Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Spectral fits with full broadening for kHz-iteration high-accuracy gas-phase spectroscopy

**Stephen G. So**

sso@sophotonics.com

SoPhotonics Inc.

**Research Article**

**Keywords:** Tunable diode laser absorption spectroscopy, spectral line shapes, Taguchi orthogonal array, HITRAN approximation, Multi-species trace gas sensing

**Posted Date:** November 6th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11310219/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** Competing interest reported. The author's company (Sophotonics Inc.) has filed provisional patent applications on the technologies described here. The author holds shares in a previous company (Aeris Technologies Inc. / International Environmental Standards PBC) dealing with laser spectrometers, but solely interacts via a third party shareholder representative.

# Spectral fits with full broadening for kHz-iteration high-accuracy gas-phase spectroscopy

**Stephen So**

Received: date / Accepted: date

**Abstract** This work presents methods enabling real-time multi-species direct absorption spectroscopic fitting at kilohertz iteration rates with high accuracy of spectral line shapes on embedded systems. Precomputation of multi-dimensional polynomials and effective pressures allows high accuracy approximation of HITRAN Hartmann-Tran absorption coefficients with full matrix broadening and low residual error, replacing computationally slow line-by-line simulations and inaccurate fixed-basis functions. A novel Levenberg–Marquardt solver augmented with Taguchi orthogonal array exploration accelerates fitting convergence and escapes local minima. Together, these methods enable simulated implementations that fit 1000-point spectra with simultaneous etalon fringe, temperature, and pressure estimation at 1 kHz using 0.8 W simulated computational power for a single laser system. This work further simulates a simultaneous dual-laser single-detector architecture exploiting co-prime spectral lengths, fitting six concentration parameters at 0.82 kHz using 1.6 W simulated computational power. Mid-infrared TDLAS simulations targeting $CH_4$, $H_2O$, $C_2H_6$, $CO_2$, and $N_2O$ provide an approach for possible atmospheric sensing deployments and chemical kinetics studies.

**Keywords** Tunable diode laser absorption spectroscopy · spectral line shapes · Taguchi orthogonal array · HITRAN approximation · Multi-species trace gas sensing

S. So
Sophotonics Inc., Princeton, New Jersey, USA
E-mail: sso@sophotonics.com

## 1 Introduction

High-resolution spectroscopy often relies on physics-informed spectral modeling and curve fitting to extract physical parameters from weak, information-rich signals in the presence of large optical backgrounds. By greatly improving the efficient real-time extraction of physical parameters such as sample concentration, temperature, and pressure from spectra, this work enables a variety of novel capabilities for spectroscopic sensing in field deployments.

In precision narrowband light source applications, spectroscopists commonly model transmission or absorption spectra by simulation of HITRAN data and models [1] (or similar spectral databases) and combine them with a system model that maps these parameters to an expected detector signal.

These models match the experimental behavior of a gas sample to a high degree of accuracy when models include all relevant physical parameters. Laboratory grade instruments that need to prove traceable accuracy continuously store how closely the signals correspond to the underlying physics model.

Such HITRAN-driven spectral modeling coupled with least-squares curve fitting algorithms frequently finds applications aiming to measure high-accuracy chemical concentrations and/or physical parameters such as pressure and temperature in air samples.

For portable high-accuracy spectrometers, post-processing of raw acquired spectra frequently replaces real-time fitting, avoiding embedded computing requirements. However, storage and transmission of the required raw data rates without losing spectral information from a remote location is costly, especially in long-term battery/solar powered wireless sensor network deployments. Additionally, some sensor control

systems require very low latency fit results to use in local feedback systems for higher accuracy (e.g. laser wavelength stabilization).

In typical real-time sensor deployments, embedded computational units continuously compare measured spectra against simulated or recorded spectra. This applies across a wide range of techniques, including RF frequency modulation spectroscopy [2], cavity-enhanced methods (e.g., cavity ring-down and integrated cavity output spectroscopy) [3,4], and optical frequency comb spectroscopy [5]. In general, these applications typically rely on HITRAN-based line-by-line simulations [1], fit to the measured spectra to retrieve trace-gas concentrations and gas properties [6,7].

This work implements simulations of molecular trace-gas TDLAS (Tunable Diode Laser Absorption Spectroscopy) in order to demonstrate the new methods developed here — though the improvements apply to the myriad other techniques for spectroscopy which integrate least squares fits of multi-parameter models with strong interparameter interaction. Beyond atmospheric monitoring, laser absorption spectroscopy sees significant use in medical applications [8,9], planetary exploration [10], and industrial process and safety monitoring of hazardous species [11].

The main difficulty in these measurements is that the spectroscopic signature is a small fractional change (frequently $10^{-3}$ to $10^{-9}$ of the optical intensity) of a larger optical power baseline coupled with other systematic effects (e.g., etalon fringes, source intensity drift, detector responsivity variation, or optical-coupling fluctuations). Proper analysis therefore requires estimation of both the slowly varying baseline and the narrowband absorption feature(s). In practice, this is often done in two steps: first, fitting a model for the baseline; then second, fitting the absorption model [12]. The sections of the spectra which have closer to zero absorption provide points for the initial baseline fit, attempting to make the baseline independent of the absorption signals. Accordingly, these sensors usually operate at lower pressures (typically $< 0.5\,\mathrm{atm}$) to reduce complicated pressure broadened multi-line spectral wing contributions in the baseline intensity.

The absorption strength and line center of each individual transition directly depend on the sample thermodynamic conditions, especially pressure. Here $T$ and $p$ denote the sample temperature and sample pressure, respectively. For high-precision-and-accuracy applications, even modest variations in $T$ and $p$ can produce substantial changes in the effective line shape, directly biasing spectroscopic parameter retrieval if not controlled or modeled in the fit. As a result, retrieval of concentration can require explicit precision measurement (and often extreme precision control) of $T$ (to $< 5\,\mathrm{mK}$) and $p$ (to $< 0.0005\,\mathrm{atm}$) to reduce the effect of these critical parameters [13].

When scaling high-accuracy point sensor networks such as the Integrated Carbon Observation System (ICOS) [14], cost and complexity of the overall sensor architecture become significant factors versus accuracy; however, the work here significantly reduces that trade-off.

## 2 Background

### 2.1 Forward model for a single isolated line

For a gas phase molecular transition, transmission models follow the Beer–Lambert law

$$I(\nu) = I_0(\nu)\, e^{-\alpha(\nu)L}, \tag{1}$$

where $I_0(\nu)$ is the incident intensity, $L$ is the path length, and $\alpha(\nu)$ is the absorption coefficient. For a single molecular species at number density $N$,

$$\alpha(\nu) = N \cdot S(T) \cdot \varphi(\nu; T, p, \mathbf{x}), \tag{2}$$

where $S(T)$ is the line intensity, $\mathbf{x}$ is a vector of species concentrations, and $\varphi$ is the normalized line-shape function.

In most field deployed systems which model the spectrum directly in real-time, $\varphi$ is a Voigt profile, i.e., a convolution of a Gaussian Doppler component (with width set by $T$ and molecular mass) and a Lorentzian collisional component (with width set by $p$ and broadening coefficients).

More advanced spectral models (e.g. Speed Dependent Voigt (SDV), Hartmann-Tran (HT)) add speed dependence and collisional narrowing, requiring augmented computational resources. Such improved accuracy is necessary in many applications to avoid additional calibration requirements when implementing higher optical-frequency-resolution spectra of shorter wavelength scan ranges. Field instruments commonly combine a Voigt line shape with air broadening only, omitting both these line shape effects and self- and cross-broadening; Figure 1 shows the combined error from these two common omissions.

Spectral fitting typically proceeds by defining a parameter vector $\boldsymbol{\beta}$ (e.g., baseline coefficients, concentration, line center, width parameters, and sometimes $T$ and/or $p$) and minimizing a least-squares objective, typically using either direct solvers or iterative non-linear least squares fitting such as Levenberg–Marquardt [15].

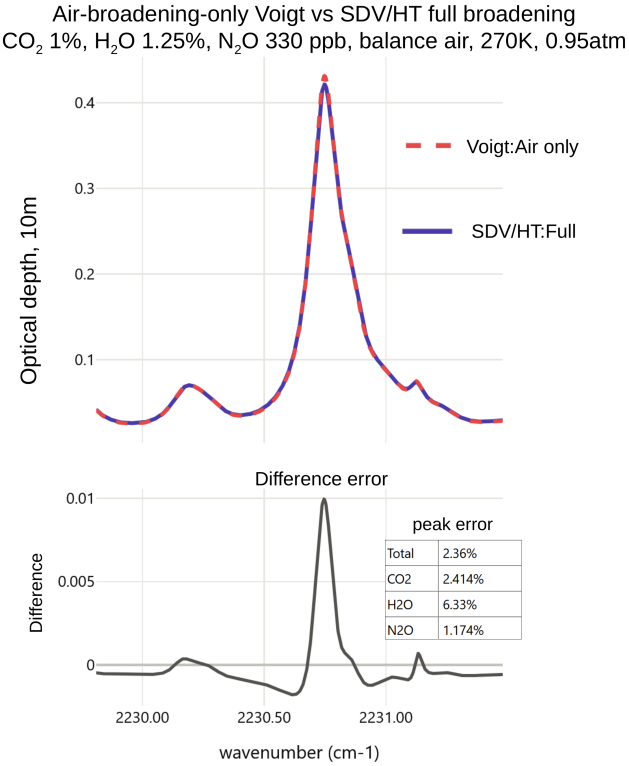



**Fig. 1 Example simulated error from omitting full-matrix broadening and SDV/HT line shapes.** Top plot: The dashed line is a Voigt profile with air broadening only; the solid line is the speed-dependent Voigt / Hartmann-Tran (SDV/HT) model with full self- and foreign gas cross-broadening. Bottom plot: Combined peak error from both omissions (Voigt line shape and air broadening only), relevant to field instruments that use air-broadening-only Voigt models.

### 2.2 Direct absorption vs. wavelength modulation spectroscopy methods

In wavelength-scanned direct absorption spectroscopy (simply "direct absorption"), a narrowband laser's wavelength tunes across the absorption lines over time for the analysis of transmitted intensity (without intentional sinusoidal wavelength modulation).

Direct absorption measurements are conceptually simple and retain a straightforward connection between the measured spectrum and the underlying absorption coefficient. However, because most of the signal content remains closer to DC when the spectral iteration rate is limited, low-frequency technical noise and baseline fluctuations often limit sensitivity, motivating migration to lock-in detection.

Wavelength modulation spectroscopy (WMS) improves noise rejection by imposing a high-frequency modulation (typ. 10-100's kHz) on the laser frequency while scanning across an absorption line, then demodulating the detector signal at the modulation frequency ($1f$) and/or its harmonics (typically $2f$) using phase-sensitive lock-in amplifier detection [16,17]. The key practical advantage is that the signal shifts away from low-frequency technical noise (e.g., $1/f$ intensity noise, baseline drifts, and slow etalon variations), allowing detection of very small absorbance changes on top of a large optical power background.

WMS/lock-in benefits come with trade-offs: the demodulated signal depends on exact modulation depth, residual amplitude modulation, higher harmonic signal power leakage, and the phase and frequency response of the electronics signal chain. Exact retrieval typically requires a full model of the modulation/demodulation process or continuous calibration rather than a direct absorbance fit [6]. Additionally, the signal may show more sensitivity to absorption line shape changes from pressure and temperature due to the mathematical derivative nature of the measurement.

However, the WMS method allows for (i) an extremely simple lock-in low-pass filter to exclude all noise outside the detection bandwidth, and (ii) multiple lasers with single detectors, by using different modulation frequencies on each laser [18], avoiding time-division multiplexing (and the associated dead-time) in single detector, multi-laser systems.

### 2.3 Matched filtering and optimum noise rejection

From a signal processing standpoint, detecting a specific spectral signature in additive noise corresponds to a matched-filter problem: the optimal signal detector directly correlates the measurement with a copy of the expected signal [19]. In practice, for spectroscopy this motivates (i) fitting with a physics-informed basis, (ii) shaping the measurement and filtering so that the signal occupies frequency bands where the noise power spectral density is low, and (iii) transformation and filtering of the basis fit result to desired outputs.

WMS and lock-in detection correspond to this principle: the spectral information moves from near-DC (where technical noise and baseline drift are strongest) into higher frequencies via sinusoidal wavelength modulation. The lock-in detection then implements a narrowband, "matched filter" local oscillator detector with low pass output filter, trading increased model complexity for improved rejection of low-frequency noise.

Some efforts aim to make WMS "calibration-free," for concentration inference without separate reference gas cells for signal normalization or empirical calibration factors [6,20]. These particular methods targeted a dominant isolated transition and directly fit a single collisional width to account for broadening, and avoided multiline overlapping spectra.

If spectrometers can fit quickly and accurately enough in direct absorption to move the signal outside the noisy sections of lower frequencies of the electronics chain, WMS offers less noise reduction benefit while the disadvantages of a complicated WMS model analysis remain.

In direct absorption, increasing the spectral iteration rate is the main method for moving the signal from near DC into a frequency with a lower noise floor. However, fitting quickly enough for the higher iteration rates directly with accuracy is not a simple task. Co-adding/averaging of spectra was therefore a necessary concession to not overtax the computational capability of portable laser spectroscopy embedded systems. However, noise in these systems from pressure control, temperature stability, and *fast broadener concentration variation* causes the spectrum to lose contrast and therefore lose significant spectral data in the averaging process.

## 3 Challenges and Results

The central computational challenge in real-time spectroscopic fitting of a useful number of spectral points at kHz rates is the evaluation of the absorption coefficient $\alpha(\nu; T, p, \mathbf{x})$ at each spectral point, where $\mathbf{x} = \{x_1, x_2, \ldots, x_M\}$ denotes the mole fractions of $M$ molecular species. A full LBL HITRAN simulation — incorporating all transitions within the spectral range, line profiles, self- and cross-broadening (termed "foreign-gas broadening" in HITRAN - this work uses "cross-broadening" to indicate the full matrix of species-to-species interactions) — is far too expensive for kHz-rate embedded system hardware fitting. A discrete GPU quickly calculated spectra in [21], but a GPU is costly in resources when used for point sensors. This work replaces exact forward models and hard-coded spectral basis functions limited to a specific $(T, p)$ operational point with a set of precomputed multi-dimensional polynomials that accurately approximate the HITRAN-simulated absorption coefficient over a wide range of the state variables $(T, p, \mathbf{x})$ to allow fast direct fitting.

This finds critical, immediate applications in high resolution atmospheric spectrometers that can now provide suitable accuracy in a wide range of sample conditions, compared to a more typical single target temperature and single pressure operation. Furthermore, this is possible in a single moderately sized FPGA, allowing for the possibility of spectrometers integrating these methods into a single application-specific integrated circuit (ASIC) combined with chip-scale photonics [22].

These approximations of HITRAN Hartmann-Tran $\alpha(\nu)$ spectra using a low order multi-dimensional polynomial provide low residual error with very fast iteration. Over large simulated operational sample conditions spanning 60 K (260–320 K) and 1.1 atm (0.2–1.3 atm), error ranges from less than 2.5% $\alpha(\nu)$–error (rms over a 3 wavenumber spectral range) using 8 polynomial coefficients — down to less than 5 ppm peak relative $\alpha(\nu)$–error using 15 polynomial coefficients. Over a narrower 43 K (270–313 K) and 0.3 atm (0.9–1.2 atm) range, rms $\alpha(\nu)$–error is less than 0.5% using 8 polynomial coefficients, suitable for high resolution and high accuracy atmospheric study in the field.

This work further integrates optical fringe, temperature, pressure, and concentration broadening into a kilohertz rate spectral fit, and such capabilities can directly improve a wide variety of high precision field applications. Because the system performs spectrum-by-spectrum fits at high speed with all lines accounted for in every iteration, the need for operating at reduced sample pressures and maintaining precision temperature control substantially relaxes; the fitter tracks $T$ and $p$ as free parameters on every iteration.

This research also accelerates general least-squares fitting and convergence of high-resolution curve fits. By implementing a novel Levenberg–Marquardt fitter coupled with Taguchi orthogonal array exploration, the fit iterations can "jump" to new locations in the parameter space without relying on the local Jacobian shape. Furthermore, the fitting system does this without dramatically increasing latency or computational resource requirements. The simulation results in Figure 3 show accurate fitting of multiple moving model parameter targets modulating at relatively high speed.

Additionally, this study demonstrates a method for a dual-laser system that shows promise for dramatically improving sensor network field deployed spectrometers and fast chemical kinetics measurements that can measure $CH_4$, $H_2O$, $CO_2$, and $N_2O$ simultaneously at scientifically relevant concentrations.

The simulations here model signals typical of a single mode diode, quantum cascade, or interband cascade laser [23], 10 meters of optical path, and noise levels which roughly correspond to those seen in Peltier-cooled-detector mid-infrared TDLAS of $10^{-4}$ to $10^{-6}/\sqrt{Hz}$ per spectrum absorption variation. The noise added to the signals is uniformly distributed and does not currently model shot noise or $1/f$ noise. Figure 2 shows a simplified block diagram of the overall spectral fitting system architecture.

The results here describe a single laser system with 16 fit parameters including: detector offset; species concentrations (3 parameters); etalon fringe parameters (3 par.); 5th order baseline polynomial (6 par.); temperature and pressure (2 par.); and interpolation shift of the entire spectrum (Figure 3).

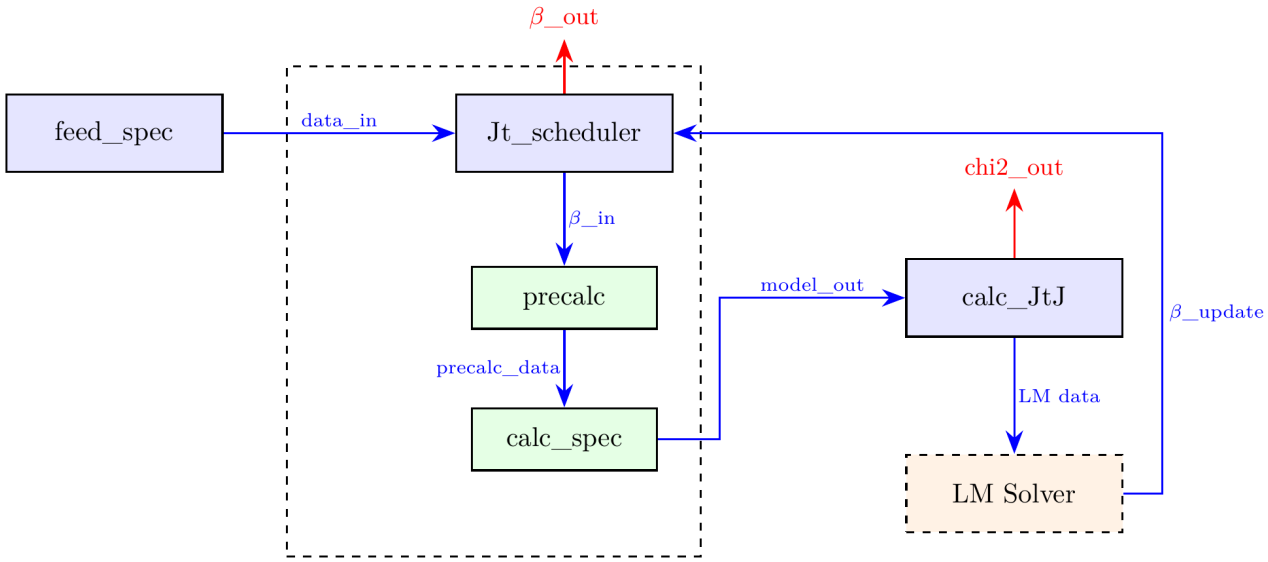



**Fig. 2 Simplified block diagram of the overall spectral fitting system architecture.** feed_spec streams the spectral data to Jt_scheduler, which sends the parameters at the correct times to precalc. Precalc performs calculations reused for each spectral point. calc_spec calculates a full spectrum for calc_JtJ to assemble data for the LM solve step and compare Taguchi OA runs. Solver then computes the LM iteration step.

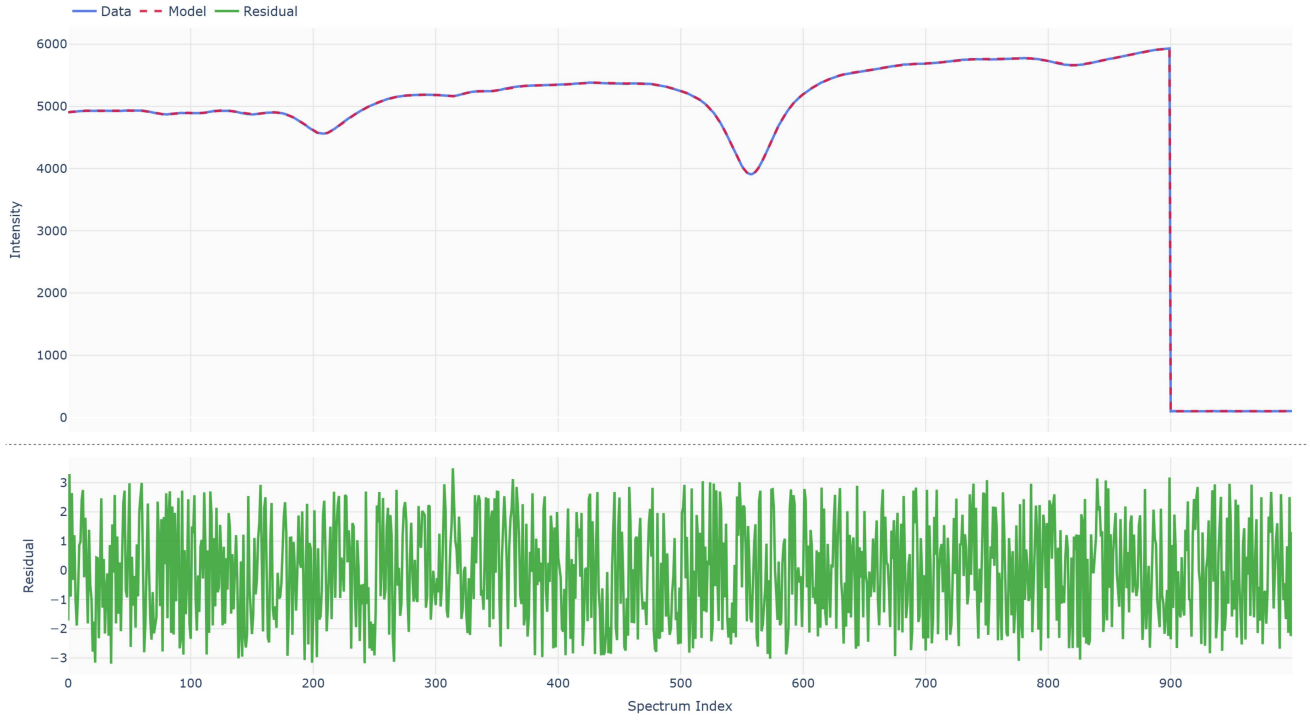
A dual system (named SYS0 and SYS1) of 29 parameters adds an additional set of concentrations, fringe, polynomial, and interpolation shift. The temperature and pressure are shared between the systems, assuming collinear beam paths. Figure 4 shows the dual-system spectrum and tracking. Figure 5 shows the dual-system polynomial approximation fit accuracy against Hartmann-Tran HITRAN data coupled with the polynomial baselines for 6 concentration parameters. All shown parameter traces over time (including $T$, $p$, concentrations, and fringe amplitudes) have some sinusoidal modulation amplitude applied, each at a differing period and phase.

### 3.1 Comparison with prior implementations

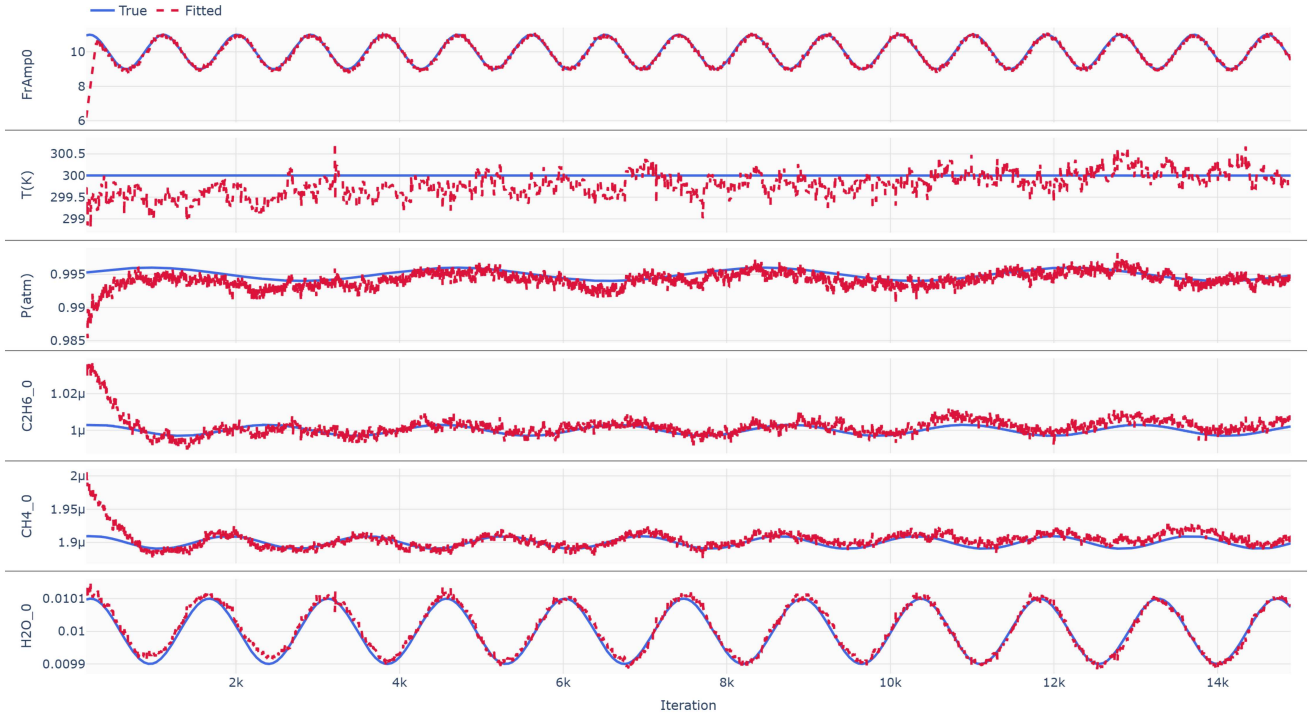
Implementations for efficient real-time spectral modeling fits reported in spectroscopic literature reduce the cost of evaluating lineshapes while maintaining physical fidelity. Common approaches include (i) precomputing Voigt $(x, y)$ arguments on a grid and interpolating via third-order Taylor expansion [24], (ii) using fast rational approximations to the complex error function (Faddeeva function) for Voigt and speed-dependent Voigt evaluation [25,26,27], and (iii) shifting portions of the spectral processing pipeline onto dedicated real-time hardware [28,16]. These approaches generally accelerate the evaluation of individual lineshapes within a line-by-line (LBL) framework; each spectral point still requires summing contributions from every relevant transition. In contrast, the all-lines polynomial approximation presented here replaces the LBL simulation and subsequent summation at each spectral point with low-order polynomial evaluation for situations where the evaluation of a full spectral model with full broadening should occur thousands of times a second.

Strategies for accurate absorption coefficient estimation from precomputed spectra exist for satellite based measurements: absorption coefficient lookup tables indexed by temperature, pressure, and water concentration are used in radiative transfer models such as ARTS [29] and in OCO-2 (Orbiting Carbon Observatory) satellite retrievals [30], and spectrally reduced integration methods have been applied to further reduce the size of data sets [31]. These approaches interpolate between the precomputed spectra from a direct grid of $(T, p, x_{H_2O})$ points. This work instead describes *effective pressure and broadening* multi-line methods to account for self- and cross-broadening developed to avoid full grid dependence on $x_{H_2O}$, and generally $\mathbf{x}$.

In [32], neural networks inferred absorption coefficients from $(T, p)$ inputs for these satellite retrievals. In this case, 3 hidden layers in the neural network perform a similar function to the multi-dimensional polynomial in this work by providing non-linearity to the model, but do not describe taking into account cross-broadening specifically. In general, there is latency from a hidden neural layer which depends on results from a previous

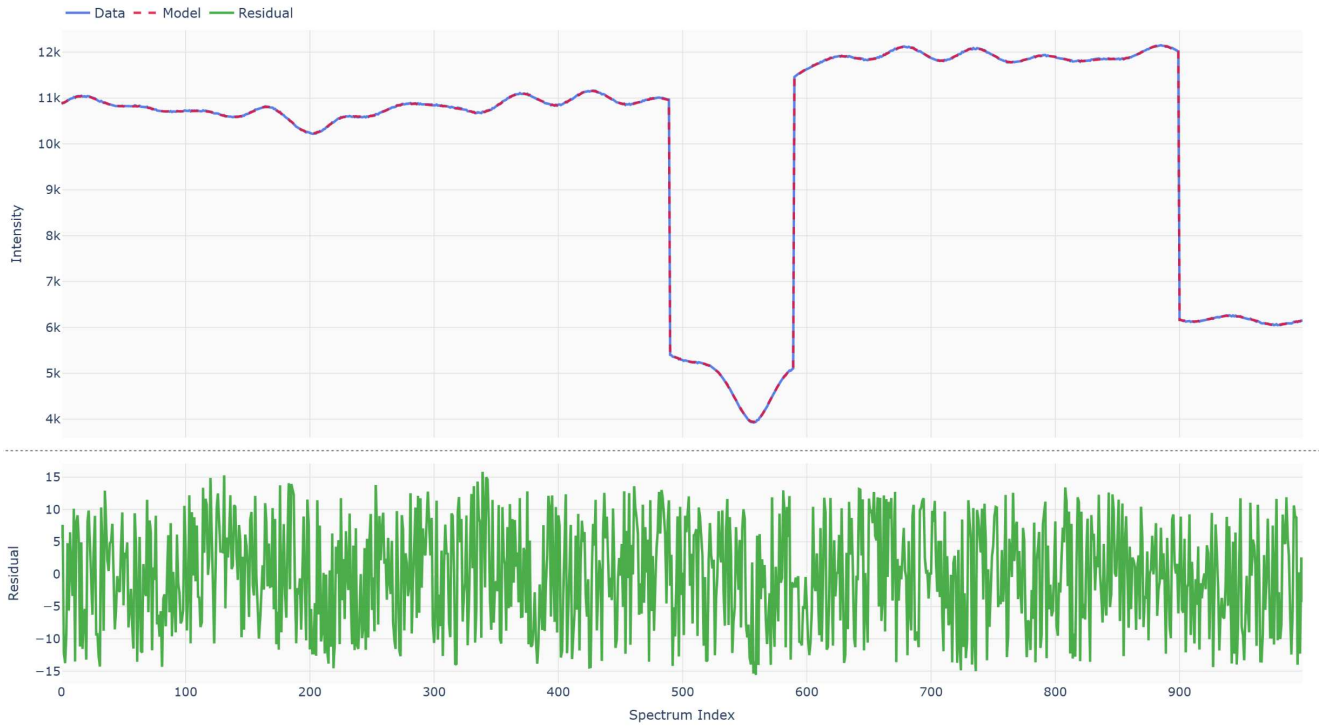



(a)

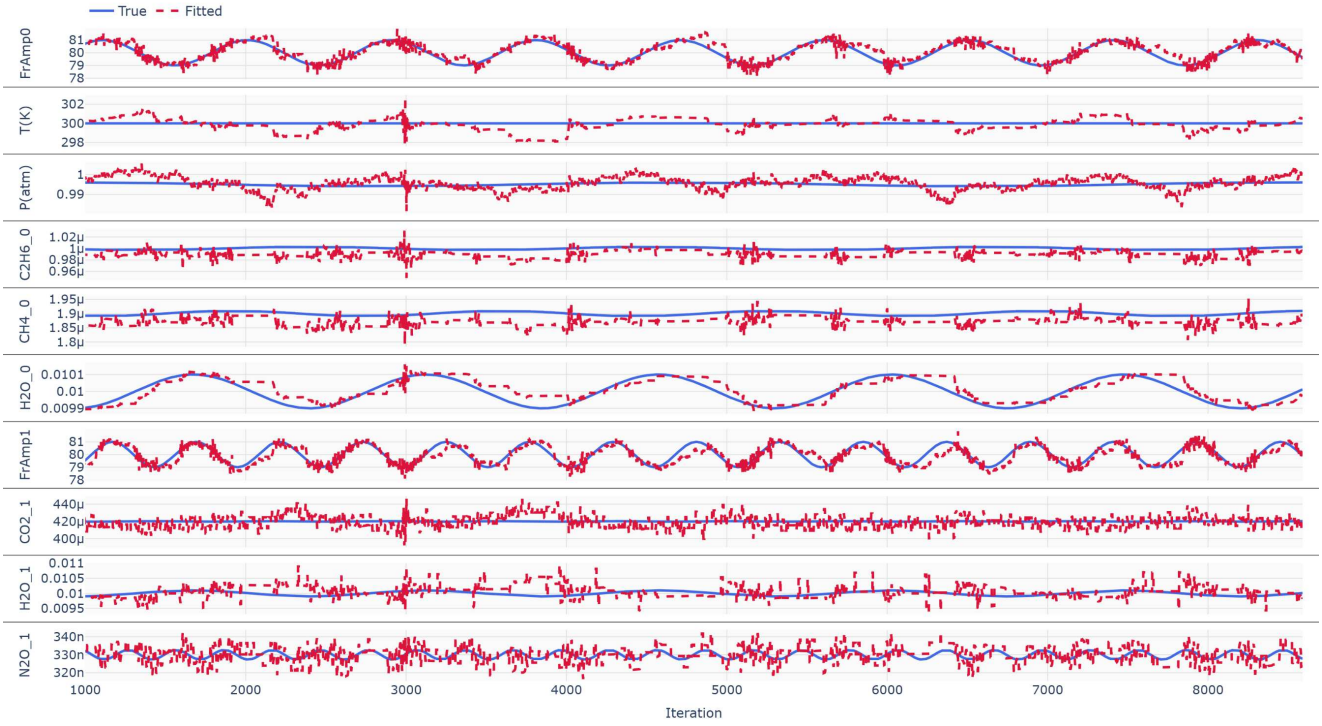



(b)

**Fig. 3 Single system spectral fitting** a) spectrum and residual (uniformly distributed intensity noise of 6 units peak-to-peak added to the detector signal model, on a baseline of ∼5000 units, at iteration 14890), b) data tracking of simulated data and fitted parameters, with sinusoidally varying etalon fringe amplitude (FrAmp0), pressure, temperature, and gas concentrations at differing periods. The signals show model-truth values (blue) and fitted values (red). One iteration point is ˜1.0 millisecond.




(a)




(b)

**Fig. 4 Dual-system simulation with noise and high fringe amplitude.** a) Combined dual-system spectrum and residual at the last iteration point. Uniform intensity noise of 27 units peak-to-peak is added to the detector signal model corresponding to a per-spectrum noise level of $\sim 5 \times 10^{-5}/\sqrt{\text{Hz}}$. b) data tracking of simulated data and fitted parameters over iterations 1000–8600. Parameters include sinusoidally varying etalon fringe amplitudes (FrAmp0 and FrAmp1), pressure, temperature, and gas concentrations at differing periods. The fringe amplitude level is set very high at ∼1.5% of the per-laser baseline. The signals show model-truth values (blue) and fitted values (red). One iteration point is ˜1.22 milliseconds.

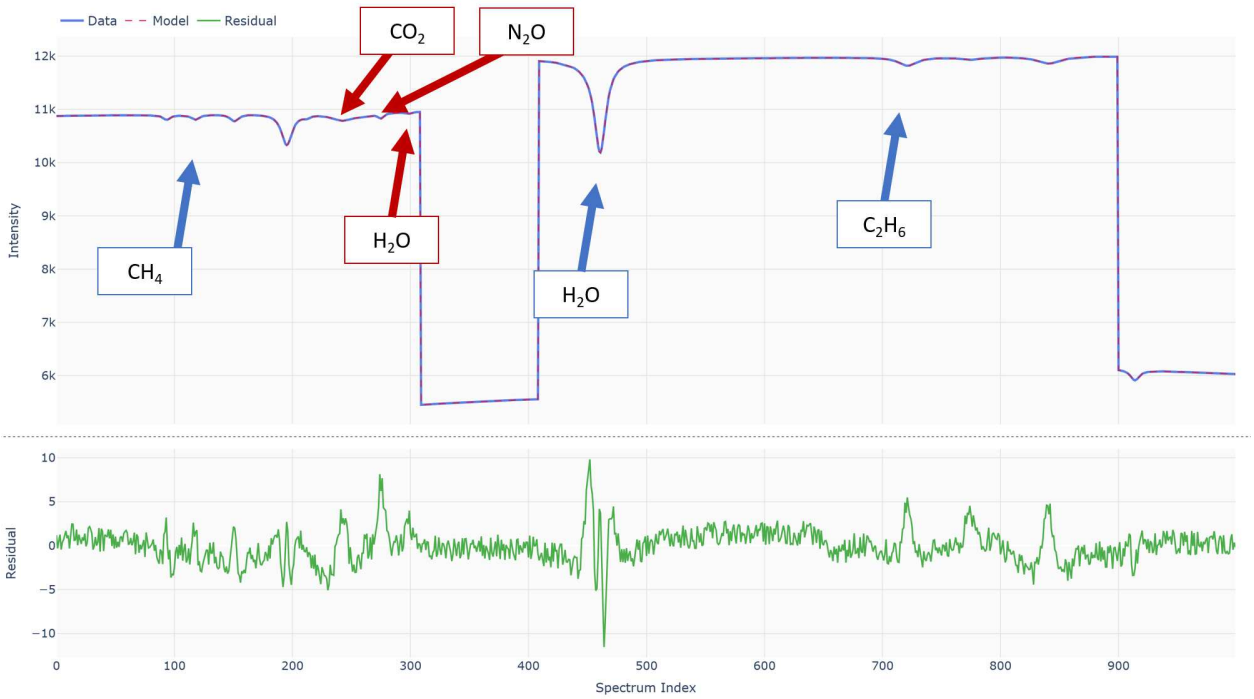



**Fig. 5 Dual system fit of HITRAN data with 6 concentration parameters** using dual 8 coefficient polynomial with $p_{\text{eff}}$ approximation at 0.3 atm and 300 K showing fit accuracy to expected sample data. Uniformly distributed intensity noise of 3 units peak-to-peak is added to the detector signal model, with zero fringe amplitude. Blue labels are dominant line positions of SYS0 and red labels are SYS1.

neural layer. The polynomial approach instead only requires simple vector multiply-accumulate with no such dependence. The latency issues compound with the more sophisticated neural network inference (not simple hidden layers) reported in [33].

On FPGA targets for laser spectroscopy, the literature focuses on real-time lock-in style demodulation and filtering [34,35,36,37], with fitting or inference offloaded to soft/hard processors or replaced by machine-learning models [33]. Table 1 compares FPGA-based spectral processing implementations against this work, including lock-in–only systems and those that perform on-FPGA spectral fitting.

All FPGA designs in Table 1 completed synthesis, implementation, and timing closure at the stated iteration rates in AMD Vivado 2025.1, followed by post-implementation behavioral simulation of the fit and post-implementation power estimation (0.8 W single, 1.6 W dual).

Fit iteration results here have 1.0 ms / 1.22 ms latency for the single and dual laser version respectively, which enables new feedback and control capabilities such as fast pressure modulation, $T/p$ sensor-free systems, or fast wide-tuning laser wavelength stabilization.

Naturally, if the application tolerates lower iteration rates, lowering the processing clock rate reduces power consumption proportional to frequency, with limitation of static power consumption. An ASIC (application specific integrated circuit) implementation generally improves both power consumption and max frequency, likely by $\sim 10\times$ in each.

The computational complexity of one LM iteration is $\mathcal{O}(n^2\, L_{\text{spec}} + n^3)$, dominated by forming $\boldsymbol{J}^T\boldsymbol{J}$ at $\mathcal{O}(n^2\, L_{\text{spec}})$ since $L_{\text{spec}} \gg n$. Doubling $L_{\text{spec}}$ or $n$ roughly doubles or quadruples the per-iteration cost, respectively. As Table 1 shows, this work simultaneously significantly increases both spectral length and parameter count, yet maintains kHz iteration rates.

## 4 Methods for fast high accuracy fitting

### 4.1 Multi-dimensional polynomial absorption coefficient fit

An approximating polynomial fit of the form

$$\hat{\alpha}_m(\nu_k; T, p) \;=\; \sum_{j=1}^{J_m} a_{j,k,m} \cdot \phi_j(T, p), \tag{3}$$

at each discrete spectral point $\nu_k$ ($k = 1 \ldots L_{\text{spec}}$) approximates the absorption coefficient for each molecule $m$. HITRAN simulations on a $(T, p)$ grid provide data for the 2 dimensional polynomial fit. Here $j$ indexes the $J_m$ monomial terms $\phi_j$ (e.g., 1, $T$, $p$, $T^2$, $T \cdot p$, $T^2 \cdot p$,

**Table 1 Comparison of FPGA real-time spectrometer processing implementations.** Logic elements (LE) and digital signal processing (DSP) slices reflect device compute capacity. $R$ is iteration/spectral rate. "Params" counts simultaneously determined parameters $n$. The $n^2 L_{\text{spec}} \cdot R$ column gives the dominant term of the per-iteration computational complexity $\mathcal{O}(n^2 L_{\text{spec}})$ multiplied by $R$; dashes indicate the metric is not applicable or not reported. *Lock-in systems with no on-FPGA HITRAN spectral fitting. †Lock-in with neural network inference on external processor. ‡Neural network based spectral fitting of 10 input points.

| Implementation | FPGA | LE / DSP | $L_{\text{spec}}$ | $R$ | Params $n$ | $n^2 L_{\text{spec}} \cdot R$ |
|---|---|---|---|---|---|---|
| Enemali & Gibson* [34] | XC7Z010-1 | 28.0k / 80 | 500 | 125 Hz | 1 ($x_{\text{H}_2\text{O}}$) | —* |
| Wang et al.* [35] (Photoacoustic) | ProASIC3EL | 35.0k max / — | — | 1 Hz | 1 ($x_{\text{CH}_4}$) | —* |
| Yang et al.* [36] | XC7S50 | 52.2k / 120 | — | 20 Hz | 1 ($x_{\text{CH}_4}$) | —* |
| Choi et al.* [37] | XC7Z020 | 53.2k / 220 | — | 10 Hz | 1 ($x_{\text{H}_2\text{O}}$) | —* |
| Xia et al.† [33] | XC7Z010 | 28.0k / 80 | 120 | 62.5 Hz | 2 ($T$, $x_{\text{H}_2\text{O}}$) | $3 \times 10^4$† |
| Xu et al. [28] | XC5VSX95T (<15% used) | 14.7k / 640 | 150 | 4 kHz | 1 ($T$) | $6 \times 10^5$ |
| Schwarm et al.‡ [16] | XC7K325T (<50% used) | 326.1k / 840 | 10 | 10 kHz | 2 ($T$, $x_{CO}$) | $4 \times 10^5$ ‡ |
| This work (single) | XCZU1CG-1 | 81.9k / 216 | 1000 | 1 kHz | 16 | $2.56 \times 10^8$ |
| This work (dual) | XCZU3CG-1 | 154.3k / 360 | 1000 / (983) | 821 Hz | 29 | $6.9 \times 10^8$ |

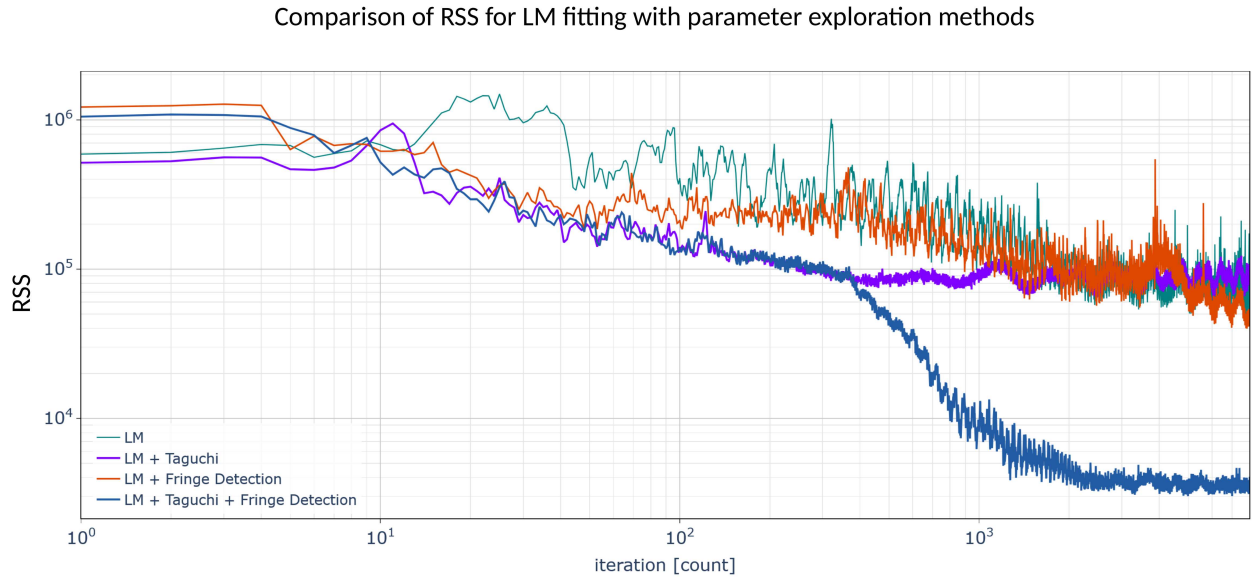



**Fig. 6 Residual sum of squares (RSS) convergence comparison for the dual system.** LM fitting with different parameter exploration methods on similar data (independently varying injected noise) and same initial guess. The target data varies similarly to Figure 4. The combined LM + Taguchi OA + fringe detection method escapes the local minimum near RSS $\approx 10^5$, converging to roughly an order of magnitude lower RSS. Fringe detection adds autocorrelation based fringe frequency estimation. Note: this is not an Allan-Werle variance; the x-axis is simply proportional to time.

…), and $a_{j,k,m}$ are the precomputed coefficients specific to spectral point $k$ and molecule $m$. At runtime, evaluating the polynomial reduces to a dot product of $a_{j,k,m}$ with the monomial vector $[\phi_1 \dots \phi_{N_m}]$. Species concentration mole fractions $\mathbf{x}$ enter via two stages: first, an effective pressure correction (Eq. 9) uses $\mathbf{x}$ to restore the correct Lorentzian broadening line shape for each molecule; then, Beer–Lambert transmission calculations occur from the summation of corrected per-molecule absorption coefficients scaled by mole fraction.

The polynomial at each $\nu_k$ captures contributions from all transition lines whose wings extend into that spectral channel — potentially hundreds to thousands of spectral lines. Figure 7 illustrates the polynomial approximation of the HITRAN absorption coefficient at a single spectral point within a scan.

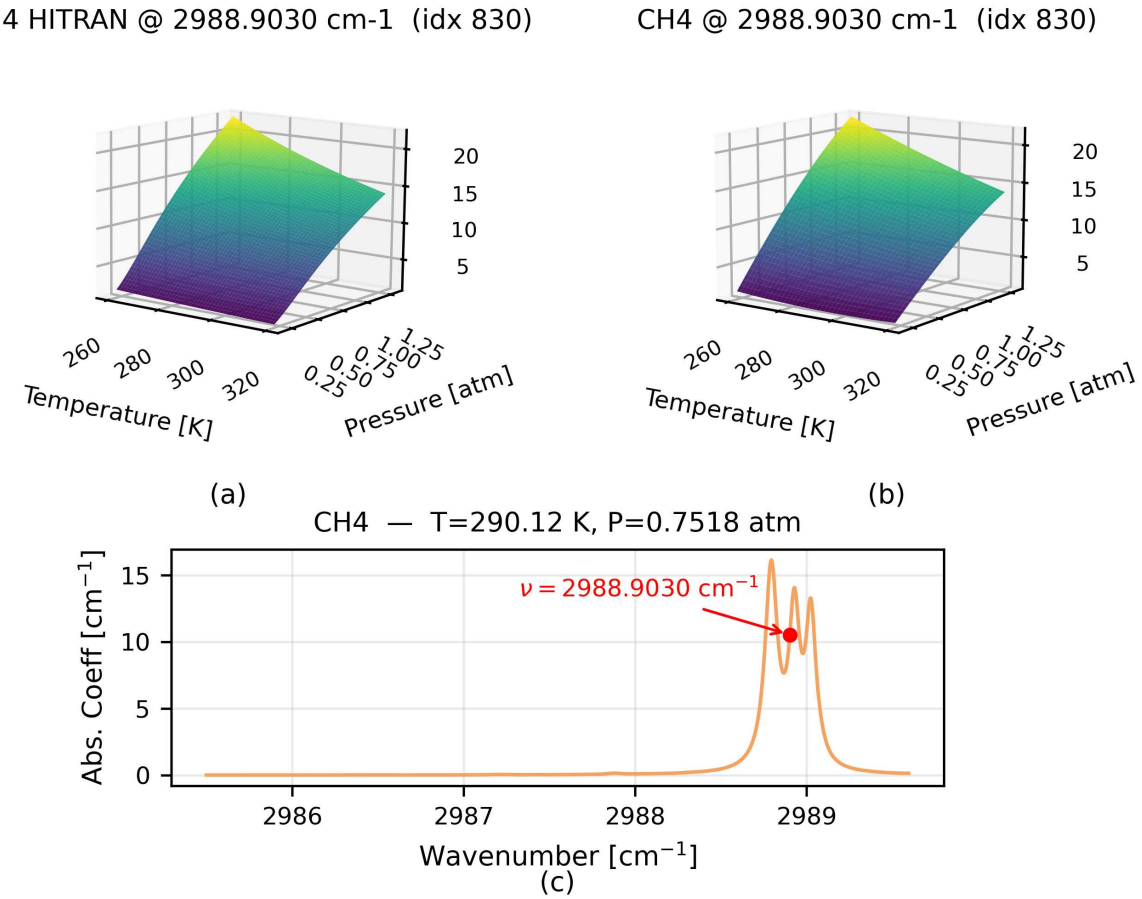



**Fig. 7 Polynomial approximation of Voigt HITRAN absorption coefficient at a single spectral point.** (a) HITRAN simulation of the $CH_4$ absorption coefficient at $\nu = 2988.903$ cm$^{-1}$ over the $(T, p)$ operating grid. (b) Multi-dimensional polynomial approximation at the same spectral point, approximating the surface with a compact set of 8 precomputed coefficients. (c) $CH_4$ absorption spectrum showing the full spectral range; the arrow indicates the wavenumber point of the surfaces in (a) and (b).

The total simulated absorption coefficient spectrum is then:

$$\hat{\alpha}_{\text{total}}(\nu_k) = \sum_m x_m \cdot \hat{\alpha}_{m,\text{corrected}}(\nu_k; T, p). \tag{4}$$

This formulation has several key advantages for real-time hardware: (i) evaluation at each spectral point reduces to a dot product of precomputed coefficients with the monomial vector, which maps efficiently onto DSP (digital signal processor) multiply-accumulate architectures; (ii) the continuous polynomial surface is inexpensive to re-evaluate, so the Jacobian is efficiently computed by finite differences of the full spectral model. Each parameter perturbs by a small step $h$ in smooth finite differences methods, avoiding the cost and discontinuities across grid-spaced interpolation in direct tabulated lookup approaches; and (iii) a balance of accuracy versus resource usage is possible by tailoring the polynomial order and included terms.

*4.1.1 Separation of concentration and broadening temperature dependence from the $(T, p)$ grid*

An obstacle to obtaining simple low-order polynomial fits in $(T, p)$ is that the Doppler and Lorentzian components must convolve to produce the Voigt line shape (the dominant spectral effects). There are additional effects from self- and cross-broadening that further depend on concentration. Doppler half-width scales as $\gamma_D \propto \sqrt{T}$, while the full Lorentzian half-width for molecule $m$ in a gas mixture is:

$$\gamma_{L,m}(T,p,\mathbf{x}) = \left[\gamma_{\mathrm{air},m}\cdot x_{\mathrm{air}}\cdot\left[\frac{T_{\mathrm{ref}}}{T}\right]^{n_{\mathrm{air},m}} + \gamma_{\mathrm{self},m}\cdot x_m\cdot\left[\frac{T_{\mathrm{ref}}}{T}\right]^{n_{\mathrm{self},m}} + \gamma_{\mathrm{H_2O},m}\cdot x_{\mathrm{H_2O}}\cdot\left[\frac{T_{\mathrm{ref}}}{T}\right]^{n_{\mathrm{H_2O},m}} + \sum_b \gamma_{b,m}\cdot x_b\cdot\left[\frac{T_{\mathrm{ref}}}{T}\right]^{n_{b,m}}\right]\cdot p, \quad (5)$$

where $\gamma_{\mathrm{air},m}$, $\gamma_{\mathrm{self},m}$, $\gamma_{\mathrm{H_2O},m}$, $\gamma_{b,m}$ are HITRAN-tabulated broadening half-widths at reference temperature $T_{\mathrm{ref}} = 296\,\mathrm{K}$ for air-, self-, $H_2O$ cross-, and other broadening species respectively, $n_{\mathrm{air},m}$, $n_{\mathrm{self},m}$, $n_{\mathrm{H_2O},m}$, $n_{b,m}$ are the corresponding temperature exponents, and $x_i$ are the mole fractions (concentration) of each broadening component. The line intensity $S(T)$ also further depends on $T$.

Convolution is a fairly expensive operation, but it is necessary to combine $\gamma_D(T)$ and $\gamma_{L,m}(T,p,\mathbf{x})$ into the line shape. On first inspection, the operation resists separation into isolated $T$, $p$, and $\mathbf{x}$ effects without post-convolution. In particular, the Lorentzian width in Eq. 5 couples all three effects through multiple distinct temperature exponents ($n_{\mathrm{air}}$, $n_{\mathrm{H_2O}}$, etc.) and broadening coefficients ($\gamma_{\mathrm{air}}$, $\gamma_{\mathrm{H_2O}}$, etc.). A polynomial in $(T,p,\mathbf{x})$ that attempts to capture this fully coupled dependence directly requires high-order cross-terms — especially for the $\mathbf{x}$-based dimensions — to represent the nonlinear $(T_{\mathrm{ref}}/T)^{n_i}$ factors for each broadener, dramatically increasing the number of polynomial terms.

This is addressed by building a simplified precomputation grid that deliberately uses air as the sole broadening partner and fixes the broadening temperature at $T_{\mathrm{broad}} = T_{\mathrm{ref}} = 296\,\mathrm{K}$. This reduces the Lorentzian width to $\gamma_{L,\mathrm{split},m} = \gamma_{\mathrm{air},m}\,p$, which depends on pressure alone. The Doppler width and line intensity retain actual $T$ ($m_{\mathrm{mol}}$ is the molecular mass, $k_B$ is the Boltzmann constant):

$$\gamma_{D,m} = \sqrt{\frac{2k_BT\ln 2}{m_{\mathrm{mol}}c^2}}\cdot\nu_0 \quad (\text{actual } T), \quad (6)$$

$$\gamma_{L,\mathrm{split},m}(p) = \gamma_{\mathrm{air},m}\cdot p \quad (\text{air-only, frozen at } T_{\mathrm{ref}}). \quad (7)$$

The polynomial at each spectral point therefore represents only the $(T,p)$ dependence (through $N$, $S$, $\gamma_D$, and $\gamma_{L,\mathrm{split}}$); the $\mathbf{x}$ dimensions and the per-broadener $(T_{\mathrm{ref}}/T)^{n_i}$ factors are removed from the grid, so low-order 2-dimensional polynomial terms capture the remaining smooth surface adequately.

The linearity of the Lorentzian width in $p$ allows compensating for this simplification. The true width (Eq. 5) and the split width (Eq. 7) are both strictly proportional to $p$, so their ratio is independent of pressure:

$$\frac{\gamma_{L,m}(T,p,\mathbf{x})}{\gamma_{L,\mathrm{split},m}(p)} = \sum_i \frac{\gamma_{i,m}}{\gamma_{\mathrm{air},m}}\cdot x_i\cdot\left[\frac{T_{\mathrm{ref}}}{T}\right]^{n_i} \equiv \eta_m(T,\mathbf{x}). \quad (8)$$

Because $p$ cancels, defining an effective pressure

$$p_{\mathrm{eff},m} = \eta_m\cdot p \quad (9)$$

and evaluating the split model polynomial at $p_{\mathrm{eff},m}$ reproduces the true Lorentzian width of a single line. Each line within and adjacent to the spectrometer tuning range has its own broadening ratios and exponents, so a single per-molecule $\eta_m$ built from weighted-average parameters (Sec. 4.1.3) closely reproduces the widths, with residual error set by the line-to-line spread of these parameters. The $\eta_m$ ratio thus allows restoration of the majority of relevant missing physics without requiring a real-time convolution — self-broadening, cross-broadening by $H_2O$ (and other species if available), and the $(T_{\mathrm{ref}}/T)^{n_i}$ temperature dependence; an associated pressure-shift residual is also addressed in Sec. 4.1.4.

### *4.1.2 Per-molecule effective pressure correction derivation*

The effective pressure $p_{\mathrm{eff},m}$ method (Eq. 9) thus allows recovery of more accurate Lorentzian line widths for the target gas mixture in a simple manner; Sec. 4.1.3 gives a compact hardware form of $\eta_m$ using per-broadener correction functions $c_i(T)$.

Evaluating the split polynomial at $p_{\mathrm{eff},m}$ instead of the actual physical pressure $p$ causes the simplified width $\gamma_{L,\mathrm{split},m} = \gamma_{\mathrm{air},m}\,p_{\mathrm{eff},m}$ to match (exactly for a single line, approximately for weighted-average parameters) the full mixture-broadened width $\gamma_{L,m}(T,p,\mathbf{x})$ (Eq. 5), thereby producing a spectrum with the correct line *shape*. However, because the absorption coefficient depends on the number density $N = p/(k_BT)$, evaluating the polynomial at $p_{\mathrm{eff},m}$ also scales the absorption magnitude by $\eta_m$. The corrected absorption coefficient is therefore:

$$\alpha_{m,\mathrm{corrected}}(\nu)\big|_p = \frac{\alpha_{m,\mathrm{split}}(\nu;p_{\mathrm{eff},m})}{\eta_m} \quad (10)$$

which restores the true number density $N$ while preserving the line shape.

A per-molecule treatment is essential for two reasons. First, the broadening parameter ratios differ across species: for $H_2O$ in the regions simulated here, $\gamma_{\mathrm{self}}/\gamma_{\mathrm{air}} \approx 4-7$, so $H_2O$'s effective pressure is dominated by self-broadening, whereas $CH_4$ is $1.3-1.6$ and predominantly air-broadened — but also experiences

$H_2O$ cross-broadening through $\gamma_{H_2O}$ parameters. Second, $\eta_m$ depends on the full concentration vector $\mathbf{x}$ — on $x_m$ through self-broadening, on $x_{H_2O}$ and $x_b$ through cross-broadening, and on all of them through $x_{air}$ — so a change in any species' concentration (e.g., $H_2O$) alters every molecule's effective pressure and contributes to the Jacobian terms $\partial\hat{\alpha}_m/\partial x_i$. Self-broadening has less effect for trace target species; however, cross-broadening by major variable constituents (e.g., percent-level $H_2O$) remains significant. With the methods described here, the reduced computational complexity allows inclusion of multi-species cross-broadening generally and expands the potential application space. Figure 8 validates the polynomial at off-grid $(T, p)$ points, and Figure 10 compares the 8-coefficient polynomial against the full Hartmann-Tran simulation.

#### *4.1.3 Fit $c(T)$ correction factors and weighted-average broadening parameters*

The effective pressure correction factor $\eta_m$ (Eq. 8) rewritten compactly using broadener-specific correction functions $c_i(T)$ allows:

$$\eta_m = 1 + \sum_i c_i(T) \cdot x_i, \tag{11}$$

where the sum runs over all broadening species $i$ (*i.e.* air, self, $H_2O$, ...), and

$$c_i(T) = \frac{\gamma_{i,m}}{\gamma_{air,m}} \cdot \left[\frac{T_{ref}}{T}\right]^{n_{i,m}} - 1. \tag{12}$$

Hence, $c_{air}(T) = (T_{ref}/T)^{n_{air}} - 1$ captures the temperature dependence of the Lorentzian width due to air removed by the split.

Over moderate temperature ranges, a linear function closely approximates $c_i(T)$ for each molecule:

$$c_i(T) \approx c_{a,i} + c_{b,i} \cdot T, \tag{13}$$

where $c_{a,i}$ and $c_{b,i}$ are least-squares fit to $c_i(T)$. This linearization enables computation of the effective pressure in real-time hardware. The coefficient of determination $R^2$ of the linear fit is typically $> 0.995$ over 60 K (Fig. 9). In cases of extended temperature ranges and stronger concentration broadening behavior, other approximations (*e.g.*, piecewise linear, polynomial) provide more accuracy.

Each individual spectral line has its own broadening parameters: ($\gamma_{air}$, $n_{air}$, $\gamma_{self}$, $n_{self}$, $\gamma_{H_2O}$, $n_{H_2O}$, ...). To simplify here, the $p_{eff}$ correction uses a single representative set per molecule. These representative values are weighted averages over all lines within the spectral range — for a $n_{air}$ example:

$$\overline{n_{air,m}} = \frac{\sum_\ell w_\ell \, n_{air,\ell}}{\sum_\ell w_\ell}, \qquad w_\ell = S_\ell \cdot \gamma_{air,\ell}, \tag{14}$$

where $S_\ell$ is the line intensity and $\gamma_{air,\ell}$ is the air-broadening half-width of line $\ell$. The weighting by $w_\ell$ ensures that stronger, broader lines — which dominate the absorption signal — contribute proportionally more to the effective temperature exponent.

For gas sample constituent matrices that deviate significantly from standard atmospheric abundance — such as combustion exhaust [38], industrial process streams, human breath samples, or planetary atmospheres — broadening by species other than air and $H_2O$ becomes especially significant, particularly $CO_2$. $O_2$ and $N_2$ may also require separate broadening coefficients compared to a single combined air-broadening coefficient. When spectral databases do not provide the relevant broadening coefficients for these combinations, experimental measurements may fill in the required gaps.

#### *4.1.4 Other issues*

The precomputation of polynomial coefficients requires evaluating the full HITRAN LBL simulation at every $(T, p)$ grid point for every spectral point. This work uses HITRAN HAPI [40] for the simulations. Speedup using other simulation methods is possible, but this work preferentially uses HAPI for calculating line shapes [39] as ground truth.

Minimax $L^\infty$ approximation at each point also reduces the maximum residual for the absorption coefficient polynomial representation and allows reweighting to target certain operational $(T, p)$, at a cost of overall $rms$ [41]. The largest $\alpha$ errors typically occur at line peaks, where the absorption signal and the leverage on concentration fit is intensified; thus, the maximum residual is typically more significant to the spectral model than $rms$, so $L^\infty$ usually improves results.

Independent grouping of isotopologues of molecular species allows for isotope-ratio measurements [42]. This is an application which benefits strongly from high accuracy because the intensity ratio of isotopologue lines with differing lower-state energies depends strongly on temperature within a typical DFB (distributed feedback) diode laser's wavelength range of a few wavenumbers [13].

Other approximations for the absorption coefficient (including rational approximations, fractional and negative exponents, or those that include transcendental functions) also improve representational accuracy, but

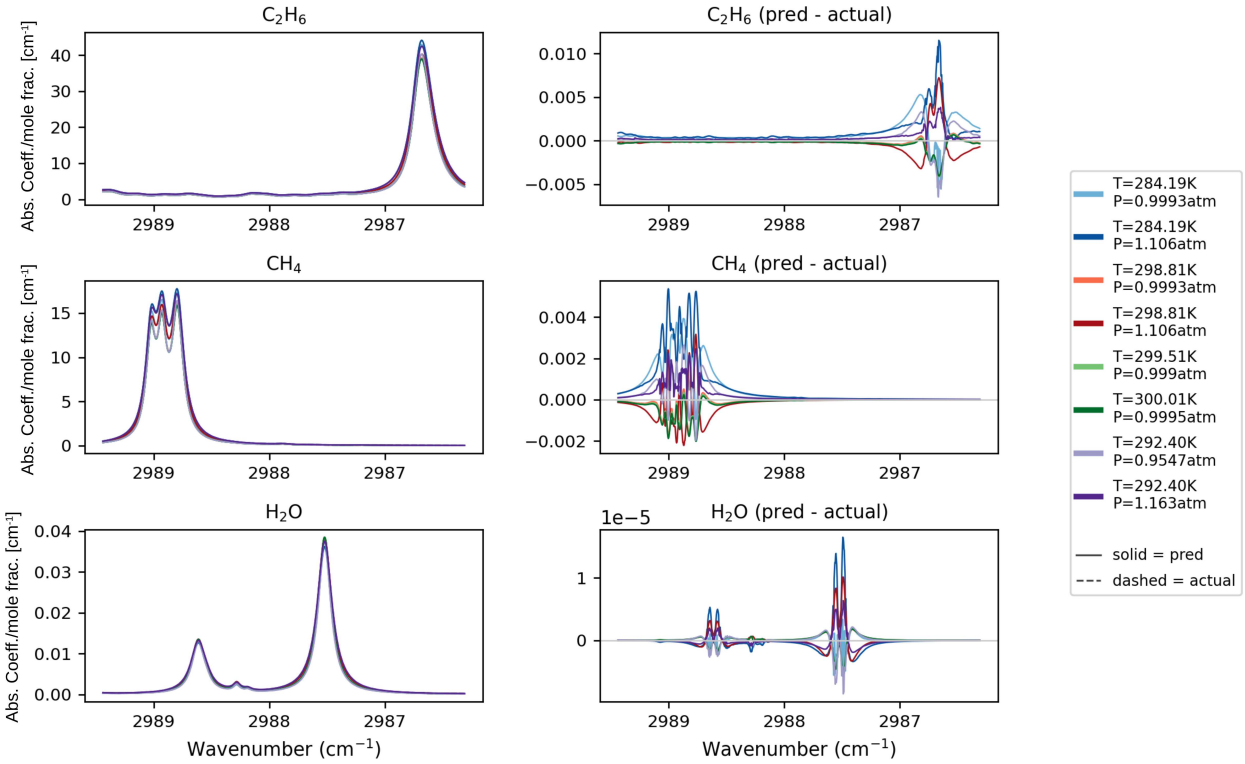



**Fig. 8 Off-grid validation.** Simulation at off grid $(T, p)$ points for an example 8-coefficient minimax polynomial targeting 270-313K and 0.9-1.2 atm.

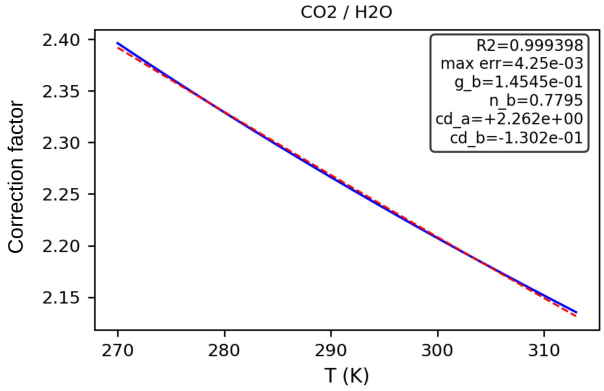



**Fig. 9 Linear $c(T)$ approximation.** Linear fit of a sample correction function $c(T)$ (Eq. 13) over an operating temperature range of 60 K.

for architectural simplicity, speed, and minimal hardware resource usage a standard polynomial form provides performance benefits.

The $(T, p)$ grid precomputation may also account for light source spectral line shape if available by using built-in instrumental line shape convolution in the HITRAN HAPI simulations or manual convolution.

The polynomial approach here inherits slight accuracy loss mechanisms arising from weighted-average approximations for $p_{\text{eff}}$. The use of a weighted-average $n_{\text{i}}$ and $\gamma_{\text{i}}$ means imperfect correction of an individual line's temperature exponents which vary line-to-line.

In addition to broadening, pressure effects cause a shift in the line center frequency: $\nu_0(p) = \nu_{0,\text{ref}} + \delta_{\text{air}}\, p$, where $\delta_{\text{air}}$ is the HITRAN pressure-shift coefficient. The polynomial approximation absorbs the majority of these induced shifts into its per-spectral-point coefficients, since the precomputation grid spans the full $(T, p)$ range and the polynomial at each $\nu$ fits that calculated absorption profile.

For the $p_{\text{eff}}$ correction, the use of $p_{\text{eff}} \neq p$ introduces a small spurious frequency shift $\delta_{\text{air}} \cdot (p_{\text{eff}} - p)$, with $\delta_{\text{air}}$ data available in HITRAN. A precomputed weighted-average pressure-shift coefficient $\overline{\delta_{\text{air}}}$ per molecule (anal-

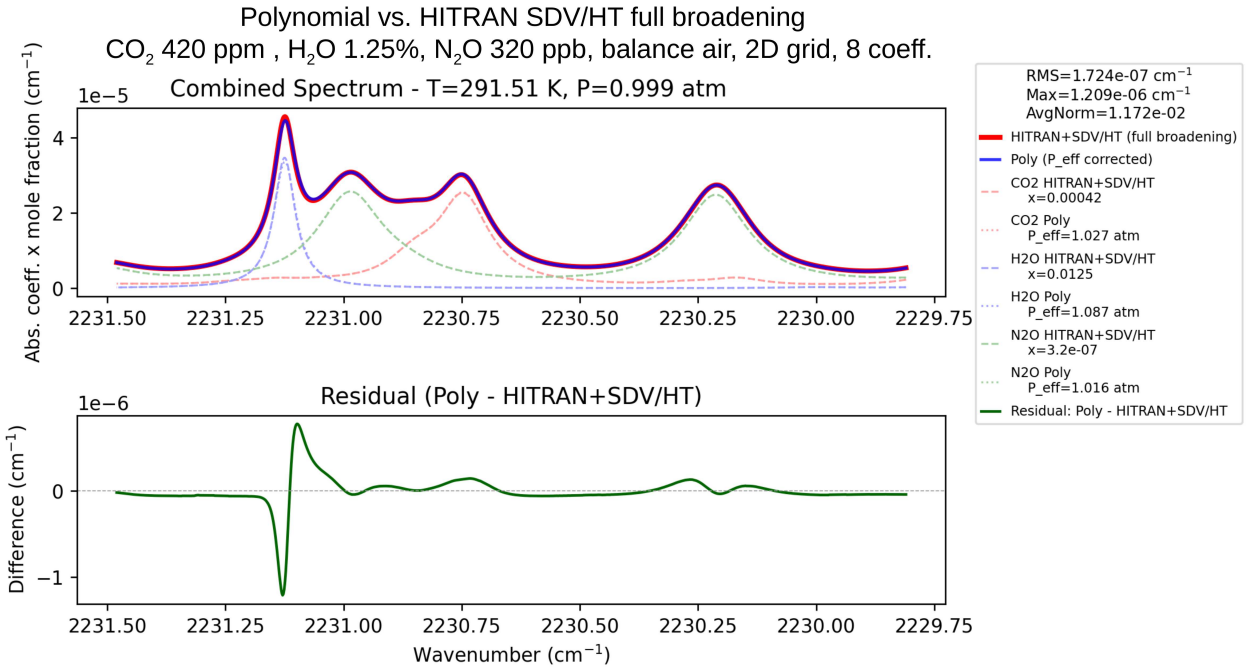



**Fig. 10 Polynomial approximation vs. full Hartmann-Tran (SDV/HT) HITRAN simulation.** Comparison of the 8-coefficient polynomial approximation with $p_{\text{eff}}$ correction for 3 molecules against the full Hartmann-Tran [39] HITRAN simulation with self- and cross-broadening. Top: combined absorption coefficient. Dashed lines are the per-species HITRAN+SDV/HT contributions; dotted lines are the $p_{\text{eff}}$-corrected polynomial approximations. Bottom: residual, with an RMS of $1.72 \times 10^{-7}\,\text{cm}^{-1}$ and a maximum of $1.21 \times 10^{-6}\,\text{cm}^{-1}$. The averaged normalized error is $1.17 \times 10^{-2}\,[\text{cm}^{-1}/\text{cm}^{-1}]$ across the spectrum.

ogous to the weighted-average broadening parameters) allows shifting that molecule's $p_{\text{eff}}$-evaluated spectrum back by $\overline{\delta_{\text{air}}} \cdot (p_{\text{eff}} - p)$ wavenumbers to compensate.

Because the shift coefficient varies line-to-line, the weighted-average correction is imperfect for individual lines (similar to $c(T)$ use of weighted-average values), but still typically improves overall accuracy.

These corrections are comparatively small. In critical accuracy applications which include a larger number of polynomial terms, these corrections become significant to lower overall error. An overall spectral shift as a fit parameter compensates for this effect to a small extent if all the lines shift in generally the same direction; however, this may skew any laser wavelength stabilization feedback. There is also a $\delta_{\text{self}}$ parameter affected by concentration but that effect is even smaller in comparison.

### *4.1.5 Extensions to additional physical effects*

The polynomial precomputation framework can extend beyond the standard HITRAN absorption model. If the underlying simulation incorporates additional physics — Zeeman splitting in the presence of an external magnetic field [43], Stark effects from electric fields, polarization-dependent absorption, dispersion, or even wavelength modulation spectroscopy (WMS) signals — a similar multi-dimensional polynomial fitting procedure can apply. The simulations (or experiments) include the additional physical parameters as another dimension if they vary during measurement (e.g., magnetic field strength $B$, modulation depth) alongside $T$ and $p$, and the polynomial extends accordingly (e.g., $\phi_j(T, p, B)$). The polynomial coefficients then encode the combined dependence on all parameters, enabling real-time evaluation of these more complicated spectral models.

### *4.1.6 Rank reduction and composition input axis*

Standard rank reduction (via principal component analysis (PCA) or singular value decomposition (SVD)) allows for more efficient higher dimensional coefficient storage, requiring decompression of the spectrum from a real-time set of transformed $(T, p)$ and optional $\mathbf{x}$ inputs. The compression and reconstruction allows use of a higher number of polynomial input terms of $T$ and $p$, improving the surface representation while allowing the rank reduction to discard basis directions with little effect on accuracy.

An "effective broadening" method similar to $p_{\text{eff}}$ can combine the concentration dependence of the individual self- and cross-broadening molecules into a composition input axis $D_m$ for each absorbing molecule $m$ (as opposed to an individual input axis for each gas mixture

molecule):

$$D_m = \sum_i x_i \cdot \left( \frac{\gamma_{i,m}(T)}{\gamma_{\mathrm{air},m}(T)} - 1 \right), \tag{15}$$

where the sum runs over the non-air broadening species $i$ (self, $H_2O$, ...) with mole fractions $x_i$, and $\gamma_{i,m}(T) = \gamma_{i,m} \cdot (T_{\mathrm{ref}}/T)^{n_{i,m}}$ uses the weighted-average broadening parameters of Eq. 14. $D_m$ is the fractional change in $m$'s average collisional line width caused by the gas composition difference from air.

Comparing this to existing approaches in literature, one method tabulates a single species' mole fraction as a direct lookup axis [29]; and another method fits an isolated transition's collisional width [6]. Here, $D_m$ collapses all broadeners for all lines into one polynomial input axis with precomputed coefficients.

Accordingly, the precomputation grid adds simulations of mixtures spanning the expected range of $D_m$. This adds $D_m$ to an expanded multidimensional polynomial $\phi_j(T, p, D_m)$ which represents the grid dependence on mole fraction. Unlike the $p_{\mathrm{eff}}$ correction, the weighted-average parameters here only provide an input axis, so the fit itself captures line-to-line differences in broadening. A single broadening input axis loses no information when only one broadener varies significantly in concentration; when several vary (as in situations with elevated levels of $CO_2$ and $H_2O$), it is an approximation; additional optional input axes of individual terms of Eq. 15 can improve accuracy further.

Regardless of inclusion of these additional axes $D$, the polynomial in $(T, p)$ with broadening compensation remains a useful method for accurate, quickly calculated representation of the grid surfaces. The rank reduction offsets any resulting growth in the number of polynomial terms in cases needing the highest accuracy.

Figure 11 shows the resulting accuracy of a 15-coefficient SVD rank-reduced representation over a multidimensional $(T, p, D_m)$ grid for the same mixture as Figure 10.

### 4.2 LM Taguchi OA fitting with RSS-scaled perturbation

This manuscript further presents a novel combination of Levenberg–Marquardt (LM) with Taguchi orthogonal array (OA) parameter space exploration for curve fitting.

Taguchi OAs [44] (and other arrays such as Plackett–Burman [45]) find use in optimization of physical industrial processes to determine the effects of changing a variety of parameters and identify a final combination that improves a predetermined signal-to-noise ratio with minimized number of experiments.

In this work, each Taguchi run $r$ perturbs multiple parameters simultaneously to explore the parameter space, each with a tunable per-parameter perturbation magnitude $\kappa_i$. This allows the LM fit to escape local minima, since the Taguchi exploration discovers the optimum direction of a large coordinated jump across all perturbed parameters — mostly unrelated to the LM iteration parameter change, which depends on the Jacobian shape and only accounts for highly localized parameter behavior.

In run $r$, $s_{i,r}$ is the Taguchi perturbation direction that row $r$ of the orthogonal array assigns to parameter $\beta_i$. The perturbation applied to $\beta_i$ during that run is

$$\Delta_{\mathrm{tag}}\beta_{i,r} = \kappa_i \cdot s_{i,r} \cdot g(\mathrm{RSS}), \tag{16}$$

where $g(\mathrm{RSS})$ is a function that adapts to the residual sum of squares (RSS), related to $\chi^2$, scaling the perturbation according to the RSS magnitude. At smaller RSS the jumps are shorter since the parameters should be closer to their optimal values.

This ensures stable perturbation magnitudes across a wide dynamic range of RSS and minimal resource usage. The winning direction for each parameter minimizes the sum of RSS values over all specific Taguchi runs in which that parameter takes a specific direction:

$$s_i^* = \arg\min_{s_i} \sum_{r\,:\,s_{i,r}=s_i} \mathrm{RSS}_r. \tag{17}$$

The fitter assembles the optimum Taguchi parameter set from the winning directions applied to the current base parameters:

$$\beta_{\mathrm{opt},i} = \beta_{\mathrm{base},i} + \kappa_i \cdot s_i^* \cdot g(\mathrm{RSS}). \tag{18}$$

The fitter then accepts whichever of the Taguchi candidate $\boldsymbol{\beta}_{\mathrm{opt}}$ and the LM step yields the lower RSS. Additional optimizations are possible such as parameter perturbation scaling by a fraction or multiple of the current value, scaling based on recent success of a parameter, or adding randomness to the perturbation — at a cost of compute resources.

The combined LM+OA+RSS scaled system converges quickly toward minima while retaining the ability to escape local minima traps.

Removing the Taguchi OA optimization immediately frees computing resources for 1 kHz operation in the dual laser system; however, 0.82 kHz operation with Taguchi OA avoids local minima and reaches roughly an order of magnitude lower RSS (Fig. 6).

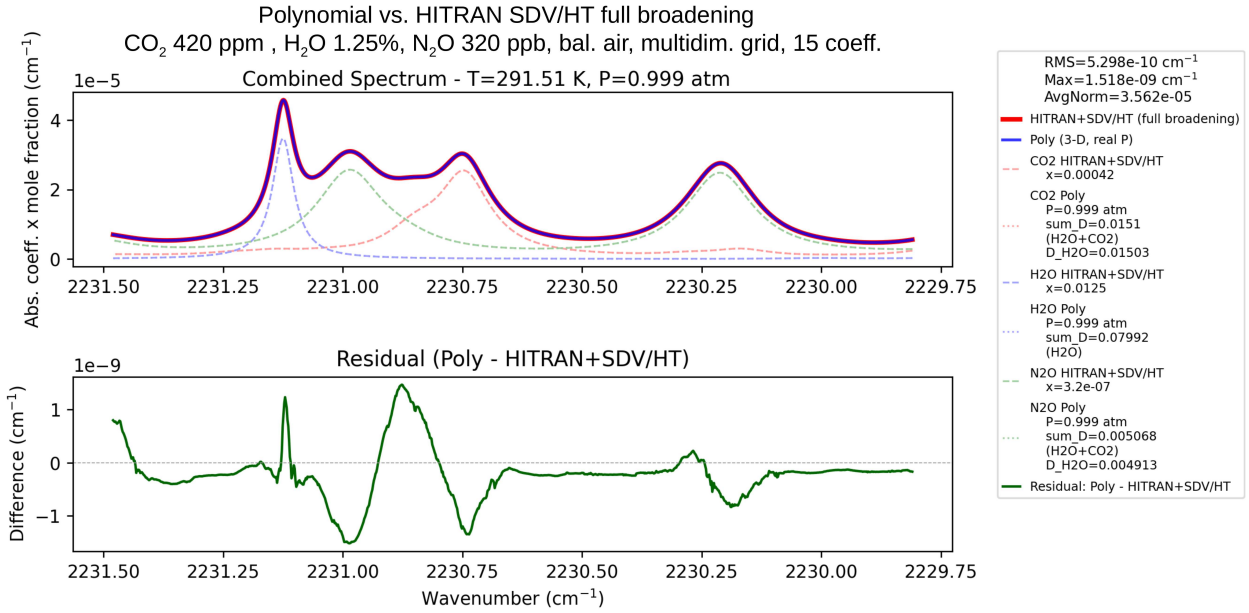



**Fig. 11 Rank-reduced approximation vs. full Hartmann-Tran (SDV/HT) HITRAN simulation.** Comparison of the rank-reduced 15-coefficient polynomial approximation over a multidimensional $(T, p, D_m)$ grid for 3 molecules against the full Hartmann-Tran [39] HITRAN simulation with self- and cross-broadening. The model combines the pressure $p$ with the composition input axis $D_m$ (Eq. 15) instead of the $p_{\text{eff}}$ correction. Top: combined absorption coefficient. Dashed lines are the per-species HITRAN+SDV/HT contributions. Bottom: residual, with an RMS of $5.30 \times 10^{-10}\,\text{cm}^{-1}$ and a maximum of $1.52 \times 10^{-9}\,\text{cm}^{-1}$. The averaged normalized error is $3.56 \times 10^{-5}$ [$\text{cm}^{-1}/\text{cm}^{-1}$] across the spectrum.

4.3 Simultaneous dual-laser single-detector architecture

Another contribution of this work is a simultaneous dual-laser single-detector spectrometer architecture that does not require time-division multiplexing. Two lasers operating at different wavelength ranges (e.g., $3.3\,\mu$m and $4.5\,\mu$m) illuminate a single detector (or multiple detectors with electronically summed signals if photonic mixing creates excessive interference), and both scan wavelength simultaneously. The fast spectral fitting described here specifically enables the separation of the combined signal in the digital domain by allowing discrimination of co-prime spectral lengths as a form of phase-sensitive detection.

*4.3.1 Co-prime spectral lengths and signal separation*

In TDLAS, sawtooth current ramps cause sweeps of laser wavelength across the spectral range. Here, two lasers scan at co-prime spectral lengths $L_1 = 1000$ and $L_2 = 983$ (i.e., $\gcd(L_1, L_2) = 1$), which also set the lengths of the current control waveforms. Because the spectral lengths are co-prime, the lasers' sweep pattern repeats only after $L_1 \times L_2$ samples, ensuring no persistent alignment between the two spectra. Both lasers contribute to every detector sample (with independent specific laser-off offset measurement periods); the fitter separates them by jointly fitting the sum of the two spectral models. The speed of spectral fit processing allows the system to avoid co-adding of the spectra — which would not align due to the co-prime lengths.

*4.3.2 Reverse ramp for baseline isolation*

To further prevent the polynomial baseline models of the two lasers from interacting, one laser may be driven with a forward (increasing) current ramp while the other is driven with a reverse (decreasing) current ramp. Since the polynomial baseline captures the laser's intensity-versus-current characteristic, the forward ramp produces a mostly positive slope while the reverse ramp produces a mostly negative slope. This opposite monotonicity provides additional isolation of the two baseline models. Noise from the two light sources will mix, but field-deployed TDLAS systems typically show noise issues dominated by either (i) detector noise, (ii) fringes which are coherent with each spectrum, or (iii) technical noise from pressure and temperature variation. Figure 12 illustrates the co-prime spectral alignment and reverse-ramp structure of the dual-laser system. Figures 4 and 5 use this configuration.

## 5 Conclusion

This work integrates several major changes to the standard methodology for real-time gas-phase laser spectro-

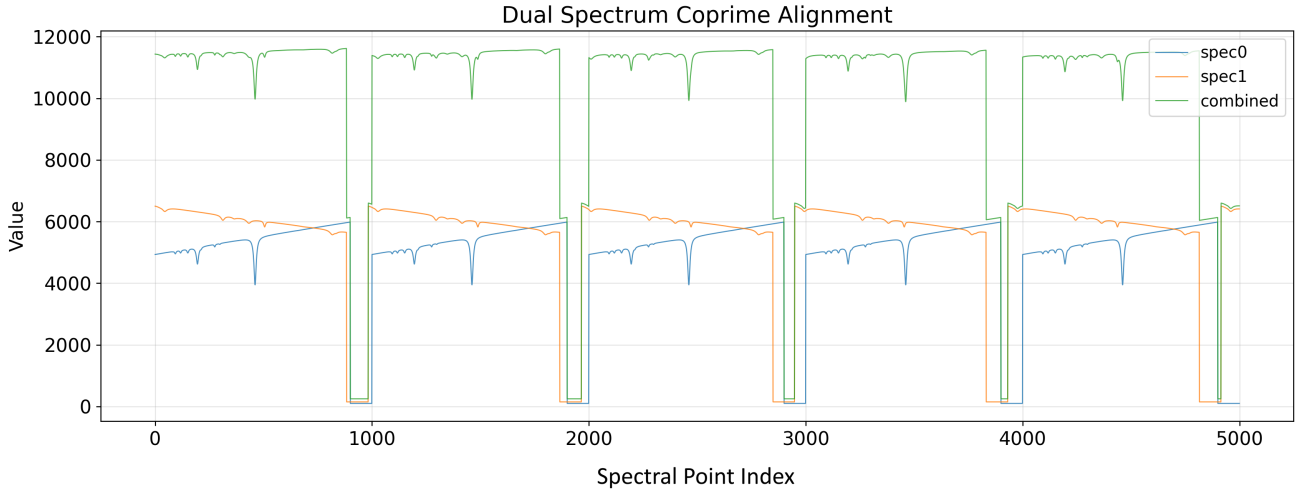



**Fig. 12 Dual-spectrum co-prime alignment.** The two laser spectra (spec0 forward ramp, spec1 reverse ramp) and their combined single-detector signal, showing the interleaved co-prime spectral lengths and opposite ramp directions.

scopic analysis. First, precomputed multi-dimensional polynomials based on inputs related to $T$ and $p$ approximate all-lines line-by-line HITRAN simulations at each spectral point, reducing absorption coefficient evaluation to a single vector multiply-accumulate operation per molecule per point while maintaining percent-level to ppm-level fidelity across wide temperature and pressure ranges. Second, effective pressure and broadening correction ($p_{\text{eff}}$, $D_m$) methods recover mixture self- and cross-broadening of the full gas matrix, which many field spectrometers omit. Third, the Levenberg–Marquardt fitting solver augmented with Taguchi orthogonal array exploration and RSS-scaled perturbation accelerates convergence and escapes local minima that trap conventional LM solvers. Fourth, a co-prime dual-laser single-detector architecture can fit six concentration parameters simultaneously at 0.82 kHz on a single mid-range FPGA consuming 1.6 W simulated computational power for the fit systems. All configurations (single and dual laser, with and without Taguchi OA, for the 8- and 15-coefficient polynomial models, with and without rank reduction) completed post-implementation timing closure, power estimation, and behavioral simulation in AMD Vivado 2025.1 at the stated iteration rates on the devices listed in Table 1. Together, these methods enable orders of magnitude higher speed direct absorption fitting of 1000-point spectra — with simultaneous temperature, pressure, etalon fringe, and multi-species concentration retrieval — at millisecond latency. This opens routes to vastly improved field-portable, battery-powered high-accuracy spectroscopic sensing.

## Declarations:

**Author Contributions:** S.S. performed all work. Base simulation code, computer architecture, and critical path FPGA code was conceived and written by S.S.. Portions of auxiliary coding and tooling used Claude Code 4-5.5 AI LLMs, with all outputs compared to HITRAN HAPI by S.S.. The author takes full responsibility for the content of the publication.
**Funding Statement:** There were no external funding sources for this work.
**Competing interests:** The author's company (Sophotonics Inc.) has filed provisional patent applications on the technologies described here. The author holds shares in a previous company (Aeris Technologies Inc. / International Environmental Standards PBC) dealing with laser spectrometers, but solely interacts via a third party shareholder representative.
**Ethical Statement:** Not applicable.
**Data Availability:** Spectral data was retrieved from HITRAN (doi:10.1016/j.jqsrt.2026.109807). Other data is available from the author on reasonable request. All data were generated via simulation.

## References

1. I. E. Gordon, L. S. Rothman, R. J. Hargreaves, F. M. Gomez, T. Bertin, C. Hill, et al., The HITRAN2024 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **353**, 109807 (2026).
2. G. C. Bjorklund, M. D. Levenson, W. Lenth, C. Ortiz, Frequency modulation (FM) spectroscopy: Theory of lineshapes and signal-to-noise analysis. *Appl. Phys. B* **32**, 145–152 (1983).
3. G. Berden, R. Peeters, G. Meijer, Cavity ring-down spectroscopy: Experimental schemes and applications. *Int. Rev. Phys. Chem.* **19**, 565–607 (2000).

4. Q. He, C. Zheng, K. Zheng, F. K. Tittel, Off-axis integrated cavity output spectroscopy for real-time methane measurements with an integrated wavelength-tunable light source. *Infrared Phys. Technol.* **115**, 103705 (2021).
5. N. Picqué, T. W. Hänsch, Frequency comb spectroscopy. *Nat. Photonics* **13**, 146–157 (2019).
6. C. S. Goldenstein, C. L. Strand, I. A. Schultz, K. Sun, J. B. Jeffries, R. K. Hanson, Fitting of calibration-free scanned-wavelength-modulation spectroscopy spectra for determination of gas properties and absorption lineshapes. *Appl. Opt.* **53**, 356–367 (2014).
7. D. I. Herman, M. Walsh, M. K. Kreider, N. Lordi, E. J. Tsao, A. J. Lind, M. Heyrich, J. Combes, J. Genest, S. A. Diddams, Squeezed dual-comb spectroscopy. *Science* **387**, 653–658 (2025).
8. A. Amann, D. Smith, *Breath Analysis for Clinical Diagnosis and Therapeutic Monitoring* (World Scientific, 2005).
9. B. Henderson, A. Khodabakhsh, M. Metsälä, I. Ventrillard, F. M. Schmidt, D. Romanini, G. A. D. Ritchie, S. te Lintel Hekkert, R. Briot, T. Risby, N. Marczin, F. J. M. Harren, S. M. Cristescu, Laser spectroscopy for breath analysis: towards clinical implementation. *Appl. Phys. B* **124**, 161 (2018).
10. C. R. Webster, G. J. Flesch, K. Mansour, R. Haberle, J. Bauman, Mars laser hygrometer. *Appl. Opt.* **43**, 4436–4445 (2004).
11. J. D. Shutter, N. T. Allen, T. F. Hanisco, G. M. Wolfe, J. M. St. Clair, F. N. Keutsch, A new laser-based and ultra-portable gas sensor for indoor and outdoor formaldehyde (HCHO) monitoring. *Atmos. Meas. Tech.* **12**, 6079–6089 (2019).
12. R. L. Herman, R. F. Troy, K. M. Aaron, I. Sanders, K. Schwarm, J. E. Klobas, A. Swanson, A. Carpenter, S. Ozog, K. Chin, L. E. Christensen, D. Fu, R. F. Jarnot, R. A. Stachnik, R. Vasudev, JLH Mark2 — An improved opto-mechanical approach to open-path in situ water vapor measurement in the upper troposphere/lower stratosphere. *Atmos. Meas. Tech.* **18**, 4593–4610 (2025).
13. E. Kerstel, L. Gianfrani, Advances in laser-based isotope ratio measurements: selected applications. *Appl. Phys. B* **92**, 439–449 (2008).
14. J. Heiskanen, C. Brümmer, N. Buchmann, C. Calfapietra, H. Chen, B. Gielen, et al., The Integrated Carbon Observation System in Europe. *Bull. Am. Meteorol. Soc.* **103**, E855–E872 (2022).
15. D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.* **11**, 431–441 (1963).
16. K. K. Schwarm, R. M. Spearrin, Real-time FPGA-based laser absorption spectroscopy using on-chip machine learning for 10 kHz intra-cycle emissions sensing towards adaptive reciprocating engines. *Appl. Energy Combust. Sci.* **16**, 100231 (2023).
17. J. Reid, D. Labrie, Second-harmonic detection with tunable diode lasers — comparison of experiment and theory. *Appl. Phys. B* **26**, 203–210 (1981).
18. G. Wysocki, Y. Bakhirkin, S. So, F. K. Tittel, C. J. Hill, R. Q. Yang, M. P. Fraser, Dual interband cascade laser based trace-gas sensor for environmental monitoring. *Appl. Opt.* **46**, 8202–8210 (2007).
19. G. L. Turin, An introduction to matched filters. *IRE Trans. Inf. Theory* **6**, 311–329 (1960).
20. A. Behera, A. Wang, Calibration-free wavelength modulation spectroscopy: Symmetry approach and residual amplitude modulation normalization. *Appl. Opt.* **55**, 4446–4455 (2016).
21. C. S. Callahan, S. M. Bresler, S. C. Coburn, D. A. Long, G. B. Rieker, GAAS: GPU accelerated absorption simulator. *J. Quant. Spectrosc. Radiat. Transf.* **333**, 109307 (2025).
22. C. Yao, W. Zhang, P. Bao, J. Ma, W. Zhuo, M. Chen, Z. Shi, J. Zhou, Y. Ye, L. Ming, T. Yan, R. Penty, Q. Cheng, Chip-scale sensor for spectroscopic metrology. *Nat. Commun.* **15**, 10305 (2024).
23. A. Kosterev, G. Wysocki, Y. Bakhirkin, S. So, R. Lewicki, F. Tittel, R. F. Curl, Application of quantum cascade lasers to trace gas analysis. *Appl. Phys. B* **90**, 165–176 (2007).
24. K. L. Letchworth, D. C. Benner, Rapid and accurate calculation of the Voigt function. *J. Quant. Spectrosc. Radiat. Transf.* **107**, 173–192 (2007).
25. J. A. C. Weideman, Computation of the complex error function. *SIAM J. Numer. Anal.* **31**, 1497–1518 (1994).
26. M. Kuntz, A new implementation of the Humlicek algorithm for the calculation of the Voigt profile function. *J. Quant. Spectrosc. Radiat. Transf.* **57**, 819–824 (1997).
27. F. Schreier, Computational aspects of speed-dependent Voigt profiles. *J. Quant. Spectrosc. Radiat. Transf.* **187**, 44–53 (2017).
28. L. Xu, R. Xue, Y. Li, H. Zhang, Z. Cao, FPGA-based real-time implementation of temperature measurement via tunable diode laser absorption spectroscopy. *IEEE Photonics J.* **10**, 6800211 (2018).
29. S. A. Buehler, P. Eriksson, O. Lemke, Absorption lookup tables in the radiative transfer model ARTS. *J. Quant. Spectrosc. Radiat. Transf.* **112**, 1559–1567 (2011).
30. V. H. Payne, et al., Absorption coefficient (ABSCO) tables for the Orbiting Carbon Observatories: Version 5.1. *J. Quant. Spectrosc. Radiat. Transf.* **255**, 107217 (2020).
31. F. R. Coelho, G. C. Fraga, A. Ziemniczak, S. P. Roy, F. H. R. França, Application of the spectrally reduced integration method to solve radiative transfer in multidimensional non-homogeneous participating media. *Int. Commun. Heat Mass Transf.* **166**, 109061 (2025).
32. F. Xie, T. Ren, Z. Zhao, C. Zhao, A machine learning based line-by-line absorption coefficient model for the application of atmospheric carbon dioxide remote sensing. *J. Quant. Spectrosc. Radiat. Transf.* **296**, 108441 (2023).
33. J. Xia, Y. Xia, D. Xiao, R. Zhang, N. Yu, C. Liu, LAS-on-edge: A real-time laser absorption spectroscopic water vapor sensor on edge computing platforms. *Sens. Actuators B Chem.* **418**, 136258 (2024).
34. G. Enemali, R. M. Gibson, An FPGA-based data acquisition system with embedded processing for real-time gas sensing applications. *Appl. Sci.* **14**, 6738 (2024).
35. J. Wang, H. Wang, X. Liu, A portable laser photoacoustic methane sensor based on FPGA. *Sensors* **16**, 1551 (2016).
36. H. Yang, X. Bu, Y. Cao, Y. Song, A methane telemetry sensor based on near-infrared laser absorption spectroscopy. *Infrared Phys. Technol.* **114**, 103670 (2021).
37. M. Choi, G. Cho, S. Hwang, Y. Kim, J. Kim, J. Lee, K. Kim, H. So, S. Choi, Programmable FPGA-based TDLAS-WMS hygrometer (PFTWH) for broad-range water vapour detection: Quantitative assessment of digital lock-in amplifier (DLIA) performance via integrated WMS system (I-WMS). *Sens. Actuators B Chem.* (2025).
38. A. Sepman, Y. Ögren, Z. Qu, H. Wiinikka, F. M. Schmidt, Tunable diode laser absorption spectroscopy diagnostics of potassium, carbon monoxide, and soot in oxygen-enriched biomass combustion close to stoichiometry. *Energy Fuels* **33**, 11795–11803 (2019).
39. N. H. Ngo, D. Lisak, H. Tran, J.-M. Hartmann, An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. *J. Quant. Spectrosc. Radiat. Transf.* **129**, 89–100 (2013).
40. R. V. Kochanov, I. E. Gordon, L. S. Rothman, P. Wcislo, C. Hill, J. S. Wilzewski, HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with

spectroscopic data. *J. Quant. Spectrosc. Radiat. Transf.* **177**, 15–30 (2016).
41. L. N. Trefethen, *Approximation Theory and Approximation Practice* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2019).
42. I. Levin, M. Cuntz, Measurement of stable isotopes in carbon dioxide, methane, and water vapor, in: T. Foken (Ed.), *Springer Handbook of Atmospheric Measurements*, Springer International Publishing, Cham, 2021, pp. 509–532.
43. S. So, E. Jeng, G. Wysocki, VCSEL based Faraday rotation spectroscopy with a modulated and static magnetic field for trace molecular oxygen detection. *Appl. Phys. B* **102**, 279–291 (2011).
44. G. Taguchi, *Introduction to Quality Engineering: Designing Quality into Products and Processes* (Asian Productivity Organization, Tokyo, 1986).
45. R. L. Plackett, J. P. Burman, The design of optimum multifactorial experiments. *Biometrika* **33**, 305–325 (1946).