



Short communication

Simulation and iterative deconvolution of residual gas spectra

Berthold Jenninger^{*}, Antoine Benoit, Paolo Chiggiato

CERN European Organization for Nuclear Research, 1211, Geneva, 23, Switzerland



ARTICLE INFO

Keywords:

Residual gas analysis
Deconvolution mass spectra
Simulation mass spectra
Ultrahigh vacuum
Contaminants

ABSTRACT

Residual gas analysis may be time consuming if the identification of all gas species and their quantitative contribution to the mass spectrum is required. This is because the partial pressures of the residual gas cover several orders of magnitudes and fragmentation patterns are in general convoluted. Noise, offset, and the limited mass ranges of the analysers used in UHV applications further reduce the sensitivity of a spectrum. In this work, the authors propose a method to simulate residual gas spectra based on fragmentation patterns and partial pressures. With an iterative algorithm, residual gas spectra can be deconvoluted in logarithmic scale. The use of indicators simplifies the identification of residual gas compositions. The authors realised this algorithm in a prototype application. The performance of this tool is encouraging and opens the path for the development of an UHV specific web-based application.

1. Introduction

The purpose of residual gas analysis (RGA) is the identification and quantification of partial pressures in a vacuum system, in particular traces of contaminations. The instruments used for this objective are quadrupole mass spectrometers (QMS) [1]. They give valuable information about the cleanliness of a vacuum system and processes under vacuum [2,3]. QMS became affordable and they are nowadays installed in many ultrahigh vacuum (UHV) systems. The typical QMS spectrum shows an ion current signal as a function of mass-charge ratio in steps that are a fraction of an atomic mass unit (amu). The residual gas species have partial pressures that can cover several orders of magnitude and the fragmentation patterns in general overlap in the mass spectra. Consequently, a thorough interpretation of the mass spectrum requires its deconvolution. There exist a number of mass spectrometry software for analytical chemistry. The web site hereafter gives an overview of existing software in this field and [4] describes the status and perspective of deconvolution methods in gas chromatography mass spectrometry-based metabolomics https://en.wikipedia.org/wiki/List_of_mass_spectrometry_software#cite_note-68 (accessed Oct 2020).

However, those algorithms are generally combined with gas separation methods and the analysers do not work at the pressure limits of the system. The methods used in analytical chemistry to deconvolute coeluted species in gas chromatography mass spectrometry are not practicable with the instruments used in UHV. Obviously, little have been published in the field of residual gas analysis. The authors did not

find publications dealing with algorithm-based automated residual gas analysis. One may mention an early study by Belič and Gyergyek [5] about using neural networks for mass spectra recognition. This paper is interesting in this context, because mass spectra are generated by random linear combinations of mass spectra signals.

In this work, we try to introduce a practical approach for an algorithm-based residual gas analysis. Residual gas spectra are simulated based on fragmentation patterns and expected gas species. In an iterative algorithm, we compare simulated with measured spectra and minimise their difference in a normalised logarithmic scale. The use of specific indicators orients users to discard or include additional species from the analysis. A detailed description about the reconstruction of residual gas spectra can be found in the CERN Accelerator School tutorial: https://indico.cern.ch/event/565314/contributions/2285748/attachments/1467497/2281208/CAS_Tutorial_RGA_Reconstructi_on_Mass_spectra.pdf.

1.1. Simulation of residual gas spectra

A residual gas analyser measures the ion current at a given mass over charge ratio (m/e), here simply referred to as mass. At a given mass, the ion current is the sum of the contributions of all species present in the residual gas [6]. The linear equations that describe a mass spectrum are expressed in Eq. (1).

^{*} Corresponding author.

E-mail address: berthold.jenninger@cern.ch (B. Jenninger).

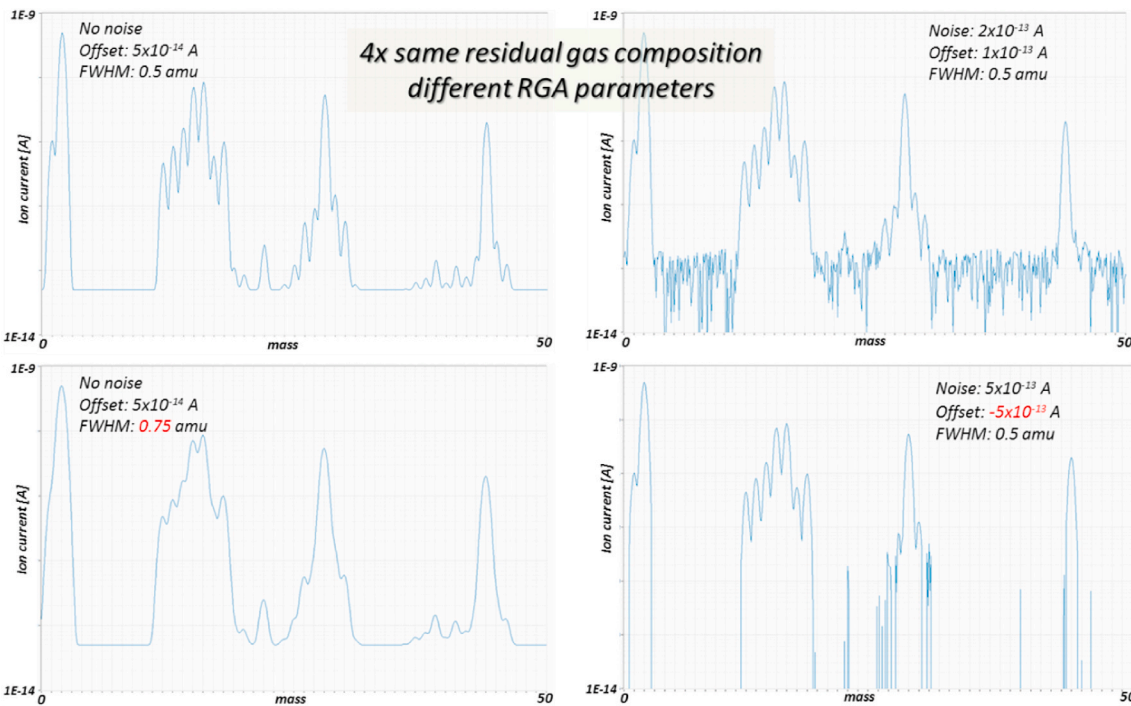


Fig. 1. Four calculated residual gas spectra. The spectra have the same composition but differ in the analyser setting. A negative offset leads to an apparent clean spectrum without noise. However, the information about small peaks up to the amount of the offset is lost.

$$\begin{aligned}
 I_1 &= c_{1,1}\alpha_1 P_1 + c_{1,2}\alpha_2 P_2 + c_{1,j}\alpha_j P_j + \dots + c_{1,N}\alpha_N P_N I_2 = c_{2,1}\alpha_1 P_1 + c_{2,2}\alpha_2 P_2 \\
 &+ c_{2,j}\alpha_j P_j + \dots + c_{2,N}\alpha_N P_N I_k = c_{k,1}\alpha_1 P_1 + c_{k,2}\alpha_2 P_2 + c_{k,j}\alpha_j P_j + \dots \\
 &+ c_{k,N}\alpha_N P_N I_z = c_{z,1}\alpha_1 P_1 + c_{z,2}\alpha_2 P_2 + c_{z,j}\alpha_j P_j + \dots + c_{z,N}\alpha_N P_N
 \end{aligned} \quad (1)$$

I_k is the ion current at mass k (ranging from 1 to z , with z being the mass range) and P_j is the partial pressure of the species j (ranging from 1 to N). α_j is the sensitivity of the residual gas analyser to the species j (2); it is the ratio of the ion current at a reference mass of a species j ($\Delta I_{j, \text{ref}}$) to the variation of the partial pressure of the same species (ΔP_j). The reference mass is the one that displays the largest signal for a given gas.

$$\alpha_j = \frac{\Delta I_{j, \text{ref}}}{\Delta P_j} \quad (2)$$

The sensitivities are instrument-specific and depend on its settings. They must be determined by specific calibrations if precise partial pressure measurements are required [7]. Fragmentation patterns are single component spectra, normalised to the reference mass. $c_{k,j}$ is the relative abundance of a fragment at mass k with respect to the reference peak of the species j . Eq. (1) can be written in a matrix form.

$$\begin{bmatrix} I_1 \\ I_k \\ \vdots \\ I_z \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,j} & \dots & c_{1,N} \\ c_{k,1} & c_{k,j} & \dots & c_{k,N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{z,1} & c_{z,j} & \dots & c_{z,N} \end{bmatrix} \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_j & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_N \end{bmatrix} \begin{bmatrix} P_1 \\ P_j \\ \vdots \\ P_N \end{bmatrix} \quad (3)$$

Ion currents, partial pressures, sensitivities and fragmentation patterns appear in separated matrix.

Most often, the analysers are not calibrated for all species in the spectrum. This is not necessary if the identification of species and an approximate estimate of its abundance is required. In that case, it is enough to use the average relative sensitivities α_j/α_i with respect to a reference species i (e.g. N_2) for a type of analyser and settings (see Eq. (4)). The relative sensitivities vary up to about 40% between analysers of different types [8]. Further, if one can accept uncertainties of up to a factor of around 2 or 3, the relative sensitivities may all be set to one and the relative sensitivity matrix of Eq. (4) disappears.

$$\begin{bmatrix} I_1 \\ I_k \\ \vdots \\ I_z \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,j} & \dots & c_{1,N} \\ c_{k,1} & c_{k,j} & \dots & c_{k,N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{z,1} & c_{z,j} & \dots & c_{z,N} \end{bmatrix} \begin{bmatrix} \frac{\alpha_1}{\alpha_i} & 0 & \dots & 0 \\ 0 & \frac{\alpha_j}{\alpha_i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\alpha_N}{\alpha_i} \end{bmatrix} \begin{bmatrix} P_1 \\ P_j \\ \vdots \\ P_N \end{bmatrix} \alpha_i \quad (4)$$

Fragmentation patterns also depend on analyser settings that modify electron impact energy and ion distribution inside the ionisation cell. For precise partial pressure measurements, fragmentation patterns should therefore be determined by individual calibrations. For common use, it is however good practice, to use the fragmentation patterns from large and publicly available libraries as those provided by NIST [9]. Once data are given to the partial pressure vector, the spectrum obtained by Eq. (4) is a bar graph in the mass range. Analogue mass scans are in general preferred as they provide information on analyser-specific parameters such as mass alignment, mass resolution, noise and offset that are not visible in bar-graph spectra.

Analogue scans can be reproduced in our model transforming bars into Gaussian curves for each of the signal I_k , and adding ion current offset and noise. In this way, the full mass spectrum is described by Eq. (5).

$$I(m) = \sum_{k=0}^z I_k e^{-\frac{(m-m_k)^2}{2\sigma^2}} + I_{\text{offs}} + I_{\text{noise}} \quad (5)$$

m_k is the integer atomic mass and m is the continuous variable. Like in real instruments, m is changed by small steps (e.g. around 0.1 amu) resulting in the analogue spectrum $I(m)$. I_{noise} is a random number calculated for each mass increment Δm . I_{offs} corresponds to the electronic offset of the instrument; it can be both positive and negative. For the mass resolution we use the definition of Full Width at Half Maximum (FWHM). The parameter c is then determined as shown in Eq. (6). Fig. 1 shows examples of calculated analogue mass scans.

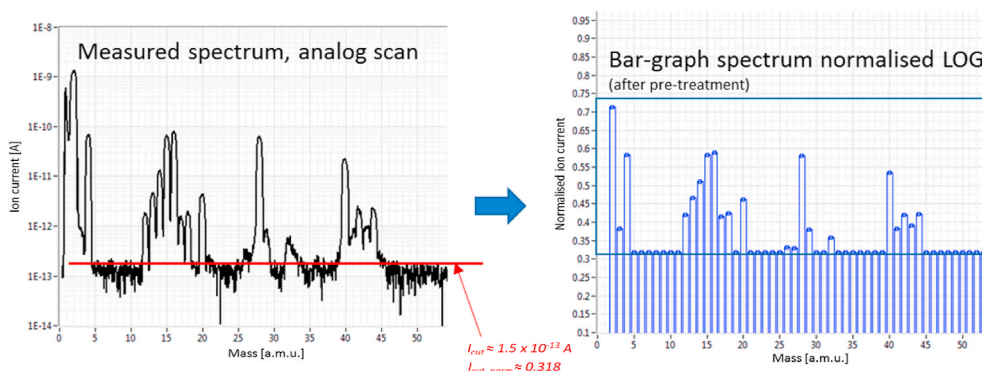


Fig. 2. Analogue raw input scan and the corresponding pre-analysis treated bar-graph spectrum.

$$c = \frac{FWHM}{2\sqrt{2\ln 2}} \quad (6)$$

It shall be noted that throughout this work we work with integer m/e . Measured analogue mass spectra, however, can also include multiple ionised species of particularly stable radicals, molecules or atoms, resulting in signals at non-integer m/e . In principle, one could also include those in the simulations. Then, the index k would no more correspond to the ion m/e . The fragmentation pattern library from NIST only provides the integer m/e . Therefore, for clarity, we preferred also to stick on integer m/e in the simulations. The pre-analysis treatment of a measured analogue mass spectrum (see further below) will reveal non-integer signals, which then give helpful hints for the identification of the corresponding gas species.

The maxima of the measured peaks are in general not exactly at integer mass. To fit the calculated spectrum in the measured one, a pre-analysis treatment is necessary. This consist of attributing the signal peaks to the correct integer mass and applying a cut-off limit. Signal peaks that are identified as multiple ionisation at non-integer masses are

omitted, as they are not part of the fragmentation pattern library. The cut-off is the limit at which a signal peak can just still be reasonably distinguished from noise. All values below that limit are set to this limit (not to zero). In the case of a negative offset with no noise visible, it is good practice to set this limit to the level of the smallest visible signal. In UHV, we are interested in traces, such as those originating from contaminations. Therefore, calculating the difference between pre-treated measured and calculated spectra are done in the logarithmic scale. Signals from traces have then the same weight in the calculations as those from the dominant species. The spectra are normalised so that the expected ion currents are limited between 0 and 1. The following normalisation is applied (7):

$$I_{norm} = \frac{\text{Log}(I) + 16}{10} \quad (7)$$

In this way 10 orders of magnitude of ion currents between 10^{-16} to 10^{-6} A are represented in the normalised spectrum. Fig. 2 illustrates the pre-analysis treatment.

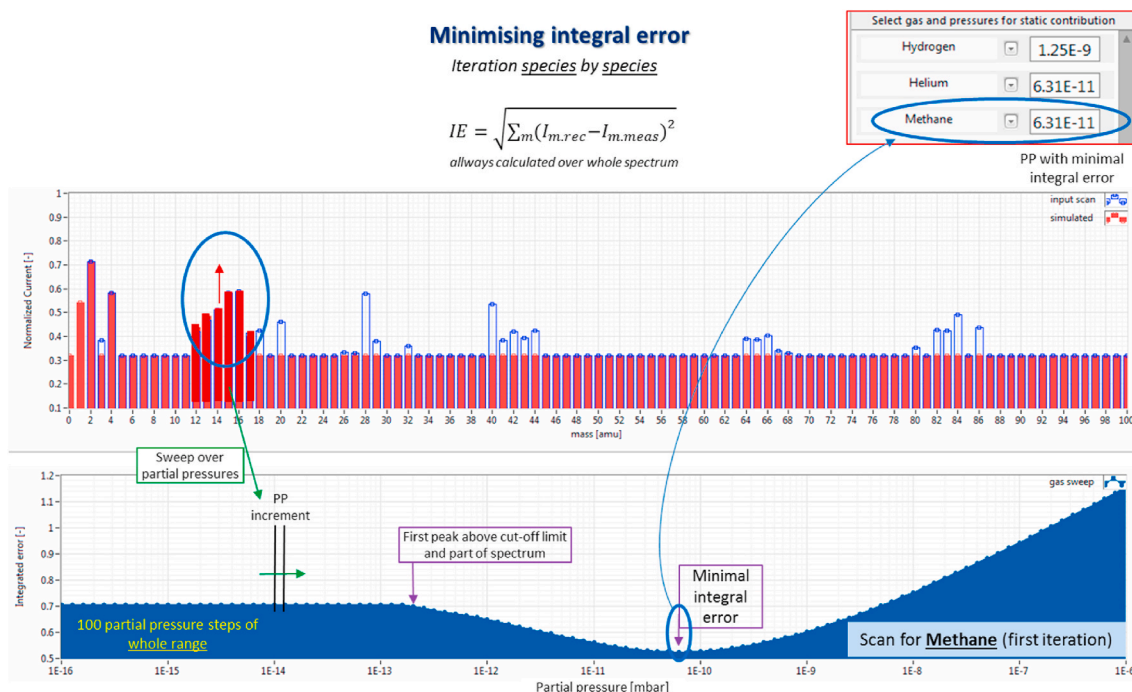


Fig. 3. The upper graph compares a pre-treated input spectrum (blue) with a calculated spectrum (red). Only three species have yet been considered (H_2 , He, CH_4). The lower graph illustrates the evolution of the integral error (IE) as the partial pressure of one species (methane) increases. The entire partial pressure range is here divided in 100 equal steps in logarithmic scale. The partial pressure at which the error is minimal is kept for the calculation with the following range. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

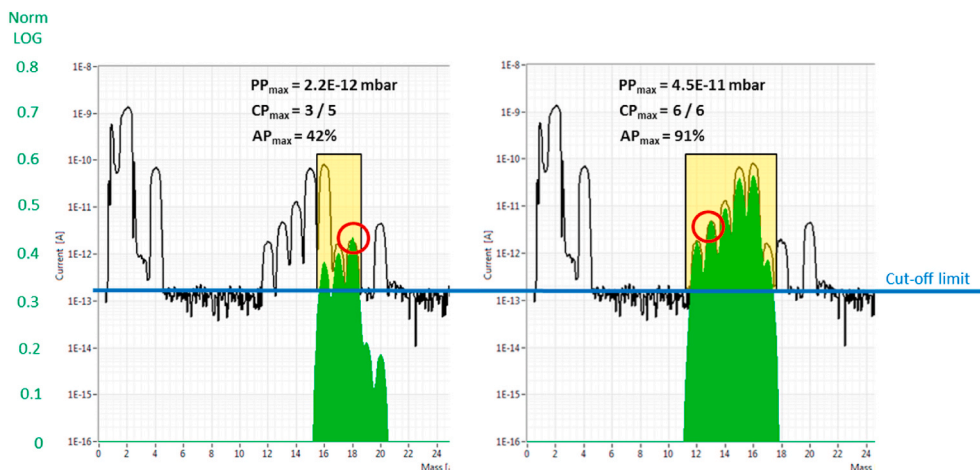


Fig. 4. Visualisation of the indicators of two species (H_2O and CH_4) in a measured scan ($\alpha = 1 \text{ A mbar}^{-1}$). The circles indicate the position where the first peak of the species exceeds the measured spectrum. The yellow rectangles show the peaks to calculate AP_{\max} . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

1.2. Deconvolution

In the case that all gas species contributing to the spectrum are known, the mass spectrum can be deconvoluted by successively varying the partial pressures through the entire range, species by species. This is done in small increments. At each increment, the mass spectrum is calculated, transformed into the normalised logarithmic scale, and the integral error IE between the calculated and pre-treated measured spectrum is determined (8). The same cut-off limit applies in both the simulated and pre-treated measured spectra.

$$IE^{\text{norm}} = \sqrt{\sum_m (j_{m_calc}^{\text{norm}} - j_{m_mes}^{\text{norm}})^2} \quad (8)$$

A full iteration process consists of several rounds of iterations. The different steps are described below. For each species, the partial pressures (PP) are always calculated in increments over the whole pressure range (PP -scan). Such a PP -scan may have one or several minima. The PP , which presents the lowest IE , corresponds to the best fitting PP at the presence of the other species selected in the process so far at their corresponding PP . This PP is selected and kept for the scan with the following species. In this way IE can only decrease or remain constant between species. The process can be automated. The iteration steps are:

1. Starting with species A , calculate IE for all PP -increments.
2. Keep the PP of species A , at which IE is minimal, and calculate IE for species B for all PP steps.
3. Keep the PP of species A and B , at which IE s were minimal, and calculate IE for species C for all PP increments. This is situation shown in Fig. 3.
4. Continue this way for all species that have been selected for the analysis.
5. Repeat steps 1 to 4 several times, but start step 1 of the new round each time with the PP s after finishing step 4 of the previous round. In case that species have convoluted fragmentation patterns, the partial pressures at which IE presents minima shift between rounds and the minimal IE further decrease. After a few full rounds of iterations, the method converges to a global minimum that does not decrease further.

The method can comprise many gas species with large fragmentation pattern to help identifying residual gas compositions. After a few rounds of iterations, the partial pressures of most species that are not present in the measured spectrum should be reduced to values that are close to the

one that corresponds to the ion current cut-off (9).

$$P_{\text{cut}} = \frac{I_{\text{cut}}}{\alpha} \quad (9)$$

α is the sensitivity of the instrument in [A mbar^{-1}]

This is, however, not always the case. We therefore introduced indicators that help reduce the selection of species to be included for the analysis. They are calculated for all species in the library. These indicators are:

PP_{\max} (maximum possible partial pressure): This is the partial pressure of a gas generating at least one peak exceeding the contour of the pre-treated input spectrum.

CP_{\max} (peak contribution at maximum partial pressure): This is the ratio of the number of peaks above the cut-off limit that would contribute to the spectrum at PP_{\max} to the number of peaks in the fragmentation pattern of the selected gas.

AP_{\max} (integral ratio at maximum partial pressure): This is the ratio of peaks' integral for the selected gas at PP_{\max} to the integral of the same peaks of the pre-treated measured spectrum.

The integral of a spectrum is calculated above the cut-off limit (10).

$$S^{\text{norm}} = \sum_m (j_k^{\text{norm}} - j_{\text{cut}}^{\text{norm}}) \quad (10)$$

The indicators, normalisation and integral of spectra are illustrated in Fig. 4.

1.3. Residual gas analysing steps in the prototype tool

In the prototype application, the measured analogue scan is uploaded in the calculation tool as a text file. A pre-analysis treatment is performed within the application. In a first step, after the pre-treatment, the program calculates the indicators for all species in the fragmentation pattern library. This allows to eliminate species for which maximum partial pressure is close to the cut-off limit or having small peak contributions. The indicators can be visualised in the measured spectrum as shown in Fig. 4. For the remaining selection, we run the iterative deconvolution process. This further eliminates some species from the selection. Species that are unlikely to be present may remain in the selection when the measured spectrum cannot be entirely reconstructed with the species in the fragmentation pattern library. Either the measured spectrum is composed with species that are not part of the library or the fragmentation patterns given by the instrument differs

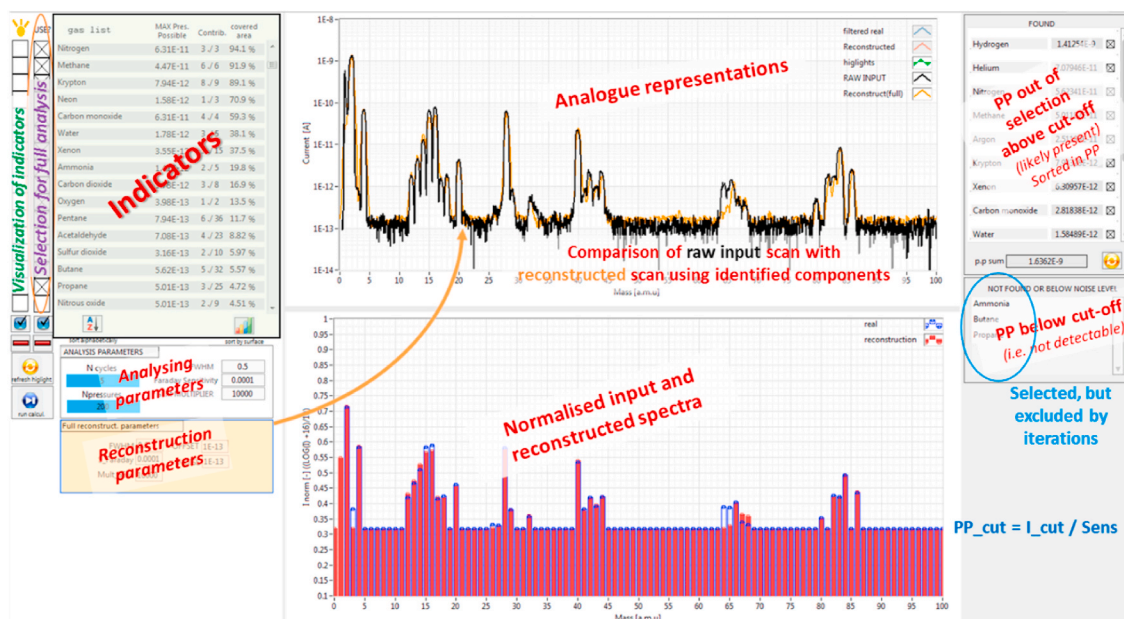


Fig. 5. Main user interface of the prototype application. The upper graph compares the raw input spectrum (black) with the reconstructed scan based on the identified components (brown). Mass resolution, noise and offset are included in the simulated scan. The lower graph shows the input spectrum after pre-analysis treatment (blue) and the calculated spectrum, both in normalised logarithmic scales. Those spectra are used for the integrated error calculations. The left top corresponds to species from the library and their corresponding indicators, sorted as decreasing surface coverage indicator AP_{\max} . The right top shows the result of the analysis with the corresponding partial pressures (PP). To see all species, one needs to scroll through the lists. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

from those of the library. Some final evaluations are then required by the user to judge whether or not the remaining species need to be removed for other reasons. The main user interface of this prototype analysing tool is shown in Fig. 5.

1.4. Time needed for the analysis

For this study, we used a rather small fragmentation pattern library of 80 species. A full iterative deconvolution process that includes all the species in this library converged within less than 5 iterations. In total, this implied 80,000 spectra simulations and error calculations (i.e. 20 partial pressure steps per species and decade, 10 decades, 80 species, 5 iterations). On our desktop PC (CPU 3.4 GHz, RAM 32 GB, 64-bit Operating System) this took around 6 s. The determination of the indicators took less than 2 s. A real-world application with a UHV-dedicated library would involve several hundred species. By extrapolation and assuming a library of 1000 species and 10 iterations, we estimate the time for a full library analysis to around 150 s (2×10^6 simulations) and 25 s to determine the indicators. This is still a reasonable time for such an analysis. If only the reconstruction with a selection of known species is required, the deconvolution process takes only a fraction of a second.

2. Conclusion and future perspective

With this study, we could demonstrate that the iterative deconvolution is a valuable option for the analysis and interpretation of residual gas spectra. The possibility to include medium-size fragmentation patterns libraries opens the way to a web-based residual gas analysing application dedicated to the UHV community. Such an application could be a platform for the exchange of experience about vacuum contaminations. Participating institutes may make available fragmentation pattern libraries with spectra of contaminations they experienced. Users of the application could then include such libraries in their analysis.

In addition, the fact that we can simulate residual gas spectra based on partial pressures and fragmentation patterns allows the training of

machine-learning applications with a huge number of simulated spectra. The number of spectra necessary for training would not be available by measurement. This option has been investigated in a separate feasibility study between CERN and Intelligent Data Analysing Laboratory IDAL, Spain [10]. The positive experience we made, encourage us to develop further this iterative deconvolution and possibly combine this application with machine-learning options.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors warmly thank Vincent Baglin, Adriaan Rijllart and Odd Oyvind Andreassen for their valuable support and trust in this work.

References

- [1] H. Peter, in: Dawson (Ed.), *Quadrupole Mass Spectrometry and its Applications*; Book, Elsevier Scientific Publishing Company, Amsterdam-Oxford-New York, 1976.
- [2] G.F. Weston, Measurement of ultra-high vacuum. Part 2. Partial pressure measurement, *Vacuum* 2 (1979), [https://doi.org/10.1016/S0042-207X\(80\)80045-5](https://doi.org/10.1016/S0042-207X(80)80045-5).
- [3] James E. Blessing, Robert E. Ellefson, Bruce A. Raby, Gerardo A. Brucker, Robert K. Waits, Recommended practice for process sampling for partial pressure analysis, *J. Vac. Sci. Technol.* 25 (2007) 167, <https://doi.org/10.1116/1.2364001>.
- [4] Xiuxia Du, Steven H. Zeisel, Spectral deconvolution for gas chromatography mass spectrometry-based metabolomics: current status and future perspectives, *Comput. Struct. Biotechnol. J.* 4 (5) (January 2013).
- [5] I. Belić, L. Gyergyek, Neural network methodologies for mass spectra recognition, *Vacuum* 48 (7–9) (1997) 633–637.
- [6] R D Craig, E H Harden; The interpretation of mass spectra in vacuum measurements; *Vacuum*, volume 16, number 2.
- [7] 2018 ISO, *Vacuum Technology Vacuum Gauges – Characterization of Quadrupole Mass Spectrometers for Partial Pressure Measurement*, 2017.

- [8] J. Miertusova, Reliability and accuracy of total and partial pressure measurements in the UHV pressure range under real experimental conditions, *Vacuum* 51 (1) (1998) 61–68.
- [9] <https://webbook.nist.gov/chemistry>.
- [10] B. Jenninger, F. Mateo, Summary report on collaboration contract KE3129/TE between CERN and IDAL, Study of the feasibility and usefulness of machine-learning methods to help identifying residual gas composition, CERN-ACC-NOTE-2020-0052, <https://cds.cern.ch/record/2740627>.