

# Methane, Ethane, and Propane Detection Using a Quartz-Enhanced Photoacoustic Sensor for Natural Gas Composition Analysis

Aldo F. P. Cantatore, Giansergio Menduni,\* Andrea Zifarelli, Pietro Patimisco, Marilena Giglio, Miguel Gonzalez,\* Huseyin R. Seren, Pan Luo, Vincenzo Spagnolo, and Angelo Sampaolo




Cite This: <https://doi.org/10.1021/acs.energyfuels.4c03726>



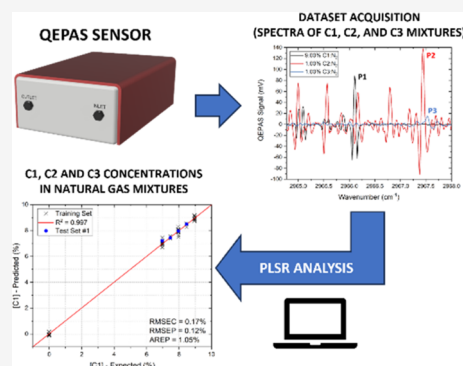
Read Online

ACCESS |

 Metrics & More

 Article Recommendations

**ABSTRACT:** A compact and portable gas sensor based on quartz-enhanced photoacoustic spectroscopy (QEPAS) for the detection of methane (C1), ethane (C2), and propane (C3) in natural gas (NG)-like mixtures is reported. An interband cascade laser (ICL) emitting at 3367 nm is employed to target absorption features of the three alkanes, and partial least-squares regression analysis is employed to filter out spectral interferences and matrix effects characterizing the examined gas mixtures. Spectra of methane, ethane, and propane mixtures diluted in nitrogen are employed to train and test the regression algorithm, achieving a prediction accuracy of ~98%, ~96%, and ~93% on C1, C2, and C3, respectively. With respect to previously reported QEPAS sensors for natural gas analysis, the high prediction accuracy as well as the capability to discriminate and detect C3 within natural gas-like complex mixtures provided by the employment of partial least-squares regression mark significant improvements. Furthermore, these results enable an improved performance of the sensor for in situ, real-time, and online natural gas composition analysis.



## 1. INTRODUCTION

Gas detection and compositional analysis are fundamental across the oil and gas industries. During the exploration and production of hydrocarbon fluids, gas may be captured for analysis from downhole tools, the wellhead, from separators, or extracted from drilling fluids at the surface through mud logging operations.<sup>1,2</sup> The relative abundance of different light hydrocarbon species (from methane to pentane, C1–C5) in a gas sample can be used to determine the properties of the gas (e.g., wetness, density), and their ratios are useful in fluid typing and fingerprinting.<sup>3,4</sup> Gas composition information also helps in the economic evaluation and designing of the facilities for production.<sup>5</sup> Once gas plants refine a commercial natural gas (NG) product, its composition is monitored to determine its calorific value, which is related to its combustibility and therefore indicative of its quality and final price.<sup>6</sup> The calorific value is then monitored during distribution as gas from different sources is blended, transported, and sold to end-users. Up to date, standard methods for the analysis of natural gas composition require the use of gas chromatographs (GCs), frequently coupled to flame ionization detectors (GC-FIDs) or mass spectrometers (GC-MSs).<sup>7,8</sup> These conventional methods typically require extensive maintenance and calibration tasks in the field. On the other hand, optical spectroscopy techniques, while being simpler, more cost-effective and reliable, and better suited for online monitoring, have been hindered by challenges in obtaining quantitative composition

information due to their overlapping spectral components in natural gas samples. However, when combined with chemometric methods, optical spectroscopy can provide improved quantitative predictions.<sup>9</sup>

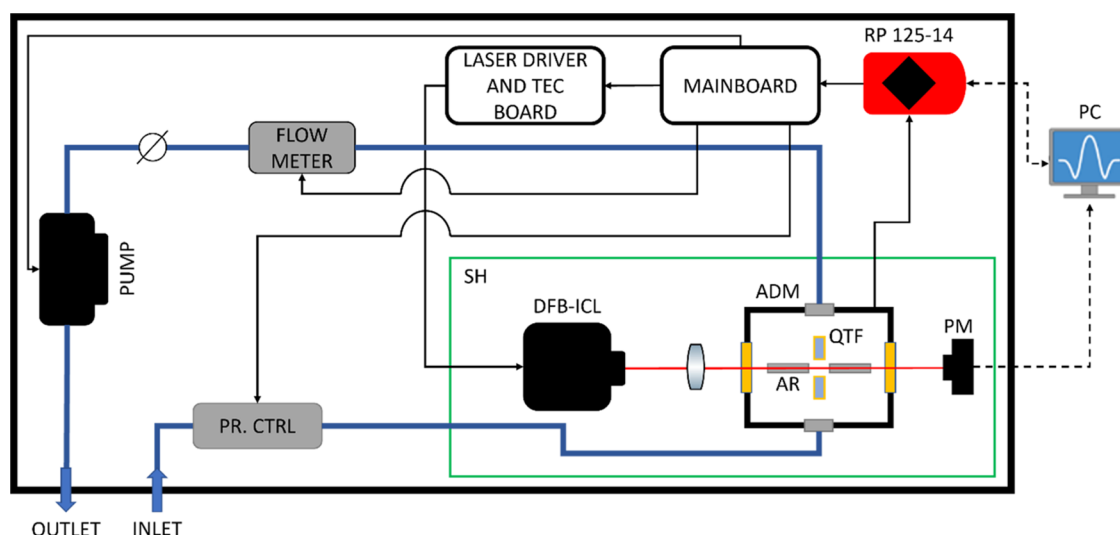
Despite being characterized by an extremely high sensitivity, GC analysis has multiple drawbacks, which ultimately depend on either the detectors or the hyphenated systems deployed.<sup>8,10</sup> For example, FIDs display an almost universal response to organic compounds, thus becoming strongly limited when applied to composition analysis.<sup>8</sup> On the other hand, the main drawbacks of GC-MS systems are represented by (i) the need for high vacuum pumps (down to  $10^{-7}$ – $10^{-6}$  Pa), (ii) the coupling interface between the two instruments, (iii) the cross-interferences in the mass spectra, (iv) the analyte separation intrinsic to the technique, and (v) the overall cost of the whole apparatus.<sup>8,9</sup>

Gas sensing systems based on optical spectroscopy have been demonstrated to be a solid and reliable alternative to traditional analytical techniques for hydrocarbon detection.<sup>11–14</sup> Among the large number of optical-based

Received: August 1, 2024

Revised: October 6, 2024

Accepted: October 30, 2024



**Figure 1.** Architecture of the QEPAS sensor. Black arrows represent electronic connections, bold blue lines represent gas handling line connections, and dashed black arrows represent USB connections. The box walls are represented as bold black rectangles. DFB-ICL - Distributed FeedBack Interband Cascade Laser, ADM - Acoustic Detection Module, QTF - Quartz Tuning Fork, AR - Acoustic Resonator, PM - Power Meter, SH - Sensor Head, TEC - ThermoElectric Cooler, PR. CTRL - PPressure ConTRoLler, RP - RedPitaya, and PC - Personal Computer.

techniques, gas sensors based on quartz-enhanced photoacoustic spectroscopy (QEPAS) represent a solid alternative to GCs for NG analysis, thanks to their compactness, ruggedness, high selectivity and sensitivity,<sup>15</sup> as well as low cost and capability of performing online, dynamic flow, and non-destructive measurements.<sup>16,17</sup> Indeed, QEPAS sensors have already been successfully deployed for the detection of lighter alkanes, i.e., C1, C2, and C3, from traces to high concentrations.<sup>18–23</sup>

QEPAS is an indirect absorption technique, which is developed as a variation of traditional photoacoustic spectroscopy (PAS), employing a quartz tuning fork (QTF) as a sharply resonant transducer. A laser beam, whose wavelength is resonant with an absorption feature of the target molecule, is focused between the QTF prongs and modulated at its resonance frequency or one of its subharmonics. Then, the weak sound waves, generated through the relaxation of the excited molecules via nonradiative energy transfer processes,<sup>17,24</sup> are detected and converted into an electric signal, exploiting the piezoelectric properties of quartz. As reported in the literature, QEPAS response is dependent on the sample composition, influencing both the photoacoustic relaxation cascade and the sound wave propagation.<sup>25–27</sup> Fluctuations in the mixture composition may lead to variations of the gas density, influencing both the sound speed and the QTF quality factor.<sup>27,28</sup> Conversely, the relaxation rate of the excited molecule is affected by the different components of the mixture, having an effect on its relaxation pathway and, in turn, on the QEPAS signal. These effects, known as “matrix effects”,<sup>19,29,30</sup> together with overlapping spectral features of different analytes, are the most challenging aspects of QEPAS signal processing.<sup>29,31,32</sup> The issue of spectral overlap is crucial in hydrocarbon detection by optical spectroscopy. In particular, the absorption bands of multiple hydrocarbons fall in the spectral range around 3.3  $\mu\text{m}$ , related to a stretching of the C–H bond common to all of the alkanes.<sup>33</sup> A first tentative quantification of C1, C2, and C3 concentrations in a diluted natural gas-like gas sample was demonstrated by Luo et al.,<sup>23</sup> employing a QEPAS sensor with a 3345 nm interband cascade

laser (ICL) as a laser source and a standard QTF as a sensitive element. The evaluation of the sample composition relied on a complex algorithm mainly based on univariate calibrations for the three main alkanes and exploiting the saturated absorption of C1 at tens of percent scale to evaluate the effect of the C2 and C3 fluctuations in terms of cross-sensitivities on methane detection. This approach has proved to have limitations in modeling the energy relaxation in multicomponent gas samples, and it does not offer a clear advantage in terms of response time, even though the signal interpretation is based on peak signal extraction rather than extended spectra acquisitions.<sup>23</sup> Recently, partial least-squares regression (PLSR) has been proposed as a multivariate analysis (MVA) technique to filter out matrix effects and spectral interferences in the QEPAS response, thus providing a more accurate prediction of every single analyte concentration in complex gas mixtures compared to the standard univariate calibration.<sup>19,30,34</sup>

In this work, we report on a compact and portable QEPAS sensor to detect C1, C2, and C3 in NG-like mixtures, employing an ICL emitting at 3.367  $\mu\text{m}$  to exploit the most intense absorption band of propane in the whole infrared region. NG-like mixtures are referred to as mixtures containing C1, C2, and C3 in typical NG concentrations but diluted in a 1:10 ratio in pure nitrogen. PLSR is used to retrieve the analyte concentrations, removing both spectral interferences and matrix effects. The data analysis was performed exploiting a machine-learning-like approach, training the algorithm using C1–C2–C3 mixtures, whose concentrations varied in the range of 7–9%, 0–1%, and 0–1% within a nitrogen matrix, respectively. The obtained results demonstrate the suitability of the developed sensor as a versatile tool for real-time and in situ monitoring of the most significant geochemical fingerprints in natural gas, even with widely variable gas matrices, paving the way toward a fully integrated QEPAS-based NG analyzer system.

## 2. SENSING ARCHITECTURE FOR METHANE, ETHANE, AND PROPANE DETECTION

A scheme of the QEPAS sensor employed to acquire the spectra associated with mixtures of C1, C2, and C3 in a pure nitrogen matrix is depicted in Figure 1. The sensor is enclosed within a portable 19"-diagonal aluminum rack box.

The sensor head (SH) hosts the optical components of the sensor, preventing possible misalignments, thus increasing its ruggedness. Light emitted by a DFB-ICL (Nanoplus GmbH), with a central emission wavelength of 3367 nm ( $2970\text{ cm}^{-1}$ ) and a peak power of  $\sim 11\text{ mW}$  at  $23\text{ }^{\circ}\text{C}$ , is focused through an acoustic detection module (ADM, Thorlabs ADM01<sup>35</sup>), i.e., a vacuum-tight gas cell equipped with two ZnSe wedged windows with a  $2\text{--}13\text{ }\mu\text{m}$  antireflection coating (Thorlabs WG70530-E4) and two connectors for gas inlet and outlet. The ADM encloses a spectrophone consisting of a T-shaped quartz tuning fork (QTF) acoustically coupled to a pair of acoustic resonator (AR) tubes, exhibiting an overall resonance frequency  $f_0 = 12461.6\text{ Hz}$  and a  $Q\text{-factor} \cong 16,500$  at an operating pressure of 200 Torr. The QTF output piezoelectric current is converted into a voltage signal by means of a transimpedance amplifier with a  $10\text{ M}\Omega$  feedback resistor (not shown in the setup scheme). A power meter (Thorlabs S120C) is placed behind the ADM for alignment purposes, and it is connected to a PC via a USB connection.

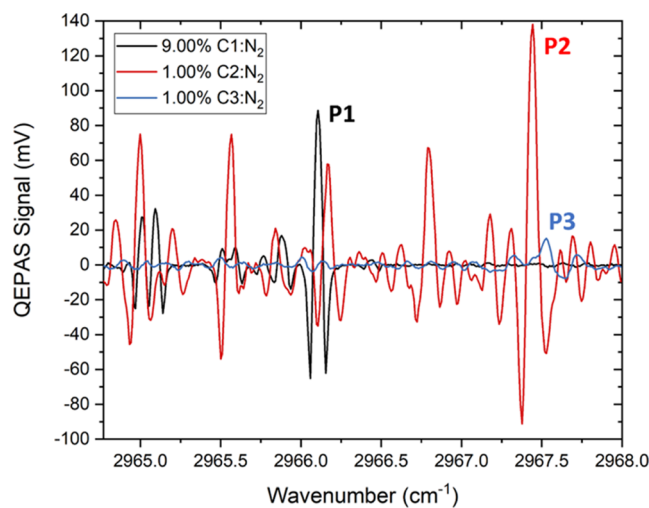
The QTF output signal is fed to a RedPitaya (RP) STEMlab 125-14 board, which is also used to control a laser driver and a thermoelectric cooler (TEC) driver chip (Thorlabs MTD1020T), mounted on a dedicated PCB. A mainboard bridges the electronic connections from the PCB to the RP board, controlled by a PC by means of dedicated LabVIEW-based software. All of the measurements were performed in wavelength modulation and second harmonic detection (WM-2f), modulating the laser current sinusoidally at half of the spectrophone resonance frequency and demodulating the spectrophone output signal at the QTF resonance frequency.<sup>36</sup> A slow ramp can be superimposed to the sinewave to scan the laser dynamic range. Both the modulation and demodulation processes are handled by the RP board and a LabVIEW-based dual-phase digital lock-in amplifier having a maximum input voltage of 1 V.

The measurements were performed at an operating pressure of 200 Torr and a gas flow of 25 sccm. These conditions were kept stable and monitored by means of a pressure controller (Alicat EPC-15PSIA-P0), a flow meter (Axetris MFM 2220-BA-U0), a needle valve, and a diaphragm pump (Thomas 1420BLDC), connected to the RP board through the mainboard. NG-like mixtures of the three alkanes diluted in a nitrogen matrix were generated by means of a gas mixer (MCQ Instruments GB-103). The gaseous samples were generated using cylinders with the following certified concentration: pure C1 (99.50%), 9.85% C1:N<sub>2</sub>, 1.00% C2:N<sub>2</sub>, and 1.00% C3:N<sub>2</sub>, each one characterized by an expanded relative uncertainty of 2%.

## 3. SINGLE ANALYTE SENSOR CALIBRATION AND MATRIX EFFECTS IN TWO GAS MIXTURES

The QEPAS sensor was first calibrated for C1, C2, and C3 separately. The explored C1 concentration range for the calibration spans from 0.99% to 9.85%, while the C2 and C3 concentration ranges span from 0.10% to 1.00%. Performing the measurements at 200 Torr is beneficial since it is a

sufficiently low pressure to ensure a sufficient spectral separation among the absorption features of the three analytes.<sup>37</sup> Examples of the single analyte QEPAS spectra acquired by scanning the laser dynamic range from  $2964.77\text{ cm}^{-1}$  to  $2968.00\text{ cm}^{-1}$  are shown in Figure 2.



**Figure 2.** Examples of the single analyte spectra acquired for a 9.00% C1 (black solid line), 1.00% C2 (red solid line), and 1.00% C3 (blue solid line) concentration in nitrogen within the  $2964.77\text{--}2968.00\text{ cm}^{-1}$  spectral range.

The spectra of the three components exhibit a high overlap in the investigated tuning range. The single analyte calibrations were performed by monitoring the P1, P2, and P3 peaks of C1, C2, and C3, respectively, at different analyte concentrations. The first two peaks correspond to the ro-vibrational transitions falling at  $2966.11\text{ cm}^{-1}$  (C1) and  $2967.49\text{ cm}^{-1}$  (C2).<sup>37</sup> P3 is the most intense peak detectable in the targeted spectral range, belonging to a broad C3 absorption band, as reported in the PNNL database.<sup>38</sup> The calibration curves obtained plotting the QEPAS signal at different C1, C2, and C3 concentrations are shown in Figure 3a–c.

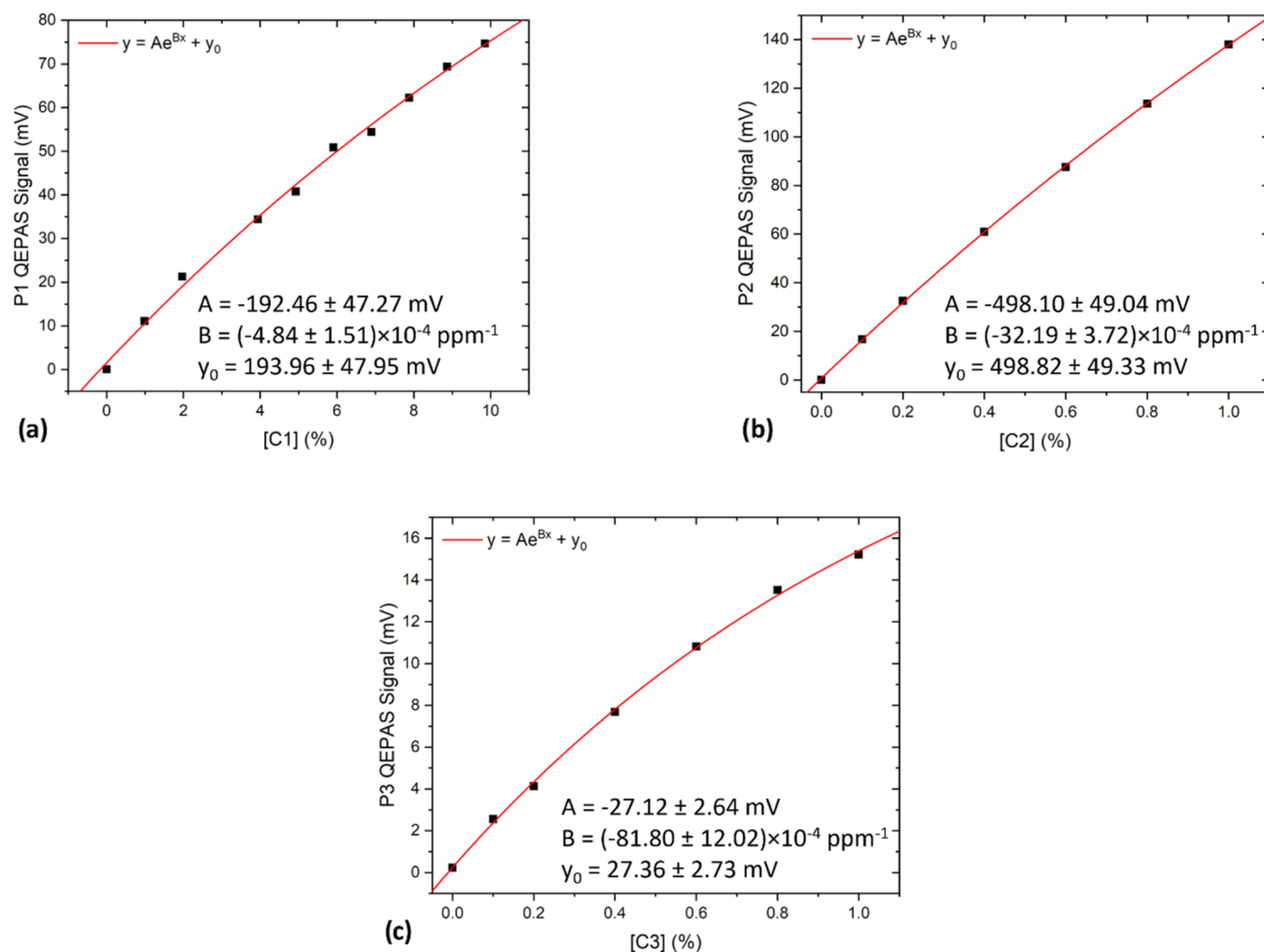
A moderate nonlinear sensor response was observed for all of the analytes, which is consistent with the behavior shown by the QEPAS signal when dealing with high concentrations of the target analytes.<sup>19</sup> Indeed, when the target analyte concentration is too high to consider the first-order approximation of the Beer–Lambert law, a nonlinear trend is expected.<sup>19</sup> Therefore, an exponential fit curve was superimposed on the measured data:

$$y = A e^{Bx} + y_0 \quad (1)$$

The retrieved fit parameters are reported in Table 1.

As reported in the literature, in the explored concentration range, the photoacoustic signal generated by the C1 absorption within a NG-like mixture is affected by the matrix composition.<sup>19</sup> For this reason, the influence of C2 and C3 concentration variations on the C1 photoacoustic signal was investigated. This analysis was carried out by generating several C1–C2 and C1–C3 mixtures and acquiring the QEPAS signal on the P1 peak. The dependence of the P1 QEPAS signal on C2 and C3 concentrations is shown in Figure 4a,b, varying the C1 concentration in the 7.00–9.00% range.

First, by looking at the trend in Figure 4a, a significant influence of C2 on the P1 values in the different mixtures can



**Figure 3.** Calibration curves of (a) C1 performed on the P1 peak in the 0.99–9.85% concentration range; (b) C2 performed on the P2 peak in the 0.10–1.00% concentration range; (c) C3 performed on the P3 peak in the 0.10–1.00% concentration range. In each case, an exponential fit (solid red line) is superimposed on the experimental points.

**Table 1. Parameters Defining the Exponential Fit Superimposed to the Single Analyte Calibrations Performed on C1, C2, and C3**

	A (mV)	B ( $10^{-4}$ ppm $^{-1}$ )	$y_0$ (mV)
C1	$-192.46 \pm 47.27$	$-4.84 \pm 1.51$	$193.96 \pm 47.95$
C2	$-498.10 \pm 49.04$	$-32.19 \pm 3.72$	$498.82 \pm 49.33$
C3	$-27.12 \pm 2.64$	$-81.80 \pm 12.02$	$27.36 \pm 2.73$

be inferred. Indeed, the promoting action of C2 on the C1 relaxation leads to a sharp increase by a factor of 2 up to 0.2% C2. In contrast to the saturation in the C1 signal previously observed,<sup>19</sup> the decrease in the P1 signal for concentrations higher than 0.6% (Figure 4a) is due to the spectral interference with the negative lobe of the C2 absorption peak at 2966.17 cm $^{-1}$  (see Figure 2).<sup>37</sup> The dependence of the C1 signal on the C3 concentration is reported in Figure 4b. In this case, a sharp increase (a factor of  $\sim 2.4$ ) in the P1 signal is observed for C3 concentrations up to 0.2%, with a roll-off for C3 concentrations higher than 0.4%. This behavior has already been predicted and observed for other gases, e.g., water vapor mixed with C1.<sup>30,32</sup> The influence of different C3 concentrations on P2 in two gas mixtures was investigated as well; however, no mutual effects on the measured QEPAS signals were observed.

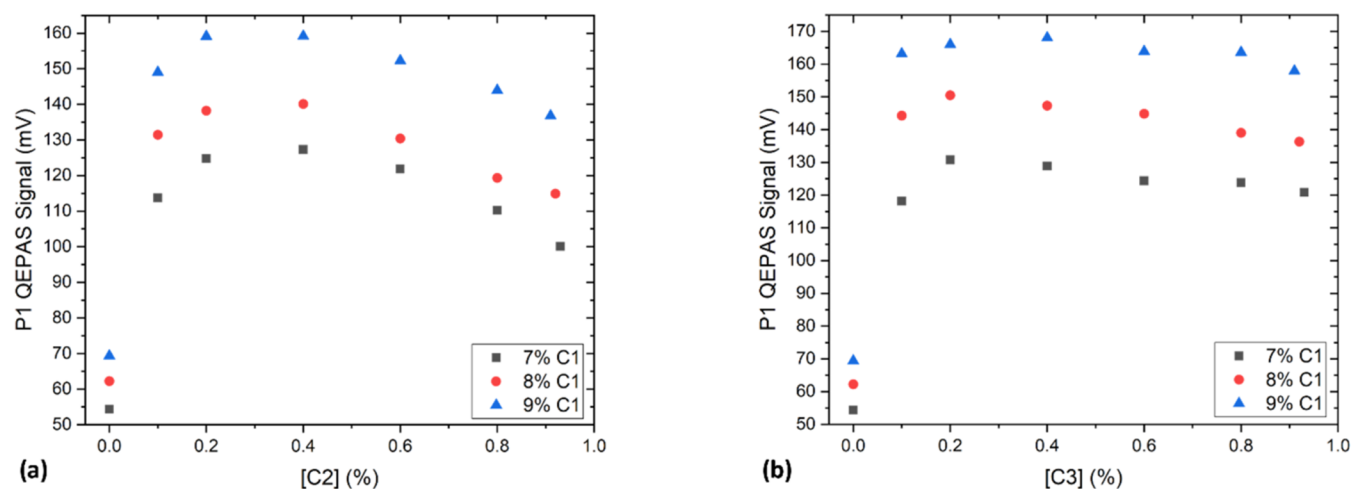
Therefore, both spectral interferences and matrix effects were observed in the analyzed gas samples, thus requiring a suitable analysis tool capable of filtering them out.

#### 4. PARTIAL LEAST-SQUARES REGRESSION IN GAS SPECTROSCOPY: FUNDAMENTALS

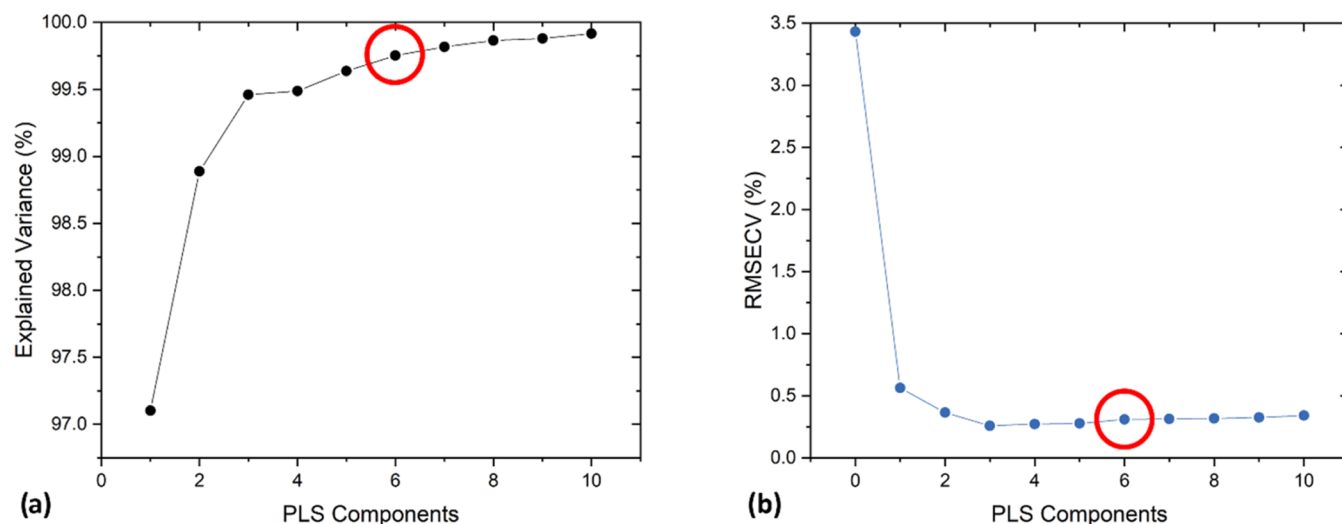
PLSR is a multivariate analysis (MVA) technique, commonly employed to handle noise, missing values, and correlations among nominally independent variables belonging to relatively small datasets.<sup>38</sup> Recently, PLSR has been effectively employed as an analysis tool for filtering out mutual spectral interferences and matrix effects in PAS.<sup>19,30,34,40</sup> This technique is based on a simple regression model, which reads in matrix form:

$$Y = XB + E \quad (2)$$

where  $X$  is the predictor matrix,  $Y$  is the response matrix,  $B$  is the regression coefficient matrix, and  $E$  is the residual (or error) matrix.<sup>39</sup> In optical spectroscopy, the rows of the  $X$  matrix consist of the acquired spectra, while the nominal concentrations of the mixtures are the rows of the  $Y$  matrix. During the training step of the algorithm,<sup>19,32</sup> the regression coefficient matrix is retrieved by maximizing the covariance matrix  $\text{cov}(X,Y)$ , with the only assumption that the analyzed system can be described in terms of a small subset of truly



**Figure 4.** P1 peak QEPAS signal for a fixed concentration of methane ranging from 7.00 to 9.00% when varying the (a) C2 and (b) C3 concentrations within two gas mixtures.



**Figure 5.** (a) Explained variance and (b) root-mean-square error of cross validation (RMSECV) evaluated for each LV (or PLS component) up to 10 employing the full input dataset. The selected number of LVs is 6, as highlighted by the red circle.

independent variables or “latent variables” (LVs).<sup>39,41</sup> From this point of view, PLSR involves the projection to a latent structure of the original system, meaning that the analysis is performed in a subspace of the original wavelength space with a reduced dimensionality, given by the number of LVs.<sup>42</sup> The assessment of the optimal LV number is crucial to avoid possible under- or overfittings of the model, and it is typically carried out by employing two figures of merit: the explained variance maximization and the cross-validation (CV) algorithm.<sup>43</sup>

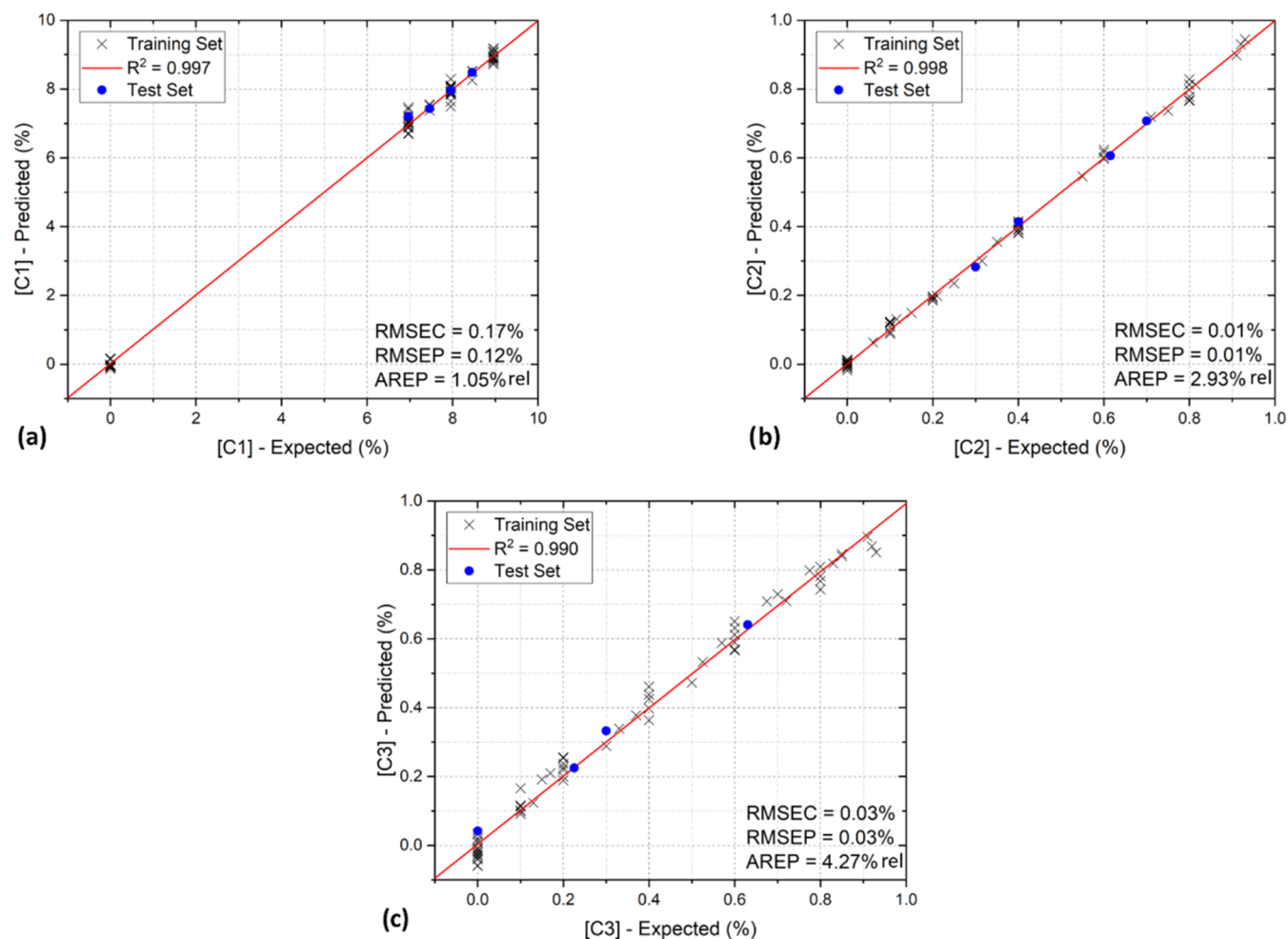
One of the main advantages of the PLSR approach, when applied to PAS, is the possibility of accurately retrieving the concentration of each analyte within a complex gas mixture by only analyzing the cross-correlations among different wavelengths of the same spectrum. Indeed, there is no need for an a priori knowledge of the photoacoustic relaxation cascade dynamics, contrarily to other methods relying on the calculation of the full relaxation cascade, which may lose their effectiveness when the matrix complexity increases.<sup>30–32,40</sup>

## 5. PLSR ANALYSIS: ALGORITHM TRAINING AND TEST

The input data set of the PLSR algorithm was built acquiring 71 spectra with different sample compositions, including 2- and 3-gas mixtures (i.e., C1–C2, C1–C3, C2–C3, and C1–C2–C3) in nitrogen. The concentration ranges considered for the data analysis process are 7.0–9.0% for C1 and 0.1–1.0% for C2 and C3. Each spectrum includes 346 data points, with a spectral sampling of  $\sim 0.0095 \text{ cm}^{-1}$  (Figure 2).

A 10-fold cross-validation (CV) algorithm was applied to the full data set to determine the optimal number of LVs for the subsequent PLS analysis. Both the explained variance and the root-mean-square error of cross validation (RMSECV) were evaluated for each LV, as shown in Figure 5a,b, respectively.

As it can be noticed, the RMSECV is minimized for three LVs, remaining substantially constant for a higher number of PLS components. At the same time, the explained variance is first maximized for three LVs, although it increases significantly up to six LVs, which was selected as the optimal number of components for analysis. Indeed, a higher number of PLS components would lead to only negligible increments. Even though the entire analysis process does not provide a physical



**Figure 6.** Outcome of the PLSR algorithm for (a) C1; (b) C2; and (c) C3 for a representative test set. For each analyte, the RMSEC, the RMSEP, and the AREP are displayed.

**Table 2.** Predicted and Expected Concentrations for Each Analyte and Each Test Set

	C1		C2		C3	
	expected (%)	predicted (%)	expected (%)	predicted (%)	expected (%)	predicted (%)
test set #1	7.96 ± 0.18	7.96 ± 0.17	0.40 ± 0.01	0.41 ± 0.01	0	0.04 ± 0.03
	6.97 ± 0.16	7.21 ± 0.17	0.30 ± 0.01	0.28 ± 0.01	0.63 ± 0.01	0.64 ± 0.03
	7.46 ± 0.17	7.43 ± 0.17	0.70 ± 0.02	0.71 ± 0.01	0.23 ± 0.01	0.22 ± 0.03
	8.46 ± 0.19	8.48 ± 0.17	0.62 ± 0.01	0.61 ± 0.01	0.30 ± 0.01	0.33 ± 0.03
test set #2	7.96 ± 0.18	7.85 ± 0.16	0.20 ± 0.01	0.18 ± 0.01	0	0.03 ± 0.03
	6.97 ± 0.16	7.28 ± 0.16	0.10 ± 0.01	0.11 ± 0.01	0.83 ± 0.02	0.80 ± 0.03
	7.46 ± 0.17	7.74 ± 0.16	0.40 ± 0.01	0.38 ± 0.01	0.53 ± 0.01	0.51 ± 0.03
	8.96 ± 0.20	9.47 ± 0.16	0.71 ± 0.02	0.72 ± 0.01	0.20 ± 0.01	0.26 ± 0.03
test set #3	7.46 ± 0.17	7.36 ± 0.17	0.25 ± 0.01	0.23 ± 0.01	0.68 ± 0.02	0.69 ± 0.03
	7.96 ± 0.18	7.93 ± 0.17	0.55 ± 0.01	0.54 ± 0.01	0.37 ± 0.01	0.37 ± 0.03
	8.46 ± 0.19	8.02 ± 0.17	0.82 ± 0.02	0.81 ± 0.01	0.10 ± 0.01	0.08 ± 0.03
	8.96 ± 0.20	8.92 ± 0.17	0.06 ± 0.01	0.06 ± 0.01	0.85 ± 0.02	0.83 ± 0.03
test set #4	6.97 ± 0.16	6.69 ± 0.16	0.93 ± 0.02	0.94 ± 0.01	0	−0.05 ± 0.03
	7.46 ± 0.17	7.57 ± 0.16	0.40 ± 0.01	0.38 ± 0.01	0.53 ± 0.01	0.52 ± 0.03
	7.96 ± 0.18	7.95 ± 0.16	0.75 ± 0.02	0.74 ± 0.01	0.17 ± 0.01	0.21 ± 0.03
	8.96 ± 0.20	8.77 ± 0.16	0.21 ± 0.01	0.20 ± 0.01	0.70 ± 0.02	0.73 ± 0.03
test set #5	8.96 ± 0.20	9.14 ± 0.16	0.60 ± 0.01	0.60 ± 0.01	0	−0.01 ± 0.03
	6.97 ± 0.16	6.87 ± 0.16	0.80 ± 0.02	0.80 ± 0.01	0.13 ± 0.01	0.13 ± 0.03
	7.96 ± 0.18	7.64 ± 0.16	0.20 ± 0.01	0.19 ± 0.01	0.72 ± 0.02	0.69 ± 0.03
	8.46 ± 0.19	8.62 ± 0.16	0.12 ± 0.01	0.14 ± 0.01	0.80 ± 0.02	0.77 ± 0.03

interpretation of the LVs, it is still possible to interpret them in a qualitative way as independent contributions to the QEPAS spectra. Namely, the first three of them may be considered as the independent contributions of the three analytes composing the mixtures; the others may be considered as the matrix effects related to (i) the photoacoustic relaxation of C1 through C2, (ii) the photoacoustic relaxation of C1 through C3, and (iii) the self-relaxation of each analyte, which may be assumed to become relevant in the explored concentration ranges.<sup>19</sup>

The PLSR analysis was carried out by means of a MATLAB-based algorithm, splitting the acquired dataset into a training set and a test set. The two sets are then employed to calibrate and test the model, respectively. The coefficient matrix retrieved from the calibration step (see eq 2) is employed to predict the C1, C2, and C3 concentrations of the corresponding test set. Five independent test sets were assembled, each one consisting of four spectra randomly picked from the full dataset. For each test set, the root-mean-square error of calibration (RMSEC, %) and root mean square error of prediction (RMSEP, %) were evaluated together with the average relative error of prediction (AREP, % rel). The RMSEC parameter provides quantitative information about the algorithm calibration, thus returning information on its precision, while the RMSEP parameter provides information about the discrepancy between the expected values and the obtained results, thus returning information about its accuracy.<sup>44</sup> The results of the calibration and the test step of the PLSR analysis for a representative test set are shown in Figure 6a–c for C1, C2, and C3, respectively. The expected and predicted concentrations calculated for each test set are reported in Table 2, together with their uncertainties. The uncertainty on the expected concentration has been evaluated starting from the expanded uncertainty on the gas cylinders and the relative uncertainty on the dilution (1%), while the RMSEC has been adopted as the uncertainty on the predicted concentrations.

The data collected as the outcome of the algorithm display a clear linear trend for each analyte, and the superimposed linear fits always fall close to the bisector of the plane, as expected, with an  $R^2$  value always higher than 0.990. Furthermore, the significant decrease in the relative prediction accuracy observed on C3 with respect to the other two analytes can be attributed to the lower number of C3 absorption features within the targeted spectral range as well as to the lower signal-to-noise ratio (SNR) of the C3 QEPAS signals, as can be noticed from the acquired spectra (e.g., Figure 2). Therefore, it can be inferred that a low SNR induces a degradation of the PLSR algorithm prediction accuracy with respect to one or more analytes composing the gas mixture.

From Table 2, it can be noticed that the RMSEC varies between 0.16% and 0.17% for C1 and 0.01% and 0.02% for C2 and C3 when changing the test set removed from the full dataset. This proves that both the calibration and the validation steps are unbiased. Furthermore, the predicted and expected concentrations are compatible within their uncertainties ( $1\sigma$ ), with only a few minor exceptions.

The mean RMSEC, RMSEP, and AREP evaluated over all five test sets are reported in Table 3. The average AREP values point out an algorithm accuracy of ~98%, ~96%, and ~93% for C1, C2, and C3 concentration predictions, respectively. It is worth noting that varying the number of LVs in a neighborhood of 6 does not lead to substantial variations of

**Table 3. RMSEC, RMSEP, and AREP Averaged Out of All of the Test Sets for Each Analyte**

	C1	C2	C3
mean RMSEC (%)	0.16	0.01	0.03
mean RMSEP (%)	0.22	0.01	0.03
mean AREP (% rel)	2.19	4.31	6.50

the average AREP values, hence confirming the validity of the choice.

## 6. CONCLUSIONS

In this work, a compact and portable QEPAS sensor for in situ and real-time detection of light alkanes, namely, C1, C2, and C3, in natural gas-like mixtures is presented. The sensor consists of a 19"-diagonal aluminum rack box, enclosing the sensor head, the electronics, and the gas handling system. Natural gas component analysis has been carried out in the laboratory generating mixtures of C1 (7.00–9.00%), C2 (0–1.00%), and C3 (0–1.00%) in nitrogen, starting from certified cylinders. PLSR analysis has been implemented by means of a MATLAB-based code to filter out both spectral interferences and matrix effects observed in the acquired QEPAS spectra, thus allowing the retrieval of C1, C2, and C3 concentrations with an accuracy of ~98%, ~96%, and ~93%, respectively. These results mark a significant improvement with respect to the previously reported demonstrations of QEPAS sensors for natural gas analysis from the point of view of<sup>19,23</sup> (i) the applicability of the multivariate approach to complex and highly variable NG gas matrices, (ii) the prediction accuracy, and (ii) the capability to detect and discriminate C3 within natural gas-like complex mixtures.

Further developments of this technology will be focused on field testing the sensor prototype and extending the detection of the natural gas components to the heavier alkanes, i.e., C4, C5, and C6, as well as to the nonhydrocarbon components, e.g., carbon dioxide and hydrogen sulfide. Indeed, these features would translate into more accurate, continuous, and reliable assessment of the most relevant gases at various stages in oil and gas operations, from exploration, drilling, and production to hydrocarbon processing and distribution.

## AUTHOR INFORMATION

### Corresponding Authors

**Giansergio Menduni** – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy; [orcid.org/0000-0001-7889-6917](https://orcid.org/0000-0001-7889-6917); Email: [giansergio.menduni@poliba.it](mailto:giansergio.menduni@poliba.it)

**Miguel Gonzalez** – Aramco Services Company, Houston, Texas 77084, United States; Email: [miguel.gonzalez@aramcoamericas.com](mailto:miguel.gonzalez@aramcoamericas.com)

### Authors

**Aldo F. P. Cantatore** – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy

**Andrea Zifarelli** – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy

**Pietro Patimisco** – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy; PolySense Innovations srl, Bari 70126, Italy

**Marilena Giglio** – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy

Huseyin R. Seren – Aramco Services Company, Houston, Texas 77084, United States  
Pan Luo – EXPEC Advanced Research Center, Saudi Aramco, Dhahran 31311, Saudi Arabia  
Vincenzo Spagnolo – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy; PolySense Innovations srl, Bari 70126, Italy  
Angelo Sampaolo – PolySense Lab, Dipartimento Interateneo di Fisica, University and Polytechnic of Bari, Bari 70126, Italy; PolySense Innovations srl, Bari 70126, Italy;  
[orcid.org/0000-0003-3790-3767](https://orcid.org/0000-0003-3790-3767)

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.energyfuels.4c03726>

### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Funding

Authors from PolySense Lab acknowledge funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 101135764 EVOQUE, funding from PNRR MUR project PE0000023-NQSTI; MUR—Dipartimenti di Eccellenza 2023–2027—Quantum Sensing and Modeling for One-Health (QuaSiModO), and also financial support from PNRR MUR project PE0000021-NEST.

### Notes

The authors declare no competing financial interest.

### ABBREVIATIONS

ADM acoustic detection module  
AR acoustic resonator  
C1 methane  
C2 ethane  
C3 propane  
C4 butane  
C5 pentane  
FID flame ionization detector  
GC gas chromatograph  
GHG greenhouse gas  
MS mass spectrometry  
NG natural gas  
MVA multivariate analysis  
PLSR partial least-squares regression  
QEPAS quartz-enhanced photoacoustic spectroscopy  
QTF quartz tuning fork  
SMR steam methane reforming

### REFERENCES

- (1) McKinney, D.; Flannery, M.; Elshahawi, H.; Stankiewicz, A.; Clarke, E.; Breviere, J.; Sharma, S. In *Advanced Mud Gas Logging in Combination with Wireline Formation Testing and Geochemical Fingerprinting for an Improved Understanding of Reservoir Architecture*, SPE Annual Technical Conference and Exhibition, SPE-109861; SPE: Anaheim, California, USA, 2007.
- (2) Caroli, E.; Lafaurie, C.; Barraud, B.; Ségolini, G. In *Quantitative Mud Gas Reconciliation with Downhole Fluid Analysis: towards a Quantitative Fluid Log*, SPE Annual Technical Conference and Exhibition, D021S019R003; SPE: New Orleans, Louisiana, USA, 2013.
- (3) Pixler, B. O. Formation evaluation by analysis of hydrocarbon ratios. *J. Pet. Technol.* **1969**, *21* (06), 665–670.
- (4) Haworth, J. H.; Sellens, M.; Whittaker, A. Interpretation of hydrocarbon shows using light (C1–C5) hydrocarbon gases from mud-log data. *AAPG Bull.* **1985**, *69* (8), 1305–1310.
- (5) Devold, H. *Oil and Gas Production Handbook: an Introduction to Oil and Gas Production*; Lulu. Com., 2013.
- (6) Jaeschke, M.; Schley, P.; Janssen–van Rosmalen, R. Thermodynamic research improves energy measurement in natural gas. *Int. J. Thermophys.* **2002**, *23*, 1013–1031.
- (7) Baylis, S. A.; Hall, K.; Jumeau, E. J. The analysis of the C1–C5 components of natural gas samples using gas chromatography-combustion-isotope ratio mass spectrometry. *Org. Geochem.* **1994**, *21* (6–7), 777–785.
- (8) *Chromatographic Analysis of the Environment*, 3rd ed.; Nollet, L. M. L., Ed.; CRC Press, 2006.
- (9) 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.
- (10) Dettmer-Wilde, K.; Engewald, W. *Practical Gas Chromatography*; Springer: Berlin, Heidelberg, 2014.
- (11) Jin, T.; Zhou, J.; Lin, P. T. Real-time and non-destructive hydrocarbon gas sensing using mid-infrared integrated photonic circuits. *RSC Adv.* **2020**, *10* (13), 7452–7459.
- (12) Zifarelli, A.; Sampaolo, A.; Patimisco, P.; Giglio, M.; Gonzalez, M.; Wu, H.; Dong, L.; Spagnolo, V. Methane and ethane detection from natural gas level down to trace concentrations using a compact mid-IR LITES sensor based on univariate calibration. *Photoacoustics* **2023**, *29*, No. 100448.
- (13) Soskind, M. G.; Li, N. P.; Moore, D. P.; Chen, Y.; Wendt, L. P.; McSpirt, J.; Zondlo, M. A.; Wysocki, G. Stationary and drone-assisted methane plume localization with dispersion spectroscopy. *Remote Sens. Environ.* **2023**, *289*, No. 113513.
- (14) Correa Pabón, R. E.; de Souza Filho, C. R.; de Oliveira, W. J. Reflectance and imaging spectroscopy applied to detection of petroleum hydrocarbon pollution in bare soils. *Sci. Total Environ.* **2019**, *649*, 1224–1236.
- (15) Wang, Z.; Wang, Q.; Zhang, H.; Borri, S.; Galli, I.; Sampaolo, A.; Ren, W.; et al. Doubly resonant sub-ppt photoacoustic gas detection with eight decades dynamic range. *Photoacoustics* **2022**, *27*, No. 100387.
- (16) Patimisco, P.; Scamarcio, G.; Tittel, F. K.; Spagnolo, V. Quartz-enhanced photoacoustic spectroscopy: a review. *Sensors* **2014**, *14* (4), 6165–6206.
- (17) Ma, Y. Review of recent advances in QEPAS-based trace gas sensing. *Appl. Sci.* **2018**, *8* (10), 1822.
- (18) Sampaolo, A.; Menduni, G.; Patimisco, P.; Giglio, M.; Passaro, V. M.; Dong, L.; Spagnolo, V.; et al. Quartz-enhanced photoacoustic spectroscopy for hydrocarbon trace gas detection and petroleum exploration. *Fuel* **2020**, *277*, No. 118118.
- (19) Menduni, G.; Zifarelli, A.; Sampaolo, A.; Patimisco, P.; Giglio, M.; Amoroso, N.; Spagnolo, V.; et al. High-concentration methane and ethane QEPAS detection employing partial least squares regression to filter out energy relaxation dependence on gas matrix composition. *Photoacoustics* **2022**, *26*, No. 100349.
- (20) Li, Y.; Wang, R.; Tittel, F. K.; Ma, Y. Sensitive methane detection based on quartz-enhanced photoacoustic spectroscopy with a high-power diode laser and wavelet filtering. *Opt. Lasers Eng.* **2020**, *132*, No. 106155.
- (21) Li, C.; Dong, L.; Zheng, C.; Lin, J.; Wang, Y.; Tittel, F. K. Ppbv-level ethane detection using quartz-enhanced photoacoustic spectroscopy with a continuous-wave, room temperature interband cascade laser. *Sensors* **2018**, *18* (3), 723.
- (22) THORLABS. QEPAS-CH4 - Quartz-Enhanced Photoacoustic Sensor for Methane. <https://www.thorlabs.com/thorproduct.cfm?partnumber=QEPAS-CH4>.
- (23) Luo, P.; Harrist, J.; Menduni, G.; Mesdour, R.; StMichel, N.; Sampaolo, A. Simultaneous detection of methane, ethane, and



propane by QEPAS sensors for on-site hydrocarbon characterization and production monitoring. *ACS Omega* **2022**, *7* (4), 3395–3406.

(24) Kosterev, A. A.; Bakhrkin, Y. A.; Curl, R. F.; Tittel, F. K. Quartz-enhanced photoacoustic spectroscopy. *Opt. Lett.* **2002**, *27* (21), 1902–1904.

(25) Barreiro, N.; Vallespi, A.; Santiago, G.; Slezak, V.; Peuriot, A. Influence of oxygen on the resonant photoacoustic signal from methane excited at the  $\nu_3$  mode. *Appl. Phys. B* **2011**, *104*, 983–987.

(26) Barreiro, N.; Peuriot, A.; Santiago, G.; Slezak, V. Water-based enhancement of the resonant photoacoustic signal from methane–air samples excited at 3.3  $\mu\text{m}$ . *Appl. Phys. B* **2012**, *108*, 369–375.

(27) Barreiro, N.; Peuriot, A.; Slezak, V.; Santiago, G. Study of the dependence of the photoacoustic signal amplitude from methane on different collisional partners. *Vib. Spectrosc.* **2013**, *68*, 158–161.

(28) Rousseau, R.; Maurin, N.; Trzpił, W.; Bahriz, M.; Vicet, A. Quartz tuning fork resonance tracking and application in quartz enhanced photoacoustics spectroscopy. *Sensors* **2019**, *19* (24), 5565.

(29) Sampaolo, A.; Patimisco, P.; Giglio, M.; Zifarelli, A.; Wu, H.; Dong, L.; Spagnolo, V. Quartz-enhanced photoacoustic spectroscopy for multi-gas detection: A review. *Anal. Chim. Acta* **2022**, *1202*, No. 338894.

(30) Zifarelli, A.; Cantatore, A. F. P.; Sampaolo, A.; Mueller, M.; Rueck, T.; Hoelzl, C.; Spagnolo, V.; et al. Multivariate analysis and digital twin modelling: Alternative approaches to evaluate molecular relaxation in photoacoustic spectroscopy. *Photoacoustics* **2023**, *33*, No. 100564.

(31) Müller, M.; Rück, T.; Jobst, S.; Pangerl, J.; Weigl, S.; Bierl, R.; Matysik, F. M. An algorithmic approach to compute the effect of non-radiative relaxation processes in photoacoustic spectroscopy. *Photoacoustics* **2022**, *26*, No. 100371.

(32) Pangerl, J.; Müller, M.; Rück, T.; Weigl, S.; Bierl, R. Characterizing a sensitive compact mid-infrared photoacoustic sensor for methane, ethane and acetylene detection considering changing ambient parameters and bulk composition ( $\text{N}_2$ ,  $\text{O}_2$  and  $\text{H}_2\text{O}$ ). *Sens. Actuators, B* **2022**, *352*, No. 130962.

(33) Boudon, V.; Rey, M.; Loëte, M. The vibrational levels of methane obtained from analyses of high-resolution spectra. *J. Quant. Spectrosc. Radiat. Transfer* **2006**, *98* (3), 394–404.

(34) Zifarelli, A.; Giglio, M.; Menduni, G.; Sampaolo, A.; Patimisco, P.; Passaro, V. M.; Spagnolo, V.; et al. Partial least-squares regression as a tool to retrieve gas concentrations in mixtures detected using quartz-enhanced photoacoustic spectroscopy. *Anal. Chem.* **2020**, *92* (16), 11035–11043.

(35) THORLABS. ADM01 - Acoustic Detection Module for QEPAS. <https://www.thorlabs.com/Thorproduct.Cfm?Partnumber=ADM01>.

(36) Patimisco, P.; Sampaolo, A.; Bidaux, Y.; Bismuto, A.; Scott, M.; Jiang, J.; Spagnolo, V.; et al. Purely wavelength-and amplitude-modulated quartz-enhanced photoacoustic spectroscopy. *Opt. Express* **2016**, *24* (23), 25943.

(37) Gordon, I. E.; Rothman, L. S.; Hargreaves, E. R.; Hashemi, R.; Karlovets, E. V.; Skinner, F. M.; Yurchenko, S. N.; et al. The HITRAN2020 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer* **2022**, *277*, No. 107949.

(38) VPL Molecular Spectroscopic Database. <https://vpl.astro.washington.edu/spectra/>.

(39) Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58* (2), 109–130.

(40) Loh, A.; Wolff, M. Multivariate analysis of photoacoustic spectra for the detection of short-chained hydrocarbon isotopologues. *Molecules* **2020**, *25* (9), 2266.

(41) Kvalheim, O. M. Interpretation of direct latent-variable projection methods and their aims and use in the analysis of multicomponent spectroscopic and chromatographic data. *Chemom. Intell. Lab. Syst.* **1988**, *4* (1), 11–25.

(42) Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdiscip. Rev.: Comput. Stat.* **2010**, *2* (1), 97–106.

(43) Olivieri, A. C. *Introduction to Multivariate Calibration*; Springer International Publishing: Cham, 2018.

(44) Barra, I.; Kharbach, M.; Qannari, E. M.; Hanafi, M.; Cherrah, Y.; Bouklouze, A. Predicting cetane number in diesel fuels using FTIR spectroscopy and PLS regression. *Vib. Spectrosc.* **2020**, *111*, No. 103157.