https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

Potential CO_2 measurement capabilities of a transportable Near Infrared Laser Heterodyne Radiometer (LHR)

Marie Thérèse El Kattar[1], Tingting Wei[1], Aditya Saxena[1], Hervé Herbin[2], and Weidong Chen[1]

[1]Laboratoire de Physico-Chimie de l'Atmosphère (LPCA), Université du Littoral Côte d'Opale, 189A, Avenue Maurice Schumann, 59140, Dunkerque
[2]UMR 8518 – LOA – Laboratoire d'Optique Atmosphérique, Univ. Lille, CNRS, F-59000, Lille, France

Correspondence to: Marie Thérèse El Kattar (marie-therese.el-kattar@univ-littoral.fr), Hervé Herbin (herve.herbin@univ-lille.fr) and Weidong Chen (chen@univ-littoral.fr)

Abstract. Heterodyne detection stands as a powerful method for enhancing sensitivity limits and attaining exceptional resolution. It also offers the advantage of being transportable and the ability to make it more compact which makes it favorable for ground-based remote sensing in field campaigns. An all-fiber coupled laser heterodyne radiometer (LHR), using a wideband tunable external cavity diode laser (1520-1580 nm) as local oscillator laser was developed for CO_2 measurements. Optimal absorption lines and transmission spectra of the LHR was achieved by using a balanced photodetector to suppress the relative intensity noise of the local oscillator laser. This work aims to quantify how the LHR contributes to measuring tropospheric CO_2 abundances in the atmospheric column from the ground. Here, we demonstrate the LHR's ability to measure CO_2 vertical profiles through a extensive analysis of information content, channel selection, and error budget estimation. This comprehensive analysis relies on the radiative transfer model ARAHMIS, developed at the Laboratoire d'Optique Atmosphérique (LOA). Additionally, we present a comparison of the LHR with other ground-based instruments, such as the EM27/SUN and the IFS125HR from the TCCON network. Furthermore, this work supports ongoing MAGIC campaigns focused on greenhouse gas monitoring and the validation of current and future space missions such as MicroCarb and FORUM.

1 Introduction

Developing robust and affordable techniques for the accurate measurement of greenhouse gas (GHG) concentrations in the atmosphere is crucial for better quantifying the radiative impacts on the atmosphere, studying the effects of climate change, and identifying anthropogenic sources. Alongside spaceborne instruments such as OCO-2 (Eldering et al. 2017), which offer global coverage and high GHG column abundance accuracy, there's a growing need for compact, portable, and cost-effective instruments that can validate satellite observations but also monitor major GHGs in the atmospheric column. In addition to compactness, high mobility and low cost, these devices must have extremely high spectral resolution to meet GHG observation requirements (IPCC 2014). The Bruker IFS125HR Fourier Transform Spectrometer (FTS), with a spectral resolution of approximately 0.02 cm^{-1}, is the main instrument used by the TCCON monitoring network (Wunch et al. 2010). However, this spectrometer's limitations for field campaigns hinder its broader use in ground-based atmospheric measurements worldwide.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

The COCCON network is an important supplement of TCCON, as the logistic requirements are low and the FTS used, the EM27/SUN, is easy to operate and, due to the fact that the spectrometers are portable, it especially contributes to the quantification of local sources (Frey et al. 2019). Datasets of these networks are widely utilized for satellite validation in order to use them in data assimilation models (Jiang et al. 2016). The drawback of being portable is that the FTS reduces the spectral resolution. In contrast, heterodyne detection offers a cost-effective and highly mobile method for enhancing sensitivity limits and achieving exceptional resolution (Weidmann 2021). This suggests that heterodyne spectro-radiometers could serve as complementary methods to TCCON measurement instruments (Palmer et al. 2019). While commercially heterodyne spectroradiometers are currently unavailable, scientific groups worldwide are presenting their achievements in the development and application of these instruments in the near-infrared (NIR) spectral range (Clarke et al. 2014, Zenevich et al. 2020). An all-fiber coupled laser heterodyne radiometer (LHR) has been developed at the Laboratoire de PhysicoChimie de l'Atmosphère (LPCA) for measuring carbon dioxide (CO_2) and water vapor (H_2O) concentrations in the atmospheric column (Wang et al. 2023). The LHR uses a broadband tunable external cavity diode laser operating between 1520-1580 nm as a local oscillator (LO) laser. To improve signal to noise ratio in LHR spectra, a balanced photodetector is employed to suppress the laser relative intensity noise (RIN) of the LO laser.

This study presents the principle of the LHR experimental setup and quantifies its potential for CO_2 retrieval. The structure of the paper is as follows: Section 2 describes the setup and technical characteristics of the instrument, Section 3 provides a detailed explanation of the forward model, state vector, and a complete error analysis. We present in Section 3.5, a comparison with the other FTS instruments for the retrieval of CO_2 building on earlier research (El Kattar 2020). Section 4 presents the channel selection methodology employed in this work, which is essential for determining the most suitable channels for measurement. The study concludes with a summary of the results and explores future directions, such as improving measurement precision, and particularly the CO_2 retrieval.

2 Experimental setup

The used LHR in the present work, depicted in **Figure 1**, is designed to measure atmospheric CO_2 and H_2O concentrations by measuring their absorptions of the sunlight in the NIR. Solar radiation is captured using a portable solar tracker (STR-21 G; EKO Instruments Co., Ltd.) to real-time track the sun's movement and is modulated with a mechanical chopper. This modulated sunlight was combined with light from a tunable external cavity diode laser (TUNICS-BT 3642 HE CL; NetTest) operating at room temperature, acting as the LO laser. The laser's wavelength can be continuously tuned between 1520–1580 nm, with a maximum power of 5 mW. Solar radiation was collected into a 2-meter-long single-mode fiber (SMF-28-J9; Thorlabs, Inc.) using a fiber collimator (F810APC-1550; Thorlabs, Inc.), which is mounted on the solar tracker and has a numerical aperture of 0.24. On sunny days, the solar power collected in the single-mode fiber can reach 7.9 μW. A mechanical chopper (MC2000B; Thorlabs, Inc.) was installed in the front of the collimator to modulate the incident solar radiation. The modulated radiation was split by a single-mode fiber splitter with a 40:60 beam splitting ratio. The 40% power was measured with a

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

photodetector (PDA20CS-EC; Thorlabs, Inc.) to monitor sunlight intensity variation during measurements. The 60% power was mixed with the LO laser for heterodyne detection. In addition, a balanced amplified photodetector (PDB425C; Thorlab Inc.) was used to reduce laser RIN resulting from the LO laser. For this purpose, the LO laser output light was split into two beams with a 50:50 fiber splitter, 50% used to mix the sunlight and the other 50% used for balanced detection. The beating signal at radio frequency (RF) from the balanced photodetector was passed through a band-pass filter with an effective bandwidth of 24– 95 MHz. Subsequently, a square-law detector with an electronic bandwidth (B_{IF} = 100 kHz to 2 GHz) were employed to perform heterodyne detection. The output signal was then demodulated through a lock-in amplifier (LIA-MV-150; FEMTO Inc.). A data acquisition card (DAQ) (USB-6366; NI Inc.) digitizes the spectral signal that is then transferred to a laptop for further data processing and retrieval.

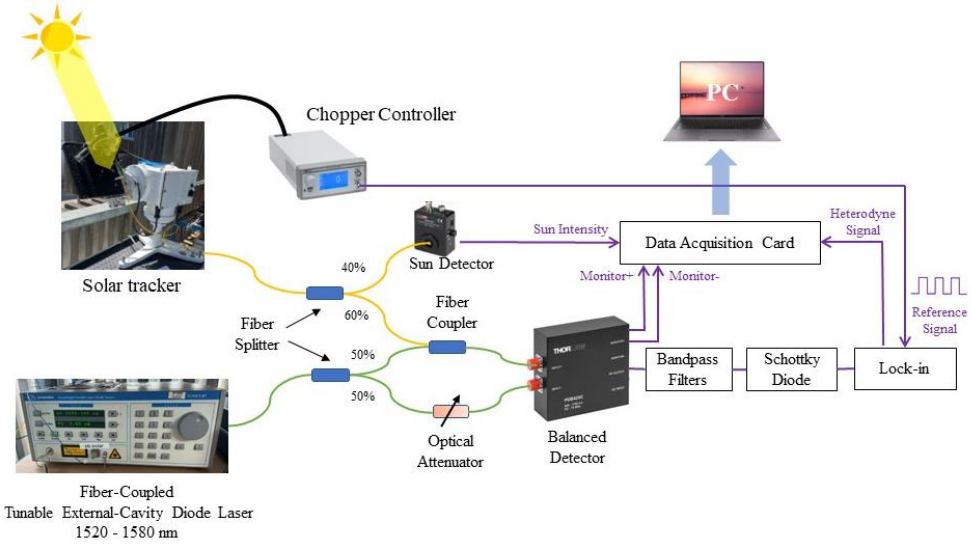

Figure 1: Schematic of the developed LHR. Each percentage represents the proportion of radiation beams split or combined.

To achieve accurate measurements of relative transmittance using the LHR, it is essential to extract spectral signals with a high Signal-to-Noise Ratio (SNR). The SNR for the measurement via coherent detection can be written as follows (Sun et al. 2024):

$$SNR = \frac{2T_0\eta\sqrt{\Delta f\tau}}{2\eta+\exp(h\upsilon/kT_S)-1}, \quad (1)$$

where τ is the integration time, Δf is the filter bandwidth, T_0 is the transmission efficiency, η is the quantum efficiency of the photodetector, T_S is the temperature of the heat source, k is the Boltzmann constant, υ is the wavenumber and h is the Planck constant. For a typical sunlight measurement, with an integration time of 100 ms, Δf of 52 MHz, T_0 of 1, a quantum efficiency of 0.81 (provided by the manufacturer) and T_S of 6011 K on average, we find an average theoretical SNR of 710 for the spectral domain covered by the LHR. The absorption spectra obtained from these measurements are illustrated in **Figure 2**.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

3 Information content analysis

In order to determine and evaluate the capacity and the performance of the developed LHR, information content (IC) study is conducted to assess its potential for GHG retrieval and compare it with other well-established instruments/techniques for worldwide observation.

3.1 The forward model

To accurately simulate the transmittances observed by the LHR, the line-by-line radiative transfer algorithm ARAHMIS was used across a broadband NIR spectrum of 1.567–1.577 µm. The absorption spectrum of gases is derived using the updated HITRAN 2020 database (Gordon et al. 2022) , with absorption lines modelled assuming a Gaussian line profile. In addition, absorption continua for water vapor (H_2O) and carbon dioxide (CO_2) are incorporated using the MT-CKD model (Clough et al. 2005). The incident solar spectrum is derived from the pseudo-transmittance spectra for direct sunlight originating from the center of the solar disk, as provided by Toon 2015, and subsequently interpolated onto the LHR's spectral range. The Planck function is calculated across the LHR's spectral domain using the LATMOS function (Meftah et al. 2018), as the effective brightness temperature shows considerable variation with wavenumber. Accurate determination of the spectrometer's line-of-sight (LOS) is crucial for determining the spectral absorption of solar radiation as it propagates through the atmosphere during the retrieval process of gases. To achieve this, the timing and duration of each measurement are recorded, allowing the calculation of the Solar Zenith Angle (SZA) using the methodology described in (Michalsky 1988).

Measurements are conducted in Dunkirk (51.035°N, 2.369°E) under clear sky conditions. The calculations depend on the concentration of the target atmospheric profile, along with associated data profile such as temperature, pressure, and relative humidity, which are obtained from a nearby PTU Vaisala radiosonde, while CO_2 and H_2O ancillary data are derived from the TCCON database. **Figure 2** displays the results of ARAHMIS's simulation compared to a typical measurement of the mid-infrared band by the LHR. Additionally, the impact of the solar spectrum, CO_2, and H_2O is shown, demonstrating the strong consistency between the forward model simulation and the LHR measurements under clear-sky conditions.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

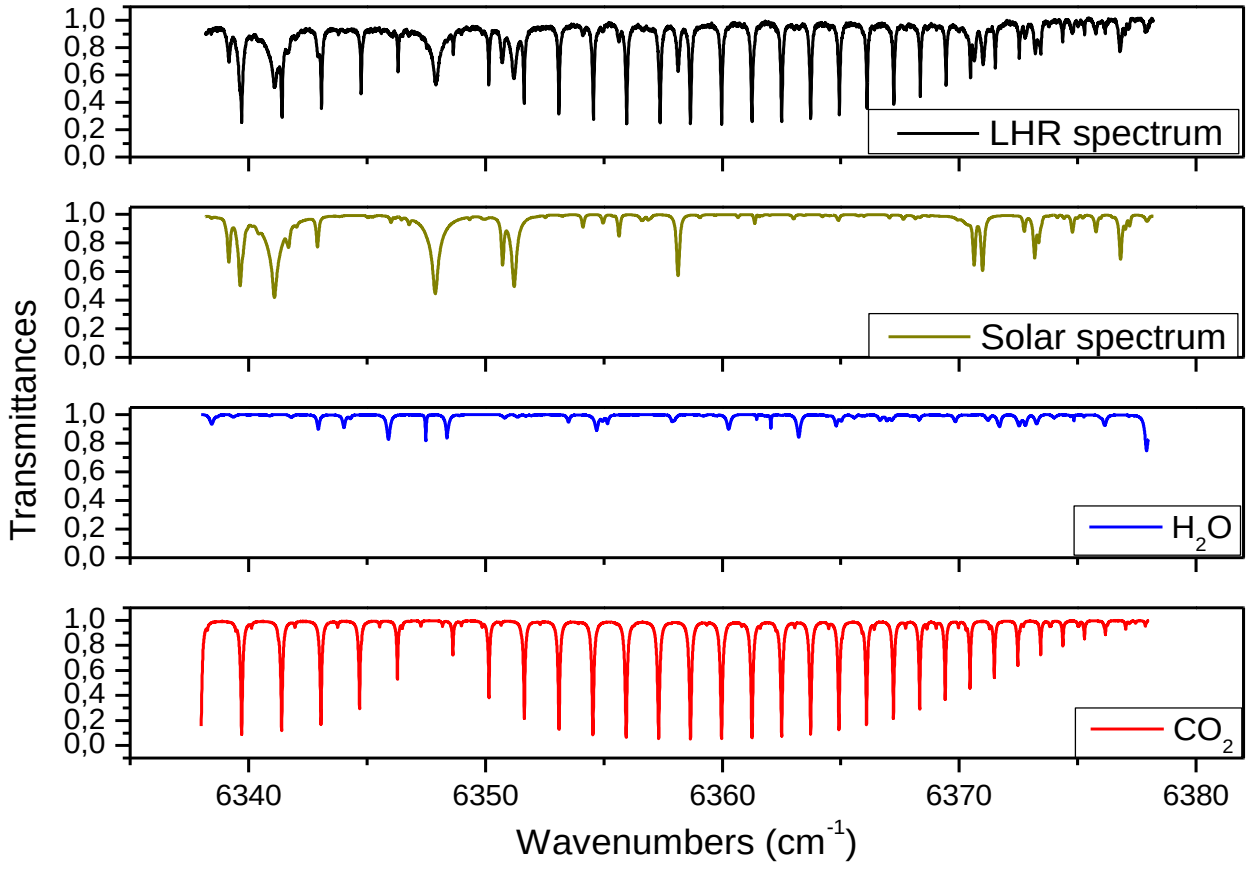

Figure 2: Comparison of measured and simulated LHR transmittance spectra under clear-sky conditions in Dunkirk. The simulations are generated using the line-by-line forward model ARAHMIS, incorporating the solar pseudo-transmittance spectra from (Toon 2015).

3.2 Theoretical basis of Information Content (IC)

Following the computation of the forward model, we apply the framework developed by Rodgers 2000, which incorporates the optimal estimation theory employed in the retrieval process. This theory has been extensively discussed in prior works (Herbin, Labonnote, and Dubuisson 2013) and briefly summarized here. As this study builds on previous research, certain sections are condensed, focusing only on essential details. For a comprehensive explanation, please refer to El Kattar et al. 2020.

In cases where the atmosphere is divided into discrete layers, the forward radiative transfer equation establishes an analytical relationship between the observation set y (radiance) and the true atmospheric parameter vector x (vertical CO_2 concentration profiles to be retrieved):

$$y = F(x; b) + \varepsilon, \quad (2)$$

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

Here, F is the forward radiative transfer function (ARAHMIS code), b represents fixed parameters influencing the measurement (e.g., atmospheric temperature, interfering species, viewing angle), and ε is the measurement error vector. For the information content analysis presented here, two matrices define the LHR's information: the averaging kernel A and the total error covariance S_x.

The averaging kernel matrix A, quantifies the sensitivity of the retrieved state to the true state and is defined as follows:

$$A = \partial\hat{x}/\partial x = GK, \tag{3}$$

where K is the weighting function (more commonly called the Jacobian matrix), where the ith row represents the partial derivatives of the ith measurement with respect to each (j) element of the state vector: $K_{ij} = (\partial F_i/\partial x_j)$, and K^T denotes its transpose.

The gain matrix G, is defined as the matrix whose rows correspond to the derivatives of the retrieved state with respect to the spectral points, as follows:

$$G = \partial\hat{x}/\partial y = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} K^T S_\varepsilon^{-1}, \tag{4}$$

where S_a stands for the a priori covariance matrix, reflecting our knowledge of the state space prior to measurement, while S_ε denotes the covariance matrix encompassing errors from both the measured signal and the forward model.

At any altitude, the peak of an averaging kernel row marks the altitude of the highest sensitivity, while its full width at half maximum (FWHM) estimates the vertical resolution. The Degrees Of Freedom (DOFs) of the signal, given by the trace of matrix A, represents the number of independent pieces of information retrievable from observations. In an ideal retrieval with an optimal inverse method, the averaging kernel matrix A would equal the identity matrix, and the DOFs would match the size of the state vector. Hence, each parameter to be retrieved corresponds to a partial degree of freedom, represented by the respective diagonal element of A.

Another key matrix in the IC study is the error covariance matrix S_x, which characterizes the state space post-measurement. Rodgers 2000 showed that this matrix can be expressed as:

$$S_x = S_{smoothing} + S_{meas.} + S_{fwd.mod.}, \tag{5}$$

In the above equation, the smoothing error covariance matrix $S_{smoothing}$ captures the vertical sensitivity of the measurements to the retrieved profile:

$$S_{smoothing} = (A - I)S_a(A - I)^T, \tag{6}$$

$S_{meas.}$ reflects the influence of the measurement error covariance matrix S_m, derived from spectral noise, on the posterior error covariance matrix S_x. S_m is calculated from the spectral noise as follows:

$$S_{meas.} = GS_mG^T, \tag{7}$$

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

Finally, $S_{fwd.mod.}$ represents the contribution to the posterior error covariance matrix via S_f the forward model error covariance matrix, which accounts for uncertainties in non-retrieved model parameters:

$$S_{fwd.mod.} = GK_bS_b(GK_b)^T = GS_fG^T, \quad (8)$$

where S_b represents the error covariance matrix of the non-retrieved parameters.

3.3 A priori information

The IC analysis uses simulated radiance spectra from the current LHR. The initial CO_2 vertical concentrations in the state vector x_a follow the criteria in Section 3.1, divided into 40 layers from ground level to 40 km at 1 km intervals. Non-retrieved parameters, such as water vapor profile, temperature, and SZA, are included as outlined in Section 3.3.3. A priori values and their variability are detailed in **Table 1** and discussed in subsequent sections.

3.3.1 A priori covariance matrix

The a priori error covariance matrix S_a can be evaluated using in-situ data or climatology, but diagonal matrices are often used for space-based retrievals (e.g., De Wachter et al. 2017). As our study prioritizes measurement information over climatology or in situ data, we assume S_a is a diagonal matrix, with each diagonal element ($S_{a,ii}$) defined as:

$$S_{a,ii} = \sigma_{a,i}^2 \text{ with } \sigma_{a,i} = x_{a,i}.\frac{p_{error}}{100}, \quad (9)$$

where $\sigma_{a,i}$ denotes the standard deviation in the Gaussian statistics framework, and the subscript i corresponds to the ith parameter of the state vector. The CO_2 profile a priori error is derived from (U. Schmidt 1991) aligns with prior studies using FTS instruments (El Kattar et al. 2020).

3.3.2 Measurement error covariance matrix

The measurement error covariance matrix is calculated based on instrument performance and accuracy, linked to the radiometric noise characterized by the SNR (discussed in Section 2). We assume that this matrix is diagonal, with the ith diagonal element computed as:

$$S_{m,ii} = \sigma_{m,i}^2 \text{ with } \sigma_{m,i} = \frac{y_i}{SNR}, \quad (10)$$

where $\sigma_{m,i}$ is the standard deviation of the ith measurement (y_i) in vector y, representing the noise equivalent spectral radiance. The LHR's theoretical SNR is estimated at ~710, with additional instrument details provided in **Table 3**.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

3.3.3 Characterization and accuracy of non-retrieved parameters

Errors from non-retrieved parameters are complex, primarily arising from water vapor and temperature effects in our scenario (see **Figure 2**). We assume vertically uniform uncertainties for both. Notably, water vapor is treated as a non-retrieved parameter in this study.

We set the H_2O column uncertainty (p_{Cmol}) at 10% instead of using a profile error. For temperature, we assumed a realistic uncertainty of $\delta T = 1\ K$ for each layer, consistent with typical ECMWF assimilation values. The SZA uncertainty is set at 0.35°, reflecting typical solar angle variations during measurements. These values are summarized in **Table 1**.

The total forward model error covariance matrix (S_f), assumed diagonal, is the sum of contributions from each diagonal element, with the ith diagonal element ($S_{f,ii}$) expressed as:

$$S_{f,ii} = \sum_{j=1}^{n\ level} \sigma_{f,T_j,i}^2 + \sigma_{f,H_2O,i}^2 + \sigma_{f,SZA,i}^2, \quad (11)$$

This section excludes spectroscopic effects like line parameters, line mixing, and continuum errors, which are discussed in Section 3.4.2 in relation to the X_G column estimation.

State vector elements	T	H_2O	SZA	CO_2
A priori values	Radiosondes	Radiosondes	10/80°	TCCON database
A priori uncertainty (p_{error})	1 K per layer	10%	0.35°	1.3%-8%

Table 1: State vector parameters.

3.4 Application of Information Content analysis to CO_2 profiles

We perform an information content analysis on the CO_2 broadband spectrum. The state vector includes gas concentrations at each level from 0 to 40 km, matching FTS and MAGIC instrument altitudes (balloons exceeding 25 km). This setup estimates each gas profile individually, with other atmospheric parameters and gas profiles known from ancillary data with specific uncertainties. Two SZAs, 10° and 80°, are chosen to illustrate the impact of solar optical path on sensitivity (depends on viewing geometry). Detailed discussions on averaging kernels, error budgets, and total column estimations follow in subsequent subsections.

3.4.1 Estimation of Averaging Kernels and error budget

Figure 3 shows the averaging kernel A and total posterior error S_x for CO_2 at a 10° SZA. Results for 80° SZA are omitted due to similar vertical distributions, though slight differences in amplitude exist. These variations are discussed to quantify the viewing geometry's impact. A, derived independently using Section 3.3 variability, reflects the partial degree of freedom at each level. Values near 1 indicate measurement-dominated information, while values near 0 suggest prior knowledge dominance. Kernels are close to 1 in the first layer and remain significant between 1 and 10 km, indicating a meaningful improvement of the information, while approaching 0 above 25 km.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

The measurement offers insights into the CO_2 levels from the ground up to 20 km, but at higher altitudes, information primarily relies on prior knowledge due to reduced sensitivity of these gases in the upper atmosphere. This contrast is clear in the error analysis: the a posteriori total error (green line) is much smaller than the a priori error (violet line) in the lower atmosphere (0-15 km), indicating an improved CO_2 profile knowledge. Above 15 km, however, the total posterior error equals the prior error, signaling reduced sensitivity at high altitudes. Additionally, the errors associated with measurement and the forward model's dependence on non-retrieved parameters are minimal compared to other errors, indicating negligible SNR error. Nevertheless, smoothing error outweighs other errors, particularly beyond 20 km, indicating strong reliance on the a priori profile at higher altitudes and minimal contribution from measurements.

In **Table 4**, the DOFs for CO_2 are presented for both 10° and 80° angles. The table indicates that, with a diagonal prior covariance matrix, four to five partial tropospheric columns for CO_2 can be retrieved. As anticipated, the DOFs are slightly higher at 80° due to the longer solar optical path through each layer. The total profile error, derived from the diagonal of S_x, is discussed in the next section. In general, it's worth noting that the LHR demonstrates high sensitivity with reduced error in the lower layers, where satellites are least sensitive.

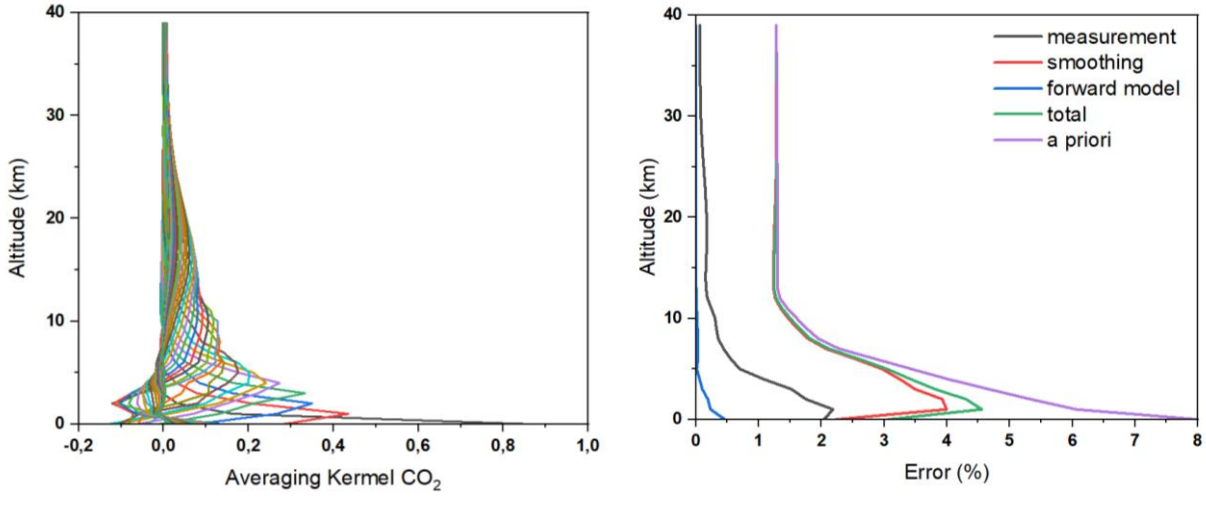

Figure 3: Averaging kernels and error budgets for CO_2 vertical profiles using the LHR. The violet and green lines stand for the prior $\boldsymbol{S_a}$ and posterior $\boldsymbol{S_x}$ errors respectively. The forward model parameters $\boldsymbol{S_{fwd.mod.}}$, smoothing $\boldsymbol{S_{smoothing}}$ and measurement $\boldsymbol{S_{meas.}}$ errors are shown in blue, red and black respectively.

3.4.2 Estimation and uncertainty of total columns

Similar to the LHR, ground-based instruments such as the IFS125HR (TCCON) and EM27/SUN (COCCON) operate in the NIR and derive column-averaged dry-air mole fractions (X_G for gas G) by observing simultaneously the O_2 columns. X_G is computed as the ratio of the gas slant column to the O_2 slant column from the same spectrum. Since the LHR is narrow-banded and does not cover the absorption lines of O_2, a different method is needed to calculate this ratio. Following the NDACC

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

network (De Mazière et al. 2018), X_G is calculated without using oxygen as a reference. Following the method outlined in Wunch et al. 2010 (used in Zhou et al. 2019), X_G for CO_2 can be calculated as follows:

$$X_G = \frac{column_G}{column\ dry\ air}, \tag{12}$$

$$column\ dry\ air = \frac{P_S}{g_{air} m_{air}^{dry}} - column_{H_2O} \frac{m_{H_2O}}{m_{air}^{dry}}, \tag{13}$$

where m_{H_2O} and m_{air}^{dry} are the mean molecular masses of water and dry air, respectively. P_S is the surface pressure and g_{air} the column-averaged gravitational acceleration. Thus, X_G can be calculated if all necessary parameters are available, particularly in field measurements with data from balloons and radiosondes (e.g., temperature, relative humidity, surface pressure). Consequently, X_G values will be computed for the LHR using ARAHMIS, and then compared to the EM27/SUN and the IFS125HR. It's important to note that TCCON's method removes systematic errors common to both the target gas and O_2 columns, which is not possible here.

Moreover, the total column uncertainty is calculated by adding up the concentration of each layer, adjusted by the dry air column (**Figure 3**). **Table 2** displays the total column uncertainties for both zenith angles using a diagonal a priori covariance matrix: 2.74% for CO_2 at 10° and 2.31% at 80°. Lower uncertainty at 80° results from better information distribution due to a longer optical path difference (OPD). The main source of uncertainty stems from smoothing, which outweighs other uncertainties and contributes significantly to the total profile error. Key forward model parameters like H_2O, temperature and SZA influence non-retrieved parameter uncertainty. Additionally, spectroscopic uncertainty, not included here, is purely systematic. Evaluating this uncertainty is complex because we use different absorption lines, each with distinct uncertainties listed in HITRAN.

Error	CO_2	
SZA	10°	80°
Smoothing	2.5	1.91
Measurement	0.99	1.05
Non-retrieved parameters	0.114	0.311
Total	2.74	2.31

Table 2: The total column errors for the CO_2 profile for the LHR for the two SZAs. The uncertainties are expressed as percentages (%).

3.5 Comparison with ground-based FTS instruments

A similar previous study was performed for ground based Fourier Transform spectrometers, including the TCCON's IFS125HR, COCCON's EM27/SUN, and another spectrometer operating in the middle infrared region called CHRIS (El Kattar et al. 2020), as part of the MAGIC campaigns. Here, we compare these instruments with the LHR, but first we present in **Table 3** the different characteristics of the various instruments involved in this study.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

	Resolution(cm^{-1})	OPD (cm)	Spectral region (cm^{-1})	SNR
LHR	0.0047	Fiber-coupled	6338-6378	710
EM27/SUN	0.5	1.8	4700-12500	1080
IFS125HR (TCCON)	0.02	45	4000-15000	~750
CHRIS	0.135	4.52	680-5200	780

Table 3: Instrumental characteristics of the LHR, CHRIS, EM27/SUN and IFS125HR of TCCON.

The methodology from Section 3.2 is applied: the state vector includes only CO_2 concentrations across 0–40 km layers, incorporating the SNR and spectral resolution specific to the FTS instruments (see **Table 3**). A comparison of averaging kernels with FTS instruments reveals sharper peaks and a more homogeneous vertical distribution than CHRIS, EM27/SUN and IFS125HR, suggesting higher sensitivity at higher altitudes though the a posteriori error S_x is significantly reduced in the lower atmosphere. This is further supported by the error budget analysis: the a posteriori total error (green line) remains distinguishable from the a priori error (violet line) even in the higher atmosphere. This discrepancy is due to LHR's higher spectral resolution compared to FTS instruments, ensuring continuous enhancement of our understanding along the atmospheric column.

Table 4 shows the DOFs for CO_2 and the total profile error for both viewing angles using the diagonal a priori covariance matrices. FTS instruments have DOFs ranging from 2.95 at 10° to 4.23 at 80°, while for the LHR, they're 4.13 and 5.15 respectively. This means the LHR can retrieve the same number of CO_2 partial columns at 10° and an extra column at 80° compared to the FTS instruments. Generally, at an 80° angle, the LHR can retrieve one to two additional CO_2 columns in the troposphere, while the profile error remains the same at 10° and improves at 80°.

	DOFs		Error	
	10°	80°	10°	80°
LHR	4.13	5.15	2.74%	2.31%
EM27/SUN	3.03*	3.31*	2.77%*	2.67%*
IFS125HR (TCCON)	3.9*	4.23*	2.82%*	2.72%*
CHRIS	2.95*	3.71*	2.89%*	2.6%*

Table 4: DOFs and column errors (%) for CO_2, per instrument and viewing angle. *from the previous study (El Kattar et al. 2020)

4 Channel selection

The time required to obtain one spectrum with the LHR depends on the chosen spectral range and step which can take a long time, so by selecting the absorption lines prior to the measurement not only we reduce the acquisition time, but we have the opportunity to register more spectra which can lead to better daily statistics while comparing with satellites. Furthermore, using all channels in retrieval significantly increases computational time and systematic errors due to species correlation,

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

complicating the evaluation of the a priori state vector x_a and the error covariance matrix S_a. Channel selection, as described by Rodgers 2000, optimizes retrievals by identifying the subset of channels offering the most information from high-resolution infrared sounders. Cooper et al. 2006 and Kuai et al. 2010 presents a depiction of this process rooted in the Shannon information content which we describe in this section.

Firstly, an "information spectrum" is constructed to assess the information content concerning the a priori state vector. The channel with the highest information content is selected, and the a posteriori covariance matrix is updated to include its contribution. Using this updated state space, a second channel is chosen to maximize information relative to the new covariance matrix. This iterative process continues until the remaining channels' information falls below the measurement noise level. As suggested by Shannon information content and Rodgers 2000, it is beneficial to work in a basis where measurement errors and prior variances are uncorrelated, enabling comparison of measurement error with prior variability. Thus, the Jacobian matrix K (see Sect. 3.2) is transformed into $\widetilde{K}$ using:

$$\widetilde{K} = S_y^{-1/2} K S_a^{1/2}, \tag{14}$$

where both the a priori and measurement covariance matrices are unit matrices. Rodgers also shows that the number of singular values of $\widetilde{K}$ greater than unity determines the effective rank of the problem, representing independent measurements exceeding the noise measurement.

Let S_i represent the error covariance matrix for the state space after i channels have been selected. The information content of channel j among the remaining unselected channels is expressed as:

$$H_j = \frac{1}{2} log_2(1 + \widetilde{k_j^T} S_i \widetilde{k_j}), \tag{15}$$

where $\widetilde{k_j}$ is the jth row of $\widetilde{K}$. H_j represents the information spectrum (expressed in bits), used to select the first channel. If channel l is chosen, the covariance matrix is updated for the next iteration using:

$$S_{i+1}^{-1} = S_i^{-1} + \widetilde{k_l} \widetilde{k_l^T}, \tag{16}$$

Channels are selected iteratively until 90% of the total information spectrum H is achieved, ensuring the measurement noise threshold is not exceeded.

DOFs	CO_2	
	90%	99%
Number of channels	258	1484
Percentage of the total number of channels	2.8%	16.13%

Table 5: Number of selected channels for the DOFs of CO_2 and their percentage of the total channels for the LHR.

After converting H to DOFs, we obtained **Figure 4**, which illustrates the evolution of the CO_2 total DOFs as a function of the number of selected channels for a SZA of 10°. Initially, in **Figure 4**, the DOFs show a sharp increase with the first selected

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

channels, followed by a more gradual rise. **Table 5** presents the number of channels needed to achieve 90% and 99% of the total information. Out of the 9196 exploitable channels in the LHR, only 2.8% (258 channels) are necessary to reach 90% of the retrieved information and 16.18% are needed for 99% of the information. In other words, using selected channels corresponding to 90% of the total information content produces results comparable to using all channels, as nearly 98% of the information is redundant.

Additionally, in **Figure 5**, the first 10 channels are shown in red, channels 11 to 20 in blue, and channels 21 to 30 in green. These channels have the highest IC for the CO_2 Jacobian. Notably, the information primarily originates from three lines in the range 6362-6365 cm^{-1}. Therefore, in future acquisitions, this range can be used for faster measurements while keeping a small step scan. For future measurements, focusing on this range can speed up data collection while maintaining a small scan step, allowing us to quickly track CO_2 concentration changes. These results emphasize the importance of identifying the best channels for CO_2, making the retrieval process easier and more efficient, which is one of the benefits of using a broadband tunable laser.

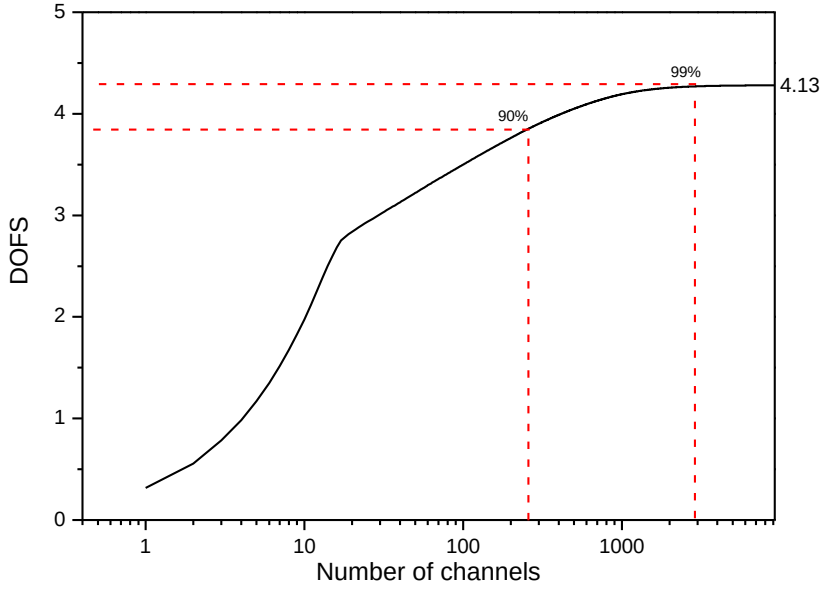

Figure 4: Evolution of the DOFs with the number of selected channels for CO_2.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

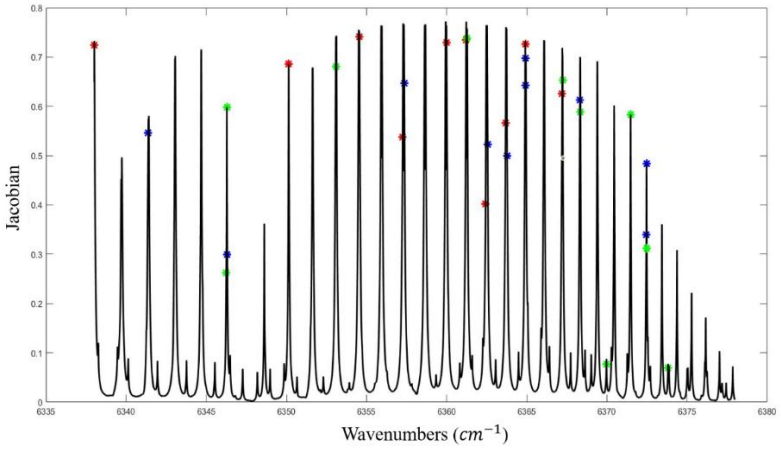

Figure 5: Channels with the highest IC for CO_2, where the first 10 channels shown in red, channels 11 to 20 in blue, and channels 21 to 30 in green

5 Conclusions

In conclusion, this paper presents the measurement capabilities of a new near-infrared laser heterodyne radiometer, which allows the retrieval of CO_2 in the atmospheric column based on ground heterodyne measurement of the sunlight. This spectroradiometer has exceptionally high spectral resolution (0.0047 cm^{-1}) and an exploitable spectral domain ranging from 6338 to 6378 cm^{-1}. An extensive information content analysis is conducted to evaluate the LHR's potential for CO_2 retrieval, using two SZAs (10° and 80°) to quantify the impact of solar optical path on information quality. The total column uncertainty is estimated, revealing a 2.74% error at 10° with a diagonal a priori covariance matrix. Furthermore, a comparison has been carried out with the referenced FTS instruments, such as the TCCON's IFS125HR and COCCON's EM27/SUN, both widely utilized in satellite validation. The LHR exhibits unique advantages over the FTIR spectrometers in retrieving gas columns with better vertical discretization. It is therefore a promising alternative instrument for local scale measurements or for satellite validation. Finally, a channel selection is implemented to eliminate redundant information and identify an optimal spectral range in order to improve daily statistics.

Author contributions

TW and AS performed the measurements; MTEK analyzed the data and wrote the manuscript draft; HH and WC validated, reviewed and edited the manuscript.

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

The authors thank the co-financial supports from the LABEX CaPPA project (ANR-10-LABX005), the CPER ECRIN program, the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant agreement No 872081.

References

Clough, S A et al. 2005. "Atmospheric Radiative Transfer Modeling: A Summary of the AER Codes." *Journal of Quantitative Spectroscopy and Radiative Transfer* 91(2): 233–44. http://www.sciencedirect.com/science/article/pii/S0022407304002158.

Cooper, Steven J. et al. 2006. "Objective Assessment of the Information Content of Visible and Infrared Radiance Measurements for Cloud Microphysical Property Retrievals over the Global Oceans. Part II: Ice Clouds." *Journal of Applied Meteorology and Climatology* 45(1): 42–62.

Eldering, Annmarie et al. 2017. "The Orbiting Carbon Observatory-2: First 18 Months of Science Data Products." *Atmospheric Measurement Techniques* 10(2): 549–63.

Frey, M et al. 2019. "Building the COllaborative Carbon Column Observing Network (COCCON): Long-Term Stability and Ensemble Performance of the EM27/SUN Fourier Transform Spectrometer." *Atmos. Meas. Tech.* 12(3): 1513–30. https://amt.copernicus.org/articles/12/1513/2019/.

Gordon, I. E. et al. 2022. "The HITRAN2020 Molecular Spectroscopic Database." *Journal of Quantitative Spectroscopy and Radiative Transfer* 277.

Herbin, H., L. C. Labonnote, and P. Dubuisson. 2013. "Multispectral Information from TANSO-FTS Instrument - Part 1: Application to Greenhouse Gases (CO2 and CH4) in Clear Sky Conditions." *Atmospheric Measurement Techniques* 6(11): 3301–11.

IPCC, 2014. 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland, 151 pp.

Jiang, X. et al. 2016. "Earth and Space Science Satellite Retrievals and Models." *Earth and Space Science* 3: 1–10.

El Kattar, M.-T., F Auriol, and H Herbin. 2020. "Instrumental Characteristics and Potential Greenhouse Gas Measurement Capabilities of the Compact High-Spectral-Resolution Infrared Spectrometer: CHRIS." *Atmos. Meas. Tech.* 13(7): 3769–86. https://amt.copernicus.org/articles/13/3769/2020/.

El Kattar, Marie-Thérèse, Frédérique Auriol, and Hervé Herbin. 2020. "Instrumental Characteristics and Potential Greenhouse

https://doi.org/10.5194/egusphere-2025-250
Preprint. Discussion started: 26 March 2025

Gas Measurement Capabilities of the Compact High-Spectral-Resolution Infrared Spectrometer: CHRIS." *Atmospheric Measurement Techniques* 13(7): 3769–86.

Kuai, Le et al. 2010. "Channel Selection Using Information Content Analysis: A Case Study of CO2 Retrieval from near Infrared Measurements." *Journal of Quantitative Spectroscopy and Radiative Transfer* 111(9): 1296–1304. https://www.sciencedirect.com/science/article/pii/S0022407310000828.

De Mazière, Martine et al. 2018. "The Network for the Detection of Atmospheric Composition Change (NDACC): History, Status and Perspectives." *Atmospheric Chemistry and Physics* 18(7): 4935–64.

Meftah, M et al. 2018. "SOLAR-ISS: A New Reference Spectrum Based on SOLAR/SOLSPEC Observations★ ." *A&A* 611. https://doi.org/10.1051/0004-6361/201731316.

Michalsky, Joseph J. 1988. "The Astronomical Almanac's Algorithm for Approximate Solar Position (1950-2050)." *Solar Energy* 40(3): 227–35.

Palmer, Paul I. et al. 2019. "Potential Improvements in Global Carbon Flux Estimates from a Network of Laser Heterodyne Radiometer Measurements of Column Carbon Dioxide." *Atmospheric Measurement Techniques* 12(4): 2579–94.

Rodgers, C.D. 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice.* Singapore: World Scientific. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1223583.

Sun, Chunyan et al. 2024. "Optimization of Signal-to-Noise Ratio of Laser Heterodyne Radiometer." *Microwave and Optical Technology Letters* 66(1): e33857. https://doi.org/10.1002/mop.33857.

Toon, Geoffrey C. 2015. "Solar Line List for the TCCON 2014 Data Release." https://data.caltech.edu/records/251.

U. Schmidt, and A. Khedim. 1991. "In Situ Measurements of Carbon Dioxide in the Winter Arctic Vortex and at Midaltitudes: An Indicator of the 'age' of Stratospheric Air." 18(4): 763–66.

De Wachter, Evelyn et al. 2017. "Retrieval and Validation of MetOp/IASI Methane." *Atmospheric Measurement Techniques* 10(12): 4623–38.

Wang, Jingjing et al. 2023. " External-Cavity Diode Laser-Based near-Infrared Broadband Laser Heterodyne Radiometer for Remote Sensing of Atmospheric CO 2 ." *Optics Express* 31(6): 9251.

Weidmann, Damien. 2021. "4 - Atmospheric Trace Gas Measurements Using Laser Heterodyne Spectroscopy." In eds. Weidong Chen, Dean S Venables, and Markus W B T - Advances in Spectroscopic Monitoring of the Atmosphere Sigrist. Elsevier, 159–223. https://www.sciencedirect.com/science/article/pii/B9780128150146000051.

Wunch, D. et al. 2010. "Calibration of the Total Carbon Column Observing Network Using Aircraft Profile Data." *Atmospheric Measurement Techniques* 3(5): 1351–62.

Zenevich, Sergei et al. 2020. " Improvement of Dark Signal Evaluation and Signal-to-Noise Ratio of Multichannel Receivers in NIR Heterodyne Spectroscopy Application for Simultaneous CO 2 and CH 4 Atmospheric Measurements ." *OSA Continuum* 3(7): 1801.

Zhou, Minqiang et al. 2019. "TCCON and NDACC XCO Measurements: Difference, Discussion and Application." *Atmospheric Measurement Techniques* 12(11): 5979–95.