

UC Davis

UC Davis Previously Published Works

Title

Hollow-Core Fiber-Enhanced Raman Instrument for Standard-Free Multi-Component Gas Analysis in Industrial Process Environments

Permalink

<https://escholarship.org/uc/item/7zn032xz>

Journal

IEEE Transactions on Instrumentation and Measurement, PP(99)

ISSN

0018-9456

Authors

Knoesen, Andre
Yankelevich, Diego Ruben
Alimadad, Parssa
[et al.](#)

Publication Date

2026

DOI

10.1109/tim.2026.3731740

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Hollow-Core Fiber-Enhanced Raman Instrument for Standard-Free Multicomponent Gas Analysis in Industrial Process Environments

André Knoesen , Senior Member, IEEE, Diego Ruben Yankelevich , Parssa Alimadad ,
Dhilan Hitesh Raman Patel , James Nelson , and Roger Brett Boulton 

Abstract—Multi-component gas analysis under industrial process conditions (elevated pressure, near-saturation humidity, dynamic composition) challenges established analytical modalities: infrared absorption cannot detect homonuclear diatomics (N_2 , O_2), gas chromatography is offline, and electrochemical sensors drift under sustained humidity. This paper presents a hollow-core anti-resonant fiber-enhanced Raman spectrometer for simultaneous, standard-free detection of five process gases (N_2 , O_2 , CO_2 , ethanol vapor, H_2O) and quantification of the three that define the bulk gas balance (N_2 , O_2 , CO_2). A three-gas normalization framework anchored to atmospheric N_2 as an internal standard yields quantitative molar fractions from peak-area ratios using relative Raman cross-sections, without external calibration, pressure sensing, or absolute intensity reference. The instrument was deployed for six days on a temperature-controlled wine fermentation, in a research winery environment and powered down between daily acquisitions. Over that period the instrument maintained wavenumber stability below 0.021 cm^{-1} (relative standard deviation $\leq 0.0013\%$), signal-to-noise ratios of 101:1 to 955:1 for the major gases, limits of detection of 0.33, 0.25 and 0.26 mol% for CO_2 , N_2 and O_2 respectively at 1000 s and 515 kPa absolute, well below the 9 to 62 mol% the three species actually occupied. The oxygen-to-nitrogen response factor governing the framework was determined experimentally from ambient air. Composition is reported on a three-gas basis; CO_2 was quantified from its 1.10 mol% limit of quantification to a maximum of 52 mol%, a 47-fold range.

Index Terms—Standard-free measurement, fiber-enhanced Raman spectroscopy, gas analysis, hollow-core anti-resonant fiber, industrial process monitoring, mass balance, multi-component sensing, Raman spectrometer, ratio metrology, signal-to-noise characterization.

I. INTRODUCTION

A. Knoesen, D. R. Yankelevich, P. Alimadad, D. H. R. Patel, and J. Nelson are with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: aknoesen@ucdavis.edu).

J. Nelson is also with the Department of Viticulture and Enology, University of California, Davis, CA 95616 USA.

R. B. Boulton is with the Department of Viticulture and Enology, University of California, Davis, CA 95616 USA.

Received 13 June 2026; revised 10 August 2026; accepted 24 August 2026. Date of publication 7 September 2026; date of current version 17 September 2026.

© 2026 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This is the author's accepted manuscript. It has not been copy-edited or typeset by IEEE. The version of record is published as: A. Knoesen, D. R. Yankelevich, P. Alimadad, D. H. R. Patel, J. Nelson, and R. B. Boulton, "Hollow-core fiber-enhanced Raman instrument for standard-free multicomponent gas analysis in industrial process environments," *IEEE Trans. Instrum. Meas.*, vol. 75, 2026, Art. no. 7006218, doi: 10.1109/TIM.2026.3731740. Deposited in eScholarship under the University of California Open Access Policy.

SIMULTANEOUS quantification of multiple gas species under industrial process conditions remains a persistent measurement challenge. Bioprocess fermentation, chemical reactor monitoring, distillation, and energy applications such as biofuel production all require instruments capable of resolving several analytes concurrently at the parts-per-million to mole-percent level, while operating under elevated pressure, high humidity (often exceeding 95% relative humidity), and dynamically evolving composition. No single established gas analysis modality satisfies all of these constraints simultaneously.

Infrared absorption methods, including non-dispersive infrared, tunable diode laser absorption spectroscopy (TDLAS), and wavelength modulation spectroscopy (WMS), offer high sensitivity and have benefited from recent advances incorporating deep learning for multi-species discrimination [1]. However, the underlying selection rule, requiring a change in molecular dipole moment during vibration, fundamentally excludes homonuclear diatomic molecules such as N_2 and O_2 . This limitation is consequential. Atmospheric N_2 is the only candidate internal standard available across the broad range of process environments encountered in semi-open industrial vessels, and quantitative O_2 tracking is essential in any application where oxygen concentration is a variable. Photoacoustic spectroscopy (PAS) reaches parts-per-billion sensitivity with a compact cell and, in its fiber-enhanced form, detects trace species inside hollow-core waveguides at sub-ppm levels. Fourier transform infrared spectroscopy (FTIR) recovers the full mid-infrared band structure in a single interferogram and is a standard multi-species process tool. Fiber-enhanced *absorption* methods, including hollow-core TDLAS and hollow-core PAS, share with the present work the long path length and small mode volume that a hollow-core fiber provides. All three, however, remain absorption techniques and are bound by the same dipole-moment selection rule: none detects N_2 or O_2 at any concentration. The fiber enhancement improves the path length available to an infrared-active analyte; that enhancement does not make a homonuclear diatomic infrared-active. The distinction matters here because the internal standard on which the standard-free framework of this paper rests is the atmospheric N_2 signal itself, not merely sensitivity to that signal. Section VI-B compares detection limits across these modalities quantitatively.

Electrochemical sensors achieve low cost and rapid response for individual species but suffer cross-sensitivity in complex mixtures and exhibit baseline drift in sustained high-humidity environments [2]. Semiconductor metal-oxide sensors face similar humidity-induced drift. Gas chromatography-mass spectrometry (GC-MS) provides definitive identification

with parts-per-billion sensitivity but operates offline with cycle times of hours, precluding its use for real-time process control [3], [4]. Liquid-phase optical methods are vulnerable to matrix interference, including the strong fluorescence from anthocyanins and phenolic compounds typical of process streams in food and beverage applications [5].

Raman spectroscopy in the liquid and solid phases is a mature industrial process analytical technology. Backscatter immersion and non-contact Raman probes are routinely deployed in pharmaceutical manufacturing, polymer processing, bioprocess monitoring, and chemical synthesis [2]. Quantitative concentration estimates in these applications are typically produced by multivariate chemometric models (most commonly partial least squares regression) trained against extensive calibration sets that span the expected operating composition range.

Gas-phase Raman is a fundamentally different measurement regime. At atmospheric pressure and room temperature, gas-phase molecular densities are roughly three orders of magnitude lower than in the corresponding liquid, and the integrated Raman signal scales linearly with molecular density. Conventional backscatter probes therefore do not produce usable gas-phase signals at ambient pressure without compensation of some kind, and the available routes are few: raise the sample density, lengthen the interaction, raise the laser power, or recover the analyte from a weak signal with a calibration model trained on reference mixtures. Laboratory work has typically taken the middle two, with extended free-space cavities, very high laser power, or both [6], while the commercial process analyzers described below take the last. The present work takes the first two and declines the fourth. Despite this sensitivity limitation, gas-phase Raman retains key fundamental advantages over both infrared absorption and liquid-phase methods for process gas monitoring: the technique directly detects homonuclear diatomics inaccessible to infrared selection rules, exhibits narrow Q-branch features amenable to simultaneous multi-species deconvolution, and operates in the visible region where water has negligible spectral interference [2], [6].

A distinction of scope should be drawn at this point. Because every Raman-active species contributes its own characteristic bands under a single excitation source, the method is well suited to establishing the simultaneous presence of several species from one measurement and to following how their relative abundances evolve.

That second capability is what motivates the present work. The gases in a fermentation headspace are coupled, since CO_2 and ethanol production follow the same metabolism, so what carries process information is their joint trajectory rather than any single value read once. Offline methods identify these species with far greater sensitivity, but a sample carried to a laboratory returns a point, and an hour-scale cycle limits the structure that can be recovered from a sequence of such points. Following the whole set in place, at a cadence set by the process rather than by a laboratory queue, is the capability this instrument is built to provide, and is the route by which headspace measurements become useful for understanding and/or controlling a fermentation rather than merely auditing

one.

One further property follows from measuring the headspace rather than the liquid. The headspace is fed by gas evolved throughout the liquid, so a single acquisition reports a vessel-averaged composition rather than the composition at a probe tip.

Converting band intensities into concentrations is a separate problem: Raman intensities are inherently relative and depend on instrument and configuration factors, so quantitative Raman analysis requires either species-specific calibration or an internal standard [7]. The chemometric practice described above takes the first route. The framework developed in this paper takes the second, using atmospheric N_2 as the internal standard. What the framework reports is therefore a set of molar fractions on a three-gas basis, obtained without external calibration, rather than externally calibrated absolute concentrations of a single targeted analyte; for the latter, at the lowest detection limits, absorption methods remain preferable wherever the analyte is infrared-active.

Commercial gas-phase Raman analyzers have been deployed in petrochemical and refining applications for some years, most notably the Endress and Hauser Raman Rxn5™ process analyzer (Kaiser Optical Systems technology), which is applied to ammonia production, hydrogen purity, and refinery syngas streams and which can detect homonuclear diatomics including N_2 and H_2 . These deployments use conventional free-space probes with single-stream sampling and rely on species-specific chemometric calibration against reference mixtures. More recently, the same probe family has been extended to bioreactor off-gas with the Raman AirHead™ probe, demonstrated for single-species (CO_2) monitoring in mammalian cell culture [8]. Existing commercial gas-phase Raman therefore establishes industrial feasibility but does not exploit hollow-core fiber enhancement, the simultaneous multi-component capability fundamental to Raman, or a standard-free quantification framework.

Hollow-core anti-resonant fiber (HC-ARF) addresses the gas-phase sensitivity gap by simultaneously providing tight optical confinement within a small core (typical mode diameters of 30 μm), extended interaction length over meter-scale fiber sections, and the ability to operate the sample gas at elevated pressure, which directly compensates for the lower molecular density [9]–[11]. Of the four routes named above, the fiber supplies the second and the pressurization the first, while the confinement recovers much of what the third would otherwise demand of the laser. Supplying three by construction is what removes the need for the fourth. The combined effect yields several orders of magnitude signal enhancement relative to free-space gas configurations [9], [10], [12]. Fiber-enhanced Raman spectroscopy (FERS) has been demonstrated in laboratory settings for trace gas detection at tens of parts per million, including for species relevant to atmospheric chemistry, energy storage, and transformer monitoring [13]–[17].

Prior FERS demonstrations have largely remained in controlled laboratory environments, and have not combined operation under elevated pressure and near-saturation humidity with repeated deployment on a working process vessel and a

standard-free reporting basis.

This paper presents an HC-ARF-enhanced Raman instrument designed to address these combined constraints. The contributions are threefold. First, an instrument architecture combining forward-scattering FERS with the thermal and gas-handling engineering needed to operate at 515 kPa absolute on a sample stream drawn at near-saturation humidity, deployable in an industrial enclosure. Second, a three-gas normalization framework using atmospheric N_2 as an internal standard and literature Raman cross-sections for N_2 , O_2 , and CO_2 to yield quantitative molar fractions from peak area ratios, eliminating the need for external concentration standards, pressure sensors, or absolute intensity calibration. Third, metrological characterization of this instrument as built and deployed, comprising wavenumber stability, signal-to-noise ratios for five concurrent species, concentration-based limits of detection and quantification for the three quantified gases, and an experimental determination of the oxygen response factor from ambient air. These limits characterize an instrument configured for composition monitoring and are not a bound on fiber-enhanced Raman itself: an implementation of the same technique optimized for sensitivity reaches the parts-per-million decade, and Section VI-B sets the two side by side and accounts for the gap.

For field validation, the instrument was deployed on a temperature-controlled red wine fermentation. The fermentor was thermostatically controlled at $22 \pm 0.5^\circ\text{C}$, but the instrument itself operated in the uncontrolled winery production environment, with ambient temperature and humidity around the enclosure varying freely; over the six-day campaign the fermentation hall temperature ranged from $27\text{--}34^\circ\text{C}$ and relative humidity from 27–48%. This choice of validation case was deliberate because the deployment concurrently exposes the instrument to the full set of target conditions: elevated headspace pressure, near-saturation humidity inside the gas path, dynamic composition in which the major component (CO_2) rises from an atmospheric starting condition to more than 50 mole percent and falls again, ethanol vapor evolution corresponding to liquid-phase concentrations of 0 to 4 mol%, once-daily unattended measurements over a six-day period with the instrument powered down between acquisitions, and uncontrolled ambient conditions surrounding the optical hardware. The field deployment is presented as validation of instrument capability rather than as a study of the biological process; chemical and biological interpretation is deliberately deferred to other work.

The remainder of the paper is organized as follows. Section II describes the instrument design. Section III details the measurement methodology, including the three-gas normalization framework. Section IV presents the metrological characterization. Section V reports the six-day field deployment. Section VI compares the instrument to alternative analytical modalities, and Section VII concludes.

II. INSTRUMENT DESIGN

The instrument is organized around five subsystems: a 532 nm forward-scattering optical engine, a hollow-core anti-resonant fiber acting as the gas interaction cell, a gas handling

and pressurization subsystem, an actively stabilized thermal subsystem for the fiber, and an industrial enclosure that mechanically and thermally isolates the optical hardware from the uncontrolled ambient of a production environment. A schematic of the overall configuration is shown in Fig. 1.

A. Optical Configuration

The optical engine implements forward-scattering FERS. For a fiber shorter than its attenuation length the collected signal is comparable in both directions, but the forward geometry rejects the fluorescence and the cladding and inlet-facet Raman background that dominate back-scattering, giving a flatter baseline and roughly four-fold higher signal-to-noise in an equivalent hollow-core configuration [17]. Excitation is a 200 mW continuous-wave 532 nm diode-pumped solid-state laser (MKS EXLSR-532-200, < 10 MHz linewidth, < 50 MHz/ $^\circ\text{C}$ frequency drift, fan-cooled), coupled into the fiber with a 15 mm focal-length anti-reflection-coated lens. At the fiber output, the small fiber numerical aperture (0.014) produces a naturally collimated beam, eliminating the need for a collimating lens between fiber and spectrometer and reducing alignment sensitivity. The 532 nm excitation is separated from the forward-scattered Raman signal by an interference longpass edge filter (Newport 10CGA-610) used at 45° , followed by a 532 nm single-notch filter (Semrock StopLineTM) at normal incidence. Operating the edge filter off normal incidence shifts its transmission edge toward shorter wavelengths, which places the edge below the shortest Stokes band of interest while leaving the filter reflective at the laser line. The 532 nm beam reflected by the edge filter is monitored by a calibrated silicon power meter (Ophir Centauri PD300) to track laser-to-fiber coupling stability.

The Raman signal is dispersed by a high-throughput $f/1.3$ spectrometer (Wasatch Photonics WP-532-C-EXR-IC, NA = 0.36, $10\times$ the light gathering of an $f/4$ spectrometer), free-space coupled through a barrel assembly with integrated collimator and 50 μm fixed entrance slit. A volume phase holographic transmission grating and longpass filter provide $150\text{--}4000\text{ cm}^{-1}$ coverage at 16 cm^{-1} FWHM resolution. The upper limit is a deliberate choice rather than a stock specification. Raman spectrographs are commonly specified only to the top of the C–H stretching region near 3200 cm^{-1} , which would exclude the water monomer at 3657 cm^{-1} and with it the ability to follow water at all. Extending coverage to 4000 cm^{-1} admits that band, and also opens the Raman-inactive interval at $3718\text{--}3918\text{ cm}^{-1}$ that serves as the noise floor throughout Section IV-B. The detector is a thermoelectrically cooled silicon CCD (-13°C). The spectrometer's factory relative-intensity correction was not applied. Each spectrum uses 1000 s integration, sized for SNR margin (Section IV-B) rather than to support a specific sampling cadence.

The whole optical system is housed in a fiberglass-reinforced polycarbonate hinged enclosure with all optical components on an opto-mechanical cage system. The fiber input and output couplers are mechanically isolated from the breadboard so that pressurization transients are not transmitted into the optical alignment. This enclosure, together with the actively stabilized fiber sub-enclosure described in Section II-D,

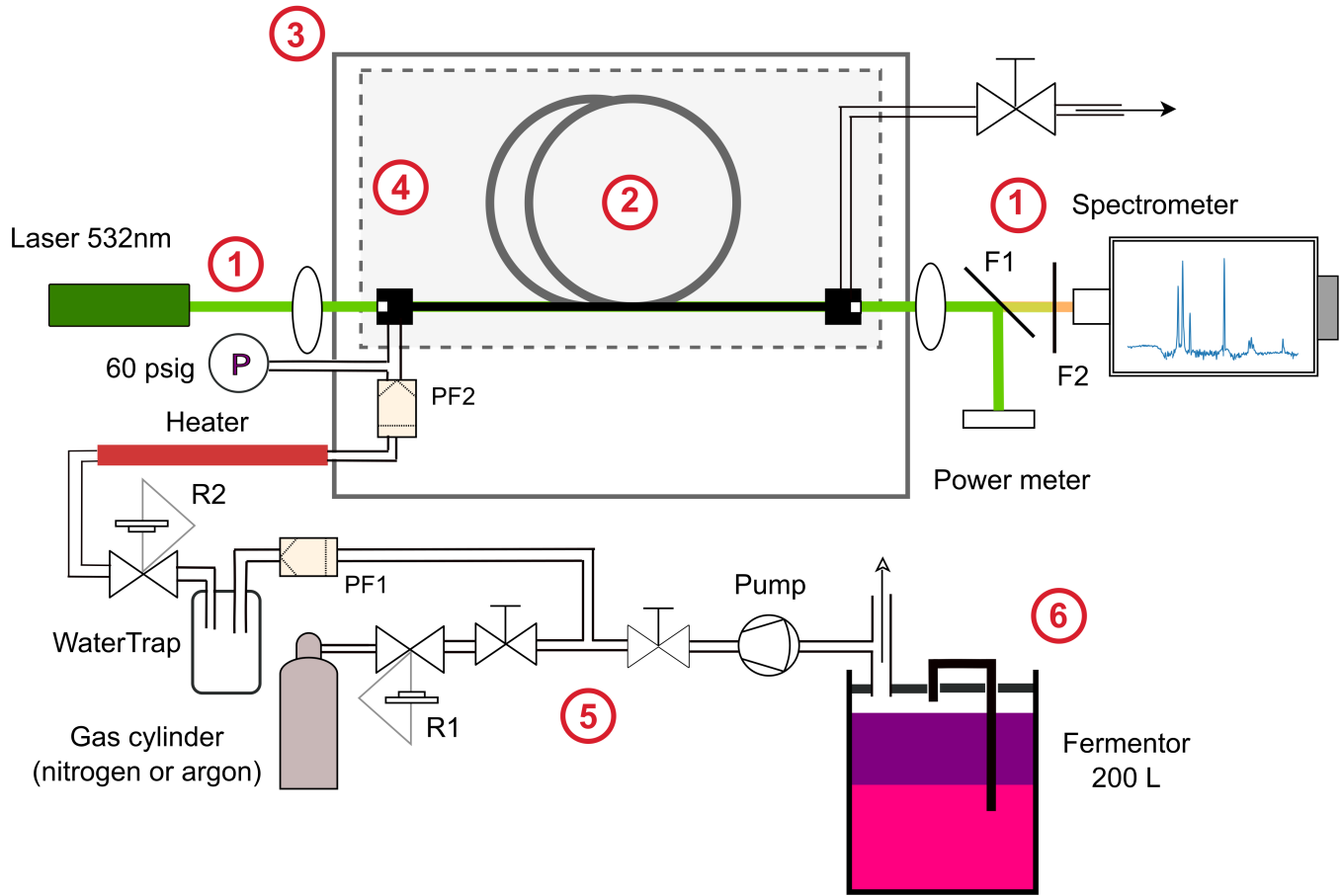


Fig. 1. Schematic of the FERS instrument. Numbered subsystems: **1** optical engine (532 nm); **2** HC-ARF gas cell; **3** industrial enclosure; **4** thermal sub-enclosure at 40 °C; **5** gas handling and pressurization; **6** fermentor. Nitrogen (for purge), argon (for the instrument-response reference) or headspace gas can be routed through the hollow-core fiber; the cylinder connection takes one bottle at a time. Headspace gas is drawn by the pump through PF1, a 5 μm particulate prefilter, and the water trap, which is unheated and fixes the water content delivered downstream (Section II-D); the heated line and PF2, a 0.1 μm stainless steel particle filter, follow immediately before the fiber. R1 and R2 are pressure regulators and P a pressure gauge. The excitation source is a continuous-wave frequency-doubled diode-pumped solid-state laser at 532 nm. The forward-scattered signal passes two filters in series: F1, a longpass edge filter at 45°, which reflects most of the 532 nm excitation to the power meter, and F2, a 532 nm notch filter at normal incidence, which removes the residual. The signal is then dispersed by a high-throughput $f/1.3$ spectrometer.

decouples the optical system from the production-floor ambient.

B. Hollow-Core Anti-Resonant Fiber

The fiber is a hollow-core anti-resonant fiber with hypocycloidal core geometry (GLOphotonics PMC 500–700; cross-section in Fig. 2). The 30 μm inner diameter and 2 m active length yield a 1.4 μL interaction volume. Light and sample gas co-propagate through the central channel; the microstructured silica cladding confines the optical mode by anti-resonant guidance while allowing the channel to be filled and pressurized. Transmission spans 500–700 nm with straight-section loss below 0.05 dB/m at 532 nm. Coiling the fiber to a 20 cm diameter to fit inside the thermally stabilized sub-enclosure (Section II-D) adds approximately 3 dB of bending loss, an accepted tradeoff for uniform thermal control along the interaction length. Light and gas are coupled into and out of

the fiber via custom stainless steel couplers rated to 690 kPa gauge (100 psig), that is 791 kPa absolute against the 515 kPa the fiber is operated at.

C. Gas Handling and Pressurization

Headspace gas is extracted through approximately 2 m of polyurethane tubing by an oil-less diaphragm vacuum pump (KNF NPK09) under programmable solenoid valve control, then passed through a two-stage particulate filter (5 μm porous polypropylene prefilter, 0.1 μm stainless steel particle filter) chemically inert to the target analytes. Although polypropylene can absorb some volatile organic compounds [18], the minimal filter mass and operational flow rate render gas-phase partitioning negligible for the species reported here.

The fiber core is pressurized to 515 kPa absolute (60 psig) during measurement. Elevated pressure increases molecular density linearly, directly enhancing Raman signal intensity and

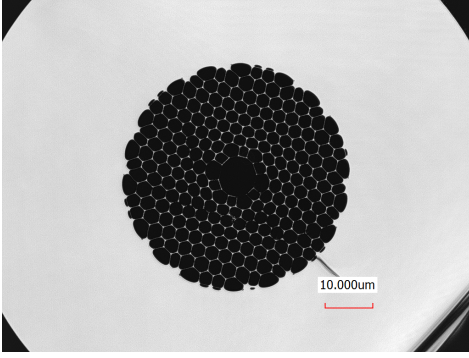


Fig. 2. Cross-section of the hollow-core anti-resonant fiber (PMC 500–700). Courtesy of GLOphotonics.

partially compensating for the gas-phase density gap discussed in Section III. The fill time from atmospheric to operating pressure is approximately 15 s [17], and pressure stability over the 15 min acquisition window is ± 14 kPa ($\pm 2.7\%$). Between measurements, the system is purged with nitrogen at 239–377 kPa absolute (20–40 psig) for 15 min. Because the framework operates in ratio mode (Section III-C), pressure fluctuations cancel and have no measurable effect on reported molar fractions.

D. Thermal Management

The fiber and a resistive heater occupy a polycarbonate sub-enclosure with forced-air circulation, actively stabilized at 40°C and maintaining $\pm 0.06^\circ\text{C}$ across the 20 cm coiled fiber.

The gas path from the pump to that stage runs: pump, particulate prefilter and water trap, coiled stainless-steel heated pipe at $38\text{--}49^\circ\text{C}$, and a $0.1\text{ }\mu\text{m}$ particle filter immediately before the fiber. Only the coiled pipe and the fiber sub-enclosure are heated; the trap and the tubing entering the main enclosure sit at room temperature.

Water governs the design of the gas path. The headspace above a fermenting must, the juice and grape skins from which the wine is made, is close to saturated, and not incidentally so: the relative humidity over the liquid is the water activity of that liquid, $a_w \simeq 0.95$ for a must at this stage, so the sampled gas arrives at roughly 95% RH whatever temperature the vessel is held at. Water is consequently the one species that is always near its condensation point somewhere in the instrument.

Compression sharpens the problem rather than easing it. At 515 kPa every partial pressure is five times its headspace value, so a dew point evaluated at atmospheric pressure does not describe the gas the fiber sees: gas taken from a 30°C headspace at 95% RH would, on compression alone, reach a dew point near 61°C , above every surface in the instrument. Liquid water in a hollow-core fiber would end the measurement outright, so the 40°C setpoint and the conditioning train ahead of it both follow from this single property of the sample.

Condensation is therefore located rather than avoided. The trap is the last point at ambient temperature under full pressure, so the gas leaving the trap is saturated at that temperature whatever happened upstream; over the $27\text{--}34^\circ\text{C}$ ambient

recorded during the campaign this fixes the delivered water content at 0.6–1.0 mol% and the corresponding dew point at $26\text{--}33^\circ\text{C}$. The unheated section that follows sits at that dew point rather than below, so no further condensation is driven, and active heating is applied only where liquid water would be destructive: the fiber and its couplers are held 7 to 14°C above the delivered dew point, a relative humidity inside the fiber of 45–68%. Because the delivered composition is set by the coldest point at full pressure rather than by the headspace, the delivered composition is insensitive to the headspace temperature.

III. MEASUREMENT METHODOLOGY

The measurement methodology comprises three layers: spectral preprocessing and peak fitting, wavenumber calibration, and a three-gas normalization framework that converts integrated peak areas into molar fractions. The framework is the core contribution. Section III-C5 collects what the three layers together do and do not require, and fixes the sense in which the instrument is called standard-free.

A. Spectral Preprocessing and Peak Fitting

Prior to deployment, the additive instrument background was characterized by filling the fiber with argon and acquiring a reference spectrum. Argon is a noble gas with no first-order Raman activity, so the resulting trace carries no analyte signal and records only sample-independent contributions: dark current, residual silica Raman from the fiber wall, residual fluorescence and stray light. The trace was acquired once and reused throughout the deployment. It is subtracted, not divided. With no Raman signal present the trace reports nothing about the wavelength-dependent efficiency with which the chain converts a scattered photon into a count, since recovering that would require a source of known spectral radiance. That efficiency is addressed separately in Section IV-C. Baselines were corrected using asymmetric least squares (ALS) with smoothing parameter $\lambda = 10^6\text{--}10^9$ and asymmetry parameter $p = 0.001$ [19].

Spectral peaks were fit with Gaussian lineshapes for an instrumental rather than molecular reason: gas-phase natural linewidths from Doppler and collisional broadening are below 1 cm^{-1} , well under the 16 cm^{-1} monochromator resolution (Section II-A). The observed lineshape is therefore dominated by the spectrometer's Gaussian point-spread function. Multi-component Gaussian decomposition with Bayesian Information Criterion component selection was applied to spectrally congested regions, including the CO_2 Fermi doublet (12 components, $1200\text{--}1500\text{ cm}^{-1}$) and the ethanol C–O/C–C manifold (6 components, $980\text{--}1160\text{ cm}^{-1}$). Non-linear least-squares optimization used the Levenberg–Marquardt algorithm. Fit quality was assessed using the coefficient of determination ($R^2 > 0.98$ for all major species), the Durbin–Watson statistic on residuals, and a signal-to-noise ratio computed from the fitted peak amplitude relative to the baseline-corrected noise in a nearby Raman-inactive region.

Gas-phase species were identified by their characteristic Raman vibrational frequencies, shown for a representative

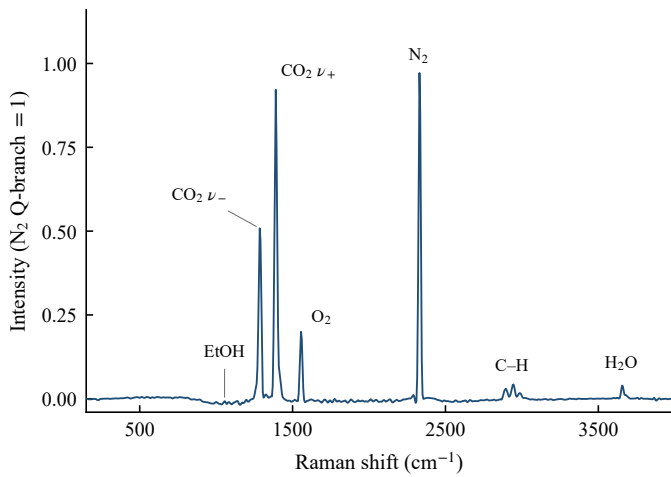


Fig. 3. Representative headspace Raman spectrum acquired on Day 2 of the field deployment, normalized so that the N_2 Q-branch has unit amplitude. Labelled bands: N_2 (2331 cm^{-1}), O_2 (1556), the two components of the CO_2 Fermi doublet, ν_- (1285) and ν_+ (1388), the ethanol C–O stretch (1056) with the C–H manifold near 2935 , and the water vapor monomer (3657). The spectrum is read in Section V-B.

spectrum in Fig. 3: N_2 Q-branch at 2331 cm^{-1} , O_2 Q-branch at 1556 cm^{-1} , CO_2 Fermi doublet at 1285 and 1388 cm^{-1} , ethanol C–O skeletal stretch at 1056 cm^{-1} , and water monomer O–H stretch at 3657 cm^{-1} . Assignments were verified against HITRAN [20] and Long [21]. The ethanol C–O stretch at 1056 cm^{-1} was preferred over the C–H (2800 – 3000 cm^{-1} , congested) and O–H regions (3600 – 3700 cm^{-1} , overlapping water vapor) because the C–O mode is internal to the molecular skeleton and spectrally isolated. Ethanol has no Raman fundamental between approximately 1500 and 2870 cm^{-1} , so ethanol does not contribute at the N_2 Q-branch used as the internal standard; the nearest CO_2 fundamental to that position, the antisymmetric stretch at 2349 cm^{-1} , is Raman-inactive by mutual exclusion because the molecule is centrosymmetric.

B. Wavenumber Calibration

Wavenumber calibration uses a two-tier approach combining the single-time argon reference described above with daily internal frequency standards drawn from the sample itself. The atmospheric N_2 Q-branch at 2331 cm^{-1} and O_2 Q-branch at 1556 cm^{-1} are present in every headspace measurement due to the semi-open vessel geometry and are spectrally isolated from analyte features. Their fitted peak positions were used to apply a linear frequency correction to each spectrum, compensating for any spectrometer thermal drift over the deployment. The metrological performance of this approach is reported in Section IV-A. The reference peaks are therefore drawn from the sample itself, and no wavenumber standard is introduced during deployment.

C. Three-Gas Normalization Framework

The framework converts integrated peak areas to molar fractions using atmospheric N_2 as an internal standard and relative Raman cross-sections for N_2 , O_2 , and CO_2 (Table I).

N_2 is unity by definition and the CO_2 factor is tabulated; the oxygen factor is measured on this instrument from ambient air (Section IV-C).

The normalization is defined over those three species and not over the five the instrument detects. The reason is quantitative and is set out here, ahead of the definition that encodes it, because it is the principal assumption on which the framework rests.

In wine fermentations the ethanol and water vapor are each known to contribute less than 2 mol% to the overall headspace composition [22], and every remaining constituent, the aroma-active esters, aldehydes and higher alcohols among them, lies orders of magnitude below even that [3], [4]. A vapor-liquid equilibrium calculation for this fermentation confirms the ethanol figure and puts water slightly higher. Headspace water is set by the water activity of the must, $a_w \simeq 0.95$ at this stage (Section II-D), together with the saturation vapor pressure at the vessel temperature: $p_{H_2O} = a_w p^{\text{sat}}(T)$. At the 22°C at which the fermentor was thermostatted this gives 2.5 mol% of a headspace near atmospheric pressure, and water falls below 2 mol% only in vessels held under 18.5°C . The headspace gas was not itself instrumented, and need not sit at the liquid setpoint: the gas stands above the cap, the most metabolically active part of the vessel, in a hall that ran 27 to 34°C . The semi-open vent limits the excess, since evolving CO_2 sweeps the space continuously with gas that has just left the liquid. Resolving the point is not required here. A 30°C headspace would carry 4.0 mol% water rather than 2.5, and would still deliver the same composition to the fiber, which is set by the coldest point at full pressure (Section II-D). The literature value therefore holds for the cooler fermentations typical of white wine, and is modestly exceeded by a red fermented at 22°C or above. Ethanol at the same temperature is 0.26 mol%, the dilute solution being close to ideal at fermentation temperature; the large activity coefficient often quoted for ethanol and water applies to boiling mixtures near 90 to 100°C and not to a vessel held at 22°C .

What the normalization consumes is not the raw headspace but the conditioned gas, and there the margin widens. N_2 , O_2 and CO_2 traverse the conditioning train unchanged. Water and ethanol do not: the compression and room-temperature trap of Section II-D remove most of the water and carry dissolved ethanol out with the condensate, which is the accepted cost of keeping liquid water out of the fiber. Where the fiber sees the gas, water is 0.6 to 1.0 mol% and ethanol 0.12 to 0.18 mol%, together a fraction $f \leq 0.012$ of the delivered gas. The three normalized species are therefore at least 98.8 mol% of what is actually measured. Carrying the boiling-mixture activity coefficient instead would raise f to 0.02, which changes no conclusion drawn here.

Two things follow, and together they justify the basis. The excluded species are small enough to bound: the resulting bias is one-sided, makes every reported fraction slightly too large, and is at most 1.2%, derived in Section IV-D. And a correction of that size is smaller than the accuracy of the quantity it would correct, the $\pm 8\%$ of Section III-D. Carrying ethanol and water in the sum would introduce two further relative cross-sections, less securely established than those of the three

retained species (Table I), in order to make a change below the level at which any fraction is reported. The basis is thus at once the quantitatively dominant choice and the metrologically defensible one. All five species are detected (Section V-B); three are quantified.

The choice of N_2 as the internal standard is a statement about the process as much as about the spectroscopy. Fermentation neither produces nor consumes molecular nitrogen: the N_2 in the headspace is atmospheric nitrogen admitted with the vessel air, and the metabolism leaves that inventory untouched. The distinction is worth drawing, because nitrogen nutrition matters a great deal to a fermentation; what the yeast assimilate, however, is nitrogen bound in must-borne compounds, not headspace N_2 , which no wine yeast can fix. Carbon dioxide and oxygen are the opposite case, CO_2 being a metabolic product and O_2 a chemical reactant involved in autoxidation, so both are process variables in a way nitrogen is not. An internal standard must be a species the process leaves alone, and in a wine fermentation headspace nitrogen is the only candidate that is simultaneously unreactive, strongly Raman-active, and present in every acquisition.

Unreactive is not the same as constant. As CO_2 accumulates, the gas already in the headspace is displaced and χ_{N_2} falls, reaching 0.39 at the CO_2 maximum and recovering to 0.62 as CO_2 production declines (Section V-C). That excursion is dilution of a conserved species rather than production or loss, which is what allows the nitrogen channel to be read as a reference while the CO_2 channel is read as a measurement.

A condition on the application follows. The framework requires a headspace in which nitrogen is a spectator and enters from the ambient. A nitrogen-purged or hermetically sealed vessel, or a process that fixes or evolves nitrogen, would not satisfy that requirement. The restriction is mild in practice, since nitrogen is 78% of the atmosphere and any vessel with passive or active atmospheric exchange carries nitrogen in abundance, but the restriction is real. Section VI-A returns to the processes that do not satisfy the condition, and to the injected inert tracer that restores the reference channel when nitrogen is unavailable or is itself a participant.

1) *Signal model:* For each species i , the integrated Raman peak area A_i is, to first order in optically thin propagation,

$$A_i = \kappa n_i \sigma_i P_{\text{laser}} \eta L \quad (1)$$

where n_i is the molecular number density of species i , σ_i is its differential Raman scattering cross-section, P_{laser} is the laser power coupled into the fiber, η is the overall optical throughput (fiber coupling, transmission, filter pass-band, detector responsivity), L is the fiber interaction length, and κ is a geometric collection factor. One species needs its area defined before the symbol is used further. For CO_2 , A_{CO_2} is the integrated area of the whole Fermi manifold, not of the dyad alone: the resonant components at 1285 and 1388 cm^{-1} together with the satellite intensity flanking them on either side. Every molar fraction reported in this paper is formed from that manifold area.

2) *Ratio cancellation:* The ratio of two species' integrated peak areas in a single acquisition is

$$\frac{A_i}{A_j} = \frac{n_i}{n_j} \cdot \frac{\sigma_i}{\sigma_j}. \quad (2)$$

The instrument-dependent and configuration-dependent factors κ , P_{laser} and L cancel exactly, because every species in the sample sees the same laser, the same fiber and the same collection optics. The collection and detection efficiency η cancels only in its species-independent part. Each species scatters at its own wavelength, and η varies across the 563 to 660 nm interval the reported bands span, so the wavelength dependence survives the ratio and is carried in the effective k_i defined in Section III-C. Section IV-C measures that surviving factor for oxygen and estimates its size for CO_2 . Total pressure also cancels under ideal-gas behavior, since pressure scales every n_i equally. This is the structural reason the framework is insensitive to laser power fluctuations, fiber coupling drift, fiber bending loss within the operating range, and the pressure stability budget quoted in Section II-C.

3) *Relative cross-sections:* Defining the relative Raman cross-section $k_i \equiv \sigma_i/\sigma_{N_2}$, the molar (or molecular) ratio relative to N_2 becomes

$$\frac{n_i}{n_{N_2}} = \frac{1}{k_i} \cdot \frac{A_i}{A_{N_2}}. \quad (3)$$

The *tabulated* cross-sections used here (Table I) are consistent with the compilations of Schr  tter and Kl  ckner [23] and Long [21], interpolated to 532 nm from argon-ion laser wavelengths (488.0 and 514.5 nm). The wavelength dependence is weak ($< 5\%$) for non-resonant gases in the visible region. The oxygen factor is not drawn from those compilations. It is measured on this instrument from ambient air (Section IV-C) and differs from the tabulated value by 11%, for reasons set out there.

The flanking intensity is retained rather than discarded, and the reason is worth stating because the choice is not neutral. The data require it: two Gaussians centered on the dyad cannot reproduce the measured intensity immediately outside the pair, which persists there on every day of the deployment. That intensity belongs to CO_2 , hot bands originating in the ν_2 bending state being expected in just those positions, and its magnitude is consistent with such an origin, amounting to 11% of the manifold area against a Boltzmann population of 9% outside the bending ground state at the 40 °C stage temperature (Section II-D). The components are not resolved at the 16 cm^{-1} instrument linewidth, so this is an accounting decision about which intensity belongs to CO_2 rather than a spectroscopic assignment. It is reported because it is not small: forming A_{CO_2} from the dyad alone would lower χ_{CO_2} by 5.5%.

4) *Molar fractions:* Defining the cross-section-weighted total intensity $I_{\text{Total}} = \sum_i A_i/k_i$ over the three-gas basis (N_2 , O_2 , CO_2), the molar fraction of each species follows directly:

$$\chi_i = \frac{A_i/k_i}{I_{\text{Total}}}, \quad I_{\text{Total}} = \sum_{i \in \{N_2, O_2, CO_2\}} \frac{A_i}{k_i}. \quad (4)$$

By construction, $\sum_i \chi_i = 1$ for the three normalized species. This identity is a property of the definition in (4), in which

TABLE I
RELATIVE RAMAN CROSS-SECTIONS k_i AT 532 NM (N_2 Q-BRANCH = 1.00).

Species	Band(s) / cm^{-1}	Description	k_i
N_2	2331	Q-branch	1.00
O_2	1556	Q-branch	1.086 ± 0.019
CO_2	1285 + 1388	Fermi doublet (summed)	1.45

N_2 is unity by normalization and CO_2 is the tabulated literature value. k_{O_2} is measured on this instrument from ambient air (Section IV-C); the measured value is an effective factor including detection-chain response.

each fraction is a share of a sum over the same three terms. The identity holds for any input areas and any cross-section set, and therefore carries no information about measurement accuracy; the sum serves in this paper only as a statement about the reporting basis. Section IV-D states what that basis excludes and bounds the resulting error.

Ethanol and water vapor are tracked on the same intensity-normalized basis, but neither is quantified.

5) *Framework Requirements and Standard-Free Definition:* The framework requires only (i) integrated peak areas measured in a single acquisition and (ii) values for the relative cross-sections k_i , taken from tabulation or, for oxygen, from ambient air (Section IV-C), whose composition is itself a tabulated constant and requires no reference material to be introduced. No absolute intensity calibration of the spectrometer is required, nor an external pressure sensor, real-time laser power measurement, knowledge of the fiber interaction length, or external concentration standards introduced during deployment. The wavenumber axis is anchored internally on the same principle, to the N_2 and O_2 Q-branches present in every spectrum (Section III-B), so no neon lamp, polystyrene reference or HeNe line is needed either. Both axes of the measurement, composition and wavenumber, are therefore referred to the sample itself.

Not required is not the same as not present, and several of these quantities are in fact measured. Pressure is transduced, the diverted beam is monitored by a calibrated power meter (Section II-A), and the fiber length is known. All three cancel from (4) by the ratio argument of Section III-C, so they inform the stability budget and confirm that coupling did not drift across power cycles, without entering any reported molar fraction. The argon reference is a separate case and should not be counted with them: its trace is subtracted as an instrument background (Section III-A) and therefore does act on the fitted areas. Argon is not a concentration standard, having no first-order Raman activity and carrying no composition information, and one trace acquired before deployment is reused throughout.

Operationally, this is what “standard-free” means in this paper: no reference material of known composition is introduced into the instrument at any point, and no chemometric model trained on labeled spectra is used. The instrument is operated and its outputs interpreted from internal measurements and tabulated physical constants alone. One instrument-specific quantity, the oxygen response factor, is determined on the instrument from ambient air (Section IV-C). The atmosphere

is not a procured standard and its composition is a tabulated constant, but that determination is a single-point response calibration, and describing the method as calibration-free would overstate what is claimed here.

D. Uncertainty Propagation

Uncertainties in the reported molar fractions were propagated from the Gaussian fit uncertainty on each integrated area, the literature uncertainty on the relative cross-sections (taken as $\delta k_i/k_i \approx 5\%$ [21], [23]), and the resulting uncertainty in the weighted sum:

$$\frac{\delta \chi_i}{\chi_i} = \sqrt{\left(\frac{\delta A_i}{A_i}\right)^2 + \left(\frac{\delta k_i}{k_i}\right)^2 + \left(\frac{\delta I_{\text{Total}}}{I_{\text{Total}}}\right)^2}. \quad (5)$$

For the dominant species at the highest signal levels recorded in Section V, the typical relative uncertainty in χ_i is $\pm 8\%$, which applies to a single reported molar fraction. The largest contributor is the $\pm 5\%$ literature uncertainty on the relative cross-sections, a property of the tabulated constants rather than of the instrument; the Gaussian-fit term is $\pm 4\%$ for the major species and larger for ethanol and the water monomer. Accuracy is therefore set by the available cross-section data, not by the instrument’s noise floor.

IV. METROLOGICAL CHARACTERIZATION

This section reports the metrological performance of the instrument characterized across six daily acquisitions spanning six days, each conducted under uncontrolled production-floor ambient conditions (Section II-A). Six performance categories are considered: wavenumber stability and resolution, signal-to-noise ratio for each measured species, experimental determination of the oxygen response factor from ambient air, limits of detection and quantification, robustness to pressure and coupling fluctuations, and temporal response. Summary values are collected in Table II.

A. Wavenumber Stability and Resolution

Wavenumber stability was assessed by fitting Gaussian profiles to the atmospheric N_2 and O_2 Q-branches in each of the six daily spectra, after the spectrometer’s internal wavenumber axis had been applied but before any post-hoc calibration correction. The fitted peak positions and full-width-at-half-maximum (FWHM) values across the deployment were:

$$\begin{aligned} \nu_{N_2} &= 2330.970 \pm 0.007 \text{ cm}^{-1} \quad (\text{RSD} = 0.0003\%), \\ \nu_{O_2} &= 1555.855 \pm 0.021 \text{ cm}^{-1} \quad (\text{RSD} = 0.0013\%), \\ W_{N_2} &= 16.735 \pm 0.095 \text{ cm}^{-1} \quad (\text{RSD} = 0.57\%), \\ W_{O_2} &= 18.669 \pm 0.142 \text{ cm}^{-1} \quad (\text{RSD} = 0.76\%), \end{aligned}$$

where uncertainties are sample standard deviations over $n = 6$ daily measurements. The peak position stability of 0.021 cm^{-1} or better is within 0.13% of the instrument resolution and demonstrates that the spectrometer wavelength scale, the laser frequency, and the fiber output coupling collectively remain stable to a small fraction of the resolution element. The sub-1% linewidth variations confirm that the spectrometer’s point-spread function is stable as well, so that integrated peak

areas are not contaminated by drifting lineshape. These results were obtained with the instrument's outer ambient temperature and humidity uncontrolled, and constitute evidence that the actively stabilized fiber sub-enclosure and the small fiber numerical aperture together decouple the wavelength scale from production-floor ambient. The frequency-axis stability also justifies using the N_2 and O_2 atmospheric peaks as internal wavenumber references during deployment (Section III-B).

B. Signal-to-Noise Characterization

Signal-to-noise ratios were computed as the fitted Gaussian amplitude of a species divided by the standard deviation of the baseline-corrected residuals in a Raman-inactive spectral region. A single noise floor was used for all species in a given spectrum, taken from the interval $3718\text{--}3918\text{ cm}^{-1}$. This region is free of Raman features for all five species observed here, lies above the water hydrogen-bonded cluster envelope, and provides 40 points on the 5.0 cm^{-1} measurement grid, sufficient for a stable variance estimate. Baseline removal used asymmetric least squares ($\lambda = 10^9$, $p = 0.001$), identical to the preprocessing applied before peak fitting.

Two integration times were used: 500 s on Day 1 and 1000 s on Days 2–6 (Table II). The two are treated as follows. A molar fraction is a ratio formed within one spectrum, so integration time cancels from (4) and the composition results need no adjustment. A signal-to-noise ratio does not cancel, scaling instead as the square root of integration time, so Day 1 is placed on the 1000 s basis by the factor $\sqrt{2}$ wherever days are compared. The representative and limiting values below are drawn from 1000 s spectra, and the ranges are unaffected by the choice either way.

The CO_2 quantity behind the signal-to-noise ratio is not the one behind the molar fraction, and the difference is deliberate on both counts. A molar fraction is built from integrated *areas* over the *whole* Fermi manifold (Section III-C). A signal-to-noise ratio is defined on fitted *amplitudes*, and uses the two resonant *dyad* components only, their amplitudes summed and the noise on that sum propagated as $\sqrt{2}\sigma$ for two independently fitted peaks. The flanking components are left out because including them would raise the numerator and the noise together, to no net effect.

On the representative Day 2 spectrum ($\sigma = 89.8$ counts) the signal-to-noise ratios were:

$$\begin{aligned} \text{SNR}(N_2) &= 469:1, \\ \text{SNR}(CO_2) &= 472:1, \\ \text{SNR}(O_2) &= 101:1. \end{aligned}$$

Across the six deployment days the ratios varied over 469–955 (N_2), 186–735 (CO_2) and 101–201 (O_2). This variation reflects two contributions of comparable size: the changing abundance of each species, largest for CO_2 , and a factor of 1.7 variation in the measured noise floor itself (54.2 to 89.8 counts among the 1000 s spectra). The spread is therefore not attributable to abundance alone.

The water vapor monomer at 3657 cm^{-1} is strong and well isolated: the ratio is 17.9:1 on the Day 2 spectrum and spans 17.7:1 to 31.6:1 across the six days, exceeding the 3σ

criterion on every day. The ethanol C–O stretch at 1056 cm^{-1} is weak in this deployment and straddles that criterion, falling on either side depending on the day and the fitting window, so no signal-to-noise ratio is quoted. The band is nonetheless identified on all six days as a reproducible local maximum at $1056.8 \pm 3.2\text{ cm}^{-1}$ (1 s.d.), within 1 cm^{-1} of the literature position. A fitting routine returns a center whether or not a band is present, so the same procedure was applied to the Raman-inactive noise window: there the centers scatter by $\pm 12\text{ cm}^{-1}$, almost four times wider, which is what separates a band from a converged fit. Ethanol is therefore reported as detected but not quantified.

The three quantified species exceed the 3σ threshold by more than a factor of 30 under field conditions, leaving margin for reduced integration time or lower analyte concentrations, a tradeoff quantified in Section IV-F.

1) *Limits of Detection and Quantification*: Converting these ratios to composition gives the concentration-based figures needed for comparison with other modalities. With the noise-equivalent molar fraction defined as $NE\chi_i = \chi_i/\text{SNR}_i$, the limit of detection is $3NE\chi_i$ and the limit of quantification is $10NE\chi_i$. Evaluated on the Day 2 spectrum under the stated operating conditions of 1000 s integration and 515 kPa absolute, these are 0.33, 0.25 and 0.26 mol% (detection) and 1.10, 0.84 and 0.87 mol% (quantification) for CO_2 , N_2 and O_2 respectively. These figures are for the three quantified species; ethanol and water vapor are reported as detected only, so no limit of quantification is claimed for them.

These limits are extrapolated from one measured operating point through the linearity of (1), not measured against prepared low concentrations, and the qualification attaches wherever they are quoted, including Tables II and III. Section VI-C identifies the dilution series that would replace it.

C. Experimental Determination of the Oxygen Response Factor

Six ambient-air spectra were acquired fourteen days before Day 1, with the gas path open to laboratory air. Air declares its own composition: the dry-air molar ratio $\chi_{N_2}/\chi_{O_2} = 0.78084/0.20946 = 3.7279$ is fixed, humidity-independent (water vapor dilutes both species equally), and unaffected by argon or CO_2 , which cancel from the ratio. The measured integrated-area ratio was

$$\frac{A_{N_2}}{A_{O_2}} = 3.433 \pm 0.060 \quad (n = 6 \text{ ambient-air spectra}),$$

reproducible to 0.2% across a 48-fold range of exposure: the four 5 s repeats give 3.430 ± 0.074 and the 120 s and 240 s acquisitions 3.438 ± 0.040 . Direct integration in place of Gaussian fitting returns 3.491 ± 0.067 , within 1.7%, confirming that a single Gaussian captures the band manifold.

Because the composition is known, inverting (4) determines the cross-section factor the instrument requires rather than testing that factor:

$$k_{O_2} = \frac{\chi_{N_2}/\chi_{O_2}}{A_{N_2}/A_{O_2}} = \frac{3.7279}{3.433} = 1.086 \pm 0.019. \quad (6)$$

This is an *effective* factor, folding the molecular cross-section together with the relative response of the detection chain at the two scattered wavelengths, because that product is what (4) consumes and what the measurement determines. The measured value differs from the tabulated 1.21 by 11%. The tabulated value is essentially the ν^4 scattering factor alone: O_2 at 1556 cm^{-1} scatters to 580 nm and N_2 at 2331 cm^{-1} to 607 nm, giving $(17241/16466)^4 = 1.202$. The difference is therefore consistent with a detection-chain response ratio, $R_{\text{O}_2}/R_{\text{N}_2} \approx 0.90$, rather than with a disputed cross-section. Grating blaze, fiber transmission and detector quantum efficiency all vary across the 580 to 607 nm interval and are the available contributors. The edge filter is not among them: used at 45° its transmission edge lies below the ethanol band at 563.7 nm, so O_2 at 580.0 nm and N_2 at 607.3 nm both sit well inside the passband. The attribution is an inference from a single interval; the value of k_{O_2} is measured and does not depend on it.

On that reading the same effect acts on CO_2 , and the consequence is stated here rather than left for the reader to infer. CO_2 scatters to 572.7 nm, 34.6 nm from the N_2 Q-branch against the 27.3 nm that separates O_2 from it, so the response ratio $R_{\text{CO}_2}/R_{\text{N}_2}$ would depart from unity by at least as much as the oxygen ratio does. Extrapolating the measured gradient across the wider interval gives roughly 0.88, which would correspond to an effective k_{CO_2} near 1.27 in place of the tabulated 1.45.

What such a factor does and does not affect should be separated, since the two differ by an order of magnitude. Applying it raises the reported χ_{CO_2} by 6 to 10% depending on the day, which is a shift in level. The trajectory is very nearly unmoved: with each series normalized to its own maximum the two differ by at most 1.8% of full scale, and no ratio between consecutive acquisitions changes by more than 1.6%. A response factor that is fixed by the optical train enters every acquisition identically and therefore cancels from the day-to-day comparison, which is what the normalization is built to support. The systematic bears on the level at which a composition is reported, not on the evolution the instrument is used to follow.

The correction is not applied, because an extrapolation from a single measured interval is not a measurement. The tabulated k_{CO_2} is therefore carried throughout with the understanding that it retains an uncorrected response factor, a systematic in a known direction and of estimable size, distinct from the random literature uncertainty of Section III-D.

The argon reference cannot remove this ratio, for the reason given in Section III-A: the trace is subtracted as an additive background, so a multiplicative response survives it untouched.

Ambient air supplies exactly the missing quantity, and at the two wavelengths where it is needed. Because the N_2/O_2 molar ratio of dry air is fixed, the measured area ratio determines the product of cross-section ratio and response ratio at 607 and 580 nm in one step. That is why k_{O_2} is reported as an effective factor: it is the quantity (4) consumes, and the ambient-air determination is the only measurement in this work that supplies it.

The determination is stable across instrument configuration.

An independent acquisition at the deployed spectral resolution (seven days before Day 1, N_2 FWHM 16.70 cm^{-1}) gives $k_{\text{O}_2} = 1.097$, and correcting the ambient session to that resolution gives 1.10–1.12. Three routes spanning a 40–65% difference in point-spread-function width agree to within 3%, which is what licenses carrying the factor into the deployment window.

No accuracy claim follows from the agreement: with k_{O_2} fixed from these spectra, the recovered $\chi_{\text{N}_2}/\chi_{\text{O}_2}$ reproduces 3.7279 by construction. What the data do establish is that the factor is reproducible and configuration-stable, and that the tabulated value would have biased the oxygen channel by 11%. One limitation cannot now be removed: the archive holds no spectrum that is both genuine ambient air and recorded in the deployed configuration, so the transfer rests on the convergence above, and the instrument has since been modified.

The determination interrogates the two atmospheric gases against each other, so it constrains k_{N_2} and k_{O_2} and does not test k_{CO_2} , which governs the dominant species throughout the deployment. What would close that gap is set out in Section VI-C.

D. Composition Basis and Excluded Species

Composition is reported on a three-gas basis. Equation (4) defines each molar fraction as a share of the weighted sum over N_2 , O_2 and CO_2 , so the three fractions sum to one by definition. That sum is an identity, not a measurement: the value is one for any peak areas and any cross-section set, and can therefore neither succeed nor fail. No accuracy claim is made from that identity, and none can be.

Ethanol vapor, water vapor and trace volatiles are present in the headspace but are not part of this sum. Their effect is not a shortfall in the total, which remains one. Instead, they make all three reported fractions slightly too large.

The size of that effect follows directly. Let f be the fraction of the delivered gas occupied by the excluded species. Each reported fraction is then too large by the factor $1/(1-f)$, so with $f \leq 0.012$ from Section III-C no reported fraction is more than 1.2% high. Section III-D gives $\pm 8\%$ as the absolute uncertainty on a single molar fraction, set mainly by the literature uncertainty on the relative cross-sections; the basis effect is the smaller of the two, and the two are independent.

Because every fraction is scaled by the same factor, ratios between N_2 , O_2 and CO_2 are free of that factor. The ambient-air area ratio of Section IV-C, from which k_{O_2} is determined, is such a ratio.

The water band supplies a check on the normalization that the closure cannot. Water reaches the fiber with its abundance fixed by the conditioning train rather than by the fermentor (Section II-D), so water carries no expected process signal over the six days. Across the deployment the water band area at 3657 cm^{-1} varies by a factor of 2.36, while I_{Total} , which is proportional to throughput, varies by 1.92. The ratio of the two varies by only 1.30, a residual scatter of 11.8% (1 s.d.). Over the same six spectra χ_{CO_2} , formed by the identical

construction, retains a 2.20-fold variation. The normalization therefore removes throughput variation from a species that should show none, while leaving intact the variation of one that carries a real signal. This is the cancellation of Section IV-E demonstrated on the data rather than argued from (2).

Unlike the closure, this outcome was not guaranteed. The normalization might have left the throughput wander in place, or the water signal might have followed the fermentation; neither occurred. No molar fraction is computed for water, and none should be: water is not on the three-gas basis, and the quantity used here is a relative variation in which the unknown water cross-section cancels.

E. Pressure and Coupling Independence

As shown in (2), any factor that scales every peak area identically within a single acquisition cancels in χ_i , making the framework insensitive to pressure, laser power, fiber coupling, and bending loss. Empirically, the $\pm 2.7\%$ pressure stability budget is absorbed by the ratio cancellation of (2), the fitted peak positions and linewidths showed no systematic correlation with day of deployment, and coupling variations recorded by the diverted-beam power monitor (Section II-A) produced no corresponding deviation in the reported molar fractions.

This cancellation is what makes the six-day trajectory interpretable at all, and it is worth stating as a condition rather than a convenience. The instrument was powered down between acquisitions, so laser power, fiber coupling and alignment were re-established before each measurement. Because the total number density in the fiber is fixed by the operating pressure and stage temperature, I_{Total} is proportional to throughput alone; across the six acquisitions that total varied by a factor of 1.92. The quantity being measured, χ_{CO_2} , varied by 2.20 over the same interval. In raw band areas these two are of comparable size and cannot be separated: a change in the instrument and a change in the headspace are the same number. Forming each molar fraction as a ratio within a single acquisition removes the instrument term entirely, so the temporal comparison of Section V rests on that cancellation and not on any assumption that the instrument returned to the same state each day.

F. Temporal Response and Integration Time

The optical integration time was 1000 s per spectrum, sized for SNR margin rather than for a particular sampling cadence. Taking the noise to scale as $\sqrt{t}$, which the single 500 s spectrum in the archive is consistent with, the most signal-limited species is the water monomer. On its weakest day of the deployment the water ratio is 17.7:1, exceeding the 3σ threshold by a factor of 5.9, so the integration could be shortened by a factor of 34 and water would still clear threshold on every day rather than merely on a typical one. The dominant species carry orders of magnitude additional margin. That factor is an extrapolation, and Section VI-B identifies the dominant noise term as background rather than detector noise, so it should be read as an upper bound on what shortening the integration would buy. The instrument

TABLE II
SUMMARY OF METROLOGICAL PERFORMANCE OVER THE SIX-DAY FIELD DEPLOYMENT.

Quantity	Result
N ₂ peak position	2330.970 ± 0.007 cm ⁻¹ (RSD 0.0003%)
O ₂ peak position	1555.855 ± 0.021 cm ⁻¹ (RSD 0.0013%)
N ₂ FWHM	16.735 ± 0.095 cm ⁻¹ (RSD 0.57%)
O ₂ FWHM	18.669 ± 0.142 cm ⁻¹ (RSD 0.76%)
SNR N ₂ (Day 2 spectrum)	469:1
SNR CO ₂ (Day 2 spectrum)	472:1
SNR O ₂ (Day 2 spectrum)	101:1
SNR H ₂ O at 3657 cm ⁻¹	22.1:1 mean (range 17.7–31.6:1)
Limit of detection (3σ)	CO ₂ 0.33, N ₂ 0.25, O ₂ 0.26 mol%
Limit of quantification (10σ)	CO ₂ 1.10, N ₂ 0.84, O ₂ 0.87 mol%
Quantifiable range	LOQ to maximum measured; 47× for CO ₂
N ₂ /O ₂ area ratio, ambient air	3.433 ± 0.060 ($n = 6$, 48× exposure range)
k_{O_2} from that ratio	1.086 ± 0.019 (tabulated value 1.21)
H ₂ O variation, Days 1–6	11.8% normalized vs. 25.7% raw (1 s.d.)
Operating pressure	515 ± 14 kPa absolute ($\pm 2.7\%$)
Fiber stage temperature	40 ± 0.06 °C
Fiber fill time	~ 15 s
Optical integration time	1000 s (Days 2–6); 500 s (Day 1)

nonetheless supports sub-hourly process sampling for all five species, with the practical lower bound on the cycle time set by the 15 s fill time and 15 min purge between samples rather than by spectroscopic SNR.

Table II consolidates the metrological results above. The ambient-air measurement of Section IV-C determines k_{O_2} rather than testing it, so no accuracy claim follows from recovering the dry-air ratio. What the table reports is reproducibility and sensitivity: the stability of the wavenumber axis and of the fitted peak shapes across six power cycles, and the limits of detection and quantification of Section IV-B. No measurement reported here constrains k_{CO_2} , which additionally carries the uncorrected response factor described in Section IV-C.

V. FIELD DEMONSTRATION

The instrument was deployed on a six-day red wine fermentation at the UC Davis Research Winery, chosen because a single test exposes the instrument to the full set of conditions motivating the design: elevated pressure, near-saturation humidity, a CO₂ molar fraction rising from an atmospheric starting condition to more than half the headspace and falling again, ethanol evolution, repeated powered-down cycling, and uncontrolled ambient. The deployment is reported here for its metrological content; chemical and biological interpretation of the headspace composition is deferred to other work.

A. Deployment Configuration

The fermentation process was a 200 L double-walled stainless steel fermentor filled with 136 kg of 2024 Cabernet Sauvignon grapes (Oakville, Napa Valley, California), with a semi-open vented headspace. The must, the fermenting juice together with the grape skins, was inoculated with *Saccharomyces cerevisiae* (Lalvin BM4X4™) at 250 mg/L.

Three features of a red fermentation matter for what follows. Evolving CO_2 buoys the skins into a floating layer, the cap, which sits above the liquid and is where colour and tannin extraction occurs. To drive that extraction the vessel performs a pump-over: juice is drawn from the middle of the fermentor and sprayed onto the cap surface, here for a median of 5.0 min at a median separation of 70.5 min. The deployment saw 83 such operations, the regular cadence ending partway through Day 4 and falling to about three per day thereafter as the fermentation finished. Each operation draws outside air into the headspace, so the pump-overs are the route by which oxygen enters and are the source of the abrupt compositional perturbations against which the instrument's response is tested. Temperature is held by cold water admitted to the jacket, the cavity between the fermentor's inner and outer walls, under thermostatic control at $22 \pm 0.5^\circ\text{C}$.

The FERS instrument was installed adjacent to the fermentor on the winery floor with no environmental conditioning of the surrounding space; ambient temperature and humidity outside the enclosure varied freely. The arrangement is drawn in Fig. 1; the vessel remained in normal production use throughout. Spectra were acquired once daily for six days; the instrument was powered down between acquisitions. Each acquisition pumped headspace gas through the fiber, pressurized to 515 kPa, integrated for 1000 s, and purged with nitrogen for 15 min.

Fig. 4 is the record of that deployment. The figure places each acquisition on the fermentation timeline and sets the measured CO_2 trajectory against the vessel conditions under which each spectrum was taken: the must, cap and jacket temperatures, and the pump-over duty cycle. The first spectrum was taken 2.92 d after inoculation and the first Brix reading 1.03 d after inoculation, so both records begin with the fermentation already advanced. Sugar consumption, tracked independently by handheld refractometer, fell from 24.3°Brix two days before the first spectrum to -1.1°Brix , establishing that the acquisitions spanned the fermentation from that advanced state through to near-completion. Brix serves here only to anchor each spectrum to a known point in the progression.

The temperature channels show a single thermal regime rather than a warm cap above cooler liquid. Over the plotted interval the cap tracks the must to within -0.3 to $+0.7^\circ\text{C}$ (mean $+0.02^\circ\text{C}$), and the must runs 0.21°C above the jacket on average. The jacket trace is a sawtooth because the cooling is applied in discrete pulses rather than modulated continuously: each admission of cold water drops the jacket by as much as 1.5°C within one 5 min sample, after which the jacket drifts back toward the must, and the pulses recur about every 50 min. The must does not follow, its own excursions being half as large, because the liquid carries far more thermal mass than the jacket. One acquisition, 2024-09-24 14:16:34, was taken 5.5 min into a pump-over and is the only one of the nine taken while the pump was running. Its N_2/O_2 area ratio is 3.53, against 3.96 to 4.27 for the six deployment acquisitions and 3.433 for ambient air, so the sample is dominated by the air drawn in by the pump-over rather than by headspace. The point is excluded from the CO_2 series on that basis and is marked separately in the figure.

A wine fermentation evolves over several days, so an autonomous hourly cadence is the intended operating mode and would resolve the headspace trajectory in detail. Section IV-F establishes that the instrument already supports that cadence. The once-daily, supervised cadence used here reflects the terms on which a first deployment was possible rather than any limit of the instrument. Fermentor access is the binding constraint: fruit arrives once a year, vessels are committed before the crush, and every ferment carries a program of its own, so what an unproven instrument is granted on a working vessel is bounded. Here that was six supervised acquisitions across the six days of active fermentation, each running unattended but with the instrument monitored between runs so that intervention remained possible. Establishing that the measurement is sound under those conditions is what earns the extended access a finer cadence requires.

One observation illustrates directly what the finer cadence would buy. The ethanol C–O band reaches its maximum one acquisition after the CO_2 maximum. That ordering is what the partitioning would predict, since CO_2 is sparingly soluble and leaves the liquid as fast as fermentation produces the gas, whereas ethanol accumulates in the must and enters the headspace with a lag. The offset is not resolved by these data. The CO_2 maximum is separated from its neighboring value by 5.3%, well inside the 10.3% repeat-to-repeat scatter established later in this section, so the two are not distinguishable, and a one-day offset is a single sampling interval. The observation is recorded as consistent with expectation, not as a measurement; establishing a lag of this size is precisely the kind of question the hourly cadence is intended to answer.

Daily sampling was nonetheless sufficient for the composition actually reported. Each of the three normalized fractions changed by less than a factor of two between consecutive acquisitions (Fig. 5), so the six points resolve that trajectory without aliasing it. That holds for the bulk gases and for nothing finer: the structure described above is not recovered at this cadence, and the case for hourly operation rests on what daily sampling cannot see rather than on what it misrepresents.

Repeatability under repeat conditions is bounded by the ambient-air spectra of Section IV-C. Six acquisitions of a fixed atmospheric composition give a relative standard deviation of 1.7% on the N_2/O_2 area ratio (3.433 ± 0.060), covering acquisition noise and gas-handling variability together.

The headspace itself was moving, so the deployment cannot reproduce that figure cleanly. Day 1 comes closest: three acquisitions of 500 s each, spanning 63 min, give a relative standard deviation of 10.3% on the normalized CO_2 doublet area, and the three values are not monotonic. At the mean rate of change between Days 1 and 2 the underlying composition would move by about 2% over that interval, so the 10.3% is dominated by measurement scatter rather than by fermentation. That scatter exceeds the roughly 2.5% uncertainty of the doublet-area fit itself, which comprises 1.5–2.1% from the peak fit and 0.5–1.5% from the choice of baseline parameters and fitting window, because gas handling and headspace sampling enter a repeat measurement and not a fit. The fit-based budget is therefore not a usable error bar, and none is drawn in Fig. 4. Days 2 to 6 contributed one 1000 s acquisition each and

place no bound on the scatter. An acquisition series at fixed composition is identified in Section VI-C as an objective for the next deployment.

B. Five-Species Detection

All five target species are detected simultaneously in the field spectrum of Fig. 3, each a well-resolved and well-isolated peak at the positions listed in Section III-A; the ethanol C–H stretching region near 2935 cm^{-1} is also visible but is not used for quantification. Recovering the water monomer under these conditions illustrates a differentiator of gas-phase Raman at visible wavelengths. The gas path ran at $> 95\%$ relative humidity, where infrared absorption faces severe O–H interference, while the small Raman cross-section of water and the spectral isolation of the 3657 cm^{-1} monomer leave a clean peak.

A second feature of Fig. 3 bears on the instrument's industrial fitness: there is no broad liquid-phase O–H feature. Condensation inside the fiber would appear as a wide background centered near 3300 cm^{-1} with liquid-water linewidth. Its absence, after the instrument had run at 515 kPa on a headspace above 95% relative humidity for the whole deployment, is direct evidence that the conditioning train and thermal stabilization of Section II-D did what they were designed to do: strand the water in the pump and trap, and hold every surface downstream above the dew point of the gas delivered to it.

C. Compositional Dynamic Range

Fig. 5 shows the daily evolution of the molar fractions χ_{N_2} , χ_{O_2} , and χ_{CO_2} derived from the three-gas framework. The dominant species, CO_2 , rose to a maximum measured molar fraction of $\chi_{\text{CO}_2} = 0.52$ on Day 2 and fell to $\chi_{\text{CO}_2} = 0.24$ by Day 6. Both N_2 and O_2 were correspondingly displaced and recovered; the measured ranges over the deployment were $0.39 \leq \chi_{\text{N}_2} \leq 0.62$ and $0.09 \leq \chi_{\text{O}_2} \leq 0.14$.

Two statements about range must be kept separate. The *quantifiable* range runs from the limit of quantification to the maximum measured value: for CO_2 this is 1.10 to $52\text{ mol}\%$, a factor of 47 . Separately, the instrument was *subjected* to a larger excursion, because the headspace began at the ambient atmospheric CO_2 level of approximately 4×10^{-4} . That lower figure is the assumed composition of ambient air at the start of the deployment, not a spectroscopic measurement: that level lies roughly an order of magnitude below the $0.33\text{ mol}\%$ detection limit established in Section IV-B, so CO_2 at that level is not detectable by this instrument, still less quantifiable. The starting value is reported here only to describe the initial condition of the vessel, and does not bound the quantifiable range.

Across the same dynamic range, the CO_2 Fermi doublet itself exhibited position stability of $\sigma < 1\text{ cm}^{-1}$, a constant splitting of $103.10 \pm 0.04\text{ cm}^{-1}$, and a doublet fit quality of $R^2 = 0.9994 \pm 0.0001$, confirming that the spectral shape parameters used in the fit were preserved across the full range of CO_2 intensities encountered during the deployment.

D. Six-Day Operational Stability

The six-day campaign, with once-daily power-on acquisitions and instrument shutdown between each, served as a repeated-operation stress test of the metrological claims of Section IV. The spectral baseline showed no degradation, the fitted N_2 and O_2 peak positions and linewidths stayed within the stability budget of Section IV-A, the diverted-beam coupling monitor showed no systematic drift, and each of the three normalized gases remained above its own limit of detection on every day of the deployment.

What makes six power cycles a test rather than a coincidence is that neither axis of the measurement depends on the instrument's state persisting between runs. The wavenumber scale is re-established on every acquisition from the N_2 and O_2 Q-branches present in that same spectrum (Section III-B), and the composition is normalized within that spectrum (Section III-C). A power cycle, an excursion in hall temperature, or a change in coupling efficiency is therefore absorbed at each acquisition rather than accumulated across the campaign.

The stability reported here is accordingly a property of the referencing scheme, not a claim that the hardware held still. The hardware did not: Section IV-B reports the noise floor moving by a factor of 1.7 across the same six days, and Section IV-E a 1.92 -fold variation in total throughput, neither of which disturbed a reported molar fraction. The instrument is on that basis suited to a campaign of this length on a production floor, with individual acquisitions running unattended and with the ambient uncontrolled throughout.

VI. DISCUSSION

Three questions follow from these results, and the subsections below take them in order. What does the standard-free framework deliver, where does it stop, and to what other processes does it carry? Where does an instrument with $\text{mol}\%$ -level detection limits then sit alongside modalities reaching parts per billion, and what can it measure that they cannot? And which of the present limitations belong to the technique rather than to this particular implementation?

A. Discussion of the Standard-Free Framework

The three-gas framework is structurally distinct from chemometric calibration. A PLS or similar regression model requires a labeled training set spanning the operating composition range and remains valid only as long as the spectroscopic system reproduces those training conditions; calibration aging is a recognized limitation of long-duration deployments. The ratio framework reported here derives molar fractions from peak areas and tabulated physical constants, with the accuracy floor set by literature cross-section uncertainty (a property of physical constants rather than of the instrument), and therefore does not exhibit calibration aging in the PLS-model sense. The atmospheric N_2/O_2 recovery of Section IV-C is the evidence supporting this claim for k_{N_2} and k_{O_2} ; the corresponding check for k_{CO_2} remains outstanding.

The framework extends beyond the species reported here. Methane, hydrogen and carbon monoxide are well characterized in the gas-phase Raman literature and are immediate

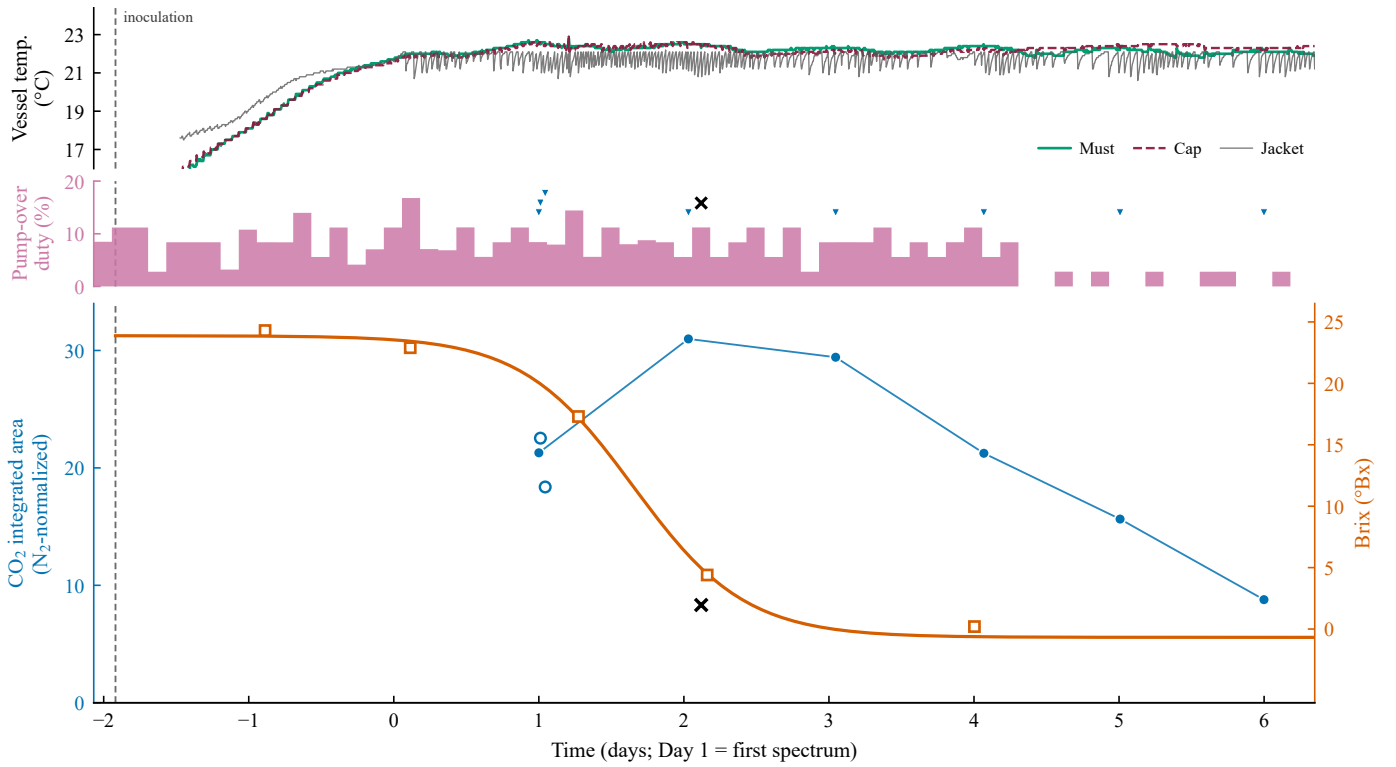


Fig. 4. Process-state reference for the six-day deployment. All three panels share a common time axis on which Day 1 is the first spectrum. **Dashed vertical line:** inoculation, at day -1.92 ; the fitted Brix curve is drawn only from inoculation onward. **Upper strip:** the three vessel temperature channels at 5 min cadence, must (green, solid), cap (red, dashed) and cooling jacket (grey, thin); the sawtooth on the jacket trace is the pulsed cooling described in Section V-A and not measurement noise. **Middle strip:** pump-over duty cycle, the percentage of each 3 h interval during which the pump-over pump was running, from the winery historian record. Individual operations are too brief to resolve at this scale. **Blue triangles:** the acquisition times, stepped vertically where they would otherwise overlap. **Black cross:** the one acquisition taken during a pump-over, excluded from the CO_2 series for the reason given in Section V-A. **Lower panel, left axis, blue filled circles:** integrated CO_2 Fermi doublet area, the 1285 and 1388 cm^{-1} components summed and normalized to the N_2 Q-branch, plotted at the acquisition time of each spectrum; **blue open circles** are the two further Day 1 acquisitions, 16 and 63 min after the first and processed identically. **Right axis, orange open squares:** Brix, measured independently of the FERS instrument by handheld refractometer, with the solid orange curve a four-parameter logistic fitted to all eight readings, which the curve reproduces within 0.8° . A curve is drawn only for a fitted physical model or a continuously sampled record: the thin segments joining the CO_2 points are guides to the eye and not a fit. Three further Brix readings at Days 7.9, 9.2 and 10.2 (-0.7 to -1.1°) lie beyond the plotted range.

candidates for biogas, hydrogen-production and combustion monitoring [2], [6]. The two conditions of Section III-C carry over to those applications and are worth checking rather than assuming. Anaerobic digestion and hydrogen production leave the headspace N_2 inventory untouched. Combustion converts a small part of it to nitrogen oxides, of order one part in a thousand, which is negligible at the mol% level reported here.

Two situations fall outside the conditions: a vessel under inert purge, where no atmospheric nitrogen enters, and a process that evolves nitrogen, where the reference is itself a product. The second is not hypothetical, since biological water treatment includes denitrification and appears among the applications proposed for this instrument. There nitrogen is abundant but is an analyte rather than a spectator.

Neither case is a dead end. An inert tracer introduced into the vessel can take the reference role, provided it takes no part in the chemistry under observation, has a tabulated Raman cross-section, and has a band clear of the analyte features. The second condition rules out the monatomic noble gases, which have no vibrational Raman spectrum whatever their inertness; sulfur hexafluoride qualifies, though its global-warming potential argues for closed use. Only the provenance

of the reference channel changes, since the tracer enters (4) as one more term with a literature cross-section.

B. Position Relative to Alternative Gas Analysis Modalities

This subsection develops quantitatively the scope stated at the outset, that the instrument is a composition monitor and not a trace analyzer. Table III summarizes the principal alternative modalities for industrial gas analysis alongside the instrument described in this paper, on metrological dimensions relevant to long-duration, multi-component, calibration-stable operation in semi-open process vessels.

Infrared methods (NDIR, TDLAS, WMS) deliver high single-species sensitivity and have benefited from deep-learning signal processing [1], but the dipole-moment selection rule excludes the homonuclear diatomics that provide both the natural atmospheric internal standard and the oxygen channel in our framework; a thousand-fold sensitivity advantage on CO_2 does not help close a nitrogen-referenced balance. GC-MS provides ppb sensitivity and definitive species identification for trace volatiles [3], [4] but is offline and hour-cycle limited, placing that technique in a complementary rather than competing role. Electrochemical and metal-oxide sensors offer

TABLE III
COMPARISON OF GAS ANALYSIS MODALITIES FOR LONG-DURATION INDUSTRIAL PROCESS MONITORING.

Modality	N ₂ , O ₂	Multi-species	Real-time	Typical LOD	Calibration
IR (NDIR, TDLAS, WMS)	No	Limited	Yes	ppb–ppm	Per-species standards, periodic recal
Photoacoustic (PAS)	No	Limited	Yes	ppb–ppm	Per-species standards
FTIR	No	Yes	Yes	ppm	Background/reference spectra
GC–MS	Partial	Sequential	No	ppb	Reference standards, offline
Electrochemical, MOS	Per-cell	No	Yes	ppm–0.1%	Periodic recal, humidity-sensitive
Liquid Raman + PLS chemometrics	N/A	Yes (liquid)	Yes	N/A	Labeled training set, model aging
HC-ARF FERS, optimized [17]	Yes	Yes	Yes	1.2–16.7 ppm·bar	Literature cross-sections only
HC-ARF FERS (this work)	Yes	Yes	Yes	1.3–1.7 × 10 ⁴ ppm·bar	Tabulated cross-sections; O ₂ factor from ambient air

GC–MS reaches the lowest detection limit in the table and is the one modality that cannot deliver it in real time. The two HC-ARF FERS rows are quoted in **ppm·bar**, the pressure-normalized limit, so that the two implementations of the same technique can be compared directly: Bai *et al.* measured at 1 bar with 60 s integration, and this work at 5.15 bar (515 kPa absolute) with 1000 s, where the limits are 2500 to 3300 ppm. The longer integration favours this work, and the limits reported here nonetheless remain more than two orders of magnitude above theirs. Section VI-B sets out the scaling, including the slit and detector differences, and finds that those differences do not account for the gap. Limits of detection for this work are SNR-derived extrapolations from that measured operating point rather than measurements from a controlled dilution series (Section IV-B). Sensitivity values for the remaining modalities are representative literature values and are not normalized for analyte, pressure, integration time, optical power, or instrument configuration.

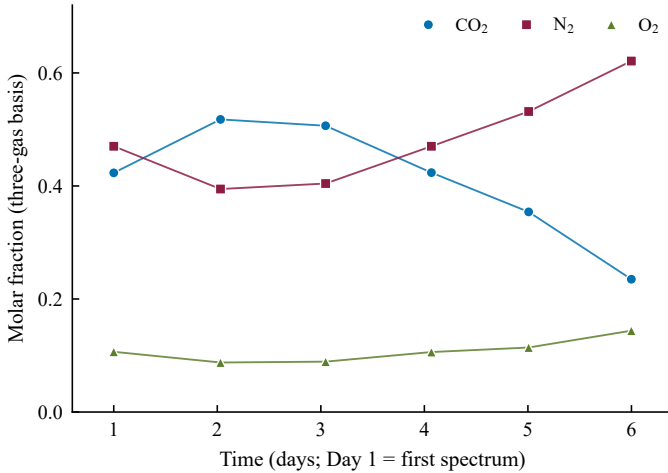


Fig. 5. Evolution of the three normalized molar fractions χ_{N_2} , χ_{O_2} and χ_{CO_2} over the six-day deployment, derived from the three-gas framework of Section III-C. The three species are plotted separately rather than stacked: on a three-gas basis the fractions sum to unity by construction, so a stacked presentation would render that identity as though the sum were a measured closure. Thin segments join the points as guides to the eye and are not a fit. The trajectory is reported as the operating range over which the framework was exercised, not as evidence of accuracy.

single-species coverage at low cost but exhibit cross-sensitivity and humidity drift, and are deployed as networks of individual sensors. Liquid-phase Raman with chemometric models is mature industrial PAT [2] but operates on dense liquid samples and does not measure gas-phase composition. Being spatially local, liquid-phase instrumentation also needs more probes as the vessel grows [24], whereas a vessel-averaged headspace measurement is independent of vessel size: a single sampling line serves a 50 to 200 kL production fermentor exactly as it serves the 200 L vessel measured here.

HC-ARF FERS with a nitrogen-referenced framework therefore occupies a position not currently filled by any of these modalities, simultaneous quantification of N₂, O₂, and CO₂ alongside detection of ethanol vapor and water vapor in real time, under elevated pressure and near-saturation humidity, without external concentration standards introduced during

deployment. That position belongs to the technique; what the present instrument contributes is a demonstration that it can be occupied outside a laboratory.

The comparison that bears on sensitivity is not with a different modality but with another implementation of the same technique, one optimized for detection limit rather than for field operation. Bai *et al.* [17] report 1.2 to 16.7 ppm·bar using forward scattering at 532 nm through a hollow-core anti-resonant fiber of comparable core diameter, at the same 200 mW coupled into the fiber. The excitation is not the difference. Their advantages are a detector at -70°C , an $f/22$ imaging spectrograph with a 10 μm slit, and digital selection of the detector rows carrying signal, which rejects background photons before they contribute shot noise.

The instrument reported here was built to a different objective, so its limits of detection are not a bound on the technique. The objective was to quantify composition, where the three reported species occupy 9 to 62 mol% and sit 9 to 75 times above their own limits of quantification throughout the deployment. Every design choice follows from that objective: a detector at -13°C , a 50 μm slit chosen for throughput rather than resolution, no spatial filtering of the detector image, and operation on a production floor rather than an optical table.

Those design differences do not account for the size of the gap. Scaling the Bai result by the ratio of dark currents between -70 and -13°C , and crediting our longer 1000 s integration and wider slit, predicts a penalty of roughly a factor of four, or an order of magnitude under the least favorable assumption about the dark-current activation energy. The measured difference is larger than that by more than two orders of magnitude, so the noise floor is not dark-limited and cooling alone would not close it. The residual is consistent with background structure, and in particular with the argon instrument-response residual that the present single-reference protocol cannot track across a six-day deployment.

Sensitivity here is therefore limited by identified and addressable terms rather than by the measurement principle. Section VI-C sets out what closing them requires.

C. Limitations and Future Work

The instrument is, by design, a relative-composition measurement, reporting molar fractions on a three-gas basis rather than absolute concentrations. The wavenumber scale is anchored to the well-established gas-phase Raman vibrational frequencies of N_2 and O_2 (HITRAN [20]), and the composition scale is anchored to the atmospheric N_2/O_2 recovery (Section IV-C); neither anchor requires national-metrology traceability or external calibration standards introduced during deployment. The accuracy floor on the reported molar fractions is set by the $\pm 5\%$ literature uncertainty on the relative Raman cross-sections themselves.

The field validation here is a single deployment with daily sampling. Both limits have the cause given in Section V-A: fruit arrives once a year and vessels are committed before the crush, so further deployments are gated by harvest and by demonstrated performance rather than by instrument readiness. Repeated deployments and the move to the autonomous sub-hourly cadence are the principal additional field-validation objectives.

The priority next measurement is a binary CO_2/N_2 mixture of known composition. Nothing reported here constrains k_{CO_2} , which governs the dominant species and therefore the whole of Fig. 5. The process record of Fig. 4 cannot substitute: the vessel vents continuously, displacing some 200 headspace volumes with a residence time that varies severalfold, so headspace CO_2 reflects a flow-through balance rather than cumulative production. A mixture prepared at a known molar ratio would constrain k_{CO_2} independently of the fermentation. Establishing that ratio gravimetrically, rather than purchasing a certified concentration standard, keeps the determination standard-free: what the framework needs is a known ratio, not a traceable concentration. We regard the measurement as the one that would fix the level at which χ_{CO_2} is reported. The composition trajectory does not wait on it, for the reason given in Section IV-C: a response factor set by the optical train is common to every acquisition and cancels from the comparison between them.

Quantification of ethanol and water vapor would require validated gas-phase Raman cross-sections for these species at 532 nm with uncertainty comparable to that of the present three-species set, which would allow them to enter I_{Total} and tighten the bound on excluded species given in Section IV-D.

1) *Route to Improved Detection Sensitivity:* Three of the terms identified in Section VI-B are addressable without altering the measurement principle, listed here in the order we judge their return on effort.

The first is the argon instrument-response reference. Acquiring it once per deployment rather than once per acquisition is what leaves the background residual untracked, so moving to a per-acquisition reference is the single change most likely to move the limits of detection. The change costs acquisition time only.

The second is spatial selection of the detector rows carrying the fiber image. Bai *et al.* [17] report that this rejects background photons before they contribute shot noise. Our detector is read in a fixed region of interest without such selection;

implementing row selection is a change to the acquisition software rather than to the optical train.

The third is detector cooling, from -13 to $-70^\circ C$, which would reduce dark current by roughly three orders of magnitude. It is placed third because the floor is not dark-limited: cooling becomes worthwhile once the background terms above are removed, not before.

The open question is therefore not whether the technique can reach substantially lower detection limits, but which of these terms dominates once the first is removed, and how much of the remaining gap is recoverable in an industrial enclosure rather than on an optical table. A dilution series prepared in the same way as the binary mixture above would also replace the extrapolated limits of detection reported here with directly measured ones.

2) *Which Additional Species Are Within Reach:* It is possible to say something more specific about which additional species are within reach, without new measurements. At a fixed noise floor the limit of detection scales inversely with the differential Raman cross-section, so the measured CO_2 limit of Section IV-B can be transferred to any species with a tabulated cross-section ratio. Applied to the $3718\text{--}3918\text{ cm}^{-1}$ noise floor at 1000 s and 515 kPa, this gives estimated detection limits of roughly 0.08 mol% (800 ppm) for CH_4 and SO_2 , 0.12 mol% for H_2S , 0.15 mol% for NO_2 , 0.19 mol% for H_2 , and 0.5 to 0.9 mol% for CO and NO , whose cross-sections are below that of the CO_2 dyad. For the fermentation volatiles relevant to this application, the corresponding figures are approximately 0.2 mol% for ethyl acetate and 0.3 mol% for acetaldehyde and the higher alcohols.

Two qualifications are essential. These are extrapolations from a single measured operating point, not measurements, and they assume a clean spectral neighborhood; for the C–H manifold in particular that assumption fails, because higher alcohols, esters and aldehydes contribute there simultaneously and no single species can be isolated without deconvolution. And the values sit at the mol% level, which is appropriate for process composition at the percent level but well above the ppm thresholds that industrial safety monitoring requires for H_2S , SO_2 or NO_2 .

The correct reading is therefore that the present implementation brings these species within reach as *process* constituents at percent level, not as *safety* monitors at threshold limit values, and that the sensitivity improvements identified above are what would separate the two.

VII. CONCLUSION

This paper has presented a hollow-core anti-resonant fiber-enhanced Raman instrument and a three-gas normalization framework for simultaneous, standard-free measurement of headspace composition, characterized over a six-day fermentation on a working winery floor rather than in a laboratory. Five species are detected and three quantified, with no external concentration standard, pressure sensor or absolute intensity reference, and with the oxygen response factor determined on the instrument from ambient air. Wavenumber stability held below 0.021 cm^{-1} , limits of detection were 0.25 to

0.33 mol%, and CO₂ was quantified over a 47-fold range; the full characterization is collected in Table II.

A fermentation releases many volatile species, and the order in which they appear, together with the dependencies among them, is not well established. A gas-phase Raman measurement records the whole spectrum in a single acquisition, so every species with a band in range is observed on one time base, without tuning to a chosen analyte or separating the mixture in time. Improved sensitivity therefore does not simply add analytes one at a time; the gain is to convert a three-gas trajectory into a simultaneous record of the volatile set, which is what studying those dependencies would require.

Establishing the normalization framework is the first step toward that, and is what the present work contributes. A spectrum acquired on one day can be compared with one acquired on another only if both are referred to a common internal standard, and Section IV-E shows that the three-gas normalization supplies one across six power cycles in an uncontrolled ambient. Sensitivity gains applied without that reference would yield more bands per spectrum but no trajectory, since nothing would license comparing one acquisition to the next.

ACKNOWLEDGMENTS

The authors thank the UC Davis Research Winery in the Robert Mondavi Institute for fermentation facilities, Leticia Rodriguez for technical assistance, and the Beringer Vineyards machine shop for fabrication of specialized components. J. Nelson was supported by the Rodgers University Fellowship in Electrical and Computer Engineering. This research was funded in part by Earthy Labs.

REFERENCES

- [1] H. Zhang, X. Zhang, J. Tang, Y. Li, Z. Zhang, and S. Zhou, "Deep learning-enhanced dual-component gas sensor based on wavelength modulation spectroscopy," *Analytical Chemistry*, vol. 97, no. 22, pp. 22 032–22 040, 2025. [Online]. Available: <https://doi.org/10.1021/acs.analchem.5c03438>
- [2] A. Chandra, V. Kumar, U. C. Garnaik, R. Dada, I. Qamar, V. K. Goel, and S. Agarwal, "Unveiling the molecular secrets: A comprehensive review of Raman spectroscopy in biological research," *ACS Omega*, vol. 9, no. 51, pp. 50 049–50 063, 2024. [Online]. Available: <https://pubs.acs.org/doi/10.1021/acsomega.4c00591>
- [3] X. Chen, R. Hu, L. Hu, Y. Huang, W. Shi, Q. Wei, and Z. Li, "Portable analytical techniques for monitoring volatile organic chemicals in biomanufacturing processes: Recent advances and limitations," *Frontiers in Chemistry*, vol. 8, p. 837, 2020. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fchem.2020.00837/full>
- [4] M. Al-Khalili, P. B. Pathare, S. Rahman, and N. Al-Habsi, "Aroma compounds in food: Analysis, characterization and flavor perception," *Measurement: Food*, vol. 18, p. 100220, 2025. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2772275925000073>
- [5] H. Fuller, C. Beaver, and J. Harbertson, "Alcoholic fermentation monitoring and pH prediction in red and white wine by combining spontaneous Raman spectroscopy and machine learning algorithms," *Beverages*, vol. 7, no. 4, p. 78, 2021. [Online]. Available: <https://www.mdpi.com/2306-5710/7/4/78>
- [6] D. Majumder, R. Janani, A. Scrimshire, A. Stone, W. Brooks, C. Holcroft, R. Werner, S. Green, N. Wheeler, and P. A. Bingham, "Raman gas sensing technology: A new horizon?" *Sensors and Actuators Reports*, vol. 9, p. 100311, June 2025. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2666053925000293>
- [7] M. J. Pelletier, "Quantitative analysis using Raman spectrometry," *Applied Spectroscopy*, vol. 57, no. 1, pp. 20A–42A, 2003.
- [8] T. Wallocha and M. Poth, "Harnessing Raman spectroscopy for enhanced bioprocess monitoring: Predictive CO₂ analysis and robust pH determination in bioreactor off-gas stream," *Fermentation*, vol. 11, no. 6, p. 317, 2025. [Online]. Available: <https://www.mdpi.com/2311-5637/11/6/317>
- [9] M. P. Buric, K. P. Chen, J. Falk, and S. D. Woodruff, "Enhanced spontaneous Raman scattering and gas composition analysis using a photonic crystal fiber," *Applied Optics*, vol. 47, no. 23, p. 4255, 2008. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=ao-47-23-4255>
- [10] S. Hanf, R. Keiner, D. Yan, J. Popp, and T. Frosch, "Fiber-enhanced Raman multigas spectroscopy: A versatile tool for environmental gas sensing and breath analysis," *Analytical Chemistry*, vol. 86, no. 11, pp. 5278–5285, 2014. [Online]. Available: <https://pubs.acs.org/doi/10.1021/ac404162w>
- [11] M. Yang, Z. Liu, L. Xiong, Q. Nie, Y. Wang, S. Gao, M. Cheng, D. Yang, S. Pei, and D. Guo, "Antiresonant fiber-enhanced Raman spectroscopy gas sensing with 1 ppm sensitivity," *Optics Express*, vol. 32, no. 3, p. 4093, 2024. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=oe-32-3-4093>
- [12] Q. Nie, Z. Liu, M. Cheng, S. Pei, D. Yang, D. Guo, and M. Yang, "Review on hollow-core fiber based multi-gas sensing using Raman spectroscopy," *Photonic Sensors*, vol. 14, no. 4, p. 240412, 2024. [Online]. Available: <https://link.springer.com/10.1007/s13320-024-0730-4>
- [13] G. Qian, F. Wan, F. Zhou, J. Wang, W. Kong, and W. Chen, "Fiber-enhanced Raman spectroscopy for trace-gas sensing in the high-concentration gas background with an anti-resonant hollow core fiber," *Frontiers in Physics*, vol. 10, p. 917688, 2022. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fphy.2022.917688/full>
- [14] W. Kong, F. Wan, R. Wang, H. Sun, and W. Chen, "Signal enhancement and noise suppression technologies in Raman spectroscopic gas sensing," *Applied Physics Reviews*, vol. 12, no. 1, p. 011334, 2025. [Online]. Available: <https://pubs.aip.org/apr/article/12/1/011334/3340109/Signal-enhancement-and-noise-suppression>
- [15] W. S. Brooks, M. Partridge, I. A. Davidson, C. Warren, G. Rushton, J. Large, M. Wharton, J. Storey, N. V. Wheeler, and M. J. Foster, "Development of a gas-phase Raman instrument using a hollow core anti-resonant tubular fibre," *Journal of Raman Spectroscopy*, vol. 52, no. 10, pp. 1772–1782, 2021. [Online]. Available: <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/jrs.6195>
- [16] M. Shirmohammad, M. A. Short, and H. Zeng, "Collision Enhanced Raman Scattering (CERS): An Ultra-High Efficient Raman Enhancement Technique for Hollow Core Photonic Crystal Fiber Based Raman Spectroscopy Gas Analyzer," *Biosensors*, vol. 13, no. 11, p. 979, 2023. [Online]. Available: <https://www.mdpi.com/2079-6374/13/11/979>
- [17] Y. Bai, D. Xiong, Z. Yao, X. Wang, and D. Zuo, "Analysis of CH₄, C₂H₆, C₂H₄, C₂H₂, H₂, CO, and H₂S by forward Raman scattering with a hollow-core anti-resonant fiber," *Journal of Raman Spectroscopy*, vol. 53, no. 5, pp. 1023–1031, 2022. [Online]. Available: <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/jrs.6320>
- [18] M. A. Morris, D. Pagonis, L. B. Algrim, J. A. De Gouw, J. L. Jimenez, and P. J. Ziemann, "Absorption of volatile organic compounds (VOCs) by polymer tubing: implications for indoor air and use as a simple gas-phase volatility separation technique," *Atmospheric Measurement Techniques*, vol. 17, pp. 1545–1559, 2024.
- [19] P. H. C. Eilers and H. F. M. Boelens, "Baseline correction with asymmetric least squares smoothing," *Leiden University Medical Centre Report*, vol. 1, no. 1, pp. 1–5, 2005.
- [20] I. E. Gordon *et al.*, "The HITRAN2020 molecular spectroscopic database," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 277, p. 107949, 2022.
- [21] D. A. Long, *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*. Chichester: John Wiley & Sons, 2002.
- [22] R. B. Boulton, V. L. Singleton, L. F. Bisson, and R. E. Kunkee, *Principles and Practices of Winemaking*. New York: Chapman & Hall, 1996.
- [23] H. W. Schrötter and H. W. Klöckner, "Raman scattering cross sections in gases and liquids," in *Raman Spectroscopy of Gases and Liquids*, ser. Topics in Current Physics, A. Weber and S. P. S. Porto, Eds. Berlin: Springer, 1979, vol. 11, pp. 123–166.
- [24] R. Boulton, J. Nelson, and A. Knoesen, "A review of sensors for the monitoring, modeling, and control of commercial wine fermentations," *Fermentation*, vol. 11, no. 6, p. 329, 2025. [Online]. Available: <https://www.mdpi.com/2311-5637/11/6/329>