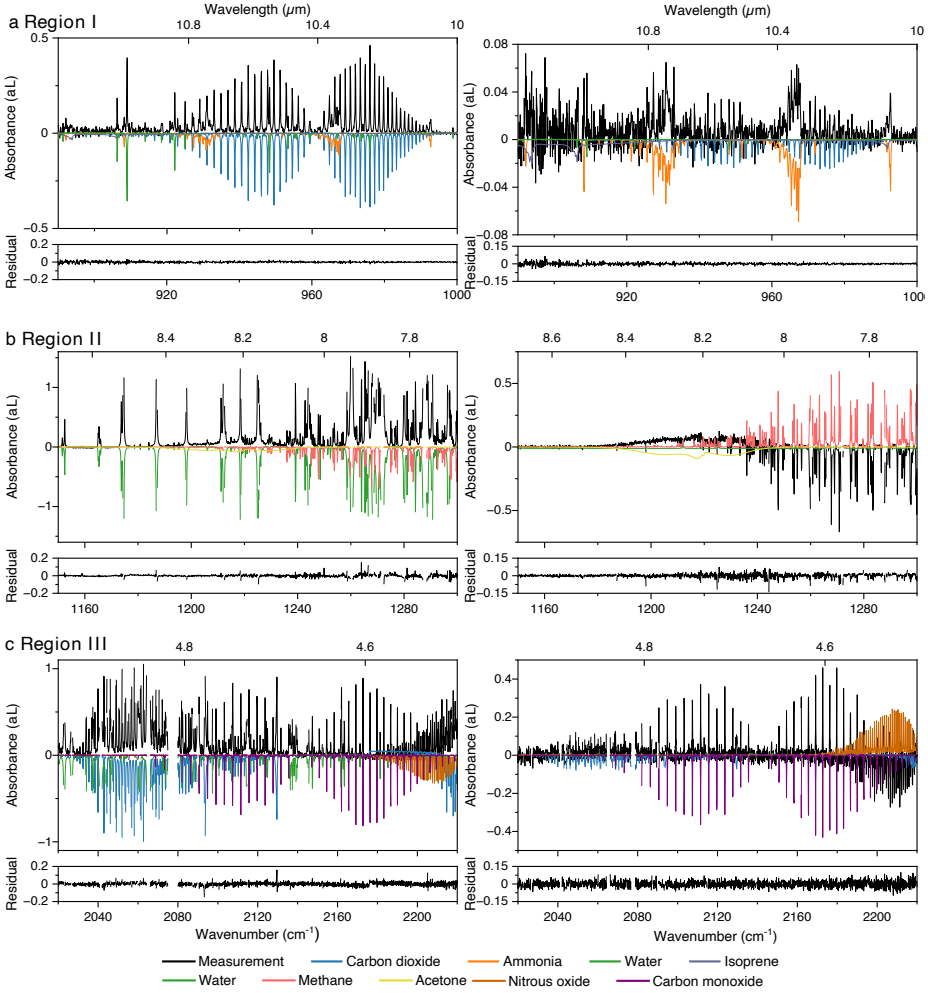


Figure 9.3: Measured absorbance spectra with fitted reference spectra and retrieved concentrations for regions I, II, and III in Figure 9.2. Absolute-concentrations absorbance spectra are shown left (normalized to blank, nitrogen gas sample). Relative-concentrations absorbance spectra are shown right (normalized to a baseline breath sample of the participant).

Relative-concentration absorbance spectra are acquired by normalizing consecutive spectra of a participant to a baseline measurement of this individual (i.e., a measurement before the intervention). This can be of particular use in tracking the effect of a certain intervention over time. As only the absorption peaks appear in the spectrum that relate to changes in concentrations, the dominant water and carbon dioxide peaks vanish. A tidy spectrum remains in which unknown or unexpected molecules can be identified and quantified. For example, in the spectra normalized

to another breath sample (Figure 9.3, right), the absorption peaks of ammonia and acetone are more visible than in the spectra normalized to nitrogen gas (Figure 9.3, left) because gas species at stable concentrations (such as CO_2 and H_2O) are removed from the spectrum. This results in more accurate fits and retrieval of ammonia and acetone concentrations in the sample.

Fits of reference spectra over the entire measured absorbance spectrum return an overview of all retrieved gas concentrations in the sample, which can be used to further analyze the composition of the exhaled breath (Figure 9.4a and 9.4b).

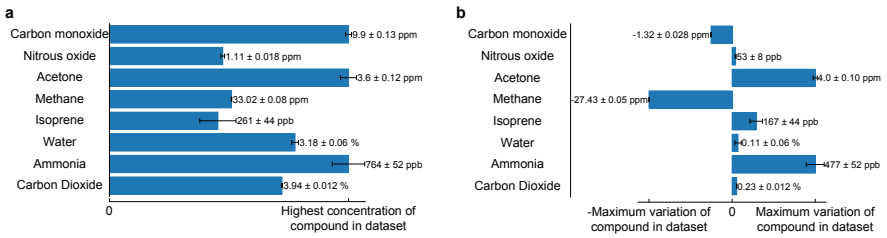


Figure 9.4: Corresponding retrieved concentrations of the spectra in Figure 9.3 for the absolute-concentrations spectra **(a)** and relative-concentrations spectra **(b)**. The length of the bar corresponds to the highest retrieved concentration of the specific compound across the 84 breath samples measured.

9.3.1 Retrieving the end-tidal concentration through carbon dioxide-normalization

A common practice within the breath analysis community is to provide concentrations of the compounds retrieved in a breath sample in relation to the carbon dioxide content of the sample (52). By doing this, any concentration differences between samples related to the sampling of the exhalation are minimized. Such differences may arise from how the participant exhaled or through instrumental sampling errors, causing a different portion of the breath to be sampled. As the carbon dioxide concentration reaches a plateau level and is highly reproducible within and between participants, normalization of the carbon dioxide content of the breath sample is a generally accepted way to compensate for such changes. The effect of this normalization is shown in Figure 9.5. The ammonia concentration that was retrieved from participants that consumed a protein-rich meal, is expected to show an upward trend. However, after 2.5 hours, the ammonia concentration of participant P2 shows an unexpected decrease. The retrieved carbon dioxide concentration (Figure 9.5b) shows a similar decrease for this specific sample. Therefore, this is very likely the result of a sampling deviation. As the ability to measure carbon dioxide simultaneously with the compounds of interest enables normalization to find the real end-tidal concentration of ammonia, the sampling error can be corrected (Figure 9.5c).

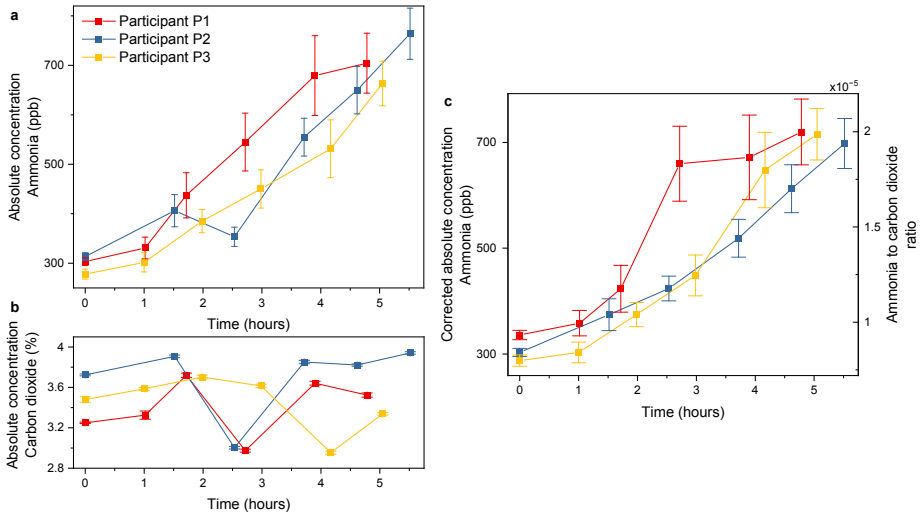


Figure 9.5: The effect of normalization to carbon dioxide on the retrieved ammonia concentration. Retrieved absolute concentrations of ammonia **(a)** and carbon dioxide **(b)** during the case study on the consumption of a protein-rich meal. Compensating the dilution of the end-tidal fraction of the sample using the retrieved carbon dioxide concentration yields a corrected concentration of ammonia in the end-tidal breath **(c)**.

9.3.2 Tracking metabolic changes and exposure: case studies

Three case studies were designed to showcase the versatility and usability of the breath analyzer. The studies were designed to highlight the detection and quantification of different classes of compounds, ranging from inorganic compounds (e.g., ammonia), hydrocarbon compounds (e.g., methane, isoprene), and ketone compounds (e.g., acetone), to oxides (e.g., carbon monoxide, nitrous oxide) and alcohols (e.g., methanol). All compounds were retrieved in a single measurement without alterations to the instrument.

In the first case study, participants were asked to consume a protein-rich meal, after which their breath was sampled hourly. Protein intake is a known source for the generation of ammonia in the body through the catabolism of amino acids (53,54). Consequently, the ammonia content is expected to increase after protein consumption (25,55). Thus, in our case study, we demonstrate the ability to verify the findings from a dedicated ammonia monitor based on thermal oxidation (used by Spacek et al. (55)) or a selected ion flow tube (SIFT) mass spectrometer (used by Smith et al. (15)). The participants demonstrated the expected increase in ammonia after protein consumption (Figure 9.6). Furthermore, we stress how, in contrast to the aforementioned techniques, we can monitor multiple molecular species simultaneously, enabling normalization (e.g., using carbon dioxide), or observing

trends in other compounds, such as the increase in isoprene concentration of participant P3. Finally, the multi-species detection also allows for breath profiling of the participants: breath samples of participant P1 can be clearly distinguished from the other participants for its increased acetone and nitrous oxide content, which can be observed in Supplementary Materials Figure S1.

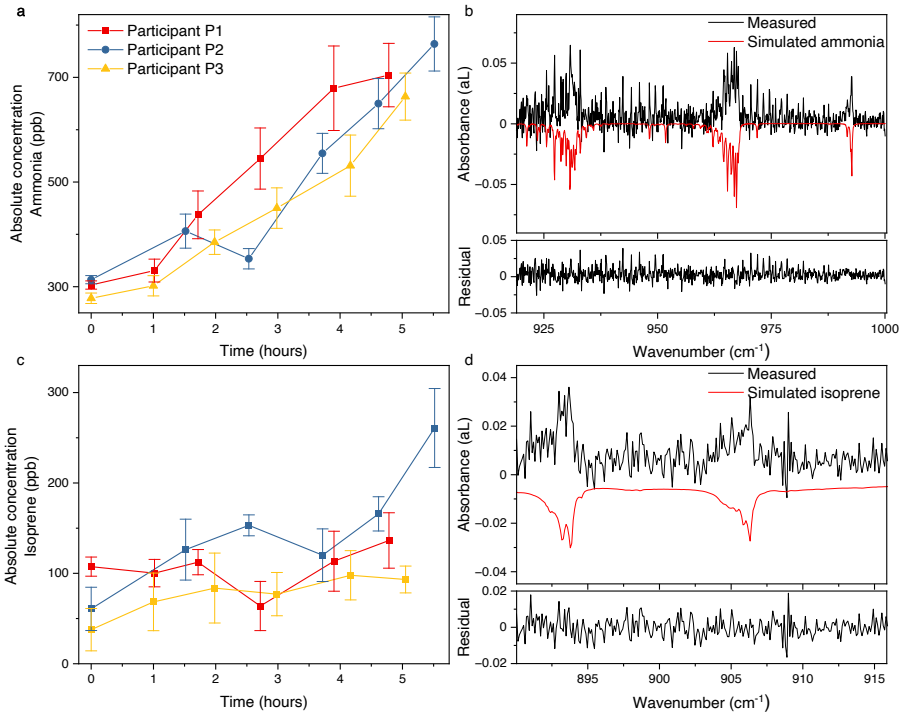


Figure 9.6: Retrieved concentrations of compounds in breath after protein intake. Retrieved absolute concentrations of participants after protein intake, demonstrating a related increase in ammonia **(a)** and isoprene **(c)**. **(b)** and **(d)** Measured (black) and simulated (red, inverted) absorbance spectra of the related compounds with associated residual of the fit.

In a second case study, three participants fasted overnight. Their breath profile was measured over the course of one working day, about 14 to 22 hours after their fasting started. During fasting, the metabolism needs to adapt to maintain the energy homeostasis. In this process, non-esterified fatty acids are oxidized generating acetoacetate, which is then converted into acetone via either spontaneous or enzymatic decarboxylation of excess acetoacetate (23,56). Consequently, an increase in acetone concentration is expected (15,57,58). As shown in Figure 9.7, two participants demonstrate a gradually increasing trend of

acetone concentration in their breath, with one participant spiking toward the end of the study, indicating the effect of fasting on their metabolism. Furthermore, the breath samples of one participant could be distinguished from the other volunteers through increased levels of methane and carbon monoxide (Supplementary Materials Figure S2).

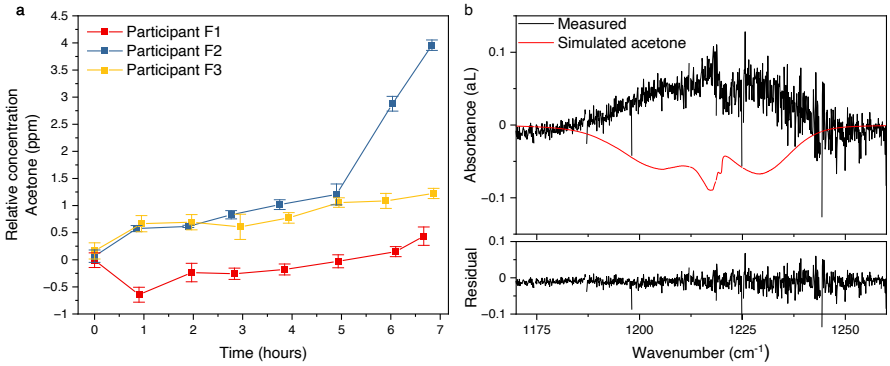


Figure 9.7: Retrieved concentrations of compounds in breath samples during fasting. **(a)** Retrieved relative change in acetone concentrations of participants during fasting. **(b)** Measured (black) and simulated (red, inverted) absorbance spectra of acetone with associated residual of the fit.

Finally, a relevant feature of our infrared IDFG-based instrument, in comparison to current state-of-the-art mass spectrometry instruments is its ability to measure (partly) exogenous compounds such as carbon monoxide, nitrous oxide, or methane in the breath samples (see Supplementary Materials Figure S1 to S3). Three regular smokers were asked to participate in a washout experiment (Figure 9.8). The long half-life of the reduction of carbon monoxide in breath indicates that this exogenous compound entered the bloodstream and bound to hemoglobin. The slow decrease of carbon monoxide after smoking is in line with other studies, indicating the half-life reduction of the hemoglobin-carbon monoxide bond to be on the order of several hours. Furthermore, despite the smokers not having smoked prior to the case study for over 10 hours, their pre-smoking carbon monoxide levels are clearly elevated compared to typical non-smoker values (< 2 ppm).

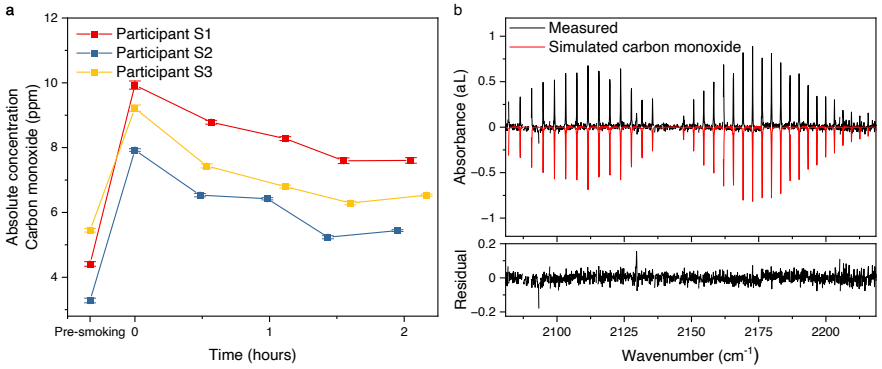


Figure 9.8: Retrieved concentrations of compounds in breath samples after smoking. **(a)** Retrieved absolute concentrations of carbon monoxide demonstrating washout over hours after smoking a cigarette. **(b)** Measured (black) and simulated (red, inverted) absorbance spectra of carbon monoxide with associated residual of the fit.

In Table 9.2, the detection limit for the species retrieved in these case studies is shown. Here, we focus on several molecular species with the strongest absorption in the wavelength range used (2-11.5 µm) showing the large dynamic range of the instrument.

Table 9.2. Compounds detectable with the current setup with detection limits and spectral range.

Volatile organic compound	Related intervention(s)	Spectral range used (cm^{-1})	Detection limit in 190 seconds
Acetone	Fasting	1320-1460	60 ppb
Ammonia	Protein intake	890-1192	4 ppb
Carbon dioxide	-	945-1079	140 ppm
Carbon monoxide	Smoking	2055-2220	16 ppb
Ethylene	Protein intake	870-1014	12 ppb
Isoprene	Fasting and protein intake	850-950	14 ppb
Methane	-	1250-1360	27 ppb
Methanol	Protein intake	970-1080	25 ppb
Nitrous oxide	Smoking	2160-2260	16 ppb
Water	-	850-1300	400 ppm

We report the detection limit as the precision of the retrieved concentrations in the breath samples. Therefore, the detection limit is specific to the wavenumber range used to fit these specific compounds. For molecules with a very high concentration in the samples, such as carbon dioxide and water (%-level concentrations), the

detection limit is in high ppm levels: relatively weak absorption lines were used that do not saturate at these high concentrations. Most molecular compounds have, in the current setup, a detection limit in the order of tens of ppb. While this is sufficient for concentrations of several compounds such as acetone, ammonia, or carbon monoxide, other compounds, like ethylene or methanol, demand improved detection limits for effective detection in breath applications.

9.4 Discussion

In this work, we present the first fully functional demonstrator for exhaled breath analysis with ultra-broadband laser-based infrared spectroscopy. The multi-species detection potential enables the instrument to be used flexibly in various studies. Therefore, we envision the role of this instrument as a broadly applicable platform to further accelerate the wide adoption of breath analysis in research and clinical practice.

In contrast to narrowband systems targeting a single or a few molecular compounds, the broadband nature of this system can selectively measure many compounds simultaneously. Being equipped with a breath sampling system to select specific phases of the exhalation, the experiments are performed in a standardized manner: measuring carbon dioxide, which is difficult to measure using mass spectrometry equipment, the end-tidal fraction of the exhalation can be computed and retrieved concentrations of the breath compounds normalized. While we chose to collect this part of the exhalation, the breath sampling system allows the collection of a desired exhalation phase depending on the research purpose. Our system proved capable of on-line monitoring of intervention-related compounds, such as acetone, ammonia, and carbon monoxide.

The instantaneous spectral bandwidth of the system presented here is effectively 2.9-11.5 µm or 2580 cm^{-1} in a single measurement. This outperforms current state-of-the-art broadband laser-based spectroscopy systems, of which most notably the recent demonstration by Liang et al. (11). Their maximum spectral coverage is 1010 cm^{-1} when effectively two different optical parametric oscillator-based laser systems are used and different measurements are combined, as their instantaneous bandwidth is maximally 180 cm^{-1}. Their unparalleled sensitivity, on the other hand, serves as an example of the potential for ultra-broadband laser-based infrared spectroscopy in breath analysis.

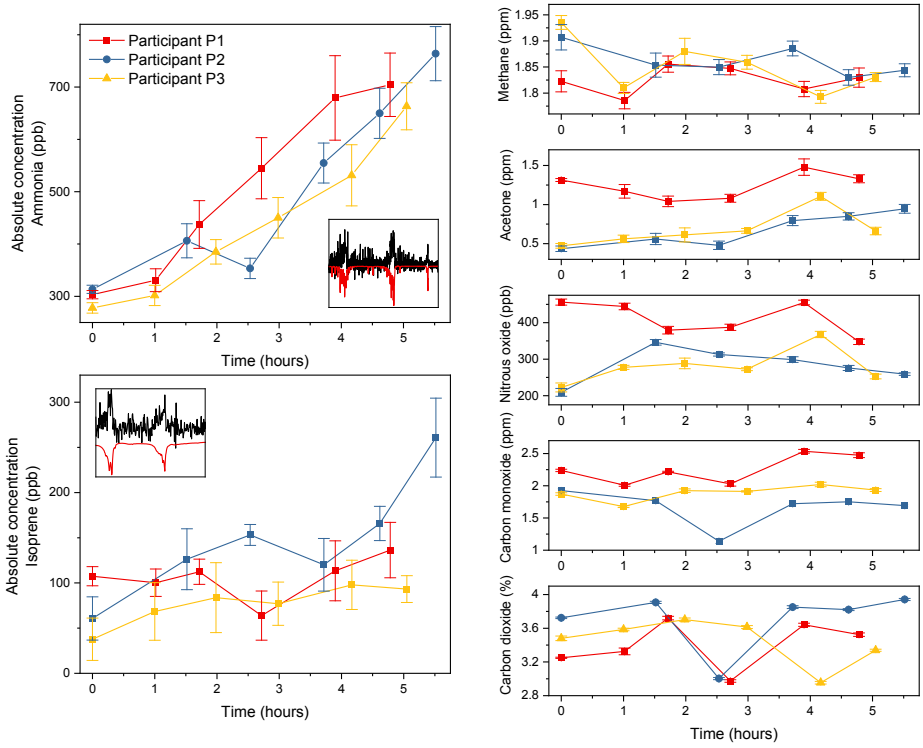
Regarding applications in breath analysis, broadband infrared spectroscopy should be regarded as a complementary technique to mass spectrometry. Both are particularly useful for the detection of specific classes of compounds (3). In

particular, the infrared IDFG-based instrument can create a full picture of those compounds not accessible by mass spectrometry, but with known biochemical pathways and thus relevant for clinical applications.

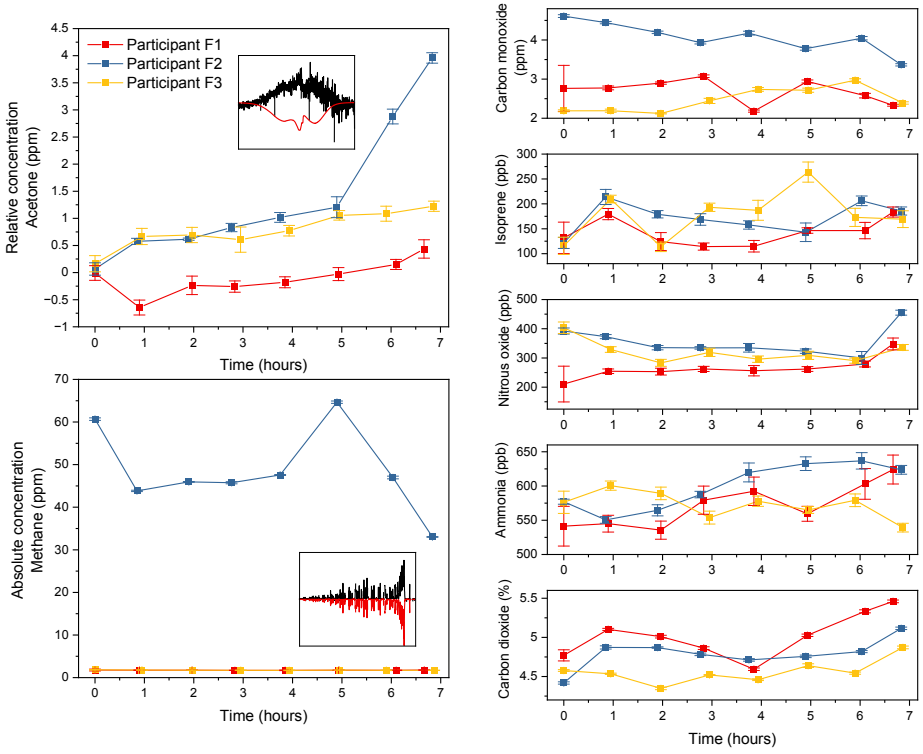
Our current system is a first step towards a versatile breath analysis platform. Despite its broad spectral coverage, further broadening of the spectrum, which is possible with the currently used laser platform (59), would incorporate the absorption regions of other compounds in the fingerprint region towards 20 µm. Improvements in the sensitivity enable more sensitive measurements of compounds with typical concentrations below 1 ppb (such as ethylene) and enable studies into the different isotopological compositions of compounds (11).

With the potential of the ultra-broadband IDFG-based laser source being used as a frequency comb, dual-frequency comb-based measurements could be foreseen as a future technique to be incorporated into the current breath analyzer. With fast measurements, this would enable real-time measurements of the breath samples, useful for time-resolved breath-to-breath analysis.

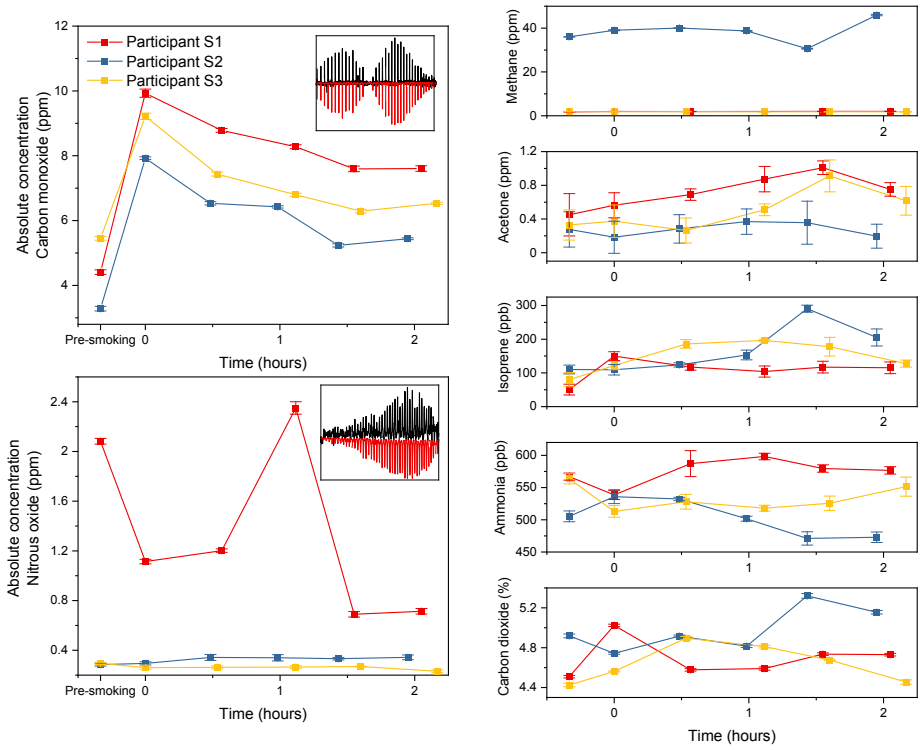
Supplementary materials



Supplementary Figure S1: Retrieved concentrations of compounds in breath after protein intake. Retrieved absolute concentrations of participants after protein intake, demonstrating a related increase in ammonia. Inserts: Measured (black) and simulated (red, inverted) spectral features of related compounds.



Supplementary Figure S2: Retrieved concentrations of compounds in breath samples during fasting. Retrieved relative change in concentrations (acetone) and absolute concentrations of participants during fasting, demonstrating a related increase in acetone in two participants. Inserts: Measured (black) and simulated (red, inverted) spectral features of related compounds.



Supplementary Figure S3: Retrieved concentrations of compounds in breath samples after smoking. Retrieved relative change in concentrations (acetone) and absolute concentrations of participants smoking a cigarette, demonstrating the washout of exogenous carbon monoxide over hours. Inserts: Measured (black) and simulated (red, inverted) spectral features of related compounds.

References

1. Smolinska A, Hauschild AC, Fijten RRR, Dallinga JW, Baumbach J, Van Schooten FJ. Current breathomics—a review on data pre-processing techniques and machine learning in metabolomics breath analysis. J Breath Res. 2014;8(2):027105.
2. Beale D, Jones O, Karpe A, Dayalan S, Oh D, Kouremenos K, et al. A Review of Analytical Techniques and Their Application in Disease Diagnosis in Breathomics and Salivaomics Research. IJMS. 2016;18(1):24.
3. Henderson B, Khodabakhsh A, Metsälä M, Ventrillard I, Schmidt FM, Romanini D, et al. Laser spectroscopy for breath analysis: towards clinical implementation. Appl Phys B. 2018;124(8):161.
4. Wang C, Sahay P. Breath Analysis Using Laser Spectroscopic Techniques: Breath Biomarkers, Spectral Fingerprints, and Detection Limits. Sensors. 2009;9(10):8230–62.
5. Pham YL, Beauchamp J. Breath Biomarkers in Diagnostic Applications. Molecules. 2021;26(18):5514.
6. Španěl P, Smith D. Quantification of volatile metabolites in exhaled breath by selected ion flow tube mass spectrometry, SIFT-MS. Clinical Mass Spectrometry. 2020;16:18–24.
7. Henderson B, Slingers G, Pedrotti M, Pugliese G, Malásková M, Bryant L, et al. The peppermint breath test benchmark for PTR-MS and SIFT-MS. J Breath Res. 2021;15(4):046005.
8. Kim KH, Jahan SA, Kabir E. A review of breath analysis for diagnosis of human health. TrAC Trends in Analytical Chemistry. 2012;33:1–8.
9. Pereira J, Porto-Figueira P, Cavaco C, Taunk K, Rapole S, Dhakne R, et al. Breath Analysis as a Potential and Non-Invasive Frontier in Disease Diagnosis: An Overview. Metabolites. 2015;5(1):3–55.
10. Maiti KS, Lewton M, Fill E, Apolonski A. Human beings as islands of stability: Monitoring body states using breath profiles. Sci Rep. 2019;9(1):16167.
11. Liang Q, Bisht A, Scheck A, Schunemann PG, Ye J. Modulated ringdown comb interferometry for sensing of highly complex gases. Nature. 2025;638(8052):941–8.
12. Henderson B, Ruszkiewicz DM, Wilkinson M, Beauchamp JD, Cristescu SM, Fowler SJ, et al. A benchmarking protocol for breath analysis: the peppermint experiment. J Breath Res. 2020;14(4):046008.
13. Wilkinson M, White I, Hamshere K, Holz O, Schuchardt S, Bellagambi FG, et al. The peppermint breath test: a benchmarking protocol for breath sampling and analysis using GC–MS. J Breath Res. 2021;15(2):026006.
14. Krebbers R, Van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, et al. Ultra-broadband spectroscopy using a 2–11.5 μm IDFG-based supercontinuum source. Opt Express. 2024;32(8):14506–14520.
15. Smith D, Spanel P, Davies S. Trace gases in breath of healthy volunteers when fasting and after a protein-calorie meal: a preliminary study. Journal of Applied Physiology. 1999;87(5):1584–8.
16. Mochalski P, King J, Mayhew CA, Unterkofler K. A review on isoprene in human breath. J Breath Res. 2023;17(3):037101.
17. Paardekooper LM, Van Den Bogaart G, Kox M, Dingjan I, Neerincx AH, Bendix MB, et al. Ethylene, an early marker of systemic inflammation in humans. Sci Rep. 2017;7(1):6889.
18. Jiang Y, Zhang T, Kusumanchi P, Han S, Yang Z, Liangpunsakul S. Alcohol Metabolizing Enzymes, Microsomal Ethanol Oxidizing System, Cytochrome P450 2E1, Catalase, and Aldehyde Dehydrogenase in Alcohol-Associated Liver Disease. Biomedicines. 2020;8(3):50.

19. Dryahina K, Pospíšilová V, Sovová K, Shestivska V, Kubišta J, Spesyvyi A, et al. Exhaled breath concentrations of acetic acid vapour in gastro-esophageal reflux disease. J Breath Res. 2014;8(3):037109.
20. Morrison DJ, Preston T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes. 2016;7(3):189–200.
21. Verma A, Bhagchandani T, Rai A, Nikita, Sardarni UK, Bhavesh NS, et al. Short-Chain Fatty Acid (SCFA) as a Connecting Link between Microbiota and Gut-Lung Axis-A Potential Therapeutic Intervention to Improve Lung Health. ACS Omega. 2024;9(13):14648–71.
22. Jones AW. Breath Acetone Concentrations in Fasting Male Volunteers: Further Studies and Effect of Alcohol Administration. Journal of Analytical Toxicology. 1988;12(2):75–9.
23. Mathew T, Pownraj P, Abdulla S, Pullithadathil B. Technologies for clinical diagnosis using expired human breath analysis. Diagnostics. 2015;5(1):27–60.
24. Španěl P, Smith D. Quantification of volatile metabolites in exhaled breath by selected ion flow tube mass spectrometry, SIFT-MS. Clin Mass Spectrom. 2020;16:18–24.
25. Smith D, Spanel P, Davies S. Trace gases in breath of healthy volunteers when fasting and after a protein-calorie meal: a preliminary study. J Appl Physiol (1985). 1999;87(5):1584–8.
26. Kroupina K, Bémeur C, Rose CF. Amino acids, ammonia, and hepatic encephalopathy. Analytical Biochemistry. 2022;649:114696.
27. Ghorbani R, Schmidt FM. Real-time breath gas analysis of CO and CO^2 using an EC-QCL. Appl Phys B. 2017;123(5):144.
28. Cheng S, Lyass A, Massaro JM, O'Connor GT, Keaney JF, Vasan RS. Exhaled carbon monoxide and risk of metabolic syndrome and cardiovascular disease in the community. Circulation. 2010;122(15):1470–7.
29. Dorey A, Scheerlinck P, Nguyen H, Albertson T. Acute and Chronic Carbon Monoxide Toxicity from Tobacco Smoking. Mil Med. 2020;185(1–2):e61–7.
30. Ryter SW, Choi AMK. Carbon monoxide in exhaled breath testing and therapeutics. J Breath Res. 2013;7(1):017111.
31. Ross BM, Glen I. Breath Ethane Concentrations in Healthy Volunteers Correlate with a Systemic Marker of Lipid Peroxidation but Not with Omega-3 Fatty Acid Availability. Metabolites. 2014;4(3):572–9.
32. Miekisch W, Schubert JK, Noeldge-Schomburg GFE. Diagnostic potential of breath analysis—focus on volatile organic compounds. Clinica Chimica Acta. 2004;347(1):25–39.
33. Bratu AM. Spectroscopic study of breath ethylene via the mouth and nose. Lasers Med Sci. 2019;34(4):773–8.
34. Törnqvist M, Gustafsson B, Kautiainen A, Harms-Ringdahl M, Granath F, Ehrenberg L. Unsaturated lipids and intestinal bacteria as sources of endogenous production of ethene and ethylene oxide. Carcinogenesis. 1989;10(1):39–41.
35. Turner C, Španěl P, Smith D. A longitudinal study of ethanol and acetaldehyde in the exhaled breath of healthy volunteers using selected-ion flow-tube mass spectrometry. Rapid Communications in Mass Spectrometry. 2006;20(1):61–8.
36. Meijnikman AS, Nieuwdorp M, Schnabl B. Endogenous ethanol production in health and disease. Nat Rev Gastroenterol Hepatol. 2024;21(8):556–71.

37. Mochalski P, King J, Mayhew CA, Unterkofler K. A review on isoprene in human breath. J Breath Res. 2023;17(3):037101.
38. Sukul P, Richter A, Schubert JK, Miekisch W. Deficiency and absence of endogenous isoprene in adults, disqualified its putative origin. Heliyon. 2021;7(1):e05922.
39. King J, Mochalski P, Unterkofler K, Teschl G, Klieber M, Stein M, et al. Breath isoprene: muscle dystrophy patients support the concept of a pool of isoprene in the periphery of the human body. Biochem Biophys Res Commun. 2012;423(3):526–30.
40. Kinoyama M, Nitta H, Ohta T, Hatase Y, Hara S, Hirakawa K, et al. Diurnal variation in the concentration of methane in the breath of methane producers. Microbial Ecology in Health and Disease. 2006;18(1):47–54.
41. Dorokhov YL, Shindyapina AV, Sheshukova EV, Komarova TV. Metabolic methanol: molecular pathways and physiological roles. Physiol Rev. 2015;95(2):603–44.
42. Vleeming W, Rambali B, Opperhuizen A. The role of nitric oxide in cigarette smoking and nicotine addiction. Nicotine Tob Res. 2002;4(3):341–8.
43. Antosova M, Mokra D, Pepucha L, Plevkova J, Buday T, Sterusky M, et al. Physiology of nitric oxide in the respiratory system. Physiol Res. 2017;66(Suppl 2):S159–72.
44. Chambers DC, Tunnicliffe WS, Ayres JG. Acute inhalation of cigarette smoke increases lower respiratory tract nitric oxide concentrations. Thorax. 1998;53(8):677–9.
45. Dawson B, Drewer J, Roberts T, Levy P, Heal M, Cowan N. Measurements of methane and nitrous oxide in human breath and the development of UK scale emissions. PLOS ONE. 2023;18(12):e0295157.
46. Zumft WG. Cell biology and molecular basis of denitrification. Microbiol Mol Biol Rev. 1997;61(4):533–616.
47. Abbas MA, Jahromi KE, Nematollahi M, Krebbers R, Liu N, Woyessa G, et al. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.
48. Vasilyev S, Smolski V, Moskalev I, Peppers J, Mirov M, Barnakov Y, et al. Multi-octave infrared femtosecond continuum generation in Cr:ZnS-GaSe and Cr:ZnS-ZGP tandems. In: Schunemann PG, Schepler KL, editors. Nonlinear Frequency Generation and Conversion: Materials and Devices XIX. San Francisco, United States: SPIE; 2020. p. 6.
49. Vasilyev S, Smolski V, Peppers J, Moskalev I, Mirov M, Barnakov Y, et al. Middle-IR frequency comb based on Cr:ZnS laser. Opt Express. 2019;27(24):35079.
50. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf. 2022;277:107949.
51. Sharpe SW, Johnson TJ, Sams RL, Chu PM, Rhoderick GC, Johnson PA. Gas-phase databases for quantitative infrared spectroscopy. Appl Spectrosc. 2004;58(12):1452–1461.
52. Risby TH, Solga SF. Current status of clinical breath analysis. Appl Phys B. 2006 Nov;85(2–3):421–6.
53. Kroupina K, Bémeur C, Rose CF. Amino acids, ammonia, and hepatic encephalopathy. Analytical Biochemistry. 2022;649:114696.
54. Solga SF, Mudalel ML, Spacek LA, Risby TH. Fast and Accurate Exhaled Breath Ammonia Measurement. JoVE. 2014;(88):51658.
55. Spacek LA, Strzepka A, Saha S, Kotula J, Gelb J, Guilmain S, et al. Repeated Measures of Blood and Breath Ammonia in Response to Control, Moderate and High Protein Dose in Healthy Men. Sci Rep. 2018;8(1):2554.
56. Laffel L. Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. Diabetes/Metabolism Research and Reviews. 1999;15(6):412–26.

57. Španěl P, Dryahina K, Rejšková A, Chippendale TWE, Smith D. Breath acetone concentration; biological variability and the influence of diet. Physiol Meas. 2011;32(8):N23–31.

58. Jones AW. Breath-Acetone Concentrations in Fasting Healthy Men: Response of Infrared Breath-Alcohol Analyzers. Journal of Analytical Toxicology. 1987;11(2):67–9.

59. Konnov D, Muraviev A, Vasilyev S, Vodopyanov K. High-resolution frequency-comb spectroscopy with electro-optic sampling and instantaneous octave-wide coverage across mid-IR to THz at a video rate. APL Photonics. 2023;8(11):110801.

Appendices

Outlook and concluding remarks

In this thesis, gas-sensing applications of novel mid-infrared supercontinuum sources have been explored. Infrared spectroscopy studies traditionally revolved around narrowband laser sources, such as quantum cascade lasers or interband cascade lasers. This results in (mostly) single-species gas-sensing solutions with high sensitivity and selectivity but also requires a distinct laser for each targeted compound. In comparison, broadband infrared sources enable multi-species detection capabilities with a single laser source.

Over the course of this PhD project, the infrared supercontinuum sources have been maturing, with a growing availability of fiber-based sources and the first difference-frequency generation-based systems appearing on the market. This enabled advances in applications revolving around broadband infrared supercontinuum-based spectroscopy. Proof-of-concept laboratory-bound gas-sensing demonstrations have now been succeeded by application-driven field studies. Supercontinuum sources have evolved sufficiently to be used in spectroscopic systems that can be compact, reliable, and robust. In this way, supercontinuum-based systems were used to drive investigations into various fields, such as *in-situ* measurements and reaction product analysis of plasmas, outdoor field deployments for open-path measurements of greenhouse gases and pollutants, and exhaled breath analysis.

Spatially coherent fiber-based sources using fluoride or ZBLAN fibers have been commercially available for a couple of years. These sources have shown reliable performance with typically hundreds of mW of output power in the infrared region between 1.5 to roughly 4 µm. In addition, chalcogenide fiber-based sources have the potential to extend further in the infrared. Even though these sources are starting to emerge on the commercial market, the availability and reliability of these specialty fibers remain a bottleneck for quick upscaling and cost reduction.

The intra-pulse difference-frequency generation (IDFG)-based sources have very recently made the step from laboratory, proof-of-concept setups to commercially available laser systems. As the necessary crystals become more readily available, IDFG-based sources are becoming strong competitors to systems that rely on meters-long high-purity chalcogenide fibers. The IDFG-based systems currently face significant costs and are, theoretically, based on less robust designs, relying on free-space optics. However, our coherent open-path spectroscopy field measurements demonstrated that this is not the case, and that these laser systems are robust and can be used reliably, even out of the laboratory. Moreover, despite their significant costs, the ability to be (both spatially and temporally) fully coherent and low noise opens possibilities for these systems to be used in cutting-edge dual-frequency comb systems, widening the scope of applications.

The high-sensitivity, high-dynamic range, broadband infrared spectra seem to unlock growing possibilities that are currently largely unexplored. In Chapter 8, the foundations were laid for advancing data processing techniques for these types of spectra. The analysis of the complex, broadband laser-based spectroscopic data can improve hugely by incorporating knowledge from established techniques in related fields, such as analytical tools from mass spectrometry-based methods, or fields such as chemometrics or data science. Current techniques rely heavily on the identification and quantification of peaks from known, expected compounds. Untargeted spectroscopic analyses using machine learning or artificial intelligence can contribute significantly to utilizing the increasingly complex broadband spectroscopic data to their full potential.

With the transition towards a circular economy still largely lying ahead of us, the demand for studies on plasma-driven conversions is expected to persist. The role of broadband infrared absorption spectroscopic methods could be consolidated as a standard technique to be applied in this field. Within further, fundamental studies on the kinetics and mechanisms inside the plasma, a demand for fast, time-resolved techniques can be foreseen. Here, the IDFG-based systems could play a role as a coherent source for time-resolved studies, for example through their use as a frequency comb. Another opportunity for these sources is to focus on increased sensitivity. Through this, reaction intermediates could be targeted that have a low concentration and are exclusively present *in situ* due to their reactivity. Monitoring such species could prove the expected pathways of conversions inside the plasma.

The demonstrator for coherent open-path spectroscopy (COPS) proved to be a useful tool for monitoring the presence of greenhouse gases or pollutants. The field measurements presented in this thesis consisted of the first out-of-laboratory deployment of this type of IDFG-based supercontinuum source. While this marked a significant milestone, stressing the potential of this technique as a robust system, further progress is required to integrate the source into a system that can reliably operate over several days or weeks.

A current challenge remains to quantify emission rates from the detected concentrations in the line-of-sight of the system. Incorporating data on wind speed and direction into modeling approaches is expected to enable this step. In contrast to existing open-path dual-frequency comb-based systems, the COPS system presented here is a significant reduction in complexity and costs as one laser source is used, and the necessity of frequency and phase stabilization is removed. Moreover, the COPS system possesses a significantly improved spectral coverage.

Further developing this system as a flexible platform for the detection of a wide range of species, enables swift deployments of this system in field measurements of different application areas. Here, the system could be used in

addition to point-measurement systems, which specialize in emission monitoring from known sources of emission or fixed emission points. In contrast, COPS systems integrate their line of sight, thus enabling gas detection in larger areas. Furthermore, the technology can be used for 2D or 3D mapping of emissions in the environment. This would be helpful for areas with heterogenous emissions, such as cattle roaming freely through a field or leaky pipelines, or widely spread emissions, such as greenhouse gas emissions from peatlands. Besides its complementary use to point-measurement systems, COPS could also be used as a reference instrument for satellite-based remote sensing techniques. The ground-level observations of compounds by the COPS instrument could calibrate or validate satellite observations of the same compounds. Finally, the COPS technology is also expected to contribute to studies on exposure. Polluting or otherwise harmful compounds could continuously be monitored in both outdoor and indoor environments, such as workplaces.

An application with great clinical potential is exhaled human breath analysis or breathomics. The current techniques for breath profiling are largely relying on mass spectrometry methods. Broadband infrared supercontinuum-based spectroscopic systems can be flexibly used for both targeted and untargeted studies on exhaled compounds that are traditionally challenging to observe and quantify through mass spectrometry methods. Especially the ability to detect metabolites with known biochemical pathways is expected to have an immediate impact on exhaled breath research. The demonstrator system shown in this thesis marks the role of the supercontinuum source in a novel platform for online breath analysis. Further enhancements in this field could be achieved by combining exhaled breath measurements with artificial intelligence-based analytical techniques. This would pave the way for untargeted analyses of breath sample spectra, with applications in disease detection and monitoring, or precision medicine.

In conclusion, this thesis has advanced gas-sensing solutions revolving around mid-infrared supercontinuum sources by exploring novel fields of applications for spectroscopy with these sources. Specifically, the development of a flexible spectroscopic platform enables progress in out-of-laboratory applications in environmental monitoring, human health, or plasma chemistry. With the demonstrations shown here, current follow-up research is currently ongoing or planned for the foreseeable future.

Research data management

The research described in this thesis has been carried out in accordance with the research data management policy of the Institute for Molecules and Materials of Radboud University, the Netherlands.[1]

The experimental data corresponding to the work presented in the chapters of this thesis are available from the promotor upon reasonable request.

- Chapter 1 – 3: No data was generated for these chapters.
- Chapter 4 – 7, and 9: Experimental data, metadata, scripts, and processed data were generated for this chapter. This data has been archived and is available from the promotor upon reasonable request.
- Chapter 8: Scripts, simulations, and processed data were generated for this chapter. The scripts used for Algorithm 8.1 can be found online at https://github.com/LaurensSluyterman/AI-for-spectroscopy.

Additional experimental data and metadata have been generated for the validation of the presented results. This data has been archived and is available from the promotor upon reasonable request.

1. https://www.ru.nl/rdm/vm/policy-documents/policy-imm/, version 2022.

About the author

Roderik Krebbers

Date of birth: February 15, 1998 (Deventer, The Netherlands)

2021 – 2025 **PhD Candidate**
Life Science Trace Detection Laboratory
Analytical Chemistry and Chemometrics – Institute for Molecules and Materials
Radboud University – Nijmegen

2019 – 2021 **MSc in Physical Chemistry (Cum Laude)**
Radboud University – Nijmegen

2016 – 2019 **BSc in Science (Physical Chemistry)**
Radboud University – Nijmegen

2010 – 2016 **Gymnasium NT-NG (Cum Laude)**
Etty Hillesum Lyceum – Deventer

List of publications

1. **Krebbers R**, Huisman M, van Kempen K, Meurs J, Khodabakhsh A, Cristescu SM. Broadband spectroscopic analyzer with unprecedented profile of biomarkers in exhaled breath. *In preparation*.

2. **Krebbers R***, Sluyterman LAÆ*, Meurs J, Khodabakhsh A, Cator EA, Cristescu SM. Optimizing Data Analysis for Broadband Mid-Infrared Absorption Spectroscopy: A Hybrid Dataset Approach. Anal Chim Acta. 2025;344303.

3. **Krebbers R**, van Kempen K, Lin Y, Meurs J, Hendriks L, Aben R, Paranaiba, JR, Fritz, C, Veraart, AJ, Khodabakhsh, A, Cristescu, SM. Ultra-Broadband Coherent Open-Path Spectroscopy for Multi-Gas Monitoring in Wastewater Treatment. Environ Sci Ecotechnology. 2025;25:100554.

4. **Krebbers R**, Liu N, Averink W, Harren FJM, Butterworth T, van Rooij G, Khodabakhsh, A, Cristescu, SM. In-situ analysis of methane plasmas with mid-infrared supercontinuum-based Fourier-transform spectroscopy. Appl Phys B. 2025;131(1):5. *Selected as* ***Editor's Choice Article Highlight*** *and* ***Issue cover image***

5. **Krebbers R**, van Kempen K, Harren FJM, Vasilyev S, Peterse IF, Lücker S, Khodabakhsh, A, Cristescu, SM. Ultra-broadband spectroscopy using a 2–11.5 μm IDFG-based supercontinuum source. Opt Express. 2024;32(8):14506–14520. *Selected as* ***Editors' pick***

6. **Krebbers R**, Liu N, Jahromi KE, Nematollahi M, Bang O, Woyessa G, Petersen CR, van Rooij G, Harren FJM, Khodabakhsh, A, Cristescu, SM. Mid-infrared supercontinuum-based Fourier transform spectroscopy for plasma analysis. Sci Rep. 2022;12(1):9642.

7. Abbas MA, Eslami Jahromi K, Nematollahi M, **Krebbers R**, Liu N, Woyessa G, Bang O, Huot L, Harren FJM, Khodabakhsh, A. Fourier transform spectrometer based on high-repetition-rate mid-infrared supercontinuum sources for trace gas detection. Opt Express. 2021;29(14):22315–22330.

8. Eslami Jahromi K, Nematollahi M, **Krebbers R**, Abbas MA, Khodabakhsh A, Harren FJM. Fourier transform and grating-based spectroscopy with a mid-infrared supercontinuum source for trace gas detection in fruit quality monitoring. Opt Express. 2021;29(8):12381–12397.

Included in this thesis, *These authors contributed equally to this work.

Presentations at scientific conferences and symposia

1. Invited talk at scientific symposium of study association Leonardo da Vinci, Radboud University, Netherlands, 2025.
2. Contributed talk "*Greenhouse Gas Monitoring at a Wastewater Treatment Plant with Ultra-Broadband Coherent Open-Path Spectroscopy*" at Field Laser Applications in Industry and Research (FLAIR), Italy, 2024.
3. Presider of the session "*Global Environmental Measurement and Monitoring (GEMM) Initiative*" at *Optica Sensing Congress 2024* (AIS, LACSEA, Sensors, QSM), France, 2024.
4. Contributed talk *"In-Situ Plasma Diagnostics Using Mid-Infrared Supercontinuum Sources Based On Nonlinear Fibers"* at *Optica Sensing Congress 2024* (AIS, LACSEA, Sensors, QSM), France, 2024. ***Winner Best Student Presentation Award***
5. Contributed talk "*Ultra-broadband mid-infrared supercontinuum-based spectroscopy for greenhouse gas monitoring at wastewater treatment plants*" at EGU General Assembly 2024, Austria, 2024.
6. Contributed talk "*Mid-infrared supercontinuum-based Fourier transform spectroscopy for multispecies open-path measurements*" at Photonics for Advanced Spectroscopy and Sensing (C-PASS), Italy, 2023.
7. Contributed talk "*Mid-infrared Supercontinuum-based Spectroscopy for Open-path Measurements*" in Conference on Lasers and Electro-Optics/Europe (CLEO/Europe), Germany 2023.
8. Invited talk "*Broadband mid-infrared spectroscopy using novel supercontinuum sources*" at sIMMposium, The Netherlands, 2022.
9. Contributed talk "*Analysing Plasmas Using Supercontinuum-based Fourier Transform Spectroscopy*" at Conference on Field Laser Applications in Industry and Research (FLAIR), France, 2022.
10. Contributed talk "*Mid-infrared Broadband Spectroscopy for Plasma Analysis*". Optical Sensors and Sensing Congress, online, 2022.
11. Contributed talk "*Analysis of Methane and Carbon Dioxide Plasmas with Supercontinuum-based Fourier Transform Spectroscopy*". At Conference on Lasers and Electro-Optics (CLEO), United States, 2022.
12. Contributed talk "*Fourier Transform Spectroscopy based on a Mid-Infrared Supercontinuum Source for Plasma Study*," at *Optica High-brightness Sources and Light-driven Interactions Congress 2022*, Hungary, 2022.

Acknowledgements

This PhD has been a great joy and an exciting journey – and it was so because of the people I got to experience it with. If there is one thing that I've learnt during this time, it is that science is a team effort and the most exciting results are obtained when working together with people in different disciplines. I would like to thank **Simona** for her guidance throughout the entire project. Her energy and vision on integrating our spectroscopy into different disciplines were truly inspiring: I got a lot of energy from the way our brainstorming on the open-path measurements ended up with us approaching people to ask whether we could bring our instrument, with the measurements at the wastewater facility in Rhenen as a beautiful result.

Starting my PhD involved some uncertainties: it was in the middle of the COVID pandemic, but also the research group was in a transition phase as **Frans** was about to retire. Though the involvement of Frans in my project evolved significantly over the years, I am very grateful for his support, mostly in the early stages of my PhD. His 'can-do' mentality was very motivating and the adventures we shared – such as attending the CLEO conference – will remain lasting memories.

It was exciting to be part of the emergence of the TDLab, in which everyone had to find his or her role in the new team. I would like to thank every (former) TDLab team member for contributing to the great atmosphere in our team, though some people deserve a specific mention. I'm very grateful for all the support from **Amir**. Whether it was hard skills, like working with optics or spectroscopic data, or soft skills, like giving presentations at conferences or project meetings, you were a beacon of knowledge throughout my PhD.

It was a great pleasure to work together with so many great colleagues. **Kees**, we have known each other for such a long time. It was great to have a close friend in the lab and I am really proud of all the great results we obtained together. **Ningwu**, it was great to work alongside a more experienced PhD candidate when I was just starting my PhD and this really helped me to kickstart my project, but more importantly: it just was a lot of fun to spend the hours in the lab with you while we were waiting for our results to roll in. **Yueyu,** I was really happy when you joined our group, as there finally was another PhD candidate in the group with whom I could share the experiences. Together we made lots of steps in the PhD and I am very thankful for that. I also really enjoyed working with and leaning a lot from **Khalil** and **Reza**. Thank you for your patience in explaining everything related to the FTS, supercontinuum source, and spectroscopic experiments.

Other colleagues I want to mention are **Brigitte** and **Magda** for their administrative support, as well as **Nick**, **Mats**, **Paul** (but also **Niek** from SCM, and the colleagues at the TechnoCentrum) for their technical support.

Various students did an internship in our group on a topic related to my project. Their input, energy, and contributions to the research are all very appreciated! Besides Kees, I would like to thank **Wessel**, **Loes**, **Paola**, and **Tim**.

For the great atmosphere in our group – and in the "office downstairs" specifically – I have to thank all people who have been or currently are part of the group, or are part of the office. Some people deserve special mention, as they have (seemingly) been part of this experience since forever: **Beate**, I have no idea where you always find all the energy to do these thousands of things you're doing. Your presence made the office a very lively place. **Joris**, dr. Panini, TDLab would not be TDLab without you. Always in good spirit, always willing to help anyone. If I could, I would ask Luc-Jan to give you another ten MWTrophies. Other former TDLab colleagues I really like to mention are **Ben**, **Dexter**, **Valia**, and **Victor**. The awesome people I got the pleasure of sharing an office with are too many to mention all here, but some specific people to mention are **Bram** and **Stroppo**. Furthermore, I enjoyed the pleasant interactions with the people from the office next door, the **SCM department**, as well as the **AC&C research group**.

I am really happy to continue working at the TDLab, as this means I will continue to work with people such as **Marleen**, who does an awesome job on the data analysis, **Jonas**, who will become the plasma expert of our group, **Pascalle**, the chicken expert, **Aliakbar**, the contamination expert, and **Chao**, the data expert.

It was a particularly interesting element of my PhD to gain so many first-hand insights into (international) scientific collaboration. I am very grateful to have been part of the TRIAGE project, through which I got to meet so many interesting researchers from various countries and scientific backgrounds, among others, **Ole**, **Christian**, **Sanghoon**, **Laurent**, **David**, **Magnus**, **Donatella**, and **Bruce**. But also closer to home, I am grateful for the collaboration with the plasma experts from Maastricht University, **Tom** and **Gerard**, the people from Radboud's Ecology and Microbiology departments, **Ida**, **Sebastian**, **Lisanne**, **Annelies**, **Quint**, **José**, **Ralf**, **Christian**, and **Sarian**, the people of HDSR who provided us access to the wastewater treatment facility, **Tonny**, and **Teunis**, and the mathematics department, **Laurens** and **Eric.**

Finally, where would I have been without the support of **Lea** and my parents, **Erik** and **Lotte**, and my brother **Cas**. Though they always joke that they are of little help because of their limited understanding of physics and chemistry, we know how invaluable their support has been.

Radboud Universiteit