## **Supporting Information**

# Comprehensive multi-gas investigation by means of fiber-enhanced Raman spectroscopy for the investigation of nitrogen cycle processes.

Annika Blohm<sup>2</sup>, Christian Domes<sup>2</sup>, Andreas Merian<sup>1</sup>, Sebastian Wolf<sup>2</sup>, Jürgen Popp<sup>2</sup>, and Torsten Frosch<sup>1,2,\*</sup>

<sup>1</sup> Biophotonics and Biomedical Engineering Group, Technical University Darmstadt, Merckstraße 25, 64283 Darmstadt, Germany
<sup>2</sup> Leibniz Institute of Photonic Technology, 07745 Jena, Germany

\*Corresponding author: <u>torsten.frosch@tu-darmstadt.de</u>, ORCID 0000-0003-3358-8878

Supporting Information

Material and Methods: Data Evaluation

## Figure SI 1

Raman spectra of the respective reference gases  $CO_2$  (1),  $N_2$  (2),  $N_2O$  (3),  $O_2$  (4), and  $NH_3$  (5).

#### Material and Methods:

#### Data Evaluation

All preprocessing and analysis of the raw Raman data were conducted using statistical programming GnuR 3.6.1.<sup>1</sup> The packages 'Peaks'<sup>2</sup>, 'minpack.Im'<sup>3</sup>, and 'MCR-ALS'<sup>4</sup> were utilized and their functions were complemented. The Raman raw data was truncated to the wavenumber region of interest (550 cm<sup>-1</sup> - 3050 cm<sup>-1</sup>, except for NH<sub>3</sub>, where the range was 1450 cm<sup>-1</sup> - 3750 cm<sup>-1</sup>), wavenumber corrected using the strongest signals of N<sub>2</sub>, O<sub>2</sub>, CO<sub>2</sub>, and N<sub>2</sub>O as reference, and baseline corrected using SNIP algorithm<sup>5</sup> (iterations = 20, order = 2).

For quantification of N<sub>2</sub>, O<sub>2</sub>, N<sub>2</sub>O, CO<sub>2</sub>, and NH<sub>3</sub>, various concentration series of the pure gases were prepared and a Lorentzian peak profile was fitted to the corresponding marker bands at 2331 cm<sup>-1</sup>, 1555 cm<sup>-1</sup>, 1284 cm<sup>-1</sup>, 1387 cm<sup>-1</sup>, and 3333 cm<sup>-1</sup>. The resulting peak areas were normalized for pressure and laser power and correlated with their set concentrations, showing a linear relation (Fig: 1). The calibration curves were used to predict the concentrations of the multi-gas samples (Tab. 3). The lab-prepared multi-gas samples with predetermined concentrations enable quantification and precision assessment using the calibration. For reproducible quantification performance, a daily standard sample of synthetic air (Linde Gas,  $80\%N_2$  and  $20\% O_2$ ) was measured. The parameters of the gas-specific calibration curves were adjusted using correction factors that were established by comparing calculated (via calibration curves) and actual (80% and 20%) N<sub>2</sub> and O<sub>2</sub> concentrations.

Furthermore, the limit of detection (LoD) was calculated for every gas by utilizing the signal-to-noise ratio of the lowest concentration step: LoD =  $3 \cdot \frac{c}{SNP} = 6$ .

Due to overlapping peaks in the spectra of CO<sub>2</sub> and N<sub>2</sub>O (Fig. 3A), calibration was improved using the MCR-ALS algorithm <sup>7, 8</sup>. Therefore, the Raman data of different CO<sub>2</sub> and N<sub>2</sub>O mixtures was baseline corrected (same parameters, *vide supra*) and truncated to the range of 1250 cm<sup>-1</sup> - 1450 cm<sup>-1</sup>. For the MCR algorithm, the highest concentrations of CO<sub>2</sub> and N<sub>2</sub>O were utilized as an initial guess, the resulting response was correlated with their respective concentrations and subsequently utilized as calibration curve, enabling the prediction of the concentrations of N<sub>2</sub>O and CO<sub>2</sub> (Fig. 3B).

Additionally, the reference spectra for all measured gases (100% for all gases except NH<sub>3</sub>, here 5%) were obtained using two grating positions, combined, baseline-corrected (same parameters, *vide supra*), logarithmically scaled, and min-max-normalized for better visualization.





Fig. S1: Raman spectra of the respective reference gases  $CO_2$  (1),  $N_2$  (2),  $N_2O$  (3),  $O_2$  (4), and  $NH_3$  (5). The peaks used for subsequent quantification are labeled and the spectra were log-scaled for better visualization.

### References

- (1) *R: A Language and Environment for Statistical Computing*; 2019. https://www.R-project.org/ (accessed.
- (2) Peaks: Peaks; 2012. https://CRAN.R-project.org/package=Peaks (accessed.
- (3) minpack.Im: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds; 2016. https://CRAN.R-project.org/package=minpack.Im (accessed.
- (4) Mullen, K. M.; Mullen, M. K. Package 'ALS'. 2015.
- (5) Ryan, C.; Clayton, E.; Griffin, W.; Sie, S.; Cousens, D. SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. *Nucl. Imstrum. Methods Phys. Res., Sect. B* **1988**, *34* (3), 396-402.
- (6) Kaiser, H. Zum Problem der Nachweisgrenze. Fresenius Z. Anal. Chem. 1965, 209 (1), 1-18.
- (7) van Stokkum, I. H.; Mullen, K. M.; Mihaleva, V. V. Global analysis of multiple gas chromatographymass spectrometry (GC/MS) data sets: A method for resolution of co-eluting components with comparison to MCR-ALS. *Chemom. Int. Lab. Sys.* **2009**, *95* (2), 150-163.
- (8) Tauler, R. Multivariate curve resolution applied to second order data. *Chemom. Int. Lab. Sys.* **1995**, *30* (1), 133-146.