



**HAL**  
open science

## Diagnosis and correction methods for spectral interference in the framework of LIBS imaging

Ludovic Duponchel, Alexandre Cugerone, Manuel Muñoz, Vincent Motto-Ros

► **To cite this version:**

Ludovic Duponchel, Alexandre Cugerone, Manuel Muñoz, Vincent Motto-Ros. Diagnosis and correction methods for spectral interference in the framework of LIBS imaging. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2023, 207, pp.106758. 10.1016/j.sab.2023.106758 . hal-04381652

**HAL Id: hal-04381652**

**<https://hal.science/hal-04381652>**

Submitted on 9 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Diagnosis and correction methods for spectral interference in the framework of LIBS imaging

Ludovic Duponchel<sup>†\*</sup>, Alexandre Cugerone<sup>II</sup>, Manuel Muñoz<sup>‡</sup>, Vincent Motto-Ros<sup>‡\*</sup>

<sup>†</sup> Univ. Lille, CNRS, UMR 8516 – LASIRE – Laboratoire de Spectroscopie pour Les Interactions, La Réactivité et L'Environnement, Lille, F-59000, France.

<sup>II</sup> Department of Earth Sciences, University of Geneva, Geneva, Switzerland.

<sup>‡</sup> Géosciences Montpellier, Université de Montpellier, CNRS, Montpellier, France

<sup>‡</sup> Institut Lumière Matière, UMR 5306, Université Lyon 1 - CNRS, Université de Lyon 69622 Villeurbanne, France.

## Corresponding Authors

\* ludovic.duponchel@univ-lille.fr / vincent.motto-ros@univ-lyon1.fr

**Keywords :** Spectral interferences, Diagnosis, Correction, Signal unmixing, Laser-Induced Breakdown Spectroscopy (LIBS), Imaging.

---

**ABSTRACT:** Laser-Induced Breakdown Spectroscopy (LIBS) has become a powerful imaging technique for elemental characterization in analytical chemistry due to its advantages over other techniques. Major, minor, and trace elements are detected with high measurement dynamic, a low limit of detection and a high acquisition rate, allowing for the quick analysis of large sample surfaces. Today, chemometric tools are commonly used to ensure the most comprehensive and unbiased exploration of such spectroscopic data. However, the integration of the signal from a wavelength assumed to be specific to the element of interest remains the basic

tool for generating a chemical distribution map from a hyperspectral dataset. This classical approach is based on a strong assumption, the specificity of the chemical information on the spectral domain being considered. Any spectral interference inevitably result in the generation of a biased distribution image. In this publication, we demonstrate how Principal Component Analysis (PCA) can diagnose the potential presence of a spectral interference and how Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) can ultimately correct it if necessary using a LIBS imaging dataset obtained from the analysis of a complex rock sample. The proposed approach combines the simplicity and effectiveness of the integration method with the diagnostic and correction capabilities of chemometric tools, providing a comprehensive solution for spectral interference in LIBS imaging.

---

## 1. INTRODUCTION

Laser-Induced Breakdown Spectroscopy (LIBS) is increasingly establishing itself as a powerful imaging technique for the elemental characterization of complex materials in analytical chemistry [1]. This is due to a set of advantages that few analytical techniques can match today. Beyond a relatively simple sample preparation, this technique allows for the detection of major, minor, and even trace elements (from % to ppm for many elements). In addition to this high measurement dynamic and the low limit of detection, its acquisition rate can now reach up to 1000 Hz, i.e. 1000 LIBS spectra acquired in one second. As a consequence, multiple square centimeters of a sample surface is possible to analyse in few hours, while maintaining a spatial resolution of approximately 10 microns. Under these conditions, we can easily end up with over a million spectra acquired on a sample, and it is only natural that chemometric tools are more and more exploited to ensure the most comprehensive and unbiased exploration. This includes unsupervised methods such as Principal Component Analysis (PCA) [2,3], clustering

(such as  $k$ -means) [4,5], and signal unmixing (such as Multivariate Curve Resolution - Alternating Least Squares, MCR-ALS) [6,7] to name just a few. The last two references cited is also an opportunity to highlight another advantage of LIBS imaging, namely its multimodality through fully optical instrumentation on a single microscope. The LIBS community is well aware today that the use of multivariate tools derived from chemometrics allows for simultaneous exploitation of all the wavelengths in the spectral domain with the least possible prior assumptions. However, it should be remembered that, as any spectroscopic imaging technique, the basic tool for generating a chemical distribution map from a hyperspectral dataset is the integration of the signal from a wavelength (or a small spectral range around it) assumed to be specific to the element in question [8,9]. It is an opportunity to emphasize that we should not oppose chemometric tools and this classical approach, but instead use them in parallel during an exploration as they provide different perspectives on our complex samples. On the other hand, the simplicity of the integration method should not make us forget that it is based on a strong assumption, which is the specificity of the chemical information on the spectral domain being considered. This assumption is all the more important in spectroscopic imaging, since any spectral interference, i.e., the presence of any unwanted chemical species in the considered spectral range, will inevitably result in the generation of a biased distribution image. Such an image could, for example, show overconcentrations of the element of interest or even, in the worst case, areas that appear to contain it when it is not present. This highlights the crucial role of spectral specificity in this imaging context. Based on the richness of emission lines contained in a LIBS spectrum and their low bandwidth, it would be a mistake to believe that an isolated and characteristic wavelength of the element of interest is always possible to find. Indeed, we remind here the excellent detection limits of LIBS but also the chemical complexity of the samples we analyze. Moreover, even if we always have a preconceived notion about the presence of certain elements in a given sample, it is difficult to know all the elements present within

it. In fact, our goal is to strive for this objective when we implement such a spectroscopic technique. This does not mean that we should abandon this signal integration method for generating our elemental maps. Indeed, it remains fast and most of the time effective and relevant, but we should combine a diagnostic method to detect the presence of any interference in the investigated spectral domain. Based on a positive diagnosis, it will then be necessary to propose a correction strategy for this interference in order to obtain a corrected image for the element of interest. The entire purpose of this publication is to address these two points, namely the diagnosis and correction of spectral interferences. The main interest of our approach lies in the fact that we do not develop new chemometric algorithms to address this issue. We propose to use only PCA and MCR-ALS, not over the whole spectral domain as it is usually done, but only over a restricted spectral range around the wavelength of the element of interest. Based on a LIBS imaging dataset obtained from the analysis of a complex rock sample composed of germanium and gallium zoned sphalerite, we will demonstrate how PCA can be used to diagnose the potential presence of spectral interference, as well as how MCR-ALS can ultimately correct it.

## **2. MATERIAL AND METHODS**

### **2.1 Origin of the sample and its preparation**

The rock sample studied was collected in the Pyrenean Axial Zone, in the Arre Pb-Zn vein deposit which was investigated to understand the concentration of associated critical elements such as Ge and Ga [10]. This rock contains sphalerite (ZnS) minerals deformed by regional metamorphism, which show heterogeneous distribution and zonation in Ge and Ga. The rock sample was prepared as a 30  $\mu\text{m}$  polished thin section. It is mostly composed of quartz, metamorphic micas and light-, to dark-brown sphalerite mineral enriched in germanium (up to  $\sim 600$  ppm Ge) and gallium (up to  $\sim 150$  ppm Ga; LA-ICP-MS analyses [11]). The zoning in

germanium and gallium appears very complex with microscale Ge-Ga rich patches, and a recent micro-, to nano-scale study shows the ubiquitous occurrence of Ge-Cu minerals in sphalerite, hosting most of the germanium (i.e., briartite; [12]).

## **2.2 LIBS instrumental setup and spectral acquisition**

The used LIBS experimental setup has already been described in several papers [13–15]. Briefly, the experiment was conducted by using a Nd:YAG laser with a 8 ns pulse duration, operating at 100 Hz at the fundamental wavelength (Centurion GRM, Quantel). A typical energy of 0.6 mJ per pulse was used. The laser beam was focused onto the sample by a 15x magnification objective. In this configuration, the crater size was typically 7  $\mu\text{m}$  in diameter for one shot. The measurements were performed at room temperature, under ambient pressure, and with an argon gas flowing the plasma region (0.8 l/min). The laser scanning was performed line by line in raster scan mode with the use of a motorized XYZ stage. The step size between two consecutive acquisitions was 13  $\mu\text{m}$ . The light emitted by the plasma was collected by a quartz lens and focused onto the entrance of a round-to-linear fiber bundle (composed of 19 fibers, each with a 200  $\mu\text{m}$  core diameter) connected to the spectrometer. The spectrometer model was a Shamrock 500 spectrometer (Andor Technology) coupled with an intensified charge-coupled device (ICCD) camera (iStar, Andor Technology). This spectrometer was equipped with a 1200 l/mm grating to cover (blazed at 300 nm) the 260-300 nm spectral range with a resolution around 0.08 nm. This range was selected for detecting in priority germanium (Ge), gallium (Ga), zinc (Zn), iron (Fe), silicon (Si) and aluminum (Al). The ICCD camera was synchronized to the Q-switch of the laser. The LIBS acquisition was performed with a delay of 700 ns and a gate of 5000 ns. Finally, a homemade software developed in the LabVIEW environment was used to control the entire setup, allowing automatic sequences of any selected regions of interest with a preset lateral resolution. The spectral dataset considered in the following represent 1.2 million of spectra (sequence of 2400x500) and was recorded in less than 4 hours.

### 2.3 Interference diagnosis and correction

In order to address the proposed issue, we suggest using PCA and MCR-ALS sequentially. Thus, PCA will be used to determine whether there is a spectral interference present in the considered spectral domain, and if so, the MCR-ALS method will enable us to separate the spectral contribution of the interferent and the element of interest. The goal here is not to develop these two very classical chemometric approaches in detail, but rather to reintroduce them briefly and focus on specific aspects of this diagnosis and interference correction.

PCA was first developed by Pearson in 1901 [16] and later by Hotelling in 1933 [17]. It is a widely used statistical method that aims to simplify the complexity of a high-dimensional data by transforming it into a lower-dimensional space. It is a technique for identifying patterns and relationships in data by reducing the dimensionality of the dataset while retaining the most important information. PCA achieves this by extracting the principal components, which are linear combinations of the original variables that capture the maximum amount of variance in the data. The resulting components are orthogonal and uncorrelated, making it easier to analyze and interpret the data. The singular value decomposition (SVD) [18], which was developed in the 1960s, is undoubtedly the most widely used method for PCA today. Given  $\mathbf{D}$ , an  $n \times m$  matrix containing  $n$  spectra, each defined by  $m$  wavelengths, its singular value decomposition is calculated as follow:

$$\mathbf{D} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^t \quad (\text{eq. 1})$$

The matrix  $\mathbf{U}$  (of size  $n \times p$ ) contains the  $p$  scores of each spectrum, which are their new coordinates in the PCA space. The maximum value that  $p$  can take is simply the minimum between the values of  $n$  and  $m$ . In particular, if we chose this maximum value for  $p$ , we could express 100% of the total variance of the matrix  $\mathbf{D}$ . However, we will see later why we never want to

extract 100% of the total variance and thus this maximum allowed number of components. In the context of spectroscopic imaging, the scores are usually represented as scores maps in order to observe their spatial distributions over the sample region of interest [2]. However, this is not really the information that we will be exploiting in this work. Actually, the diagnosis of a potential spectral interference will focus on the utilization of matrices  $\mathbf{S}$  and  $\mathbf{V}$ . The matrix  $\mathbf{S}$  (of size  $p \times p$ ) contains the singular values arranged in decreasing order along its diagonal. Then, squaring the first singular value gives the variance expressed by the first principal component. The same apply to the successive eigenvalues associated with their principal component. Our goal, when we use PCA, is to find the number of significant principal components that best describe our original data. Thus, we are searching for the intrinsic dimension of our dataset, known as the rank of matrix  $\mathbf{D}$ , which actually corresponds to the number of independent spectral sources present in it. This rank is obtained through the observation of all the calculated eigenvalues, as represented in what is called a scree plot. If the matrix  $\mathbf{D}$  contained noiseless data, the evaluation of its rank would be trivial since it would simply correspond to the number of non-zero eigenvalues. This specific situation is depicted in the eigenvalues scree plot of Figure 1a, for which we clearly have an observable rank of 2.

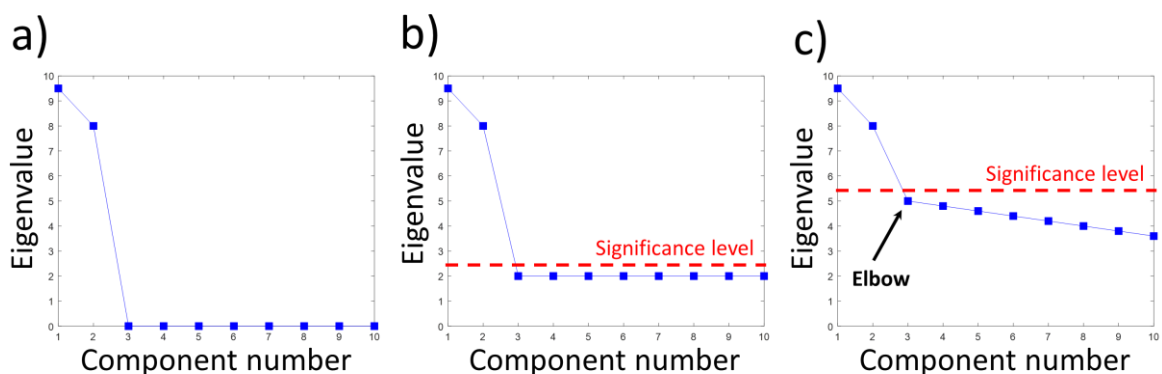


Figure 1: Eigenvalues scree plots considering different levels of noise: a) noiseless data, b) a good signal to noise ratio, c) a lower signal to noise ratio.



As part of a PCA applied to a small wavelength range, this would indicate here the presence of two spectral contributions and thus a potential interference between two signals. When we explore experimental data, they inevitably contain noise. This noise accounts for a certain explained variance that will be associated with its own eigenvalues. This situation is shown in Figure 1b, for which we have relatively low noise on our spectral data, indicating a good signal-to-noise ratio. In these actual conditions, there are no longer any zero eigenvalues, and we must find a threshold eigenvalue below which we will consider that the principal components are no longer significant. This task remains fairly simple here since the eigenvalues associated with the noise are found on a distinct plateau separated from the first two significant singular values. When the signal-to-noise ratio further decreases, the eigenvalues associated with the noise will increase to approach those corresponding to the chemical information. The plateau will no longer be observed as shown in Figure 1c but the "elbow" of the graph where the eigenvalues seem to level off will help us to select the threshold. As we can see, this technique allows us to estimate the rank of a matrix and therefore the number of spectral contributions present in the explored spectral domain. The final choice of this rank will of course not be based solely on this observation of eigenvalues, but also on a consistent spectral confirmation found in the principal components contained in the  $\mathbf{V}^t$  matrix (of size  $p \times m$ ). So to summarize, if the estimated rank of the  $\mathbf{D}$  matrix is 1, there will be no interference and we can proceed directly and unambiguously with signal integration for generating an elemental map. Otherwise, interference will occur and we will need to apply a source separation method (such as MCR-ALS) to generate a corrected elemental map free from this perturbation, which we describe in the next section.

The MCR-ALS method was developed in the 90s to unmix complex signals acquired in analytical chemistry [19,20]. The goal of a signal unmixing technique is quite simple as from only a set spectra (potentially containing mixtures), it is possible to simultaneously extract all

spectra of pure species present in the chemical system being studied and their relative contributions in each spectrum. It should be emphasized here that this unmixing is achieved without any prior knowledge of the chemical system or even the particular shape of the spectra. So based on the previous rank estimation  $k > 1$ , this source separation method allows to decompose the data matrix  $\mathbf{D}$  (of size  $n \times m$ ) considering a bilinear model:

$$\mathbf{D} = \mathbf{C} \cdot \mathbf{S}^t \quad (\text{eq. 2})$$

with  $\mathbf{C}$  (of size  $n \times k$ ) the pure contribution matrix and  $\mathbf{S}^t$  (of size  $k \times m$ ) the matrix containing pure spectra. MCR uses constrained alternating regressions to obtain this decomposition. Only the non-negativity constrain on  $\mathbf{C}$  and  $\mathbf{S}$  is used during the calculations, which is natural for LIBS spectral data. For instance, if we had an element that could potentially interfere with another element (i.e., a rank equal to 2), matrix  $\mathbf{C}$  would allow us to obtain specific distribution images for each element, while matrix  $\mathbf{S}^t$  would help us confirm the consistency of the spectral information taken into account. The computations necessary for PCA and MCR-ALS were executed using the MATLAB environment. However, readers who are intrigued by this methodology can readily explore other programming languages since these algorithms are classical in nature. It should be noted that only raw data will be used in proposed chemometric pipeline.

### 3. RESULTS AND DISCUSSION

We have selected this complex rock sample here to illustrate different interference situations that can be observed on both major and minor elements, as well as different signal structures. The mean spectrum of the dataset in question is shown in Figure 2a, where we can see the elements Si, Zn, Fe, Mg, Al, and Ge. The first element of interest in this study is silicon, whose most intense emission line is located around 288.158 nm.

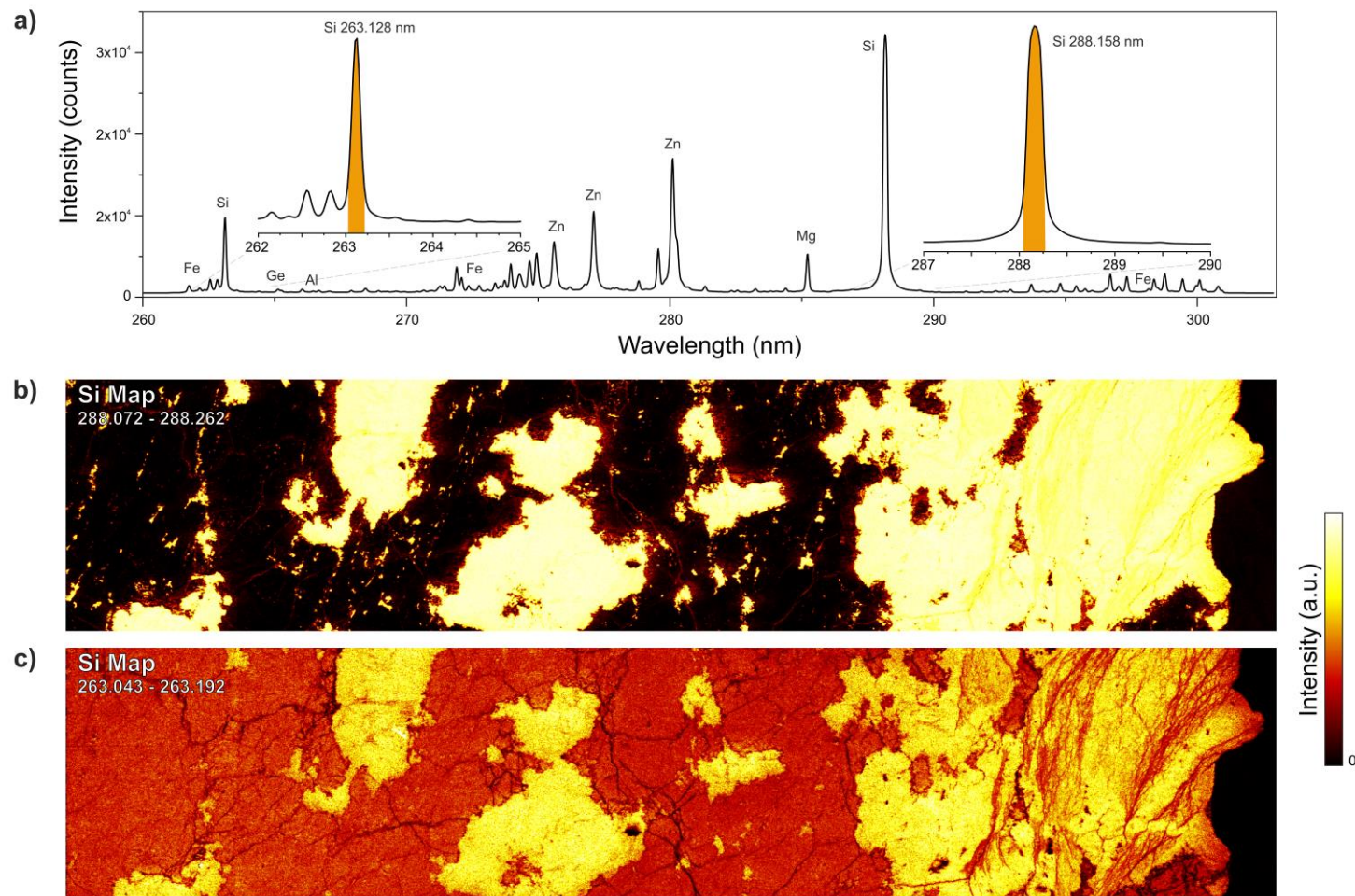


Figure 2: A first exploration of the dataset. a) The mean spectrum of the dataset over the whole spectral range and the two spectral integration regions considered for silicon. b) The Si distribution map obtained from the signal integration over the 288.0721 – 288.2621 nm spectral range. c) The Si distribution map obtained from the signal integration over the 263.0432 - 263.1922 nm spectral range.

Unfortunately, from the mean spectrum, it is apparent that numerous spectra in the dataset are certainly saturated at this wavelength. This saturation is related to the detector's dynamic range, and therefore to the maximum value it can effectively measure under these given acquisition conditions. It is, of course, unacceptable to use a saturated signal since the acquired value is no longer proportional to the concentration of the element in question. We could of course implement a correction procedure for this saturation, as we have done in our previous work by considering a statistical imputation [21], but that is not the strategy that interests us here. Figure 2b provides an integration image obtained by considering a spectral domain from 288.0721 to 288.2621 nm (i.e. the most intense line of Si). Although the spectra may be saturated, such an image allows observation of the areas of the sample that contain silicon and mostly correspond in our case to quartz and micas. However, this saturation phenomenon breaks the concentration dynamic that we should potentially observe with a representation that is quite binary, meaning without too much detail that doesn't really account for the fine variations in Si concentration between the pixels of the image. Just like in photography, we could also say that the image is saturated. Of course, such a situation is unsuitable if our goal is to accurately report the full complexity of our sample, and it is only natural that we need to work on another emission line of silicon, specifically centered at 263 nm. We can zoom in on this spectral region in the mean spectrum of Figure 2a. It can be observed that the average intensity of the silicon emission line is much lower, which avoids the previous saturation problem. As a consequence, the new Si distribution map obtained from the signal integration over the 263.0432 - 263.1922 nm spectral range (Figure 2c) seems to highlight more details in the Si-rich mineral phases. On the other hand, a new observation of the mean spectrum in Figure 2a shows that we might have another issue because numerous iron contributions are observed in this spectral sub-domain, and there is a probability that it could interfere with the considered silicon line. This hypothesis is far from unfounded, as shown in Figure 2c where a certain concentration of Si, outside the Si-rich

mineral phases seems now to be observed, whereas it was almost absent in the same area in Figure 2b. So, let's get to the heart of the matter of this publication as we need to diagnose this potential interference between Fe and Si. The goal here is not to perform a PCA on the entire spectral domain as usual, but only on a range from 261.2526 nm to 263.9793 nm that encompasses the contribution of silicon which is of interest to us. Figure 3 shows the eigenvalues scree plot resulting from this analysis.

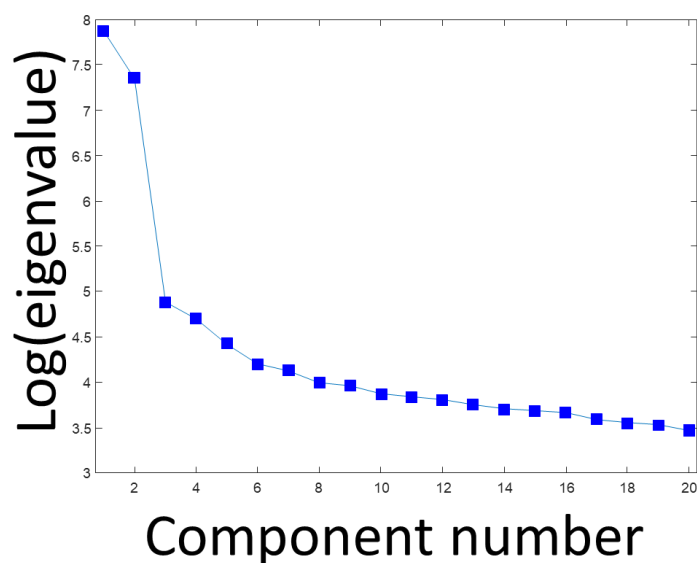


Figure 3: The eigenvalues scree plot obtained from PCA applied on the 261.2526 - 263.9793 nm spectral range.

Distinguishing the eigenvalues related to noise from the first two that represent chemical information is a straightforward task. This rank of 2 indicates that there is indeed an interference between two signals in this reduced spectral domain. This information is corroborated by the information contained in the first two principal components and their associated scores maps (Figure 4). Firstly, it should be noted that the first two principal components effectively capture the chemical information contained within the dataset, as they together account for approximately 99.68 % of the total variance. Overall, the first principal component highlights the areas

of the sample where there is relatively more Si and at the same time less Fe. The second principal component, on the other hand, highlights the phenomenon of interference with a derivative shape observed around 263 nm, and more specifically a positive contribution located at 263.0858 nm and another negative one at 263.1709 nm. Based on this diagnosis of interference between Si and Fe in this spectral domain, it is now appropriate to attempt to separate their respective spectral contributions in order to obtain an image of Si distribution that is consistent with the analytical reality of the sample. It is certainly a task that cannot be performed by PCA since by construction the principal components are perpendicular, which is never the case between two elemental spectra. This is how we will use a spectral unmixing method in the next step, specifically the MCR-ALS method based on a rank equal to two and a non-negativity constraint on the profiles during the decomposition optimization. Moreover, default values for the number of iterations and convergence criterion have been considered in this work because they are suitable for the vast majority of LIBS spectral datasets. Figure 5 displays the two pure spectra extracted from matrix  $S^t$  and the two associated distribution maps extracted from matrix  $C$ . The pure component 1 corresponds perfectly to the spectrum of silicon, which indeed exhibits a single emission line in this spectral range centered around 263.13 nm. The pure component 2, on the other hand, perfectly extracts the spectrum of iron with its numerous emission lines, the most intense one being located around 263.11 nm. It is obvious that the interference between these two elements could not be more important, since there is an almost complete overlap between this last emission line and that of silicon, the two being separated by 0.0212 nm in these measurement conditions. This result is all the more remarkable as we make no assumptions about the shape of the pure spectra or even the profile of the emission lines during this decomposition of the data set. There should be no ambiguity on this point, as MCR-ALS is a data-driven analysis that has nothing to do with a spectral fitting procedure. Thus, the extrac-

tion of the first pure map allows us to see the distribution of silicon without the spectral interference of iron. By comparing this map to that of Figure 2c, it is easy to understand that in the latter we were simultaneously observing the contributions of both elements, while we thought we were only observing silicon.

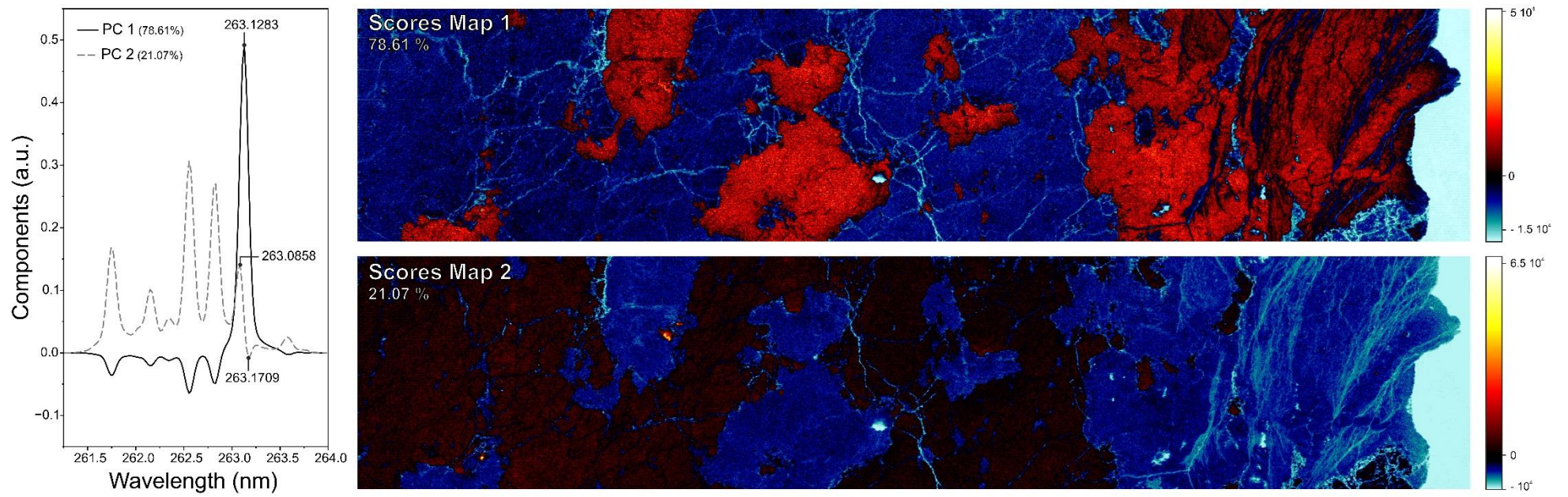


Figure 4: The first two principal components and the associated scores maps



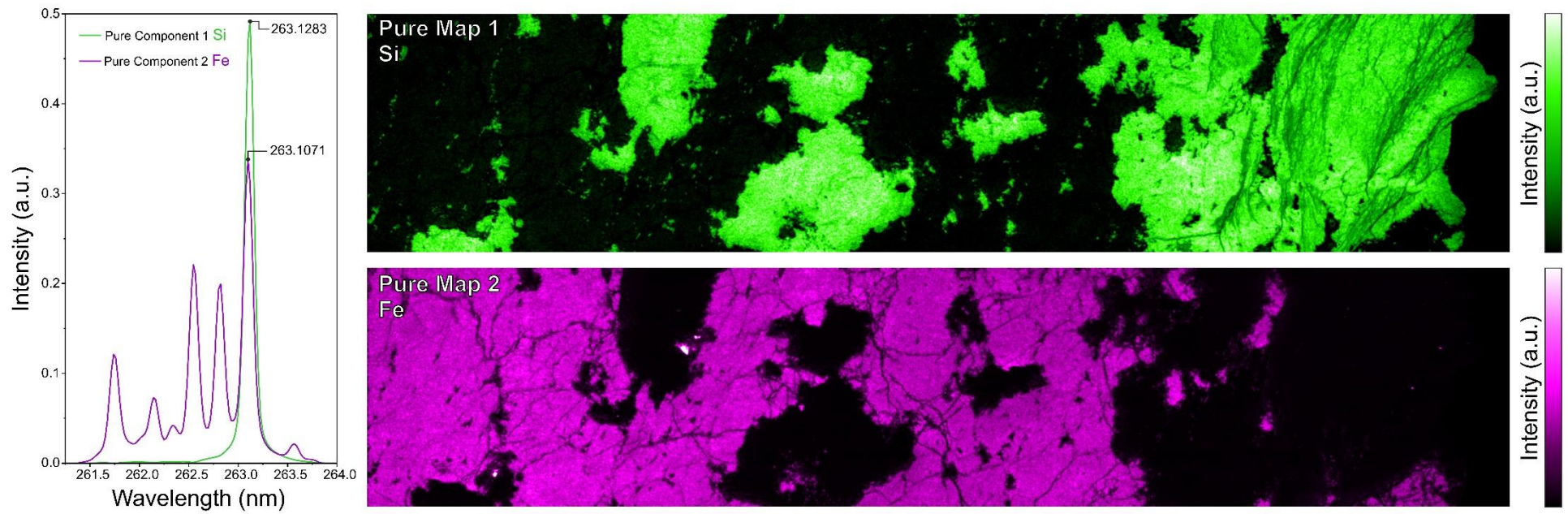


Figure 5: Extractions of the two pure spectral profiles and the two associated distribution maps.

Finally, if we revisit the Si distribution map extracted by MCR-ALS, we can still observe that there is a low concentration of Si outside the Si-rich mineral phases, which corresponds to trace amounts of Si in sphalerite minerals (i.e., the Fe-rich areas), and which could not have been detected otherwise than by this procedure. From this first example, we can thus see the simplicity and effectiveness of the proposed methodology, first in diagnosing interference and also in its correction, as well.

Let us now turn our attention to a second scenario focused on the search for a minor element in this same sample, namely gallium. Actually, based on the knowledge that geologists have about this particular sample, gallium is an element that could potentially be present in small quantities. This element is known to have an emission line around 394.42 nm, but it potentially interferes with iron which is particularly present in this sample. Figure 6a shows the mean spectrum of the dataset between 294.3042 nm and 294.6399 nm, on which a single line appears to be observed. Figure 6b provides a distribution map obtained by integrating the signal on this spectral range. We notice that this distribution map is quite similar to the iron one extracted by MCR-ALS, as depicted in Figure 5. At first glance, everything suggests that only iron is present on this spectral window.

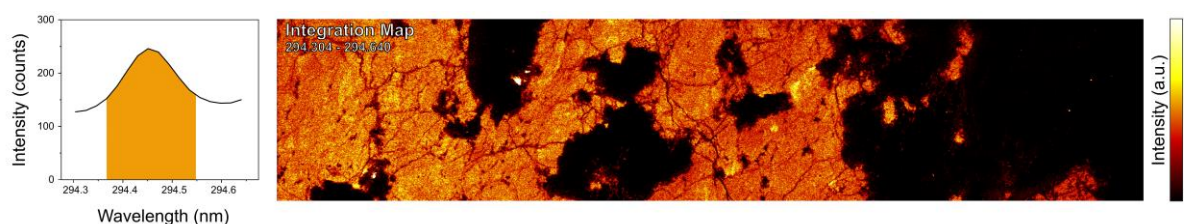


Figure 6: a) The mean spectrum of the dataset between 294.3 and 294.65 nm and, b) , nm and the integration map over the same spectral range.

In order to perform a more thorough diagnosis, we conduct a PCA on this spectral domain, with Figure 7a displaying the calculated eigenvalues.

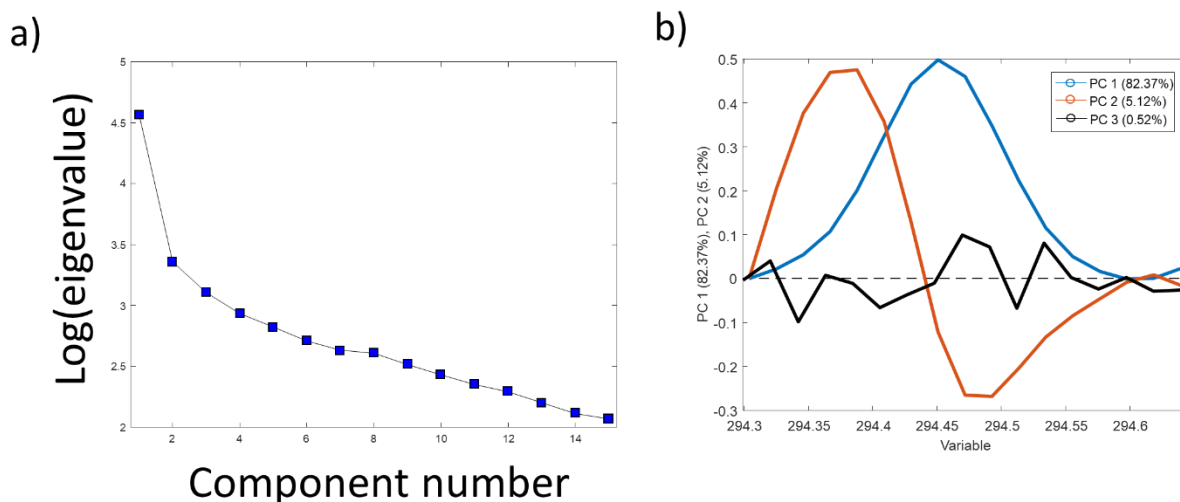


Figure 7: a) The eigenvalues scree plot obtained from PCA applied on 294.3042 - 294.6399 nm spectral range and, b) the two first principal components.

In this scenario, only one eigenvalue appears to be significant, which seems to support our previous observations. However, this is not consistent with the first two principal components shown in Figure 7b. In fact, if the rank was indeed 1, the second component should only show noise and certainly not be structured as it is here. Moreover, we can see the specific derivative shape characteristic of an interference between two peaks in the second principal component. The rank is therefore once again equal to 2. At first glance, one might think that using the scree plot is not a good strategy for evaluating the rank of the spectral data matrix. Actually, this technique is very effective when the variance of the chemical species is much greater than the noise one. It was particularly the case in the first scenario between Si and Fe, where principal components 1 and 2 expressed 76.55 and 23.12 % of the variance, respectively, with 0.33 % of residual variance associated with noise. In this new case, components 1 and 2 expressed 82.37 and 5.12 % of the variance, respectively, with 7.51 % of residual variance associated with noise. Actually, we do have a contribution from gallium expressed in the second principal component, but its low concentration results in a very small variance compared to the noise, making it impossible to detect it solely from the scree plot. In summary, rank evaluation has therefore to be

done by associating the observation of the eigenvalues scree plot and principal components. As a final step, the MCR-ALS extractions considering this rank of 2 are given in Figure 8. So the first pure component extracts an iron emission line, while the second extracts the gallium one, both separated by only 0.063 nm. The first extracted map effectively shows the distribution of iron as already seen before. More interestingly here, we can observe the distribution of gallium in the second extracted map even though its concentration is very low but nevertheless well localized in certain areas of the mineral phase considered. By observing the extracted spectra, it could be surprising to have emission lines for Fe and Ga that are quite comparable in intensity. In fact when using the MCR-ALS method, there is naturally what is called an intensity ambiguity on the extracted pure spectra. In order to remove it, we enforce the extracted spectral profiles to have a norm equal to 1 during the optimization. That is why the extracted intensities for Fe and Ga in Figure 8 appear very close. This is not very serious in this specific context, as the main goal is to identify emission lines by their maximum. However, if we wanted to give a quantitative character for a comparison of the two maps, we would need to rescale them as these two emission lines have different spectroscopic properties. The latter result further highlights the value of the proposed approach, as it is impossible to generate an unbiased distribution map of gallium using the classical integration method.

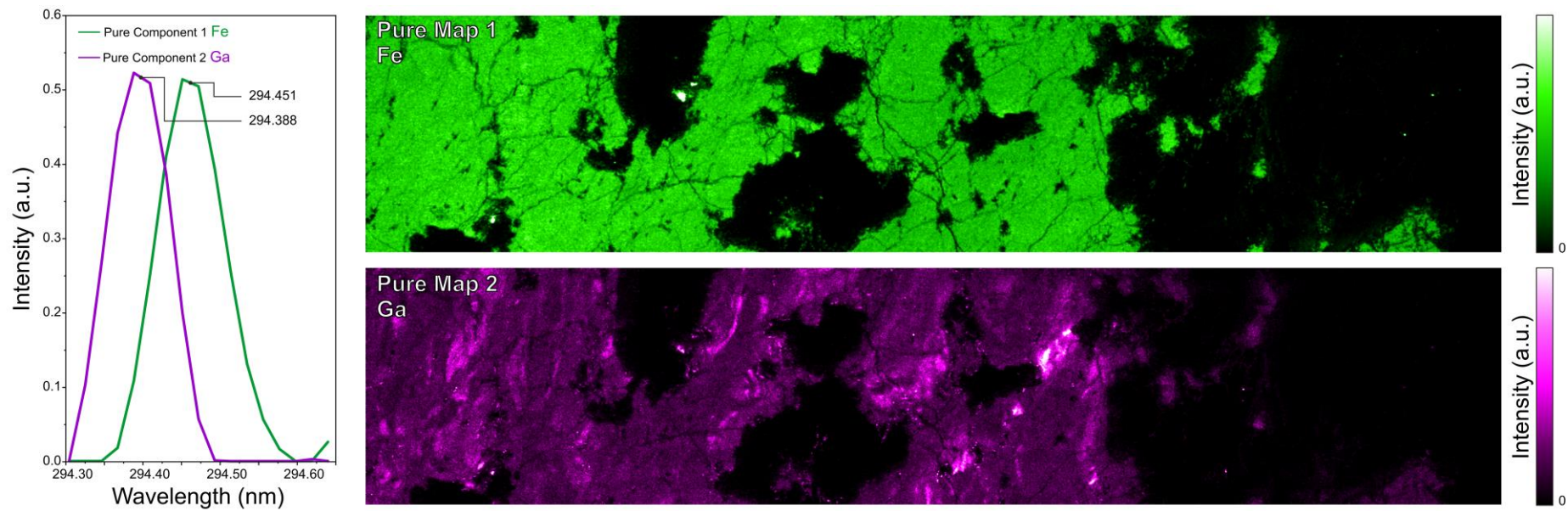


Figure 8: Extractions of the two pure spectral profiles for Fe and Ga, and the two associated distribution maps.

#### 4. CONCLUSIONS

LIBS imaging, thanks to its high acquisition speed and sensitivity, is now an elemental imaging technique that has its place among the analytical characterization tools for complex materials. Similar to other spectroscopic imaging techniques, its fundamental principle is based on integrating the signal at a particular wavelength of the element to be imaged. Although a LIBS emission line has a limited bandwidth, the complexity of the explored samples and the richness of LIBS spectral information inevitably induce interference problems. In order to be able to generate representative compound distribution images that are consistent with analytical reality, it is therefore necessary to check if such interference is present, and if so, propose an appropriate signal correction. Through two different scenarios, we first demonstrated how to use PCA to estimate the rank of a spectral matrix, which is the number of independent signals it can contain. The rank was then used to diagnose the presence of interference around the emission line that we wanted to exploit. We also observed that an effective evaluation of this rank can only be achieved through simultaneous interpretation of the eigenvalues scree plot and the principal components, especially when the compound of interest has a low concentration and/or the signal-to-noise ratio is limited. It was subsequently demonstrated that based on confirmed interference, the MCR-ALS method could resolve this issue by separating the contributions of the target element and the interferent both spectrally and spatially.

## 5. REFERENCES

- [1] V. Gardette, V. Motto-Ros, C. Alvarez-Llamas, L. Sancey, L. Duponchel, B. Busser, Laser-Induced Breakdown Spectroscopy Imaging for Material and Biomedical Applications: Recent Advances and Future Perspectives, *Anal. Chem.* 95 (2023) 49–69. <https://doi.org/10.1021/acs.analchem.2c04910>.
- [2] S. Moncayo, L. Duponchel, N. Mousavipak, G. Panczer, F. Trichard, B. Bousquet, F. Pelascini, V. Motto-Ros, Exploration of megapixel hyperspectral LIBS images using principal component analysis, *J. Anal. At. Spectrom.* 33 (2018) 210–220. <https://doi.org/10.1039/C7JA00398F>.
- [3] R. Finotello, M. Tamaazousti, J.-B. Sirven, HyperPCA: A powerful tool to extract elemental maps from noisy data obtained in LIBS mapping of materials, *Spectrochim. Acta Part B At. Spectrosc.* 192 (2022) 106418. <https://doi.org/10.1016/j.sab.2022.106418>.
- [4] A. Nardecchia, C. Fabre, J. Cauzid, F. Pelascini, V. Motto-Ros, L. Duponchel, Detection of minor compounds in complex mineral samples from millions of spectra: A new data analysis strategy in LIBS imaging, *Anal. Chim. Acta.* 1114 (2020) 66–73. <https://doi.org/10.1016/j.aca.2020.04.005>.
- [5] K. Kiss, A. Šindelářová, L. Krbal, V. Stejskal, K. Mrázová, J. Vrábel, M. Kaška, P. Modlitbová, P. Pořízka, J. Kaiser, Imaging margins of skin tumors using laser-induced breakdown spectroscopy and machine learning, *J. Anal. At. Spectrom.* 36 (2021) 909–916. <https://doi.org/10.1039/D0JA00469C>.
- [6] A. Nardecchia, A. de Juan, V. Motto-Ros, M. Gaft, L. Duponchel, Data fusion of LIBS and PIL hyperspectral imaging: Understanding the luminescence phenomenon of a complex mineral sample, *Anal. Chim. Acta.* 1192 (2022) 339368. <https://doi.org/10.1016/j.aca.2021.339368>.

- [7] A. Nardecchia, A. de Juan, V. Motto-Ros, C. Fabre, L. Duponchel, LIBS and Raman image fusion: An original approach based on the use of chemometric methodologies, *Spectrochim. Acta Part B At. Spectrosc.* 198 (2022) 106571. <https://doi.org/10.1016/j.sab.2022.106571>.
- [8] C. Fabre, K. Trebus, A. Tarantola, J. Cauzid, V. Motto-Ros, P. Voudouris, Advances on microLIBS and microXRF mineralogical and elemental quantitative imaging, *Spectrochim. Acta Part B At. Spectrosc.* 194 (2022) 106470. <https://doi.org/10.1016/j.sab.2022.106470>.
- [9] C. Marina-Montes, V. Motto-Ros, L.V. Pérez-Arribas, J. Anzano, M. Millán-Martínez, J.O. Cáceres, Aerosol analysis by micro laser-induced breakdown spectroscopy: A new protocol for particulate matter characterization in filters, *Anal. Chim. Acta.* 1181 (2021) 338947. <https://doi.org/10.1016/j.aca.2021.338947>.
- [10] A. Cugerone, B. Cenko-Tok, E. Olliot, M. Muñoz, F. Barou, V. Motto-Ros, E. Le Goff, Redistribution of germanium during dynamic recrystallization of sphalerite, *Geology.* 48 (2020) 236–241. <https://doi.org/10.1130/G46791.1>.
- [11] A. Cugerone, B. Cenko-Tok, M. Muñoz, K. Kouzmanov, E. Olliot, V. Motto-Ros, E. Le Goff, Behavior of critical metals in metamorphosed Pb-Zn ore deposits: example from the Pyrenean Axial Zone, *Miner. Deposita.* 56 (2021) 685–705. <https://doi.org/10.1007/s00126-020-01000-9>.
- [12] D. Fougereuse, A. Cugerone, S.M. Reddy, K. Luo, V. Motto-Ros, Nanoscale distribution of Ge in Cu-rich sphalerite, *Geochim. Cosmochim. Acta.* 346 (2023) 223–230. <https://doi.org/10.1016/j.gca.2023.02.011>.
- [13] L. Sancey, V. Motto-Ros, B. Busser, S. Kotb, J.M. Benoit, A. Piednoir, F. Lux, O. Tillement, G. Panczer, J. Yu, Laser spectrometry for multi-elemental imaging of biological tissues, *Sci. Rep.* 4 (2014) 6065. <https://doi.org/10.1038/srep06065>.
- [14] J.O. Cáceres, F. Pelascini, V. Motto-Ros, S. Moncayo, F. Trichard, G. Panczer, A. Marín-Roldán, J.A. Cruz, I. Coronado, J. Martín-Chivelet, Megapixel multi-elemental imaging



- by Laser-Induced Breakdown Spectroscopy, a technology with considerable potential for paleoclimate studies, *Sci. Rep.* 7 (2017) 5080. <https://doi.org/10.1038/s41598-017-05437-3>.
- [15] V. Motto-Ros, E. Negre, F. Pelascini, G. Panczer, J. Yu, Precise alignment of the collection fiber assisted by real-time plasma imaging in laser-induced breakdown spectroscopy, *Spectrochim. Acta Part B At. Spectrosc.* 92 (2014) 60–69. <https://doi.org/10.1016/j.sab.2013.12.008>.
- [16] K. Pearson, On lines and planes of closest fit to systems of points in space, *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 2 (1901) 559–572. <https://doi.org/10.1080/14786440109462720>.
- [17] H. Hotelling, Analysis of a complex of statistical variables into principal components., *J. Educ. Psychol.* 24 (1933) 417–441. <https://doi.org/10.1037/h0071325>.
- [18] G. Golub, W. Kahan, Calculating the Singular Values and Pseudo-Inverse of a Matrix, *J. Soc. Ind. Appl. Math. Ser. B Numer. Anal.* 2 (1965) 205–224. <https://doi.org/10.1137/0702016>.
- [19] R. Tauler, Multivariate curve resolution applied to second order data, *Chemom. Intell. Lab. Syst.* 30 (1995) 133–146. [https://doi.org/10.1016/0169-7439\(95\)00047-X](https://doi.org/10.1016/0169-7439(95)00047-X).
- [20] R. Tauler, A. Smilde, B. Kowalski, Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution, *J. Chemom.* 9 (1995) 31–58. <https://doi.org/10.1002/cem.1180090105>.
- [21] A. Nardecchia, V. Motto-Ros, L. Duponchel, Saturated signals in spectroscopic imaging: why and how should we deal with this regularly observed phenomenon?, *Anal. Chim. Acta.* 1157 (2021) 338389. <https://doi.org/10.1016/j.aca.2021.338389>.