# Semi-Automatic Elemental Identification and Quantification by Laser–Induced Breakdown Spectroscopy.

## DOCTORADO EN CIENCIAS (OPTICA)

*Estudiante: M en O. Ignacio Raúl Rosas Román*
*Asesor: Dr. Marco Antonio Meneses Nava*

*Septiembre de 2017*
*León, Guanajuato, México*
*Versión definitiva. Incluye cambios sugeridos por revisores*

*To my sons, the brightest stars in obscure times.*

# Acknowledgments

# Abstract

Chemical analysis has deep effects in our everyday lives, it is not only used as a tool for quality control, but also for designing better materials to meet everyday needs. One of the optical techniques used in recent times for chemical analysis has been the Laser-Induced Breakdown Spectroscopy (LIBS); an atomic emission spectroscopy technique with many advantages being its low cost, lack of sample preparation and reduced acquisition times among the most important. LIBS makes use of a strong focused electromagnetic field to produce a plasma, whose spectral emission is a signature that uniquely identifies every chemical element contained in the sample. Automatic elemental identification is a task that has not been resolved completely since human intervention is still required. In this dissertation, a new procedure for semi-automatic identification of emission species in LIBS spectra was developed, where major elements are identified automatically. Although minor emission elements and traces still need human intervention, it is enough to relax the search conditions in the proposed procedure to obtain a high identification rate.

LIBS is not only employed to determine sample's elemental content but also to perform quantitative analysis. Since its origins, LIBS qualitative analysis has relied on external standards; i.e. materials with known composition, which are used to obtain calibration curves. Although this method is still the most precise in terms of quantification limits, there are situations where standards are very expensive or nearly impossible to obtain. An alternative route is provided by the so called standarless quantification procedures, such as Calibration-Free (CF), that relieves the requirement of an external standard at expenses of elaborated calculations. One of the most complicated computation procedures required by the CF method is the spectral deconvolution, due to the high number of spectral emission lines measured in a LIBS spectrum. A novel automatic deconvolution procedure was proposed in this work, which is aimed to overcome some weakness of the traditional methods used for deconvolution, as the Levenberg–Marquardt algorithm.

The original work of CF states that, in order to make a proper quantification, every element present in the sample must be measured. However, it is shown in this work that the cited requirement is not necessary, because the concentration ratio for two distinct elements, measured by means of a LIBS spectrum, is independent of the other emission species.

# Contents

# Chapter 1

# Introduction

Materials are the key stone in the technological development of every civilization that has existed through the history in every corner around the world. They are so important that distinct ages are associated with them, as the stone age, bronze age and iron age just to mention a few. The domain of a new material brings relevant military and commercial advantages, which drives the constant search of human kind for understanding why materials behave the way they do.

The combination of chemistry and quantum mechanics, which began around 1930's, opened the path not only to a deeper material's comprehension but also to design custom materials that are better suited to the new challenges imposed by the technological advances. One of the essential tools for materials research is the *chemical analysis*, a well-established field of chemistry that comprises a whole world of techniques, which can be grouped in two categories.[1]

The first category is called *classical techniques*, which can be distinguished by signals whose response is proportional to the absolute amount of the constituents of interest in the sample, or *analytes*; gravimetry, coulometry and titrimetry are typical examples. Classical techniques are usually based on chemical reactions, therefore are destructive tests.

A second category embraces all the *instrumental techniques* that respond to the relative concentration of the analyte. They are based by the measuring of physical properties of the sample as resistivity, light absorption, etc., and can be considered as nondestructive tests. Potentiometry, voltametry and spectroscopic techniques fall into this category.

Raman spectroscopy and Laser-Induced Breakdown Spectroscopy (LIBS) are instrumental techniques studied along this work, making a grater emphasis on LIBS technique. Raman and LIBS are complementary techniques; while the former provides information about the analyte's molecular bondings and its polymorphs,[2] the latter can detect the atomic composition in concentrations as low as ppm. These techniques are nondestructive tests that can be applied on samples without any preparation, does not need any consumable and offer real time responses as well as reduced costs, because unlike other methods as X-ray diffraction or Electron Microscopy, they do not require special facilities.

Each chapter on this work was written with the aim of being a self-consistent unit; thus requiring very little knowledge on previous chapters to be understood. In the first chapter a brief description of the structure of this document is given. The state of art for elemental identification, elemental quantification and combined Raman-LIBS instruments are also included in this chapter.

LIBS and Raman share many instrumental components. Therefore, it was found convenient to make a basic description of both techniques and the equipment used through this work in a single chapter. These subjects are covered in chapter 2. Not only instrumentation is shared by Raman and LIBS, but they also share many signal preprocessing steps that are used extensively in later chapters. Signal preprocessing is also covered in chapter 2.

The fundamental information contained in a LIBS spectrum is the elemental composition, a subject covered in detail in chapter 3. In order to detect a particular chemical element in a given sample, an analyst looks for its corresponding most prominent emission lines in the measured spectrum. This work requires a lot of experience, although due to its repetitive nature the process could, in principle, be automated.

Elemental identification methods for LIBS spectra are scarce in literature. Some factors that make this task difficult are the overwhelming amount of emission lines in databases, line overlapping, poor repeatability of LIBS spectra and line shifting due to stark effect.[3] Many of the automatic identification procedures used in LIBS are borrowed from other atomic emission techniques. Neural networks[4] and fuzzy logic[5] are examples of artificial intelligence based methods. Other similarity tests, i.e. procedures to quantify the likeness of two entities, are those based on statistics, such as FFT,[6] Dot–product[7] and correlation.[8]

Another factor with great impact on the identification process is the reference spectrum choice. Two main strategies for reference spectrum generation can be identified in the literature. The first strategy is to generate the reference with the superposition of emission lines whose intensity is taken from NIST[9] database. The second strategy makes use of plasma models.[10]

In this work a sample's elemental composition is done through the Wavelength Similarity Coefficient (WSC)[11], a new procedure developed in this dissertation. WSC automatically identifies the mid and high concentration elements within a LIBS spectrum and provides a very reduced set of candidates for traces. WSC takes into account factors not included in other works as the variations in the relative intensities produced by plasma temperature or amplitude differences between the measured lines and those available in the literature. WSC is discussed on chapter 3.

On the other hand, automated material classification methods based on LIBS spectra are abundant in the literature. In its basic form, a group of samples of known composition are measured and used as a *spectral library* for identifying an unknown material. This procedure was employed by Jurado-López et al.[12] while identifying alloys used in jewelry; the comparison criterion was the Spearman rank correlation coefficient. Mineral identification has also been successfully done using a mineral library; e.g. Koujelev et al.[13] use neural networks and a mineral library measured under approximate Martian conditions aimed to identify minerals in the

ExoMars mission planned by ESA. Glass classification in forensic applications was reported by Rodriguez-Celis et al.[14] using Pearson correlation as the comparison criterion. Other reported machine identification criteria are Partial Least-Squares Discriminant Analysis (PLS-DA)[15] and Principal Component Analysis (PCA).[16,17]

Among the automated classification techniques available PCA is the preferred approach used in this work, as will be shown on chapter 4, because it is not only able to classify but also it gives information of the spectral features that has a greater impact in the classification process. PCA spectral analysis was used in this thesis to correctly classify copper slags by its geographical origin, which could be associated with the raw materials source or the processing methods employed. In addition, an analysis of archaeological objects made of copper, has provided evidence to support the hypothesis that the copper purity in a given sample could be related with its historical manufacturing period. Moreover, sample's Ca content could be the key to identify the mining site of the raw materials used to produce such objects.

The last part of chapter 4 presents the multivariate analysis applied to classify stainless steels by its crystalline structure. Although LIBS only detects elements, steel alloys have some elements added that help to stabilize a specific phase at room temperature that can be detected and used to classify stainless steel samples. Each work reported in chapter 4 was the result of an interdisciplinary collaboration with archaeologists and metallurgic specialists.

These works have also shown that the LIBS experimental setup employed can detect elements whose concentration is below 1 per cent; these findings have been confirmed with values reported in the literature. A low-energy picosecond laser was employed during these experiments with the aim to do the least sample's damage. The results obtained are encouraging, suggesting the viability for implementing a mobile Raman-LIBS Setup (RLS).

As mentioned earlier, LIBS also has the capability to determine the amount of one or more constituents in a sample, a process termed *quantitative analysis* that will be covered in chapter 5. Since its beginning, LIBS has been used with this purpose. The basic technique measures the height of a selected emission line in a set of samples with known concentration, as result, a calibration curve is obtained for a given element, which can be used to estimate the elemental concentration of an unknown sample with similar chemical composition.

In many situations, a set of standard samples is too expensive or very difficult to obtain. For such cases, an alternative quantification procedure called Calibration-Free (CF) can be applied, and it is the preferred approach used in this work. CF was proposed by the Pisa group in 1999 and it is based in a mathematical plasma description in Local Thermodynamic Equilibrium (LTE).[18] CF method uses the matrix element as an internal standard in order to quantify; thus, CF is often seen in the literature as a standardless method, highlighting the fact that does not require *external standards*.

CF is not the only option for standardless analysis. Gomba et al.[19] proposed a method similar to CF based on the balance of electron density, while Yaroshchyk et al.[20] fitted a reference spectrum, obtained by plasma modeling, to experimental measurements.

In order to improve CF analysis, several hybrid approaches have been proposed in recent years. The first uses the results of CF approach for training a neural network.[21] A second approach estimates unknown experimental factor $F$ based on a standard of known composition in order to provide correction factors that improve the results; this method is called the One-Point-Calibration method.[22] Finally, Aragón et al. introduced a method called Csigma, which can be considered as a multi-elemental calibration curve computed from a known standard that is valid for spectral lines with moderate self-absorption.[23]

Improvements to the original CF algorithm have also been proposed mainly in ensuring LTE conditions,[24] proper line selection,[25] self-absorption effects[26,27] and plasma non-homogeneity.[28] CF analysis accuracy heavily depends on the species concentration; errors higher than 50% can be obtained for low content and trace elements. Despite of this fact, CF results are useful in applications were the exact amount of minor and trace elements is not relevant.

LTE plasma equations predict the integral line for a given emission band; thus, knowing the individual area of each band measured in the spectrum is essential in every LIBS quantification procedure. A novel automated deconvolution procedure called Simultaneous Height Adjust (SHA)[29] was developed in this work, which it is covered in chapter 5. SHA is a much faster alternative to Levenberg–Marquardt algorithm, it is easy to implement and is well suited to high resolution wide range spectra as those generated in Echelle spectrometers. Automated line selection based on the definition of optical thin lines was also implemented in this chapter.

One of the requirements stated in the original CF work was to measure every element in the sample in order to obtain accurate results. As it will be demonstrated this requirement can be relaxed if weight concentration ratios among elements are calculated instead of weight proportions. This allows to quantify only elements of interest.

Finally, as an additional benefit a combined Raman-LIBS Spectrometer (RLS) was implemented. A full description can be found in appendix A. In its simplest conception, a RLS has only one laser source, one spectrometer and one detector. A removable notch filter is placed in the collection optics path when performing Raman measurements. This experimental setup was reported by both Giakoumaki et. al.[30] and by Osticioli et. al.,[31] for the study of pigments. Another interesting alternative was introduced by Matroodi et. al.,[32] for the study of minerals. They used a Glan–Taylor prism to divide a 532 nm laser in two beams, which are focused at distinct distances for exciting Raman and LIBS at the same time.

A wider spectral range is a vital feature for LIBS and desirable for Raman setups. The classical solution is to use scanning spectrometers, even though this procedure precludes repetitive probing from the same unperturbed spot and lay down large acquisition times. Bruder et. al.[33] shows an alternative solution, making use of three separate spectrometers which runs independently. In this setup, the light is collected on a 7–fiber optical bundle; three of the fibers were connected to the entrance slits of each spectrometer that cover distinct adjacent spectral ranges.

Multiple spectrometers can also be used to better meet the requirements of each technique. Wiens et. al.,[34] reported a RLS that uses two different laser sources and spectrometers. A Czerny-Turner spectrometer is used for LIBS experiments while a transmission spectrometer is used for Raman. A more recent proposal was demonstrated by Hoehse et. al.,[35] while studying the viability of diode pumped solid-state lasers in RLS instruments.

Echelle spectrometers would enlarge the spectral range for every single event, keeping a high resolution without moving parts, thus reducing acquisition times. This kind of dispersive instrument has a broad acceptance in LIBS setups but is scarcely employed on Raman, mainly due to their poor sensitivity and stray light rejection.[36] One of the first Raman setups reported that made use of an Echelle spectrometer was carried out by Pelletier[37] in 1990. More than a decade later, Hoehse et. al.[38] reported a RLS based on a modified two arm Echelle spectrograph and a single CCD detector, where each arm was designed to provide the best Raman or LIBS signal.

Most of the notch filters used for blocking the intense reflected laser light in Raman experiments have bandwidths of tens of wave numbers; therefore, they block most of the Raman signal below 200 cm$^{-1}$.[39] However, the low-frequency Raman bands convey information about the crystalline structure of materials,[40] which is a particularly important feature in the pharmaceutical industry.[41] To the best author's knowledge, low-frequency Raman capability has not been reported yet in a RLS setup.

# Chapter 2

# Fundamentals

Spectroscopic techniques boost new discoveries and applications every day in such diverse areas as medicine, chemistry, biology, etc. Spectroscopy is a science field aimed to study the interaction between matter and energy that can be measured as transmission, emission or scattering.[42] According to Ball *spectroscopy is the study of matter using electromagnetic radiation*.[43]

Origins of spectroscopy can be traced back to 1666 when Isaac Newton realize that white light is composed of the rainbow's colors. He coined the term *spectrum* for this phenomenon. Almost a century later in 1752, Thomas Melville analyzes the spectrum emitted by substances burned on a flame, finding that every substance has different spectrum, although he was not able to explain this phenomenon. In 1802 William Wollaston observes dark lines in the solar spectrum, which were later rediscovered independently by Joseph von Fraunhofer in 1814. He started a systematic study of such lines, designing and fabricating the first diffraction gratings in 1821, thus improving the spectroscope. Using Fraunhofer's technology, Gustav Kirchoff and Robert Bunsen find two new atomic elements: cesium and rubidium. Moreover, they provide an explanation about the origin of spectral lines in 1859, enunciating *each element has a unique spectrum*.

The twentieth century has witnessed the development of many technological advances, but two of them deserve a special mention in the spectroscopic field. Since its origins in the 1960's decade, the laser named "a solution in search of a problem", has been successfully applied in all science fields, including chemistry, physics, medicine, communications, mechanics and life sciences. Laser unique properties as coherence and high energies have led to selective interaction with atoms and molecules resulting in new spectroscopic methodologies. The other technological breakthrough are the advances in detectors as photomultiplier tubes, intensified CCD cameras and CMOS detectors, which begins in the 1980's decade. These devices have replaced the photographic film giving a reduction in acquisition times and improved sensitivities.

Based in all these advances, now is possible to inquire about the inner structure of atoms and molecules or to identify the chemical composition of celestial bodies like the sun, planets and stars.

The following sections describe the principles of Raman and LIBS spectroscopy. A formal treatment of Raman effect from classical and quantum physics perspective can be found in Andrews's textbook.[44]

## 2.1    Laser-Induced Breakdown Spectroscopy

When a few crystals of common salt are sprinkled into a Bunser burner, the flame turns in to a color that is characteristic of the sodium ion, this phenomenon is termed *atomic emission*. If the flame's light is analyzed, the spectrum will show the well know doublet of sodium located at 589 nm. The experiment just described receives the name of Flame Emission Spectroscopy (FES), used for first time by Kirchoff and Bunsen. FES is part of a most general area called Atomic Emission Spectroscopy (AES), which is defined as *a general method for measuring elements by the study of the radiation emitted by their atoms present in an excited state.*[45]

AES equipment must perform the following steps:[46]

- Atomization or vaporization of the sample to produce free atomic species.
- Excitation of the atoms.
- Detection of the emitted light.
- Calibration of the intensity to concentration or mass relationship.
- Determination of concentrations, masses or other information.

The most common vaporization and excitation methods include microwaves, inductively coupled plasma, electric arc sparks, glow discharge and lasers.

Laser-Induced Breakdown Spectroscopy (LIBS) uses a high-power focused laser pulse to vaporize and excite the sample, resulting in a plasma. To the naked eye plasma appears as a bright white light emanating from the focal volume. Cremers, Radziemski and co-workers coined the term LIBS in the 1980s during their work at Los Alamos National Laboratory.

Contrasting with other AES methods, LIBS does not require special facilities or special sample preparation steps. LIBS spectra acquisition last just a few seconds, and can measure solids, liquids as well as gases. Therefore, LIBS is a well suited method for in-line processes measurements and remote applications like deep-sea exploration[47] or Martian environmental analysis.[48] Quantitative measurements are also possible with the aid of calibration curves or using standard-less algorithms.

Figure 2.1 shows schematically the plasma formation process. Laser energy is absorbed by the sample surface (1) and the material starts to evaporate (2). Due to optical breakdown induced by the laser energy, plasma is formed within material

**Figure 2.1**

*Principle of Laser-induced Breakdown in phases.*



| Absorption | Sublimation | Continuum Emission | Element Specific Emission |
|:---:|:---:|:---:|:---:|
| (1) | (2) | (3) | (4) |
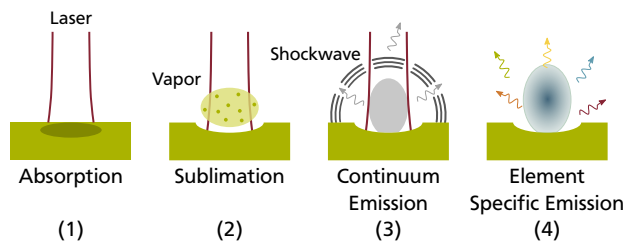
**Figure 2.2**

*Typical plasma evolution.*
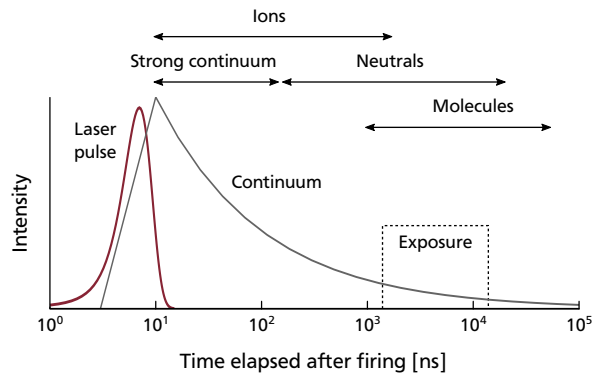


vapor and the surrounding gas atmosphere, accompanied with a loud snapping sound called shock wave (3). Then while plasma emission decays, emits element specific radiation (4), leaving a crater on the sample surface.

Initially, plasma is dominated by a "white light" named *continuum emission* that has low intensity variation as function of wavelength. Continuum is originated by bremsstrahlung and radiation coming from recombination between ions and free electrons during plasma cooling.

Typical plasma evolution is illustrated in figure 2.2. Low emission from minor and trace elements can be easily hidden by strong continuum radiation, making time-resolved measurements a must. This implies a delay time between the firing signal of the laser and the starting of the exposure time, in such way that white light has a minor contribution but atomic emissions are still detectable.

The main parts of a LIBS detection system are shown in figure 2.3.[49] In order to produce sample breakdown, values of irradiance above 10 MW/cm$^2$ are required.[50] This is accomplished by focusing the laser pulse on the sample surface. The plasma light emitted is collected and guided to a spectrograph for analysis. A time control synchronizes the camera delays and exposition time, as well as the fire laser signal.

Laser repetition rates used in LIBS experiments depends on laser technology employed, ranging from 10 to 1000 kHz. Usually, frequencies above 100 Hz and energies below of 50 mJ are used for small focal lengths (<100 mm); while energies above 50 mJ are used for focal lengths above 200 mm and low repetition rates, like remote applications.

**Figure 2.3**

*Schematic diagram of LIBS experimental setup.*
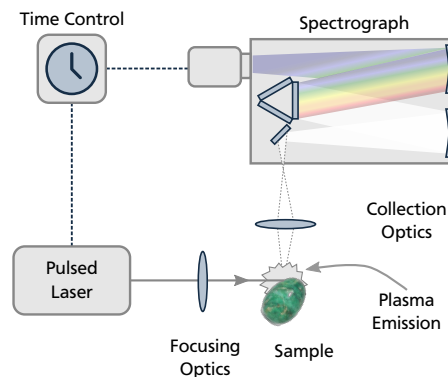
Laser wavelength choice may also benefit some LIBS experiments. Due to its high evaporation threshold, Fantoni[51] suggests the use of UV wavelengths for stones, ceramic and metals. Near IR-VIS is recommended for in water analyses.

Finally, laser pulse duration may also affect LIBS experiments. Stoichiometric evaporation is preserved when ultra-short laser pulses are used; however, this can lead to distortions on line profiles due to the Stark effect.

## 2.2 Raman Spectroscopy

Inelastic light-scattering is a phenomenon in which a material scatters photons with a frequency different from the initial illumination, even in absence of absorption. This is known as Raman effect, and was reported for first time in 1928 by Raman and Krishnan.

From a phenomenological point of view, Raman scattering is originated in an inelastic collision between an incident photon and a molecule.[52] The energy takes the molecule to a high electronic level $v$, formally described as a *virtual* level, to highlight the fact that is not necessarily a stationary eigenstate of the molecule or atom. In particular, the *resonant Raman effect* is when the virtual state coincides with a real molecular eigenstate.[53] The virtual level is related to vibrational, rotational or electronic energy of the molecule, converting Raman spectroscopy in a valuable tool for molecular study.

The virtual state is considered to have an infinitely short lifetime, therefore, the molecule releases a photon instantaneously. It is possible to identify two cases for inelastic scattering: when a photon with lower frequency that the absorbed photon is emitted, we talk of Stokes scattering; while if the emission has higher frequency the process is termed anti-Stokes scattering. Because intensity ratios of Stokes and anti-Stokes Raman bands are related by the Boltzmann's law, they can be employed to measure sample's temperature.

A graphical description of light scattering processes is shown at top of figure 2.4 with the aid of Jablonski diagram.[54] Their corresponding Raman band emission is illustrated in the bottom part of the figure. The spectra representation follows some conventions that are commonly available in the literature. The signal amplitude is plotted in the ordinate while the horizontal axis is used to show the energy in wavenumbers. Usually, abscissa is given in $cm^{-1}$ units and normalized according to equation 2.1, in order to shift the excitation emission at zero. $\lambda_L$ is the excitation wavelength and $\lambda$ is the Raman spectrum wavelength,

$$\Delta\omega = \left( \frac{1}{\lambda_L} - \frac{1}{\lambda} \right) \tag{2.1}$$

Raman spectroscopy is a non-destructive test whose spectra can be acquired in real time from very small volumes (less than 1 $\mu$m in diameter), in solid, liquid or gaseous state samples without any preparation. Raman spectrum is very specific for a given sample providing a unique finger print for each material. Moreover, scattering is not interfered by water or glass, so samples can be analyzed through a

**Figure 2.4**

*Raman spectra and their corresponding Jablonski diagrams. Solid horizontal lines represent real molecule eigenstates, while dashed horizontal line $v$ is a virtual electronic state.*
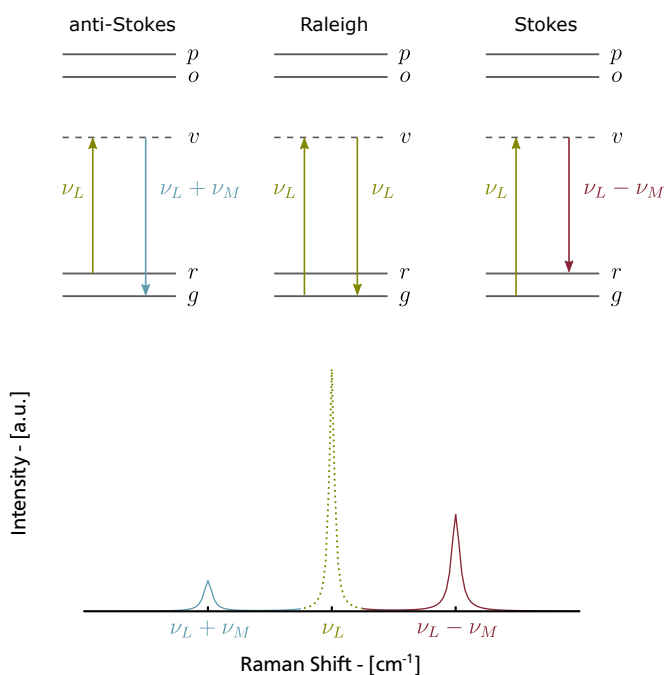


glass cell and dissolved. Another advantage is that it is easier to analyze inorganic materials by Raman than by infrared spectroscopy. In opposition, metals or alloys cannot be analyzed with this technique. All these benefits have some trade-offs that are imposed by Raman's effect practical limitations, which will be discussed after a brief description of a spontaneous emission Raman setup.

Raman setups simplicity has an inherent appeal, as shown in figure 2.5. It consists of a monochromatic light source, collection optics, a monochromator and a detector. However, Raman effect is an extremely low efficiency process: only $10^{-7}$ to $10^{-15}$ of the incident light is inelastic scattered. Besides, Raman bands line widths depends on the spectral width of the illumination source, and its amplitude is a fourth power function of the excitation light frequency, as Rayleigh shown in 1899.

**Figure 2.5**

*Schematic diagram of a spontaneous Raman spectroscopy setup*
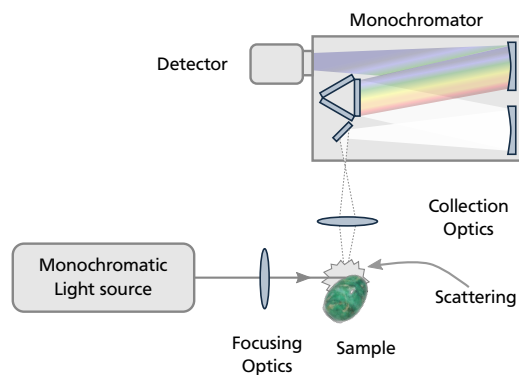
Today, the vast variety of laser sources has reduced the constraints imposed by natural laws. Their narrow band width has reduced the scattered band widths, and the high output power provides higher intensity signals. IR wavelength laser has also help to avoid fluorescence events that once dimmed the Raman bands.

Raman signal gains from better illumination sources are only half of the picture; improved detectors also plays a key role. Multichannel intensified detectors have made real-time Raman measurements a reality. In combination with pulsed lasers, conforms an alternative route to deal with fluorescence called *time-resolved Raman* spectroscopy.

## 2.3   Instrumentation

If in a time-resolved Raman spectroscopy setup the monochromator-detector set is substituted with a spectrograph, an almost exact setup for LIBS spectroscopy emerges; which in principle, would allow to perform Raman and LIBS experiments in a single instrument. However some minor aspects must be cared, particularly with the focusing and collection optics.

This section will briefly mention the hardware used along the test trials of LIBS and Raman setups that led as a final result the combined Raman-LIBS experimental setup described in appendix A.

### 2.3.1   Laser

The two pulsed laser sources used during this work are listed on table 2.1. Spectra physics laser can operate in 1064 nm, 532 nm or 355 nm. A high output power along an active Q-switch makes this device a great illumination source for many kinds of LIBS measurements. However, the facilities required for proper operation limits its use in mobile applications. In contrast, Teem Photonics laser is a compact device that can be used in applications where the sample cannot be transported out its location, such as some cultural heritage pieces. Despite its low power output, Teem Photonics laser induce breakdown in solid samples as steel and bronze. As a drawback it is necessary to mention the passive Q-switch, which generates a variable jitter delay time that excludes its use as a second laser source in double pulse applications.

**Table 2.1**   *Specifications of lasers used in Raman and LIBS test trials.*

| Feature | Spectra Physics | Teem Photonics |
|---|:---:|:---:|
| Model | Quanta-Ray Pro 230-10H | Power Chip PNG-M02010-1x0 |
| Repetition Rate | 10 Hz | 1000 Hz |
| Energy at 532 nm | 650 mJ | > 20 uJ |
| Pulse Width | 7-10 nm | < 400 ps |
| Beam diameter | < 10 mm | |
| Q-switch | Active | Passive |

## 2.3.2 Focusing Optics

High power output provided by Spectra Physics laser can easily damage mirror's reflective surfaces; therefore, high energy mirrors are required for guiding this kind of light. Moreover, UV radiation is absorbed by common BK7 optical glass,[55] as figure 2.6 shows. Therefore, synthetic fused silica lens must be used for focusing energy into the sample.

***Figure 2.6***

*Transmittance for uncoated surface in a 10mm slab. UV Grade Synthetic Fused Silica (UCGSFS), Optical Grade Synthetic Fused Silica (OQSFS) and BK-7*
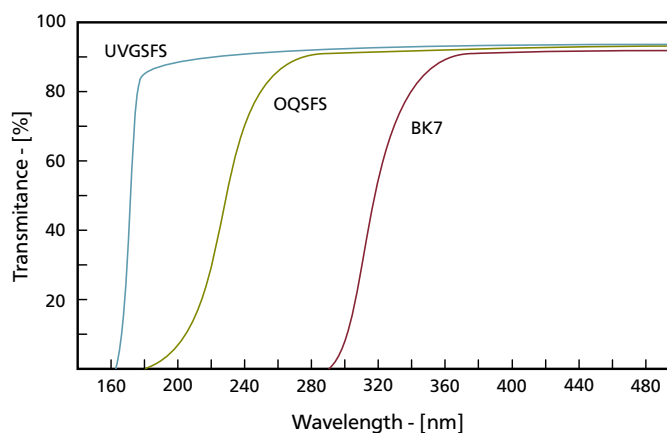


A narrow beam can also produce damage over the mirrors' surface, even in the low energy case. Thus, a beam expander was designed to preserve mirror's coatings while Teem Photonics laser was in use. A Galilean beam expander configuration was selected, as depicted in figure 2.7. This optical design allows a compact device and is well suited for high energies, as opposed to Keplerian beam expander in which air laser breakdown can occur.[56]

***Figure 2.7***

*Galilean beam expander schematic*
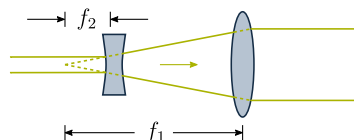


Lens separation $t$ and magnification $M$ are given by

$$t = f_1 + f_2 \qquad M = -\frac{f_2}{f_1} \qquad (2.2)$$

A magnification of $M = 10$ results when a $f_1 = -25$ mm and $f_2 = 250$ mm are used, with a lenses separation equal to 225 mm. Both lens have an anti-reflective coating to reduce laser energy looses.

Laser power is not an issue on Raman experiments, because they make use of low energies when compared with LIBS, although other consideration comes at hand. In particular, wavelength laser bandwidth can dim low frequency Raman bands, besides, dichroic mirrors are mandatory when on-axis collection geometry is used. One solution comes with some lasers that have a very narrow spectral response with high cost as a counterpart. A cheaper solution is employing laser line filters, but some energy is loosed in the process.

### 2.3.3   Collection Optics

Synthetic fused silica lens also must be used in order to collect UV emission lines on LIBS experiments. The plasma light could be guided with mirrors and lenses to the spectrograph entrance slit, but a more extended and flexible solution comes with optical fiber use as a guiding media. Fiber optic solves space restrictions in many LIBS setups, allowing to collect plasmas generated several meters away from the spectrograph.

Regarding Raman setups, Raleigh scattering reduction reveals low frequency Raman bands that would pass unnoticed otherwise. Notch filters are the preferred optical devices to this end, because allows to simultaneously measuring Stokes and Anti-Stokes bands. Long Wave Pass (LWP) edge filters are used to detect only Stokes scattering, while if Anti-Stokes scattering detection is desired, Short Wave Pass (SWP) edge filters are the solution.

### 2.3.4   Spectrograph and Detection Device

One key constituent of a time-resolved experimental setup is the timing control. Detector manufacturers provides time control features as embedded electronics or as separate units, which in turns, includes a external trigger input signal used for synchronizing external events. Timing control is handled by a software interface, where detector's delay and exposure times can be specified and synchronized with external trigger signal.

During this work three distinct spectrographs were used, whose specifications are listed on table 2.2. Acton Research and Princeton Instruments spectrographs are Czerny-Turner optical designs, which generate a spectrum with high spectral resolution in a narrow spectral region, about 40 nm wide. On the other hand, Echelle spectrographs also have a high spectral resolution but offer a much wider spectral region, although requires higher input signals, and are more noise sensitive than Czerny-Turner instruments.

Spectrum file formats are different for each spectrograph. An option to export spectral data as text files is available in all programs, though some information is lost

**Table 2.2**   *Specifications of spectrographs used in Raman and LIBS test trials.*

| Feature | Acton Research | Catalina Scientific Instruments | Princeton Instruments |
|---|---|---|---|
| Model | Spectra Pro 500i | SE 200 | IsoPlane 160 |
| Focal Length | 500 mm | 200 mm | 203 mm |
| Aperture Ratio | f/6.5 | | f/3.88 |
| Resolution | 0.05 nm | $\lambda/\Delta\lambda$ 1700 | 0.15 nm |
| Optical Design | Czerny-Turner | Echelle | Czerny-Turner |
| Operating Range | 180 - 900 nm | 180 - 850 nm | 180 - 900 nm |
| Detector | ICCD 1024 X 256 | ICCD 1024 X 1024 | emICCD 512X512 |
| Software | WinSpec/32 | KestrelSpec | Light Field |
| File format | SPE 1.6 | CRV | SPE 3.x |

in the converted file, like delays and exposition time used during the acquisition. In addition, some file support characteristics as automatic file name generation are not available when using this format. In order to preserve all the experimental information provided by the detector's vendor, several file readers functions were written based on the software documentation. [57,58,59,60]

### 2.3.5   Safety Components

Depending on its output power, lasers can produce skin burns and damage to the retina. Skin burns risk is especially high when a sample is manipulated to achieve an adequate focusing on its surface when LIBS measurements are performed. With safety in mind, two devices were included in all experimental setups: a computer controlled laser shutter, [61] and a $xyz$ motorized stage. [62]

A laser shutter blocks the incoming laser light when the instrument is not in use allowing handling the sample without risk. When the shutter is inactive, a $xyz$ stage moves the sample to a desired location or changes the focal distance to induce laser breakdown or improve the Raman bands intensity.

Having three distinct control programs for spectrograph operation, positioning stage and shutter control, adds a complexity level to the instrument operator. By fortune, the positioning stage is controlled with serial communications, and the shutter can be controlled by means of a Dynamic Load Libraries or `dll`.

A master control was programmed in C++, whose user interface is shown in figure 2.8. C++ programs have direct access to the operative system Application Programming Interface (API); [63] as a consequence, using external `dll` is an easy task. As mentioned before, serial communications are required for controlling the $xyz$ positioning system. Universal Asynchronous Receiver-Transmitter (UART) is the basic component of every serial communication system, a good reference for its internal structure can be found in Campbell's book. [64] Also, in Windows operative system, there is an API devoted to serial communications that was used into the control program. [65]

An interesting feature of the master control program is its ability to interact with `WinSpec/32` and `KestrelSpec` spectrometer's software camera. The trick is to obtain the window handle of the desired programs, searching the system with the function API `EnumWindows`, from which the window name and handle can be read. Once the window handle is known, event messages are sent to the software camera window with a simple `SendMessage` call.

***Figure 2.8***

*Hardware control program interface.*
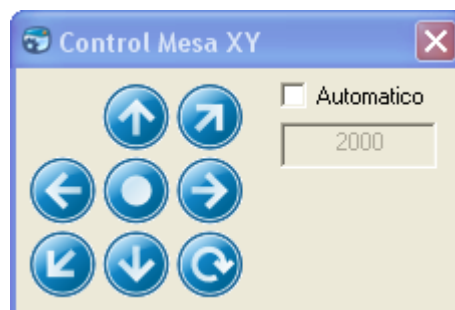
Obviously, the messages identifiers and their corresponding meaning are not freely accessible to the user in proprietary programs as `WinSpec`, however, a good program insight can be obtained with tools like the `winspy` program, delivered with some versions of the Microsoft visual C++ compiler and similar programs spread around the Internet.

The whole acquisition process can be synchronized with the button identified with a circular arrow—called acquire button. This is accomplished in two modes: automatic or manual. Automatic mode is established when the `automatic` check box in the program interface is activated. In this mode, the shutter is opened, the positioning system is continuously moving in a square of one square millimeter, and the capture mode of camera software is activated. After the time in milliseconds that was configured in the text box has elapsed, the system closes the shutter and stops the positioning system. In the manual mode, all the hardware components are activated o deactivated when the acquire button is pressed.

The shutter can also be operated independently by means of the central button with a circle in it, which opens and closes the shutter. The $xyz$ positioning system is controlled pressing the arrow buttons.

## 2.4   Signal Processing

All the steps used to extract useful information from a signal are known as *signal processing*. In general terms, a spectrum is composed of three main parts: the band emissions originated by the sample, as well as low and high frequency noise. *Noise* is the name given for undesired variations that occur within a signal.

High frequency noise is mainly introduced by instrument electronics circuits, although also exist external natural sources as cosmic rays. Some examples for this kind of noise are thermal CCD noise, shot noise and digitalization errors.[66] High quality detectors incorporate specialized electronics and temperature control devices aimed to reduce signal's high frequency noise. Digital filtering can also be applied to achieve the same purpose, being single pass and double pass moving average filters the selected techniques implemented in this work.[67]

Low frequency noise is common to nearly all spectroscopic techniques. Its origin can be traced back as a side effect of the physical variable measured in the sample. In LIBS experiments the main source is plasma continuum generation, while in Raman is fluorescence. Time-resolved measurements can decrease low frequency noise in some degree, but even so, is common to apply digital signal processing techniques to isolate emission bands.

Undesired low frequency noise is commonly known as continuum, background or baseline signal. Baseline removal techniques are being continuously developed, with special attention in the search of an automatic baseline correction procedure that provides optimum results; an objective that is not reached yet. Among the vast variety of baseline removal procedures, the Peak Clipping Algorithm[68] was implemented.

### 2.4.1 Wavelength Calibration

Every spectrograph must be wavelength calibrated prior to use. The involved steps are device dependent but in general, wavelength calibration procedures employ a light source whose emission lines are well known. A calibration lamp model 6025 Hg(Ar) manufactured by Oriel was employed to this end.

In some cases, inaccuracies emerge from measurements despite following the calibration method recommended by the manufacturer. To alleviate this situation two strategies were implemented; the first compares a reference spectrum against a shifted version of the measured spectrum until the highest correlation is found. The second strategy, uses the most prominent wavelength peaks position in both spectra to make a polynomial adjust of the wavelength axis in the measured spectrum.

Spectral comparison is done with Pearson correlation, while peak location uses the first derivative criterion. Details on both subjects will be covered in chapter 3.

### 2.4.2 Amplitude Calibration

The amplitude of the luminous signal collected from experiments is modified by the frequency response of each optical element involved during collection and guiding steps. One way to compensate such variations is determine the frequency response of whole experiment with an amplitude calibrated source. A SL1_CAL StellarNet calibrated lamp was used for this purpose.

The calibrated lamp is placed where the plasma emission will take place, and a measuring of the full operating range of the spectrograph is done. The next step is dividing the instrument's spectral response by the calibrated response spectrum provided with the lamp. The correction factor obtained on this way is used to multiply every measured point in the $x$ axis by their corresponding correction factor to achieve a calibrated spectrum.

Implementation of amplitude calibration procedure is straightforward, though some cautions are required. In particular, calibrated data is not a smooth line, instead, a noise spectrum is provided as figure 2.9 shows. If lamp's raw calibrated data are using for corrections, a resulting noisy signal will result. Moreover, the provided sampling wavelengths will not match those measured in the experiment.

**Figure 2.9**

*Raw amplitude calibrated spectral data and fitting with a 9th degree polynomial*
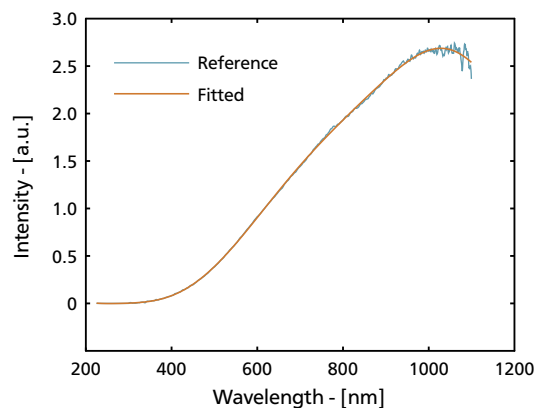
As solution, calibration data are fitted with a 9th degree polynomial getting rid of noise. As an added benefit, the function can be used to estimate the correction factor in any desired wavelength.

### 2.4.3 Additional Notes

Usually, one or more of the signal processing methods described until now must be applied to a spectrum before elemental identification or quantification steps could be performed. Even though calling distinct functions is quite simple in a Matlab script, loading and conditioning signals instructions can occupy a great bunch of code. Thus, this work provides a function that combines the spectrum load and preprocessing steps, called `Spec_Treatment`. This function also selects the proper spectral file format loading routine according to the file extension. Supported verbs for the `Spec_Treatment` routine are listed on table 2.3.

Extracting the portion between 200-250 nm and correcting the baseline with a 30 point window second order snip in a spectrum stored in the file `test.spe`, is as easy as:

```
[x, y] = Spec_Treatment
( ['c:\user\spec\'], ['test.spe'], 'Trim', [200 250], 'SnipBase', [2 30] );
```

note that verbs can be applied in any desired order

**Table 2.3**  *Description of supported verbs in* `Spec_Treatment` *routine*

| Verb | Description | Parameters |
|------|-------------|------------|
| Trim | Extract the signal between high and low points | [high, low] |
| SnipBase | Baseline correction (descending SNIP algorithm) | [order, iterations] |
| LinearBase | Baseline correction with a constant value | [y] |
| ShiftCalib | Wavelength calibration by the shifting method | [direction] |
| PolyCalib | Wavelength calibration applying a polynomial | [poly coefficients...] |
| HeightCalib | Amplitude calibration | [] |
| NoiseReduction | Smooths signal with moving average filter | [order, length] |
| Interpolate | Interpolate spectra within a given set of $x$ points | [x vector...] |

# Chapter 3

# Elemental Identification

Laser-Induced Breakdown Spectroscopy (LIBS) emerged as a promising alternative technique for determining sample's elemental composition. Although other very sensitive techniques are well established, as X-ray spectroscopy, mass spectrometry, chromatography, etc., all of them requires some sort of sample pretreatment, special facilities and expensive equipment. In contrast, in just a few seconds LIBS can analyze samples remotely or *in situ*, without any kind of preparation. The ablated mass in this kind of measurements are a few $\mu$g of material, converting LIBS in a semi-nondestructive test, which can be applied in places that are far out the reach of any other analytical techniques, as inside a steel induction furnace[69] or at the post-accident site of the Fukushima Daiichi nuclear power plant.[70]

A LIBS plasma emits light composed of two main signals: a continuum plasma emission and several emission lines whose origin can be ionic, atomic or molecular species. Continuum spectral contribution can be diminished with proper delay acquisition times, leading to a spectrum dominated by atomic emission bands as seen on Fundamentals chapter. Wavelength bands' position constitutes a "finger print" that uniquely identifies the atomic and molecular species contained in the sample. This assumption form the basis for the current chapter content, devoted to answer the question: *What elements does a sample contain?*

Elemental identification may seem as an easy task, but this is far from reality. As pointed by Tognoni et al.[71] "An automatic routine able to identify the major elemental components of a LIBS spectrum, to be joined to CF-LIBS quantitative analysis, would be extremely useful."

The origin of the complexity for a proper identification of atomic emission lines in an unknown sample, can be attributed to the huge number of emission lines available in the literature. In the wavelength range comprised between 200 and 900 nm exist more than 92000 atomic emission lines from hydrogen to bismuth, many of them, separated by a few thousandths of a nanometer. In other cases, when some lines are combined in a multi-elemental spectrum, the resulting emission line will present a wavelength shift.

*Spectral portion of Fe I,
where the superposition of
some lines generates a
unique shifted line.*
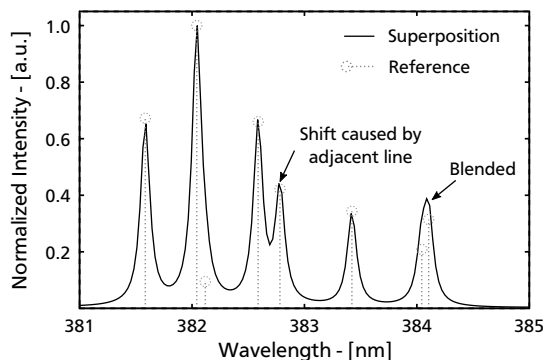


For example, consider figure 3.1 where a portion of Fe I spectrum is calculated assuming a Lorentzian line shape. The 382.58 nm line central position shift is caused by its neighbor located at 382.78 nm, while lines at 384.04 and 384.10 nm are blended in a single shifted line. Thus, is evident that in a spectrum composed of many elements, higher complexity will be introduced by shifting and blending.

Several approaches addressing line emission identification have been reported. Neural networks[4] and fuzzy logic[5] are examples of artificial intelligence based methods. Other similarity tests, i.e. procedures to quantify the likeness of two entities, are those based on statistics, such as FFT,[6] Dot–product,[7] correlation[8] among others.

Artificial intelligence methods employed for elemental identification requires advanced programming techniques and heavily depends on an adequate training set generation. On the other hand, deterministic identification methods are easier to implement. In particular, correlation—which needs a reference spectrum for comparison—has been widely studied. For example, Gornushkin[72] found that linear correlations is better suited for identification than rank correlation. Spatial masking and filtering were proposed by Gornushkin,[73] in order to improve the correlation performance; while Griffiths[74] used a weighted spectrum schema.

Before giving further details about the line identification approach implemented in this work, a brief explanation is required, concerning sources of atomic emission line information that serves as the basis for reference spectra generation.

## 3.1   NIST Database

Spectral atomic emission reference values are available in several databases, which contains data for every single element in the periodic table, including wavelength position and relative intensity of every emission line. Among the most popular in the specialized literature are the National Institute of Standards and Technology (NIST)[75] and Kuruczk database.[76] This information is collected from many sources that employs different experimental conditions and equipment; because of this, NIST suggests that "relative intensities are not basic data and must be used with caution".[77] Conversely, lines' emission wavelength position constitute reliable data.

***Figure 3.2*** *NIST database FrontEnd for atomic emission line parameters query.*

**Figure 3.3**  *Web page results displayed after a data request on NIST database.*



Kurucz database is composed of text files spread along a directory structure that makes it difficult to use. NIST database, on the other hand, provides a friendly user interface that can be consulted on the Internet. Figure 3.2 illustrates the NIST database FrontEnd interface in its current version 5.3, while figure 3.3 depicts a portion of searching results for all Fe emission species.

The information required for the identification and quantification process must be downloaded and preserved into a format easier to use. Uncompressed Level 5 MATLAB file format was chosen as the target format. Gathering spectral data is accomplished in two steps: retrieving data and translation.

Html retrieve data function is based on the code written by Alex James[78] at Lawrence Livermore National Laboratory. Results' web page automatic retrieving is done in Matlab 2010a and previous versions with the function `urlwrite`, while for higher versions, `webwrite` function exists. In essence `urlwrite` requires two parameters: the FrontEnd web page url, as well as a configuration list that contains the variable names and values needed for the options fields.

Plasma parameters requested by the retrieving routine are listed on table 3.1, its corresponding names and value fields are embedded at NIST web's page source code. In particular on Firefox navigator, a mouse right click over the desired input control shows a menu, where an item called `Inspect Element` can be found. This option displays in a separate window the control's source code, where the variable `name` and its associated `value` are the required parameters to configure the `urlwrite` function.

It is necessary to highlight that the output format requested is an ASCII table, which will be embedded in the downloaded `html` code, as shown in figure 3.3. As a consequence, a text extraction routine was written with Phyton 2.7[79] programing language, because includes a library that allows easy navigation through html tags. In order to relax software requirements to final user, the Phyton code is compiled

**Table 3.1**

*Main plasma parameters requested in NIST database FrontEnd.*

| Parameter | Value |
|---|---|
| Format | ASCII Table |
| Java Script | yes |
| Energy Units | eV |
| Lines | Only with observed wavelengths |
| Wavelengths | Air 200 - 2000 nm |
| Transition Strength | $g_k A_{ki}$ |
| Transition | Allowed and Forbidden |
| Level info | Configurations, Terms, Energies, J |

into an executable file, named `htm2ascii`, that does not depends on the interpreter. Binary conversion of scripts was done with the `py2exe` program.[80]

Conversion from ASCII emmision tables to version 6 Matlab file format[81] was implemented in C++, whose result is a command-line program named `ascii2mat`. Translated files contain only one structure with the corresponding species' atomic symbol as root name. Structure's fields contents are described on table 3.2.

Each field is an array of parameters that shares the same data type, in such way that every single row's array uses a 32-bit multiple storage. All fields have a row index relationship.

`Wavelength`, `Emission` and `IonInfo` fields are used during line identification and quantification process. `MergeKey` field is employed by a program written in C++, called `merge`, that sorts by wavelength position all the lines of every atomic species downloaded. The resulting output file includes wavelength, amplitude, atomic symbol and ion number for each line. This information is preserved into a file named `merge.mat`, whose information is very useful to make a manual identification for a given line.

**Table 3.2**  *Mat file content converted from NIST database*

| Field Name | Data Type | Column | Description |
|---|---|---|---|
| Wavelength | Real Number | 1 | Central band position |
| Emission | Real Number | 1 | Amplitude |
| | | 2 | Einstein coefficient $g_k A_{ki}$ |
| | | 3 | Lower energy value [eV] |
| | | 4 | Upper energy value [eV] |
| IonInfo | 8-bit integer | 1 | Atomic symbol |
| | | 2 | Ion Number |
| | | 3 | Lower total angular momentum J |
| | | 4 | Upper total angular momentum J |
| MergeKey | 16-bit integer | 1 | Amplitude normalized to 50,000 |
| | | 2 | Consecutive index fom 1 to $n$ lines |
| Oscilator | Real Number | 1 | Oscilator strength |
| | | 2 | line strength |

*Figure 3.4*

*Database retrieve and translation steps*
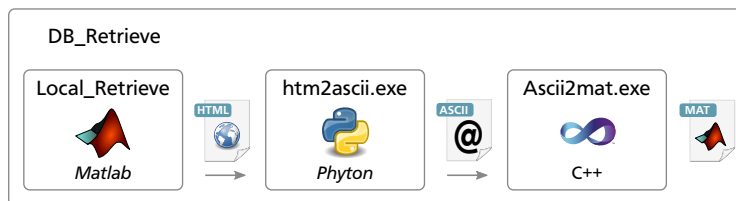
All the knowledge required to use all the related database software written thus far has been compacted into a Matlab routine called `DB_Retrieve`. Just provide a target path with the desired portion of the spectrum as the final subdirectory and `DB_Retrieve` will perform all the necessary steps to download and translate spectral information. A schematic of the internal work of `DB_Retrieve` is depicted in figure 3.4.

## 3.2   Reference Spectrum

Along this work the reference spectrum will be obtained as the superposition of emission lines with Lorentzian shape, whose mathematical expression is:

$$L(\lambda) = \frac{I}{1 + \left(\dfrac{\lambda - \lambda_o}{\gamma/2}\right)^2} \tag{3.1}$$

where $I$ is the peak intensity at the central emission line wavelength $\lambda_o$, and $\gamma$ is the full-width half-maximum value. Intensity and $\lambda_o$ values in equation 3.1 can be taken directly from NIST database, while $\gamma$ can be set equal to the instrumental broadening as a first approximation.[20]

When comparing a measured plasma emission against a reference spectrum calculated with NIST intensities, figure 3.5, the need for a better way to predict the relative intensities is clear. The measured spectrum obtained from the mineral ore shows a dramatic difference in relative intensities against the reference spectrum.

As an alternative, the use of a plasma model at local thermal equilibrium (LTE) conditions[82] is proposed in this work for the reference spectrum calculation. The line intensity is proportional to the energy level population, and can be estimated from the transition probability between the upper $E_n$ and lower $E_m$ energy levels:

*Figure 3.5*

*Measured spectrum of calcium found on a mineral cooper ore (bottom part of figure) and a reference spectrum calculated with NIST intensities.*
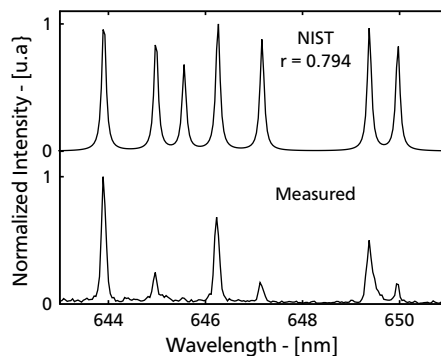
$$\overline{I_{nm}} = FC_s \frac{A_{nm}\, g_n\, e^{-E_n/Tk_B}}{U^s(T)} \tag{3.2}$$

$\overline{I_{nm}}$ represents the measured integral line intensity, $A_{nm}$ stands for the transition probability for the given line, $g_n$ is the $n$ degeneracy level, $k_B$ is the Boltzmann constant, $U^s(T)$ is the partition function for the emitting species corresponding to plasma temperature $T$; $C_s$ is the atomic species concentration and $F$ is used as an experimental parameter that groups optical efficiency of the collecting system, as well as plasma volume and density. For the sake of simplicity, a normalization procedure against any other line of the same species can be performed:

$$\frac{\overline{I_{ij}}}{\overline{I_{nm}}} = \frac{A_{ij}\, g_i}{A_{nm}\, g_n} e^{-(E_i - E_n)\,/\,Tk_B} \tag{3.3}$$

Equation 3.3 does not require a beforehand knowledge of partition function, species concentration or the experimental factor. All the other values are available on databases with the notable exception of $T$. Temperature can be estimated with the aid of LIBS experiments—using the same laser power and repetition rates—on samples with known composition, making use of the Boltzmann plot method.[83] This is a very important step, because the relative intensities of emission lines are also affected by temperature, although during our tests, a difference of 10 percent in temperature does not produce meaningful variations.

There is an evident impediment to substitute the intensity given by equation 3.3 in expression 3.1, because this last equation requires the Full Height Intensity (FHI). With the aim to solve this hassle, consider the line integral of an emission line with a Lorentzian profile represented by equation 3.1:

$$\overline{I_{jk}} = \int_{-\infty}^{\infty} L(\lambda)\, d\lambda = I_{jk}\, \pi\gamma/2 \tag{3.4}$$

where $I_{jk}$ is the FHI. Normalizing this result against another reference line, and assuming the instrumental line broadening as an approximation of $\gamma$ produces:

$$\frac{\overline{I_{jk}}}{\overline{I_{mn}}} = \frac{I_{jk}}{I_{mn}} \tag{3.5}$$

This shows that during normalized reference calculation, it is safe to use FHI found on NIST databases or the line integral intensity values obtained with the simplified plasma model of expression 3.2.

Reference spectrum calculation can be a time consuming task for elements with many emission lines; however, a shortcut is open when the omission of very low emission lines is considered. Table 3.3 shows the proportion of NIST emission lines that are below a determined amplitude threshold relative to the maximum amplitude. These data were obtained normalizing the amplitude of every atomic species in the considered wavelength range.

Due to its very low amplitude contribution to the final spectrum, a lot of lines can be thrown out from the reference computing procedure, significantly reducing computation time.

**Table 3.3**

*Proportion of lines below an amplitude threshold in the 200–900 nm region from H to Fr*

| Threshold as % of Proportion of lines | Maximum intensity below threshold [%] |
|:---:|:---:|
| 0.1 | 55.896 |
| 0.5 | 72.017 |
| 1.0 | 77.879 |
| 2.0 | 83.366 |
| 5.0 | 88.987 |
| 10.0 | 92.521 |

## 3.3   Likeness Between Two Signals

Pearson product-moment correlation coefficient provides a single number $r$, called *correlation coefficient*, that measures how much alike two signals are;[84] $r$ can be computed with the following expression:

$$r = \frac{\sum(x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum(x_i - \bar{x}_i)^2}\sqrt{\sum(y_i - \bar{y}_i)^2}} \tag{3.6}$$

where $x_i$ and $y_i$ are the measured and reference spectrum respectively. A correlation value of one indicates an exact signal correspondence, while zero value means that signals are not related. A negative correlation is obtained when one signal increases while the other tends to decrease.

A further improvement called *parametric correlation*, was proposed by Tzamali.[85] In this procedure, correlation values are enhanced through a masked equation that only preserves relevant portions in the compared signals. Parametric correlation can be calculated as:

$$r = \frac{\sum a_i(x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum a_i(x_i - \bar{x}_i)^2}\sqrt{\sum a_i(y_i - \bar{y}_i)^2}} \tag{3.7}$$

where $a_i$ is a binary mask, where the zeros corresponds to those amplitude values below a threshold that takes into account the line widths on the reference spectrum.

**Figure 3.6**

*Mask calculation from a Al I reference spectrum. Measured spectra was taken form a copper benefit slag.*
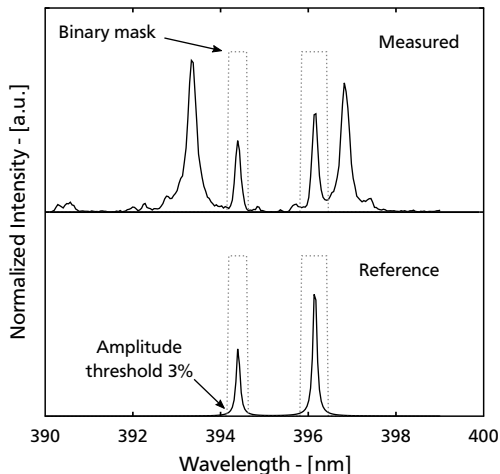
Figure 3.6 shows a binary mask calculated at a threshold of 3% maximum amplitude for an Al I reference spectrum. Measured signal comes from slags of an ore copper extraction process. When equation 3.6 is applied on both signals, a low likeness is indicated by a $r = 0.382$ value. If parametric correlation is applied a value of 0.962 emerges, which better reflects the fact that reference signal is part of the measured signal.

## 3.4   Wavelength Similarity Coefficient

Scaling differences between measured and reference spectra is irrelevant, because the correlation procedure is invariant under linear transformations. However, this also implies that, as long as relative line intensity is preserved, both low emission signals and signal portions with high emission will be ranked equal. This fact is a common source of false positives when species with a great number of emission lines are considered such as Fe, Ti, and W.

When a researcher analyses a spectrum to identify a species, more attention is paid to lines with highest amplitude, i.e., peak relevance is proportional to peak amplitude. This fact is very similar to positional notation method for encoding numbers,[86] where the symbol position determines its magnitude. In the most general case, any fractional number $N$ can be represented on an arbitrary base $b$ as

$$N = \sum_{i=1}^{n} a_i\, b^{-i} \tag{3.8}$$

where $a$ is an integer of the base, and $n$ is the number of symbols required.

The wavelength comparison procedure can only generates two states: presence or absence of a particular emission line. Therefore, the binary base is the obvious choice. With the aim of assigning more relevance to higher intensity lines, matching emission lines must be ordered in descending intensity order. Furthermore, not all the species have the same number of $n$ lines in the evaluated region, so equation 3.8 must be normalized:

$$\text{WSC} = \sum_{i=1}^{n} a_i\, 2^{-i} \bigg/ \sum_{i=1}^{n} 2^{-i} \tag{3.9}$$

this expression, will be refereed as the *Wavelength Similarity Coefficient* (WSC).

As an example consider a hypothetical measured spectrum compared against two different reference spectra, as shown in figure 3.7. The numbers represent the relative relevance of each line in the corresponding reference spectrum, and came from the sorting procedure. Species A, according to equation 3.8, has $a = [1, 1]$ and $n = 2$; while for species B, $a = [1, 1, 0]$ with $n = 3$. If equation 3.8 is applied the same value of 0.75 would be obtained in either case. However, after normalization the WSC value is 1 for species A, and 0.86 for species B, indicating a better wavelength match for the species A.

*Example of WSC calculation with an hypothetical measured spectrum and two species. WCS is 1 for species A and 0.86 for species B.*
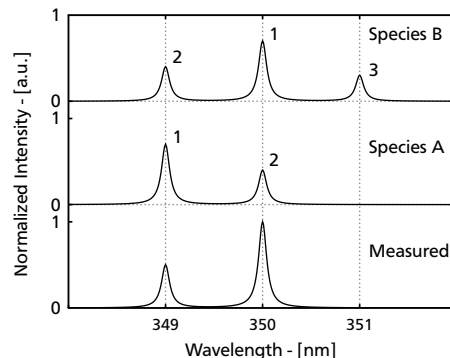


Note that WSC coefficient will not be affected by the height differences between the lines of the same species as long as the reference lines appear in the measured spectrum. For example, species A in figure 3.7 has a WSC of 1, despite of significant height differences with respect to the measured spectrum. Such dramatic effect could be observed in a self-absorbed line. In those cases, one could implement a more realistic model for reference spectrum generation, that takes in to account effects as self-absorption, Doppler line broadening, among others; nonetheless, this will only improve the correlation scores without any impact on WSC values.

As mentioned before blending of several emission lines can occur, so from a comparison point of view, is more reliable to use the calculated reference spectrum band's position than those available at the databases. Maximum peak location can be determined with the first derivative criterion employing the symmetrical form given by equation 3.10, which preserves symmetry and posses a low truncation error.[87]

$$f'(x) = \frac{f(x + \Delta) - f(x - \Delta)}{2\Delta} \qquad (3.10)$$

## 3.5   Elemental Identification in Action

Parametric correlation alone is not a sufficient criterion to identify the elemental composition of a sample. As an example, a full analysis of a D2 steel sample is presented, which has been previously analyzed by electron microprobe technique for validating purposes.

The experimental setup for this particular example employs a 532 nm Nd:YAG laser, with maximum energy of 30 $\mu$J in a 400 ps pulse, and adjustable repetition rate from 10–1000 Hz. Plasma emission is collected by a 2f lens system and guided with an optical fiber to a Czerny-Turner spectrometer, which has a 500 mm focal length and 1200 lines/mm grating, accomplishing a 0.02 nm resolution. Finally, an intensified CCD is used as a detector. With the aim of reducing surface damage and measure larger areas, samples rest on an electronic controlled $xy$ position system.

Within this setup spectral windows of approximately 40 nm are obtained, so it is necessary to make several acquisitions with the central wavelength displaced, to cover the whole instrumental range which goes from 260 to 920 nm.

Every measured spectrum comprises 30 software accumulations, where each accumulation consists of 500 shots integrated on chip. This procedure provides a very repeatable acquired spectrum with a signal to noise ratio of 95 for the Fe I line located at 302.107 nm. At a 1000 Hz laser repetition frequency, the whole instrumental spectral range can be measured in a few minutes. Spectrum baseline was removed with the Peak Clipping Algorithm.

The reference spectrum is calculated as a superposition of Lorentizan shapes, with normalized intensity given by equation 3.3. In those cases where not enough information exist for plasma modeling, the normalized NIST relative intensities are used instead. All the reference spectra are normalized respect to the highest line emission of the corresponding species in the measured spectral window range.

LTE condition is necessary to apply the plasma model. A commonly parameter used as a proof of LTE existence is the McWhriter criterion[88] that is based on plasma electron density estimation. Measuring plasma electron density requires the examination of reference emission lines,[89] which cannot be established until elemental composition has been determined. The approach used in this work, is to set the experimental parameters, as window sampling delays and acquisition times, to values where LTE conditions are fulfilled on samples with known composition.

### Table 3.4

*Parametric correlation results, using a 0.5 value as an identifying criterion as suggested by Tzamali*

| Element | Parametric correlation $r$ | |
| --- | --- | --- |
| | **NIST** | **Plasma Modeling** |
| **True positives** | | |
| Fe | 0.9092 | 0.9276 |
| Mn | 0.9204 | 0.5857 |
| Cr | 0.6093 | 0.8607 |
| Mo | 0.5338 | 0.5338 |
| Ni | 0.5260 | 0.5923 |
| Cu | | 0.6971 |
| **False positives** | | |
| Rn | 0.8145 | 0.8145 |
| Po | 0.7712 | 0.7712 |
| O | 0.6828 | 0.6828 |
| Hg | 0.6822 | 0.6822 |
| Li | 0.6193 | 0.7171 |
| Dy | 0.5835 | 0.6132 |
| S | 0.5711 | 0.5711 |
| Yb | 0.5598 | 0.5598 |
| Ho | 0.5068 | 0.6146 |
| Cl | 0.5292 | |
| K | | 0.6501 |
| N | | 0.5698 |
| Tl | | 0.5624 |

The full spectral instrumental range is measured in several windows. On each window, parametric correlation was applied for every neutral species ranging from H to Pb. Both methods for reference calculation are reported in Table 3.4. NIST labeled column gives correlation values using the amplitude values reported in literature, while plasma modeling column presents correlation values calculated with intensities given by equation 3.3. For a particular species, only the highest correlation value is presented.

According with the guideline provided by Tzamali, species with correlations higher than 0.5 were taken as a positive match. Top elements on table 3.4; i.e. Fe to Cu, are those whose presence was confirmed by microprobe technique, while all other elements are false positives, with the exception of S.

A rapid examination of table 3.4 shows a higher correlation when intensity is calculated by the plasma model, with the exception of Mn. Moreover, Cu was not detected using NIST intensities. This behavior is better explained with the aid of figure 3.8, where the Ca I emission lines measured from a mineral ore. This lines can be compared against the reference spectrum generated with the simplified plasma model. A good similarity is evident to the naked eye, which is confirmed with a $r$ value of 0.969 for plasma model versus 0.794 for NIST intensities.

Elements present in the sample and false identifications exhibit very similar correlation values, making impossible to discriminate between them. However, when WSC is used as a complementary decision factor better discrimination can be obtained, see table 3.5. WCS calculation is a fast and straightforward procedure, that quickly removes emission species whose most prominent lines does not match the measured spectrum in wavelength terms.

Table 3.5 shows that all the species with higher concentrations are correctly identified and false positives are no longer present. Elemental discrimination was obtained under the following conditions: WSC $\geq$ 0.8, $r \geq$ 0.5, and *the sampling region must include the maximum predicted intensity for that particular species*. A WSC value of at least 0.8 guarantees the identification of the two most prominent lines, or the higher intensity line along with several lines below the second most prominent.

Two main factors that affect the wavelength position of the maximum predicted intensity for a particular species can be identified. The first factor is the origin of

**Figure 3.8**

*Comparison between plasma model reference (top part of figure) and measured spectrum of calcium, found on a mineral ore (bottom part).*
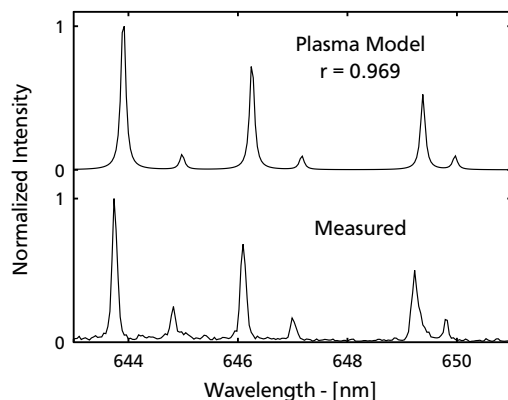
**Table 3.5**

*Identifying results using WSC as a complementary criterion to parametric correlation.*

| Element | WSC | r | Weight concentration | |
| --- | --- | --- | --- | --- |
| | | | **Microprobe** | **Literature** |
| Fe | 1.0000 | 0.6949 | 83.20 | 83.59 |
| Cr | 1.0000 | 0.5972 | 14.25 | 12.00 |
| Ni | 0.9998 | 0.5119 | 0.24 | 0.30 |
| Mo | 0.8571 | 0.5410 | 0.82 | 0.80 |
| Cu | 1.0000 | 0.7077 | 0.09 | 0.25 |
| Mn | 1.0000 | 0.5886 | 0.33 | 0.35 |

intensities employed to evaluate the reference spectrum; e.g., in the range of 260 to 1000 nm, maximum NIST intensity for Fe I is reported at 374.5561 nm, while for a temperature of 10,000 °K the simplified plasma model predicts that the highest line will be located at 358.1193 nm. Temperature used in reference calculation is the second factor, and only applies to plasma model.

Plasma temperatures for LIBS experiments ranges from 6000–23000 °K.[49] This wide range not only affects the highest line wavelength position, but also the relative intensities in the reference spectrum. In figure 3.9, the relative intensity for some of the highest lines for Fe I is plotted against temperature. For temperatures below 8500 °K, the highest relative intensity is located at 373.4864 nm, and is replaced by the line at 358.1193 nm for higher temperatures.

This fact highlights the importance of having an adequate estimation of the plasma temperature produced by the experimental setup. Once this value is known, a calculation of relative intensities for each species can be attained in order to obtain the corresponding highest emission wavelength position.

There are other elements present in the sample as Si, V, C, S and P that were not detected, and require special treatment. Detection of lighter elements as C and S in steel, requires measuring the sample in a vacuum chamber, or a rare gas atmosphere as suggested by Khater.[90] On the other hand, Si and V are elements, in this particular sample material, with very low intensities; although, they can be detected by relaxing the searching conditions. To this end, all the species with

**Figure 3.9**

*Fe I plasma model reference spectrum normalized intensity for some prominent lines*
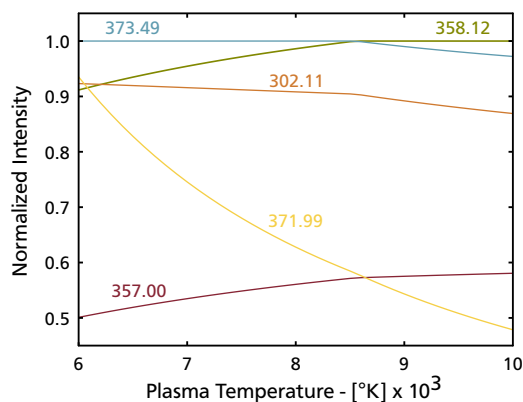
**Table 3.6**

*Detection of very low emitting species, achieved when WCS and correlation requirements are relaxed.*

| Element | WSC | r | Weight concentration | |
|---|---|---|---|---|
| | | | **Microprobe** | **Literature** |
| **Trace True Positives** | | | | |
| Si | 1.0000 | 0.1787 | 0.23 | 0.25 |
| V | 0.5806 | 0.0571 | 0.80 | 0.90 |
| **Trace False Positives** | | | | |
| Te | 0.6667 | 0.1092 | | |
| Os | 0.5709 | 0.1139 | | |
| W | 0.6378 | 0.0512 | | |
| Tc | 0.8571 | 0.2378 | | |

positive correlation and WSC $\geq$ 0.5 in the region of maximum emission were taken as a correct match. Results are shown in table 3.6 where only the additional matches are included.

Besides Si and V there are also four false positives. Although user intervention was needed to discriminate this very low emitting species, the possible candidates were reduced considerably. The reduction of candidates species with low emission is a feature not commonly available in literature. All the false positives were due to emission overlaps between the principal lines of reference spectrum of Te, Os, W, and Tc with lines pertaining to another species present in the sample.

Elemental identification in this particular material was achieved matching only neutral species, but in other cases, as elements whose highest emission pertains to ionic species, or for quantification procedures such as Calibration-Free,[82] the identification of ionic species needs to be considered. Under such circumstances, neutral and ionic species are processed separately when needed.

The lack of complete information to calculate the plasma model is a downside of this procedure and NIST intensities must be used instead when no data are available, until new emission probabilities and energy level values are reported in the literature for those species.

# Chapter 4

# Multivariate Analysis

Information age has led to an impressive amount of data stored in data warehouses around the world. Data come from very distinct sources, such as stock market registers, automated laboratory equipment, marketing studies, Population Census, etc. When these data are arranged in tables a common property emerges: multiple features were measured on data or individuals;[91] in other words, data came from *multivariate measurements*.

Some basic information can be extracted from multivariate measurements using simple statistics,[92] but obtaining a clear idea of the main parameters that affects the process under study is still a challenge. To this end, *Multivariate Analysis* or MA for short, are a set of techniques that essentially model the reality with a reduced set of variables. MA techniques have been available since the dawn of twentieth century, although its application was severe limited by the tedious calculations involved. This scenario dramatically changed with the advent of computers, making very fast analysis on large datasets a reality.

One of the goals of multivariate analysis is to generate a *variate*, which is a linear combination of variables with weights determined empirically. The mathematical expression of variate is shown in equation 4.1, where $x_n$ are the measured variables and $p_n$ are weights determined with a multivariate technique.

$$z = p_1 x_1 + p_2 x_2 + p_3 x_3 + \ldots + p_n x_n \tag{4.1}$$

Equation 4.1 could be used to obtain a maximum correlation between multiple independent variables and a single dependent variable. This process is known as *multiple regression*. Another use is in *discriminant analysis*, which identifies groups of observations. Finally, finding the underlying structure formed by the existing relationship among two or more variables—a process termed as *factor analysis*—can also be done.

A MA technique called *Principal Component Analysis* or PCA is presented in this chapter, which will be used as both factor and discriminant analysis. PCA was developed by Karl Pearson in 1901; although the general procedure as we know it today appears in a paper written by Harold Hotelling in 1933.[93]

As will be shown, LIBS spectra PCA can often eliminate the quantification step for obtain a relationship among a group of related samples, thus simplifying the involved work. A full study on metallic objects of archaeological interest will be shown. This is a first effort, done in collaboration with archaeologists, to associate objects with its raw material source and historical manufacturing period. PCA will also be used to classify a variety of steels by its crystalline structure. But before, a brief explanation for computing and interpreting PCA is introduced.

## 4.1   **Fundamentals of Principal Component Analysis**

Basic statistical techniques can be applied on multivariate data; for example, given a data set composed of several measurements of the same variable, it is possible to calculate its *mean*, a single value that represents the distribution's data center; while the *variance* is a measure of the typical dispersion of the data around the mean. Relationships among two variables also can be inferred with the *covariance*, which is a measure of the dependencies between the variables; i.e., what happens to a variable when the other is modified. Covariance can be calculated with:

$$\text{cov}(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1} \tag{4.2}$$

where $\bar{x}$ and $\bar{y}$ are the mean values corresponding to the variables denoted by $x_i$ and $y_i$. Positive covariance values indicate a direct linear relationship among $\mathbf{X}$ and $\mathbf{Y}$, while negative values reveals variables that are inversely related; zero values indicate no relationship. It must be noted that the variance is equivalent to the covariance of a variable with itself.

In this work, multivariate data is represented as a matrix $\mathbf{X}$ whose columns represent the $n$ measured variables and its rows correspond to the $m$ observations or events. The *dispersion matrix*, also known as *covariance matrix* or *correlation matrix*, is a convenient way to summarize the relationships among variables from multivariate data, which can be computed as:

$$\mathbf{C} = \begin{bmatrix} \text{cov}(\mathbf{X}_1, \mathbf{X}_1) & \text{cov}(\mathbf{X}_2, \mathbf{X}_1) & \ldots & \text{cov}(\mathbf{X}_n, \mathbf{X}_1) \\ \text{cov}(\mathbf{X}_1, \mathbf{X}_2) & \text{cov}(\mathbf{X}_2, \mathbf{X}_2) & \ldots & \text{cov}(\mathbf{X}_n, \mathbf{X}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(\mathbf{X}_1, X_n) & \text{cov}(\mathbf{X}_2, X_n) & \ldots & \text{cov}(\mathbf{X}_n, X_n) \end{bmatrix} \tag{4.3}$$

If every $\mathbf{X}_i$ is *mean centered*, said in other words, each variable is subtracted from its corresponding mean, the correlation matrix can be expressed as

$$\mathbf{C} = \frac{\mathbf{X}^T \mathbf{X}}{n-1} \tag{4.4}$$

Now, consider the number of estimated parameters $N$ that must be computed and interpreted for the multivariate case with $n$ variables. Taking into account only mean, variance and covariance $N$ will be[94]

$$N = \frac{n^2 - n}{2} + 2n \tag{4.5}$$

*Amount of parameters to be interpreted $N$ according to the variable number $n$ in the model*
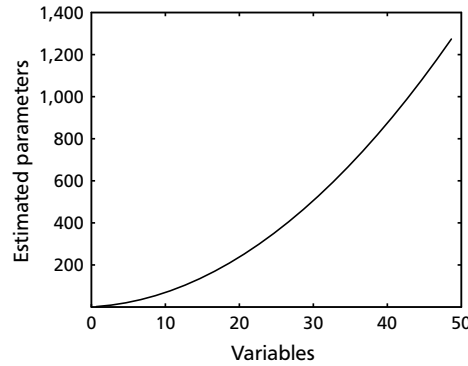


Equation 4.5 has an exponential nature as figure 4.1 shows. Multivariate data analysis with only 25 variables would require interpreting 350 parameters. PCA solves this dilemma with an *orthogonalizing transformation* that converts original correlated variables into a set of uncorrelated variables. This is equivalent to say that the correlation matrix of the transformed variables will be diagonal; which implies that only a handful of values are left for interpretation. Moreover, as will be shown later, only a reduced set of these transformed variables can explain the most of model's variability.

Before explaining PCA computation procedure, it is worth to introduce a very useful result from linear algebra[95]

> *Given an arbitrary real $m \times n$ matrix* **A**, *there are orthogonal matrices* **U** *and* **V** *and a diagonal matrix, this time denoted* $\Sigma$, *such that*

$$\mathbf{A} = \mathbf{U}\,\Sigma\,\mathbf{V}^{T} \tag{4.6}$$

> *i)* $\Sigma$ *are the square roots of the eigenvalues of* $\mathbf{A}^{T}\mathbf{A}$
> *ii) The columns of* **V** *are eigenvectors of* $\mathbf{A}^{T}\mathbf{A}$
> *iii) The columns of* **U** *are eigenvectors of* $\mathbf{A}\mathbf{A}^{T}$

Equation 4.6 is commonly known as *Singular Value Decomposition* or SVD, while a *singular value* is the name given to each value of $\Sigma$. SVD can be used to calculate PCA in a straight forward manner. To this end, it is necessary to apply SVD to a set of mean centered data **X** and using the fact that $\mathbf{V}^{T} = \mathbf{V}^{-1}$ is fulfilled in an orthogonal square matrix, which leads to

$$\mathbf{Z} = \mathbf{X}\mathbf{V} = \mathbf{U}\Sigma \tag{4.7}$$

This last expression is the matrix form of a variate introduced in equation 4.1, where **Z** represents the transformed variable that is commonly referred as *scores*,[96] while the columns of the linear transformation **V** are known as *principal components* (PC) or *loadings*[†] of **X**. From the definition of SVD, is easy to see that the **V** columns are the eigenvectors of $\mathbf{X}^{T}\mathbf{X}$.

---

[†]*This work follows the naming convention suggested by Eriksson et al.,[96] but some authors name the transformed variable as principal components, e.g., Jackson.[93]*

With the aim to prove that the transformed set of variables is uncorrelated, the dispersion matrix of $\mathbf{Z}$ will be calculated

$$\frac{\mathbf{Z}^T \mathbf{Z}}{n-1} = \frac{1}{n-1} \, (\mathbf{X}\mathbf{V})^T \mathbf{X}\mathbf{V}$$

$$C(\mathbf{Z}) = \frac{1}{n-1} \, \mathbf{V}^T \mathbf{X}^T \mathbf{X}\mathbf{V}$$

is a well-known fact that $\mathbf{X}^T\mathbf{X}$ is a real, square and symmetric matrix,[97] such matrix can always be decomposed as the product of $\mathbf{V}\Lambda\mathbf{V}^T$, where $\mathbf{V} = [\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}]$ are the eigenvectors of $\mathbf{X^T X}$ and $\Lambda$ is a diagonal matrix with the corresponding eigenvalues of $\mathbf{V}$; therefore

$$C(\mathbf{Z}) = \frac{1}{n-1} \, \mathbf{V}^T \mathbf{V}\Lambda\mathbf{V}^T \mathbf{V} = \frac{1}{n-1} \, \Lambda$$

$$= \frac{\Sigma^2}{n-1} \tag{4.8}$$

therefore

$$\Lambda = \Sigma^2 \tag{4.9}$$

The diagonal elements in $\Sigma^2$ are named *characteristic roots*.

PCA is sensitive to scaling, which means that importance attributed to a given variable in the model can be altered artificially modifying its variance. Eriksson recommends a preprocessing data step called *unitary variance scaling* that simply divides every variable by its associated standard deviation. It must be remembered that mean centering is also a prerequisite in order to compute PCA from SVD.

An equivalent procedure, termed *Standard Normal Variate* or SNV, was proposed by Barnes et al.[98] that also produces variables with unitary variance and zero mean; this is the preferred approach used in this work. SNV mathematical expression is

$$\hat{x} = \frac{x - \bar{x}}{\sigma} \tag{4.10}$$

### 4.1.1 PCA Properties

To close with PCA discussion some of its properties will be presented, which along equation 4.7 conform the core of PCA routines implemented in this work.

**Preserves total variability.** The original covariance matrix determinant $|\mathbf{X}|$ equals the determinant of the transformed variable $\mathbf{Z}$, which implies that trace of $\mathbf{Z}$ is equal to the sum of the original variances. This is particularly useful because the ratio of each characteristic root to the trace of $\Sigma^2$, indicates the proportion of total variability accounted for each principal component. The mathematical expression is

$$\sigma_{z_i}^2 = \frac{\Sigma_i^2}{\mathbf{tr}(\Sigma^2)} \tag{4.11}$$

**Signal reconstruction.** Original data set **D** can be reconstructed from principal components by solving **X** in equation 4.7

$$\mathbf{D} = \bar{x} + \mathbf{X}\mathbf{V} + \mathbf{E} \tag{4.12}$$

where the residual matrix **E** contains the noise produced when only a few principal components are used while reconstructing. A plot with the partial reconstruction transformed variables is called *score plot*, while a *loading plot* is made representing two or three PC within a Cartesian plane.

**Groups similar behavior events.** Near points in score plots indicates events with similar properties, while those far from each other show events with distinct behavior. Average property events are represented with points near to zero from any axis.

**Expose variables relationships.** A plot of principal components loadings, shows the relationships between all the variables of **X** at the same time. If two or more variables contributes with similar information, will be plotted closely together; conversely, inverse relationship is indicated with points in opposite quadrants. Distance to origin also conveys information; the further away a variable is, the more impact in the model has.

### 4.1.2  PCA Implementation

SVD is available as a Matlab routine, which sorts vectors of matrices **U**, $\Sigma$ and **V** in descending order relative to its eigenvalue magnitude. PCA calculation was written in a function named `PCA_Do`, which decomposes the original matrix using SVD. `PCA_Do` computes a zero mean and unitary variance matrix before obtaining eigenvectors and eigenvalues. Scaled data matrix columns represents the intensity measured in a specific wavelength. As an added benefit, SVD is a more stable algorithm numerically speaking than the common procedures for eigenvector and eigenvalue determination.

`PCA_PlotData` is an auxiliary function that presents in a single graph all the analyzed spectra. This function is very useful for determining the preprocessing steps required by each spectrum. The importance of a good signal conditioning should not be underestimated; in most of the cases, these steps make the difference between an useless group and a clear grouping set in the score plot.

A score plot is done with the `PCA_PlotScores` function that is used to highlight groups; it can make 2D or 3D plots using any score combination. On the other hand `PCA_PlotPC` helps to identify the most influential variables in the model. The number of variables to display can be defined by the user. Each point is labeled according to its corresponding wavelength position.

Admittedly in most of the cases the wavelength position can be confusing, so an alternative labeling can be shown with `PCA_WaveToSymbol`. This function uses the information provided by the `Find` function to substitute each wavelength with its corresponding emission species, as seen on figure 4.2.

**Figure 4.2**  *Elemental composition and loading plot for a copper slags in the 295 nm region. a) Synthetic spectrum provided by `Find` function. b) Resulting loading plot after applying `PCA_PlotPC`. c) Label substitution done with `PCA_WaveToSymbol`.*
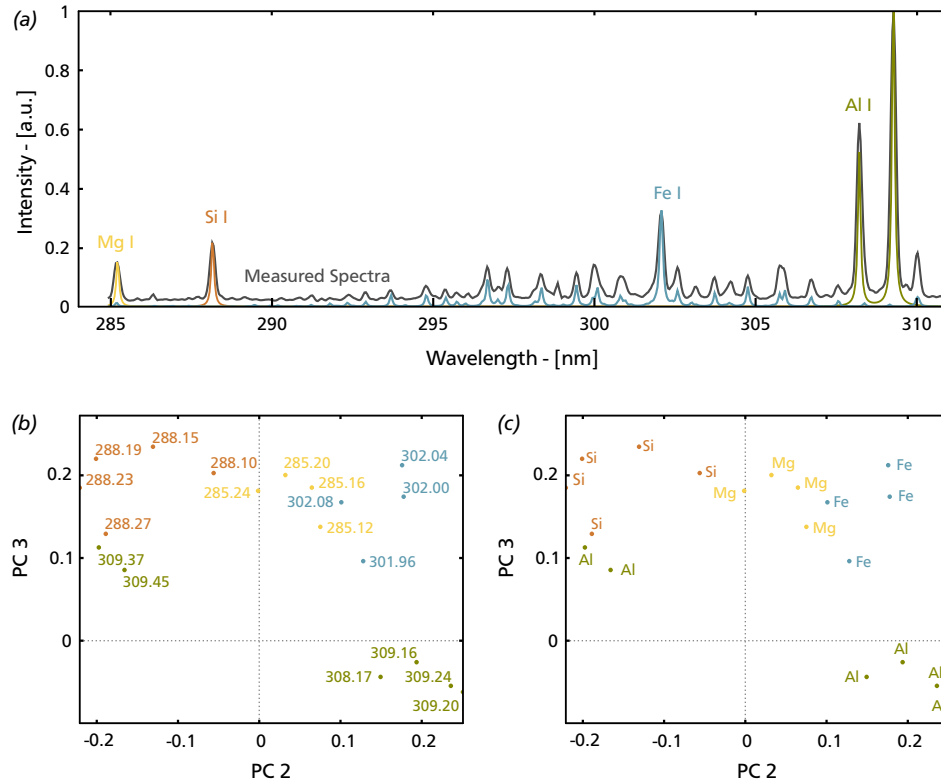


Spectral data is often stored in distinct paths on the hard drive; organized by sample type, experimental setup, etc. While is true that a full path name can always be indicated, this also implies a lot of work just to provide the correct file location. Moreover, PC analysis on spectra acquired with a Czerny-Turner spectrograph can involve from tens to hundreds of files.

In order to manage all this complexity, an auxiliary function named `PCA_Source` was written, which accepts a set of file names and assigns them to a specified path. This routine also allows making file groups for analysis. The output of `PCA_Source` can be fed to `PCA_Load` to make a matrix with the contents of the specified files, it can be used directly with `PCA_Do`, or it can be employed in the `Spec_Treatment` function, introduced in section 2.4.3, in order to apply the same preprocessing steps in a group of related files that share conditioning requisites. Groups may also be created manually using regular expressions.

Finally, `PCA_Ellipse` uses Hotelling's $T^2$ in order to highlight outliers with any confidence interval within a particular group. $T^2$ is a multivariate generalization of Student $t$-test employed to verify if a process is under control. The computation of this statistic distribution is adapted from the code written by A.J. Johnson[99] and verified with the tables published by Jones.[100] `PCA_Ellipse` can plot in two and three dimensions.

## 4.2   Archaeological Copper Slag Analysis

The office of Instituto Nacional de Antropología e Historia (INAH) at Michoacán is studying Prehispanic and early colonial metallurgy in the area of the Balsas river basin, located in the aforementioned state in México. This research is conducted by Dr. José Luis Punzo Diáz, who seeks a better understanding of the metallurgic processes carried out by the Tarascan culture more than 500 years ago.

The so called "Señorío Tarasco" is considered as the most important center of copper production in Mesoamerica during the late Post classic period. The last Tarascan capital before the arrival of the Spaniards "Tzintzuntzan", is located within the region of Pátzcuaro lake.

Although many copper objects have been found in Tarascan settlements, the copper extraction must be done in other places. The reason behind this claim is the absence of copper mines near to Pátzcuaro lake; nearest copper deposits are located to the south around 80 km in a region known as "Tierra Caliente", in places as Inguarán, Churumuco and Cuitzián (see figure 4.3).

Inguarán mine and surrounding regions are of archaeological interest because it was the most important mine in the Prehispanic period, and remained as the main copper supply in the New Spain. Common copper minerals extracted in this region are chalcopyrite ($CuFeS_2$), bornite ($Cu_5FeS_4$), and arsenopyrite ($FeAsS$). Bastan is another copper mine, but there are not evidences of Prehispanic activity in it. In this case bornite and chalcopyrite are the main extracted minerals.

**Figure 4.3**

*Copper mines location and main Tarascan settlements with evidence of copper objects manufacturing. 19° 10′ N 101° 25′ W*[101]

It seems that the first step in the Tarascan process of copper extraction was carried out in the vicinity of the mines, where workers obtained copper ingots of about 2.9 kg weight. However high scale foundry was carried out at places very distant from mines, as is the case of Santa Clara del Cobre region; see figure 4.4.

Little Prehispanic foundry evidence has been discovered at the present day; the main vestiges are *slags*. Casting debris or slags are residuals coming from the metallurgical process of copper extraction. A copper slag resembles to compact black stones, as can be seen in figure 4.5a.

Ores are concentrations of minerals in rock, high enough to be economically extracted for use. All ores are minerals, but all minerals are not necessarily ores. The surrounding rock is called *host rock* and has a great influence in the mining method. The age of a mineral deposit can also be identified with the host rock. Figure 4.5b is a picture of an ore coming from Inguarán mine.

**Figure 4.4**

*Prehispanic/Colonial settlements for the production of copper objects near to Santa Clara del Cobre. 19° 23′ N 101° 38′ W [101]*
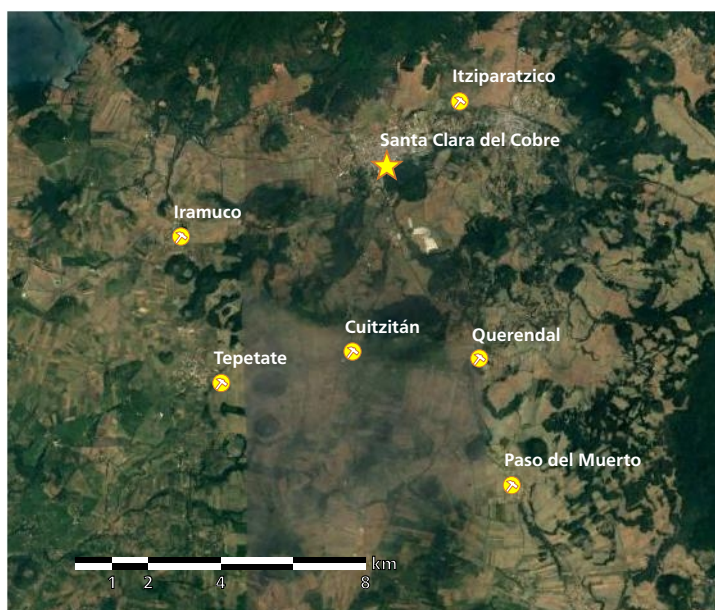


**Figure 4.5**

*a) Copper slags resemble black solid stones.*
*b) Copper ore obtained from Inguarán mine*



*(a)*                    *(b)*

### 4.2.1  Experimental Procedure

LIBS technique was applied on 11 copper slags collected in several archaeological sites. Table 4.1 shows the origin region for each sample, as well as the historical period associated to each piece. Dating was obtained with an archaeomagnetic technique, as reported by Morales et al.[102]

The experimental arrangement was described in section 3.5. Each spectrum is the superposition of 30 measurements, where each measurement integrates on-chip 500 shots. During LIBS measurements, the sample is moved automatically in steps of 100 $\mu$m every tenth of second. Therefore, each spectrum measures a surface of 0.15 mm$^2$. Despite the great number of shots involved, the sample damage infringed is low because the low energy carried out by each laser pulse.

LIBS measurements with windows centered at 295, 340, 375, 410, 440, 595 and 660 nm were performed on each slag, with a gate width of 250 ns and delay of 80 ns. Spectra were wavelength calibrated using both shifting and polynomial methods. A combination of the synthetic spectrum of Li I, Na I, Mg I, Al I, Si I, Ca I and Cu I was used as reference, depending on the spectral region. Additionally, signal overlapping was resolved by trimming the spectra in a valley when necessary.

Negative intensity values appear after applying SNV, so, an offset had to be added in each spectrum before performing a baseline removal with a second order 50 points descendant SNIP algorithm. The last preprocessing step was to join the seven spectral regions, thus obtaining a single spectrum for each sample. Analysis on wider spectral regions helps to improve the elemental identification process, diminishing false positives in trace species.

The elemental search procedure explained in chapter 3 was applied for each slag, looking for neutral species. The wavelength tolerance was 0.05 nm, all the lines with intensities below 4% were excluded from reference calculation and only experimental emission lines above 4% were used for WSC computing. A WSC value of $\geq 0.5$ and $r \geq 0.2$ were used as selection criterion, which in this particular case, allows identifying low emission trace elements as Mn in a single step.

*Table 4.1*

*Analyzed samples region of origin and dating, which was determined by Morales et al., through the use of an archaeomagnetic technique.*

| Code | Site | Region | Dating |
|------|------|--------|--------|
| cutz | Cuitzitán | Sta. Clara | Prehispanic |
| itzi | Itziparatzíco | Sta. Clara | Prehispanic |
| iram | Irámuco | Sta. Clara | Colonial |
| pamu | Paso del Muerto | Sta. Clara | Colonial |
| pasm | Paso del Muerto | Sta. Clara | Colonial |
| qdal | Querendal | Sta. Clara | Colonial |
| quer | Querendal | Sta. Clara | Colonial |
| tepe | Tepetate | Sta. Clara | Colonial |
| tept | Tepetate | Sta. Clara | Colonial |
| bast | Bastán | Bastán | Colonial |
| hueq | Cuandarillo | Huetamo | Prehispanic |

### 4.2.2 Results and Discussion

According to the results presented on table 4.2, Santa Clara region samples have similar chemical composition. These results have good qualitative agreement with those values published by Morales et al.—reproduced for comparison purposes in table 4.3 and 4.4—which summarizes the chemical composition obtained by spectral X-ray mapping analysis on `cutz`, `itzi`, `iram` and `pasm` samples.

   Light elements as C, P, and S, are difficult to detect with the experimental LIBS setup employed. In these species, the highest emission line is emitted in a zone that coincides with a region of low sensitivity of the detector. However, some trace elements with high emission can be detected as Na, Ca and Mn. Also note that Li could not be detected with the spectral X-ray mapping technique.

   It must be said that Mo was also detected as a trace element; however, the lack of information on NIST database for computing the synthetic spectrum with the

**Table 4.2**

*Slag elemental analysis obtained with LIBS spectral analysis for neutral species. A dash (-) is employed to indicate that the species was not found due to the lack of spectral information.*

| Code | Li | Na | Mg | Al | Si | Ca | Mn | Fe | Cu | Ag |
|------|----|----|----|----|----|----|----|----|----|----|
| cutz | ★ | ★ | ★ | ★ | ★ | ★ | | ★ | ★ | |
| itzi | ★ | ★ | ★ | ★ | ★ | ★ | | ★ | ★ | |
| iram | ★ | ★ | ★ | ★ | ★ | ★ | ★ | ★ | ★ | |
| pamu | ★ | ★ | ★ | ★ | ★ | ★ | | ★ | ★ | |
| pasm | ★ | ★ | ★ | ★ | ★ | ★ | | ★ | ★ | |
| qdal | ★ | ★ | ★ | ★ | ★ | ★ | | ★ | ★ | |
| quer | ★ | ★ | ★ | ★ | ★ | ★ | ★ | ★ | ★ | |
| tepe | ★ | ★ | ★ | ★ | ★ | ★ | | ★ | ★ | |
| tept | ★ | ★ | | ★ | ★ | ★ | | ★ | ★ | |
| bast | ★ | ★ | ★ | ★ | ★ | ★ | ★ | ★ | ★ | |
| hueq | - | - | ★ | ★ | ★ | ★ | | ★ | ★ | ★ |

**Table 4.3**

*Semi-quantitative elemental analysis of Prehispanic / Colonial slags, reported by Morales et al.*

| Atom | itzi1 | itzi2 | itzi3 | cutz1 | cutz2 | cutz3 |
|------|-------|-------|-------|-------|-------|-------|
| C | 8.45 | 7.90 | 7.62 | 14.99 | 13.05 | 11.13 |
| O | 32.19 | 32.31 | 32.44 | 41.64 | 33.56 | 33.83 |
| Na | 0.11 | | | 0.19 | 0.10 | 0.10 |
| Mg | 0.57 | 0.57 | | | 1.03 | 1.05 |
| Al | 1.58 | 1.49 | 1.56 | 3.09 | 3.27 | 3.45 |
| Si | 1.65 | 1.66 | 1.71 | 2.62 | 2.61 | 2.73 |
| P | | 0.08 | | 0.07 | 0.09 | 0.10 |
| S | | | | | | 0.03 |
| K | 0.16 | 0.19 | 0.24 | 0.11 | 0.06 | 0.09 |
| Ca | 0.17 | 0.18 | 0.20 | 0.98 | 1.23 | 1.27 |
| Ti | 0.09 | 0.06 | 0.06 | 0.06 | 0.08 | 0.06 |
| Mn | | | | | | 0.03 |
| Fe | 54.45 | 54.83 | 54.79 | 34.47 | 42.64 | 43.69 |
| Cu | 0.59 | 0.73 | 1.39 | 1.78 | 2.28 | 2.41 |

simplified plasma model, produces lower correlation coefficients but higher WSC values when compared with Mg. Besides, the main lines of Mo I overlap with Fe I lines, which makes it difficult to find clear evidence of its presence. The highest lines for O I and K I can be located in a region around 760-780 nm, which was not measured during the experimental procedure.

Is notable the absence of As in `bast` slag, because one of the main ores in Bastán mine is arsenopyrite. However, As ebullition point is 613 °C far below Cu fusion point of 1083 °C, meaning that most of As may be evaporated from slags during the Cu extraction process. Cu fusion point is also higher than Na ebullition point of 883 °C, which is 270 °C above the ebullition point for As. This fact combined with the high Na reactivity could explain why Na is found in slags but not As.

One may ask if distinct copper extraction process leave significant chemical composition differences in slags, such that, allows the identification of the historical period to which belongs. To handle this concern, contents of table 4.3 and 4.4 were analyzed with PCA; the resulting score plot and loadings plot are shown in figure 4.6.

By means of equation 4.11, is possible to find that the first two PC's explain 97.7% of model's total variance. In the PC plot, Mg, Na and Ca are near to the origin indicating that these elements have a marginal effect; on the contrary Fe and Cu have the biggest influence on the model. A comparison between the fourth quadrant of the score plot and the PC plot shows points a high content of Fe as the responsible for grouping `itzi`. The same reasoning can be applied to third quadrant, where the Al and Si content groups `pasm` measurements.

An average content of Fe and Cu is present in both `cutz` and `iram`. As is known from the table 4.1, they do not belong to the same historical period, so it must be concluded that chemical composition cannot be used for dating samples.

**Table 4.4**

*Semi-quantitative elemental analysis of Colonial slags.*

| Atom | pasm1 | pasm2 | pasm3 | iram1 | iram2 | iram3 |
|------|-------|-------|-------|-------|-------|-------|
| C  | 12.13 | 12.20 | 9.11  | 12.75 | 7.89  | 13.38 |
| O  | 40.85 | 38.65 | 37.73 | 39.66 | 33.68 | 32.85 |
| Na | 0.44  | 0.24  | 0.23  | 0.00  | 1.83  |       |
| Mg | 0.79  | 1.03  | 1.11  | 0.15  |       | 1.61  |
| Al | 6.85  | 6.61  | 6.47  | 2.22  | 2.61  | 2.93  |
| Si | 11.34 | 9.72  | 10.44 | 2.40  | 3.49  | 2.99  |
| P  | 0.32  | 0.32  | 0.31  | 0.18  | 0.06  | 0.06  |
| S  | 0.20  | 0.32  | 0.27  | 1.03  | 0.24  | 0.13  |
| K  | 0.38  | 0.46  |       | 0.03  | 0.05  |       |
| Ca | 1.07  | 1.05  | 0.99  | 0.11  | 0.65  | 0.59  |
| Ti | 0.27  |       | 0.28  |       | 0.09  | 0.09  |
| Mn |       | 0.12  | 0.14  |       |       |       |
| Fe | 22.29 | 24.87 | 27.74 | 32.85 | 39.38 | 36.95 |
| Cu | 3.05  | 4.33  | 5.19  | 8.61  | 9.81  | 8.38  |
| Zn |       | 0.09  |       |       |       |       |
| Mo |       |       |       |       | 0.22  | 0.05  |

**Figure 4.6**   *PCA of slag's chemical composition reported by Morales et al. The most influential elements in the model are Fe and Cu. Orange color in the score plot is assigned to Colonial samples. Roman numeral identifies the quadrant, see text.*
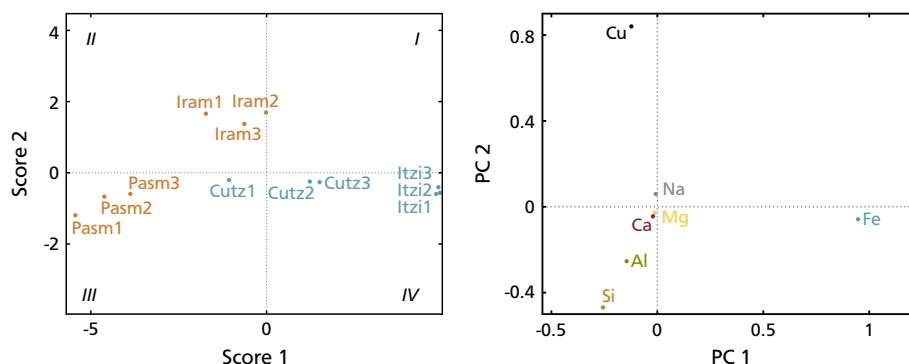


**Figure 4.7**   *LIBS spectra PCA for* `cutz`*,* `iram`*,* `itzi` *and* `pasm` *samples, in the spectral region comprised between 270-680 nm. Orange color in the score plot is assigned to Colonial samples.*
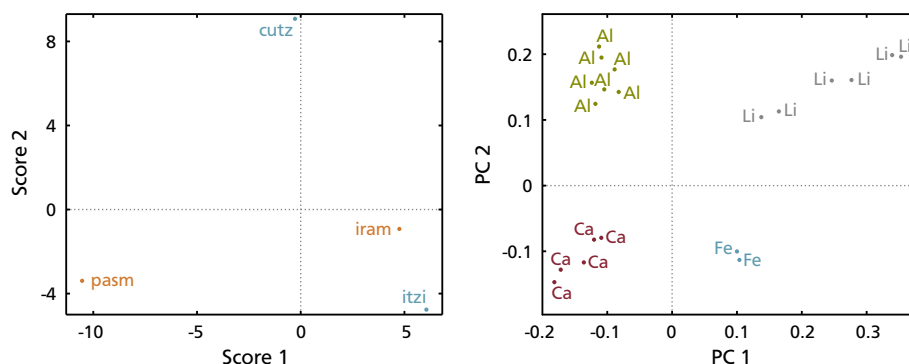


The PCA of the LIBS spectrum for the samples considered in tables 4.4 and 4.3 leads to similar conclusions. Their corresponding score and loadings plot are shown in figure 4.7. In this model, the first two PC explains 56.7% and 39.1% of the total variance respectively, accounting for 95.8% of model's variations.

Some differences must be expected between the PCA results of LIBS spectra and chemical composition. First of all, spectral X-ray mapping measures areas of a few $\mu$m, while the implemented LIBS setup analyzes larger areas in orders of magnitude, delivering measurements that are more representative of the sample's chemical composition. Secondly, although the intensity emitted by a given species is proportional to its concentration. Some elements are easier to excite. This means that in some cases a medium content element may have a higher emission than other elements with more concentration.

The relevance of Fe in the model, which is the highest concentrated species, is preserved for `itzi`, `cutz` and `pasm` samples, while it is higher for `iram`. Cu has lost its importance in the model. Ca, a minor concentration element, still differentiates the `pasm` sample. Ca is a stronger emitter than Al, which explains the differences in the ordinate axis for `cutz` and `iram` samples.

**Figure 4.8** *LIBS spectra PCA for the full slags set of table 4.1. Santa Clara del Cobre region's slags were grouped within a 95% confidence ellipse.*
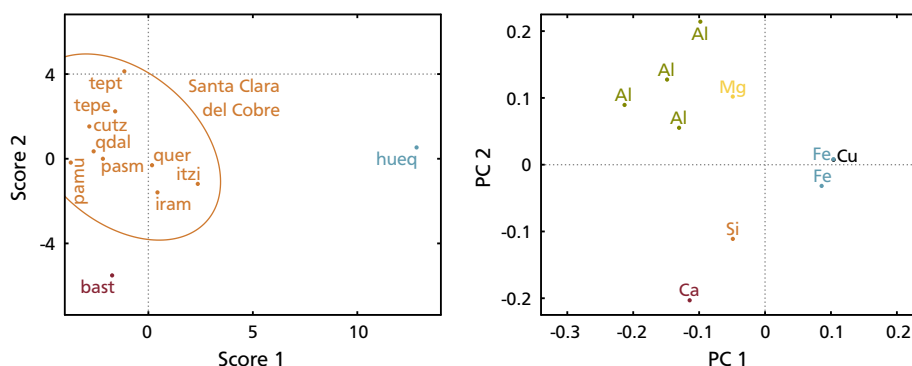


The score plot of figure 4.7 contains in its fourth quadrant samples of distinct historical periods, therefore, is useless as a dating technique. However, when the PC spectrum analysis is extended to all the available slags, new groups emerge, as shown in figure 4.8.

A 97.7% of model's variance is accounted for the first two principal components with 91% and 6.7% respectively. The slags pertaining to Santa Clara del Cobre region are surrounded by a 95% confidence ellipse; as no point falls outside the ellipse all the slags composition variations are statistically expected. Conversely, bast and hueq samples are clearly differentiated; the first has a higher Ca and Si content, while the second has a higher content of Fe and Cu.

Andesite, the host rock for Bastán mine, contains a higher Ca proportion than granodiorite,[103] which is the host rock for Inguarán site;[104] this may account for the direct relationship among bast and Ca. On the other hand, more Cu residuals on slags are expected with the use of rudimentary extraction techniques; explaining the hueq sample separation in the score plot, which is located in the opposite octant relative to bast sample and is fully associated with a Prehispanic context.

An experimental confirmation for the above claims was obtained with the LIBS analysis of ores coming from Inguarán and Bastán mines. Table 4.5 shows chemical composition given by LIBS spectral analysis, in regions centered at 295, 340, 375, 410 and 440 nm; it must be noted that the spectral regions for Na and Li were left out from the analysis. The lack of Si is due to the sampling method, which was realized on ore's metallic portions, taking cautions for not measuring the host rock. Ag presence on ming sample is not surprising because the commercial exploiting of Ag is realized in this region.[105] As expected, the Ca line at 422.66 nm had higher amplitude in mbst than ming, being 37.81% and 24.71% respectively, within a normalized spectrum.

**Table 4.5**

*Ore's chemical content obtained with LIBS spectral analysis.*

| Code | Mine | Mg | Si | Ca | Fe | Cu | Ag |
|------|------|----|----|----|----|----|----|
| ming | Inguarán | | ⋆ | ⋆ | ⋆ | ⋆ | ⋆ |
| mbst | Bastán | ⋆ | | ⋆ | ⋆ | ⋆ | |

All the available ore and slag spectra were gathered in a single PCA with the aim to identify slags's ore sources; figure 4.9 presents these results. Adding ores to the model moves the average elemental content in such a way that three PC are needed to account for 89.7% of total variance, with 70.3% for the first, 10.2% for the second and 9.2% for the third PC.

The high Ca concentration in `bast` sample is now explained by the negative portion of the third PC. As Al was not found in the ore's metallic component, it is possible to conclude that Al presence in slags is mainly due to the host rock which was not measured. Albite (NaAlSi$_3$O$_8$) and anorthite (CaAl$_2$Si$_2$O$_8$) are the main components that distinguish between granodiorite and andesite. These host rocks contain both albite and anorthithe but in different ratios. Andesite, the host rock of Bastán mine is richer in anorthite;[106] thus ores coming from Bastán mine will have a slightly higher Al content.

This seems to contradict figure 4.9 results where `bast` sample is separated by a low Al content. However, granodiorite has a greater density when compared with andesite, 2.6 g/cm$^3$ and 2.1 g/cm$^3$ respectively; implying that the separation of granodiorite by gravity concentration methods is harder because is closer to copper

### Figure 4.9

*Ore and slag PCA analysis from LIBS spectra. With ores addition, the model needs three PC for accounting 89.7% of total variation.*
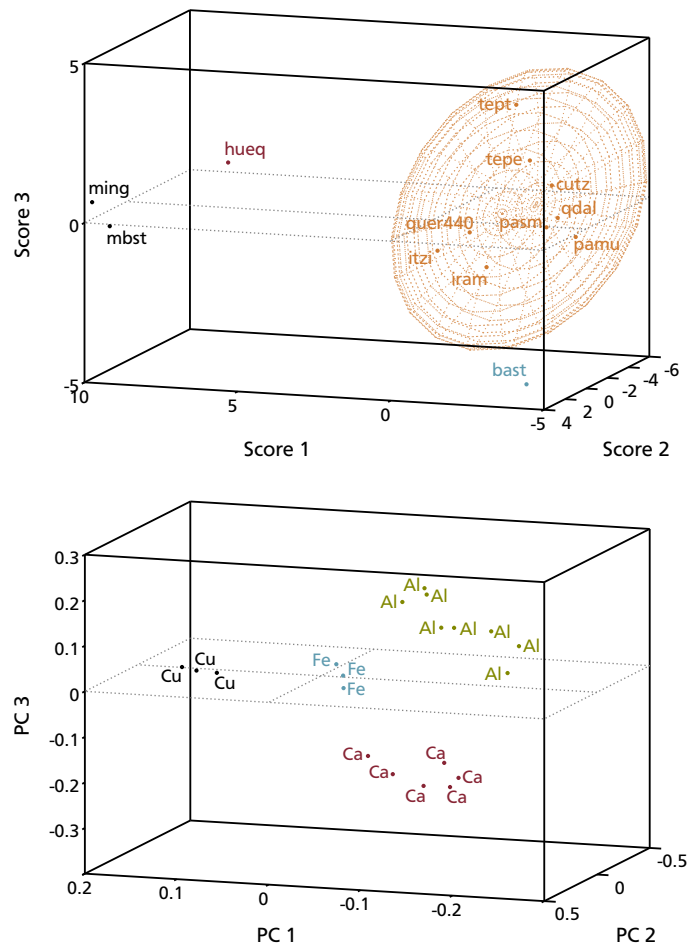
mineral density, which is above 4 g/cm$^3$. Under this circumstances, a highest Al concentration within slags coming from Inguarán mine should be expected.

Ore samples and `hueq` slag were grouped in a models' portion explained by high Cu content; although, `hueq` is separated by a higher Fe content, which is a natural consequence of Cu extraction process.

### 4.2.3  Conclusions

LIBS has proven to be a competitive technique compared to spectral X-ray mapping analysis; but offering added benefits as reduced cost and a faster acquisition. On the other hand, PCA is a powerful tool that allows the extraction of data relationships; in particular, when applied to spectroscopy measurements it can substitute the quantification process.

Not only high concentration species were detected by LIBS measurements but also trace elements. It is not possible to infer species concentration in a sample just by looking at the relative amplitude of line emission, because some elements are better emitters than others. However the slag chemical concentration reported by Morales et. al., allows to reach this conclusion.

Some light elements as C, S and P are difficult to detect with the LIBS setup employed. This issue is not exclusive for copper slags LIBS analysis but it is also shared within steel analysis. Arsenic has been reported as an alloy element used by Tarascan culture, although it was not detected in any of the slag or ore samples.

As is a component of arsenopyrite, a mineral that can be found at Inguarán mine; the analysis on `ming` sample does not show As traces but this finding is not conclusive. Other minerals can also be found at Inguarán mine and it is highly probable that the provided sample pertains to this class. It must be remarked that the reported analysis by Moreno et. al., did not find As. The As lack in slag samples could be due to the high temperature difference between the As ebullition point and Cu melting point, which is almost doubled by the second.

The accumulated experience during the realization of this work, indicates that not only the ore's metallic portions are relevant, but also the host rock. Moreover; the so called geological *plagioclase* classification system is based on the rock's Ca-Na ratio, which suggests that new information may be obtained extending the spectral sampling range to 670 nm region.

Along this work has been shown that chemical composition does not provide enough evidence to allow sample's dating, although provides valuable clues about the extraction process and opens the possibility for associating slags with its ore source. For example, the high statistical resembling of Santa Clara del Cobre slags indicates that they share similar raw materials and extraction process, which are clearly differentiated from those employed in `hueq` and `bast` slags. These results must be confirmed gathering more statistical evidence.

Finally, better copper extraction technologies such as the introduction of the covered furnace after the Spaniard arrival, decreased Cu residuals on slags; which could explain why `hueq` and `bast` slags are placed in opposite octants in the score plot space in figure 4.9; relating a higher slag Cu content with older pieces.

## 4.3   **Archaeological Metallic Object Analysis**

This section presents the results of the combined LIBS-PCA analysis performed on Prehispanic metallic objects, manufactured by the Tarascan culture, which were found in four archaeological excavation sites: "Tzintzuntzan" the main city of the Tarascan culture during the Spaniard arrival, "Zacapu" located about 40 km NW of Tzintzuntzan, "los Tamarindos" and "Cupandario" both sites belong to Huetamo region, see figure 4.3.

The group of studied objects is composed by a clothespin fragment, needles, rings, and sleigh-bells; their typical shape is depicted in figure 4.10, while the archaeological site where they were found is shown in table 4.6. These objects were found in a full Prehispanic associated contexts; except a sleigh-bell from Zacapu, a ring coming from Cupandario and another ring found in Tzintzuntzan. The first two were donated and probably belong to an early colonial period.

**Figure 4.10**

*Typical shape of the four kind of archaeological objects analyzed.*



**Table 4.6**

*Excavation site where each object was found.*

| Code | Site | Description | Depth |
|------|------|-------------|-------|
| tzn1 | Tzintzuntzan | Sleigh-bell [†] | 324 cm |
| tzn2 | Tzintzuntzan | Needle | 149 cm |
| tzn3 | Tzintzuntzan | Needle | 153 cm |
| tzn4 | Tzintzuntzan | Ring | Unknown |
| tzn5 | Tzintzuntzan | Clothespin [†] | 50-141 cm |
| hue1 | Los Tamarindos | Sleigh-bell | Pit 2 layer A level 1 |
| hue2 | Los Tamarindos | Sleigh-bell | Pit 2 layer AB level 2 |
| hue3 | Los Tamarindos | Sleigh-bell | Pit R layer AB level 2 |
| hue4 | Cupandario | Ring | Unknown |
| hue5 | Cupandario | Ring | Pit 4 level 1 |
| zac1 | El palacio | Sleigh-bell | Unknown |

[†] *Only fragments were found.*

### 4.3.1   Experimental Procedure

The experimental setup and spectral preprocessing steps discussed in section 4.2.1 were also used in this research. For each object, two spectral regions centered at 290 and 325 nm were measured, with a gate width of 1 $\mu$s and delay of 40 ns. The studied pieces are part of the Mexican historical heritage, so the damage must be kept at minimum; this is the main reason for measuring only two spectral regions.

As explained before, a little portion of the sample must be destroyed in order to obtain chemical information. Sample's damage is directly related with the laser energy used to irradiate. The teem-photonics laser provides a maximum energy of 30 $\mu$J. Although more than 30,000 shots were fired on the sample's surface, damage is negligible as shown in figure 4.11. Each crater is created by 100 laser shots, and has an average diameter of 32 $\mu$m.

The chemical composition obtained with LIBS spectral analysis is reported in table 4.7; this search only looked for neutral species emission. Due to the reduced delay time employed in the experiment, which produced a spectrum out of local thermal equilibrium, both plasma modeling at $10^4$ °K and NIST spectral reference computation methods had to be used. The measured lines below 4% were excluded for WSC computation and all the species lines above 1% were used for reference computation. The finding criterion was a WSC $\geq 0.5$ and $r \geq 0.01$.

**Figure 4.11**

*Superficial damage after 30,000 laser shots on* hue2 *sample.*
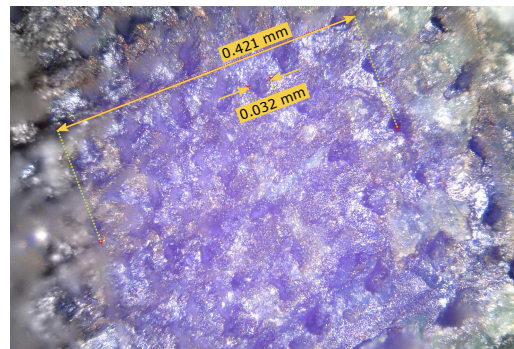


**Table 4.7**

*LIBS chemical analysis of archaeological objects.*

| Code | Mg | Al | Si | Mn | Fe | Cu | Ag | Sn | Pb |
|------|----|----|----|----|----|----|----|----|----|
| tzn1 | ★ | ★ | ★ |   | ★ | ★ |   |   |   |
| tzn2 |   |   | ★ |   |   | ★ | ★ |   |   |
| tzn3 | ★ | ★ | ★ |   |   | ★ |   |   |   |
| tzn4 | ★ | ★ | ★ |   | ★ | ★ |   |   |   |
| tzn5 |   |   |   |   |   | ★ | ★ | ★ |   |
| hue1 |   | ★ |   |   | ★ | ★ | ★ |   |   |
| hue2 | ★ | ★ | ★ |   | ★ | ★ |   |   |   |
| hue3 |   | ★ | ★ |   |   | ★ | ★ |   | ★ |
| hue4 | ★ | ★ | ★ | ★ | ★ | ★ |   |   | ★ |
| hue5 | ★ | ★ | ★ | ★ | ★ | ★ |   |   |   |
| zac1 | ★ | ★ | ★ | ★ | ★ | ★ |   |   | ★ |

For all objects, Cu line emission is the dominating spectrum's signal, while the weakest emissions are those pertaining to Mn, Fe, Ag and Pb. Lines of moderated intensity were measured for Mg, Si and Sn.

### 4.3.2 Discussion and Results

Except for Sn and Pb, all the emitting species found in the objects, were also found in slags. LIBS spectra were analyzed with PCA, their results are shown in figure 4.12. In this model, the variances explained by each of the first three PC are 57.5% 17.4% and 10.2% respectively, accounting for 85.1% of total variation. As a note, the fourth PC explains 6.9%.

The first fact to highlight is the clear separation of the clothespin identified as `tzn5` that has a high Sn content, as shown in figure 4.12. This confirms the findings made by Hosler who reported the use of As and Sn copper alloys in Prehispanic times at Balsas river basin region[107].

A pair of well defined object groups can be observed in the score plot of the first two principal components. The rightmost group in figure 4.12a contains a high concentration of Al. The samples in this group are `tzn1`, `tzn4`, `zac1` and `hue5`; this last sample is the only one coming from a foreign region. A high Al concentration could be related to pieces whose raw material comes from Inguarán mine, as seen on section 4.2.2.

**Figure 4.12** *LIBS spectra PCA analysis for archaeological objects.*
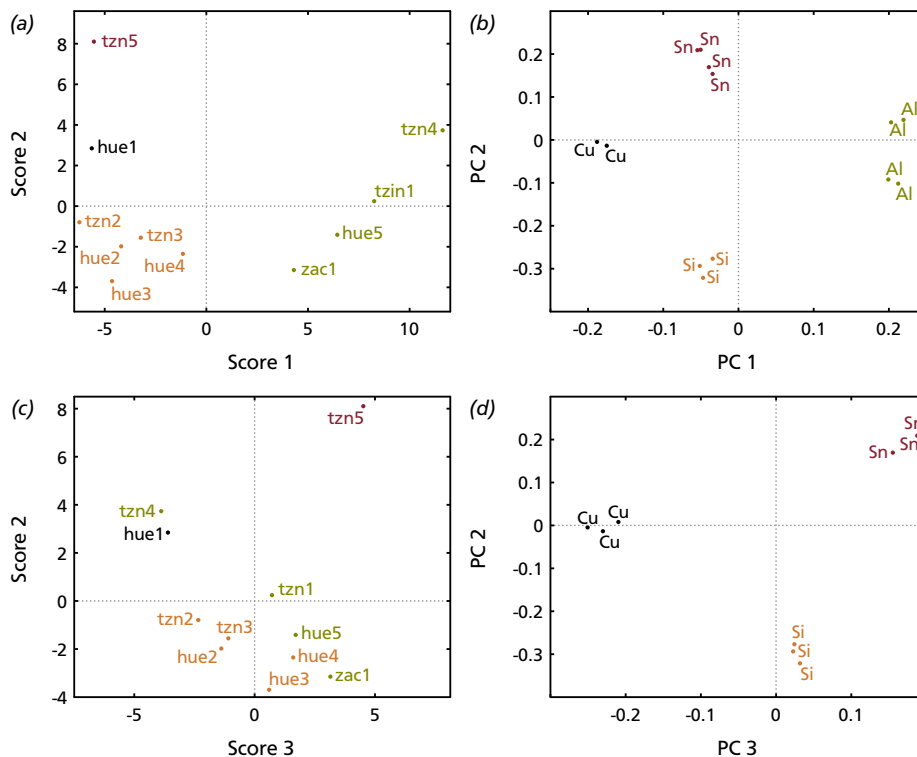
The other group has objects with high concentration of Cu and Si, located in the third quadrant of figure 4.12a. The samples identified as tzn2, tzn3, hue2 and hue3 were found at deeper layer levels in excavation sites; thus, they should be older. Only the depth where hue4 was located is unknown. The low Al content also suggests that these pieces were made with raw materials from Huetamo region.

The sample with the purest concentration of Cu is hue1 that is located in the negative portion of scores 1 and 3, which are directly related with a high copper concentration, as their corresponding PC quadrants reveals; see figure 4.12, parts *(b)* and *(d)*. As discussed on section 4.2.2, a Cu richer slag or equivalently a poorer Cu object may be related with an older Cu extraction procedure; this would indicate that tzin4 and hue1 would be related with a colonial period. This claim is also supported by the fact that hue1 was found at a surface level.

Finally, objects are not clearly differentiated by its excavation site. This could be expected because of the commercial relationships among distinct cities, and the tributes that should be paid by subordinates to the ruling regime.

### 4.3.3  Conclusions

It has been proven that LIBS measurements do not need any sample preparation and the infringed damage under adequate conditions is very low, as this study demonstrates. This contrast with X-ray mapping analysis which needs to cut a little portion of the sample, altering potentially the shape of the studied object. With LIBS this is no longer a concern because the infringed damage is scarcely seen at naked eye. Laser ablation can also be used for cleaning purposes, diminishing the chemical alterations measured due to aging and environment.

Although the fabrication at Prehispanic times of objects with copper alloys has been reported, the chemical evidence shows that not all objects were crafted with the same process. Only tzn5, the clothespin fragment, is made of an alloy, while rings, sleigh-bells and needles are almost pure Cu. The hypothesis of an slightly lower Cu content in Prehispanic objects was introduced in section 4.2.3. Assuming this claim as true, it is possible to relate the manufacturing of hue1 and tzin1 objects within the colonial period.

Moreover, the raw material source, as suggested by the presence of Al in the samples, is associated with Santa Clara del Cobre region, while the opposite would be true for pieces made with Huetamo's raw materials. The possibility for obtaining a more interesting classification scheme exists, by the inclusion of spectral regions for Ca and Na, which are the chemical footprints that identify the host rock at Huetamo and Inguarán mining sites.

## 4.4 Stainless Steel Classification

This section deals with a combined LIBS-PC analysis on a variety of stainless steels, whose corrosion resistance is being studied with the aim to find suitable candidates for constructing sea oil extraction facilities. This is an interdisciplinary effort, made in conjunction with Dr. Alejandro Zaragoza Ayala from Universidad Autónoma de Campeche in México. Following, a brief background of steels is provided.

The human kind has used steel since more than 4000 years ago; the earliest known pieces of ironware were excavated in Asia Minor. Steel has been employed in a wide variety of uses, ranging from simple nails to spacecrafts. Despite of the rapid grow of Al as a substitute of steel in some applications, more than 1,620 million of metric tons of steel were worldwide produced in 2015,[108] which is a little more than 20 times the accumulated total production of all non-ferrous metals.

Steel is a Fe alloy with another elements, being C, the most important among them. A ferrous alloy with less than 2% C by weight is classified as steel. Steels present a huge variety of properties that are due to the distinct forms, in which, Fe atoms can be bonded together for the same state of matter; a property called *allotropism*. There are three allotropes for Fe,[109] which are shown in figure 4.13.

The material made with Fe body-centered cubic crystals is known as *ferrite* or $\alpha$ Fe, while the name $\gamma$ Fe or *austenite* is employed when face-centered cubic crystals are formed. Hexagonal-close packed Fe of $\epsilon$ Fe is not studied in metallurgy because is stable only under high pressures; however, is widely studied in earth's core geologic research. In Fe-C alloys, austenite phase is obtained at temperatures above 800 °C, but it can be stabilized at room temperature using *austenite stabilizers* elements as Ni, N, Mn, and Cu. Other elements like Cr, Mo and Si promotes ferrite formation; therefore, they are called ferrite stabilizers.[110]

Stabilizers are also used with the aim to improve alloy's physical properties; for example, stainless steel contains at least 10.5 % wt of Cr. In this proportion, Cr forms an insoluble film that prevents oxygen diffusion protecting the iron matrix from oxidation; higher protection can be achieved increasing Cr content.

Stainless steels can be divided in four groups: martensitic and precipitation hardening, ferritic, austenitic and duplex stainless steels. Duplex steels contains both $\alpha$ and $\gamma$ phases.[111] Some times the term superaustenitic could be found in the literature, which refers to alloys with a high Cr, Mo, Ni and N content; the total proportion of this elements can be as high as 50% by weight.

**Figure 4.13**

*Fe allotropes.
All spheres represents Fe atoms*



Body Centerd Cubic
$\alpha$

Face Center Cubic
$\gamma$

Hexagonal Closest Packed
$\epsilon$
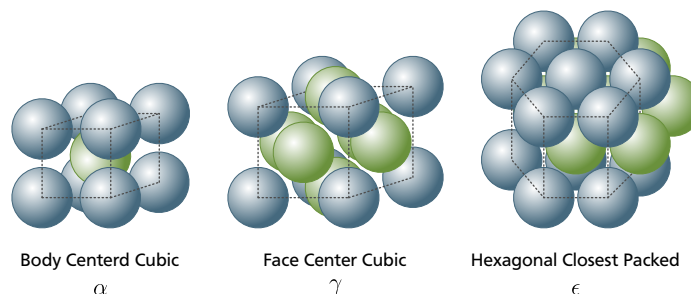
*Table 4.8*

*Stainless steels studied*

| Alloy | Grade | Samples |
|-------|-------|---------|
| 430 | Ferritic | 1 |
| UR 35N | Duplex | 10 |
| UR 45N | Duplex | 12 |
| UR 52N+ | Duplex | 6 |
| 316L | Austenitic | 11 |
| UR B66 | Super Austenitic | 4 |
| UR B26 | Super Austenitic | 11 |
| 904L | Super Austenitic | 8 |

*Figure 4.14*

*Typical shape of the stainless steel specimens studied.*



Eight distinct stainless steel grades were analyzed in this research, whose crevice corrosion properties were studied at Universidad Autónoma de Campeche. Alloys name and its corresponding grade is provided in table 4.8, along the number of available samples per grade. Specimens' typical shape is depicted in figure 4.14.

### 4.4.1  Experimental Procedure

The LIBS experimental arrangement is the same discussed in section 3.5. The delay relative to laser pulse firing signal was set to 1 $\mu$s, as well as the exposition time. In this case, 500 laser shots are integrated on detector's CCD chip, and each of this measurements is accumulated 30 times to obtain a single spectrum. The caused sample damage after the whole procedure can be seen at the upper left corner on figure 4.14 as a dark narrow stripe of few millimeters long.

In this work, the approximated chemical composition of each steel grade is known before hand; this allows to select a set of spectral regions that includes the highest emission lines for the required species with a minimum of windows; which reduces sample's damage. Thus, each specimen was sampled five times in distinct areas, on windows centered at 340 and 388 nm.

Spectral data were wavelength calibrated, employing both shift and polynomial methods. As reference, the synthetic spectrum of Fe and Cr were used. Background signal removal was done by means of SNIP algorithm; each signal was scaled with SNV prior to PC computation. A great deal of Fe emission lines were excluded from PCA analysis, using spectral regions between 335–360 nm and 378–403.7 nm, with the aim to reinforce the relevance of minor and medium species in the model.

The species acceptance criterion for the elemental searching routine was set to WSC $\geq$ 0.5 and $r \geq$ 0.01; just neutral species were looked for. The used wavelength tolerance was 0.05 nm; lines above 1% were employed for reference computation while measured emission lines below 6% were excluded from WSC calculation. The resulting elemental analysis is presented in table 4.9.

**Table 4.9**

*Stainless steel elemental composition determined with LIBS analysis.*

| Alloy | Fe | Cd | Cr | Mn | Ni | Cu | Mo |
|---|---|---|---|---|---|---|---|
| 430 | ★ | ★ | ★ | ★ | | | |
| UR 35N | ★ | ★ | ★ | ★ | ★ | | |
| UR 45N | ★ | ★ | ★ | ★ | ★ | | ★ |
| UR 52N+ | ★ | ★ | ★ | ★ | ★ | ★ | ★ |
| 316L | ★ | ★ | ★ | ★ | ★ | ★ | ★ |
| UR B66 | ★ | | ★ | ★ | ★ | ★ | ★ |
| UR B26 | ★ | ★ | ★ | ★ | ★ | ★ | ★ |
| 904L | ★ | ★ | ★ | ★ | ★ | ★ | ★ |

## 4.4.2   Results and Discussion

Sample's elemental composition obtained with the spectral analysis of LIBS signals has a good qualitative agreement with the composition reported in the literature, [112] shown in table 4.10 The presence of C and N cannot be detected because their strongest emission occurs outside the measured spectral region. All the remaining elements are detected, along the matrix element and Cd, which should be a trace element due to its very low relative amplitude emission.

Despite Mn should be only found in the UR B66 grade, Mn spectral fingerprint was detected in every sample. For comparison purposes, the measured spectrum for 316L, 430, UR B66 and UR 45N is shown in figure 4.15, where it is clear the Mn presence in every signal; although the highest emission corresponds to UR B66. The same can be said about Cu presence in 316L grade.

Considering that Mn content in UR B66 approximately corresponds to 3% by weight and comparing the highest Mn emission line against the same line in other grades, it is possible to conclude that Mn is a trace element; possibly a remaining impurity of the foundry process. By the same reasoning, a similar conclusion can be drawn for Cu content in 316L grade, as well as Cd presence in other grades. This proves that LIBS is able to detect Mn, Mo, Cd and Cu in proportions below 1% wt on steels.

A explained variance about 93% is given by the PCA of the spectral region centered in 340 nm; however, the resulting separation of super asutenitic grades is poor, mainly due to the lack of Mo emission lines in this spectral region.

**Table 4.10**

*Studied stainless steel chemical composition reported in literature.*

| Alloy | C | N | Cr | Mn | Ni | Cu | Mo |
|---|---|---|---|---|---|---|---|
| 430 | 0.03 | | 16.5 | | | | |
| UR35N | 0.03 | 0.14 | 23.0 | | 4.0 | | |
| UR45N | 0.03 | 0.18 | 22.5 | | 5.3 | | 2.7 |
| UR52N+ | 0.03 | 0.23 | 25.0 | | 7.0 | 1.5 | 3.5 |
| 316L | 0.03 | | 17.0 | | 11.5 | | 2.6 |
| URB66 | | 0.45 | 24.0 | 3.0 | 22.0 | 1.5 | 6.0 |
| 904L | 0.03 | 0.10 | 20.0 | | 24.0 | 1.5 | 4.0 |
| URB26 | 0.03 | 0.23 | 20.0 | | 25.0 | 1.0 | 6.0 |

**Figure 4.15** *Measured LIBS spectra for grades 316L, 430, UR B66 and UR 45N; only neutral species are marked. Fe emission are not shown for clarity.*
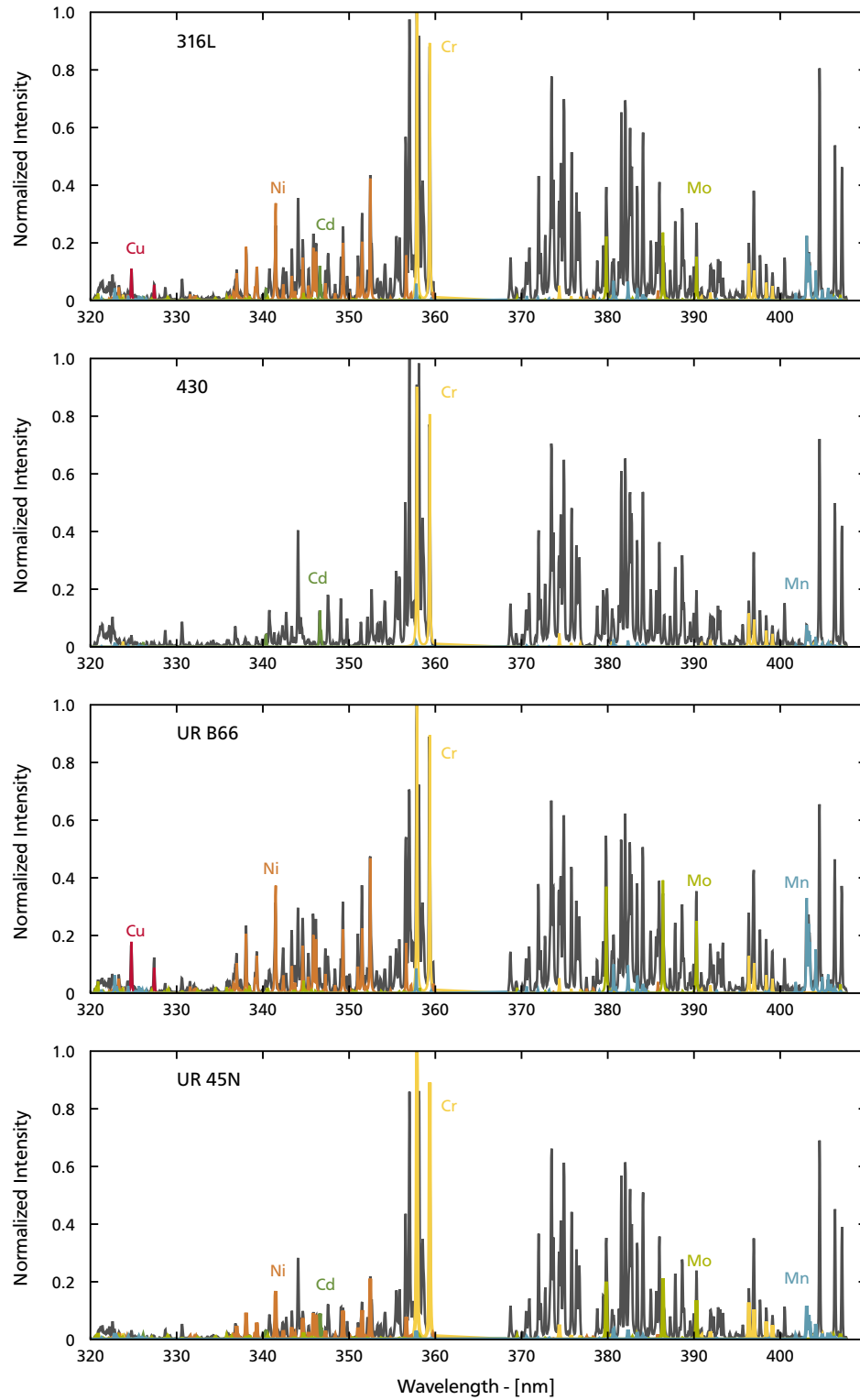
**Figure 4.16**   *Score plot and PC plot of PCA on stainless steel LIBS spectra. Quadrants in part (a) separates Ferritic (II), austenitic(III), and super austenitic (I, IV) steels.*
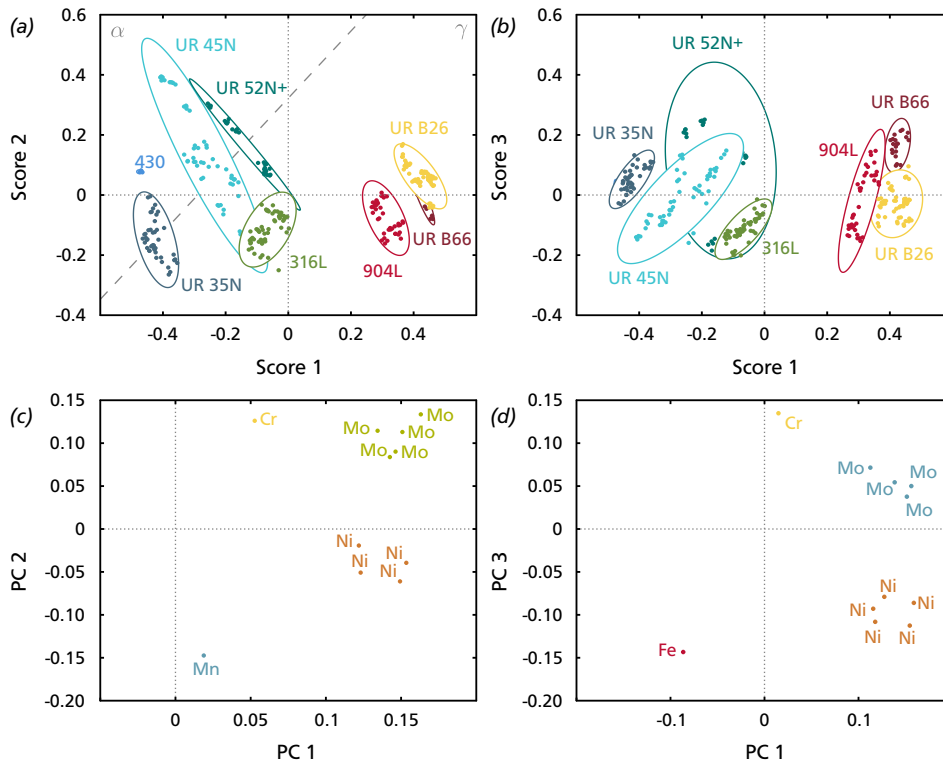


When Mo emission lines are included, the whole model's variance explained by the first three principal component decreases to 85.4% but a better grade separation is accomplished. The first PC accounts for 65.6% of total variance, while the second and third PC explains 12.7% and 7.1% respectively. The corresponding score plot and PC plot is depicted in figure 4.16.

The score plot for the first two PCs shown in figure 4.16a, demonstrates that PCA was able to group correctly the distinct stainless steel grades accordingly with its crystalline structure. The first and fourth quadrant contains super asutenitic grades. Super austenitic grades are separated due to their high Ni and Mo content, as their distance from origin, in the corresponding PC plot suggest; see figure 4.16c.

Data sheets for UR 35N, UR 45N and UN 52N+ duplex stainless specifies that both $\alpha$ and $\gamma$ phases are near to equilibrium in these alloys.[113,114,115] Therefore, an imaginary line that cross ellipses mid points can be traced for duplex stainless steels, in order to establish a frontier among austenitic and ferritic grades, where the first will be placed at right hand of such line, and the last at left hand. Phase balance in duplex stainless steels could also explain why UR 45N and UN 52N+ alloys' points have a bigger dispersion.

In literature, UR 35N is described as a *lean duplex* stainless steel; according to Bergstrom,[116] this means that it contains a low proportion of Ni and Mo when compared against other duplex grades. As UR 35N has no Mo, then is placed at the left end of the graph in figure 4.16a.

***Figure 4.17***  *Theoretical chemical composition PC analysis of stainless steel grades.*
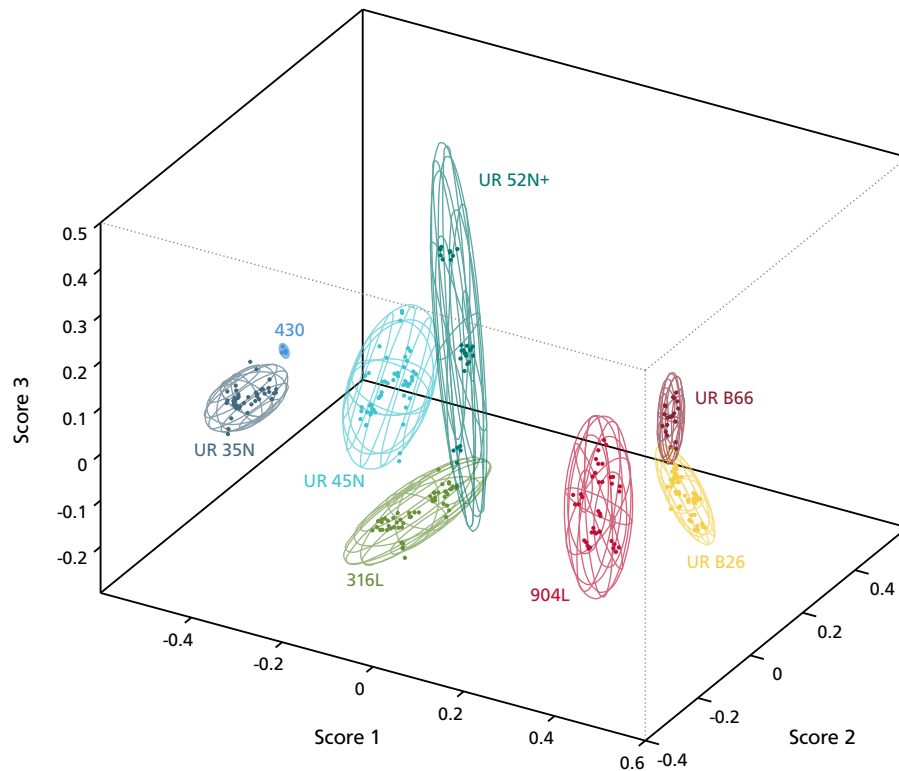


The score plot shown in figure 4.16b, for the first and third PCs reveals that UR B26 and UR B66 are not overlapped; this separation was promoted by their different Fe content, see figure 4.16d. Finally, figure 4.17 is a 3D version of the score plot, where is easier to see the separation among the distinct grades.

### 4.4.3   Conclusions

As this work demonstrates, if a related group of samples is available, the combined LIBS-PC analysis can reveal properties that are not straightforward to deduce from chemical analysis alone; in this case, the crystalline structure of stainless steels. Contrasting with microprobe technique, the LIBS analysis for the complete set of 63 samples can be acquired in a few hours instead of days; moreover, LIBS could be implemented for in-line quality control purposes.

Elements that are part of the chemical composition of some samples, as C an N were not measured. Its absence from PCA model does not possess a barrier, because PCA highlights variables with the highest variance, and the excluded elements are expected to have minimum variations among the distinct stainless steel grades. It has also become evident that LIBS is capable to detect traces of some alloy metals, specifically Cu, Cd an Mn, whose spectral footprint was found as impurities in many of the stainless steel samples studied.

## 4.5   Concluding Remarks

LIBS is a very sensitive tool for chemical analysis that can detect elemental species ranging from high content to trace quantities. This fact is confirmed by comparing the presented results against those published in literature; which were measured with well established techniques as X-ray mapping analysis and microprobe.

Elemental identification is not longer a cumbersome and tedious work, instead is performed easily using the WSC metric presented in chapter 3. It can correctly identify atomic emission species, from high content to traces, in complex signals as stainless steel spectra; its results have been corroborated with other analytical techniques.

On the other hand, LIBS-PC analysis is a valuable tool to perform qualitative analysis for a group of related samples. Samples relationships can provide valuable clues to archaeologist to make a connection between an archaeological object and its raw material source, as well as the level of technology involved in the copper extraction process which could be related to a historical period. For stainless steels, LIBS-PC analysis was able to group distinct grades according with its crystalline structure.

In many applications, having a group of pieces is not possible and the only way to extract information is to know its chemical composition from a quantitative point of view. Although a direct relationship between the amplitude of an emission line and its relative content in the sample exists, obtaining a numerical figure from it depends on many factors. This complicated task will be presented in the next part of this thesis.

# Chapter 5

# Quantitative LIBS Measurements

Several methods exists for determining the concentrations of chemical elements in a given sample that are based on LIBS spectrum analysis. These methods can be divided in those that make use of a set of calibration standards and those that do not make use of them. According to the literature, the best analytical results have been obtained with calibration curve methods;[50] however, there are many practical situations where a set of calibrated standards are difficult to prepare, too expensive, or nearly impossible to get. In such cases a standardless procedure as Calibration-Free method is preferable.[82] This chapter, discuss the implementation details of the Calibration-Free method.

The quantification in LIBS analysis is viewed as one of its weakest points,[117] but also as a fertile field for research. Quantitatively speaking, some advantages of the LIBS experiments are also the source of its limitations; for example, the small amount of ablated mass compared with ICP-AES, make LIBS results very dependent on sample's homogeneity. On the other hand, samples measured without previous preparation may be contaminated on its surface, leading to wrong results. Finally the pulsed nature of the laser yields lower signals when compared with ICP-AES reducing the plasma reproducibility.

Despite of these disadvantages, a LIBS experiment carefully done can lead to useful results for many applications,[118] because some of the mentioned drawbacks can be easily solved for many samples. Specimen's inhomogeneity can be reduced by sampling larger areas, while surface contaminants can be eliminated with a faint ablation on the surface prior to measurements.

Every chemical analytical method is based on a functional dependence between the concentration of an atomic element and a physical parameter. Regarding LIBS experiments, the intensity of an emission line is related with the number of atoms contained into the plasma for the corresponding specie. The word *plasma* was proposed by Irving Langmuir in 1922, naming the so called fourth state of the matter. When a material is increasingly heated, its molecules are transformed in the sequence: solid, liquid, gas and plasma. This latter state contains a mixture of molecules, atom, ions, and electrons, which moves in random directions. Although plasmas contains charged particles, the average charge is nearly zero.[119]

The ability to provide a numerical description of the plasma process involved in LIBS experiments is very important for the quantitative analysis. Due to the complex nature of the variables that affects the plasma behavior as pressure, state of aggregation, magnetic and optical properties of the material, etc.; plasma modeling is a whole active research field, which could lead to the development of nuclear fission in the near future.

From a classical point of view, plasmas are described by using the statistical mechanics laws. Plasma models used for quantification purposes assume that a Local Thermodynamic Equilibrium or LTE state exists. Under this condition, all the particles within a plasma volume are considered to have the same temperature, making valid the descriptions given by Maxwell's distribution for particle velocity, the Boltzmann's energy particle distribution and Saha's ionization equation. LTE is reached after a sufficient number of collisions within plasma have occurred, which distributes the energy across species and plasma volume. LTE is present in small regions of space, and may be somewhat different from region to region.[120]

One of the most widespread methods to test the LTE existence is the McWhirter criterion, which possess a lower electronic density figure to ensure that the radiative process is dominated by electron collisions. It is expressed as:

$$n_e \geq 1.6 \times 10^{12} \; T^{1/2} \, (\Delta E)^3 \tag{5.1}$$

where $n_e$ [cm$^{-1}$] represents the electron density, $T$ [°K] is the plasma temperature and $\Delta E$ [eV] is the highest energy transition for which condition holds. McWhirter criterion is a necessary but not sufficient condition for LTE existence, which is usually fulfilled in the first stages of the plasma lifetime.

Saha found that the ratio between the density of two successive ionization stages of the same element depends on electron density and electron temperature

$$n_e \frac{n_s{}^1}{n_s{}^0} = \frac{2(2\pi m_e kT)^{3/2}}{h^3} \; e^{-\chi/kT} \tag{5.2}$$

$\chi$ stands for the ionization potential of the neutral species in its ground state, $m_e$ is the electron mass, $h$ is the Planck constant, and $k$ is the Boltzmann constant, $n_s{}^0$ and $n_s{}^1$ represent the density of species $s$ in its ground stated and its first ionization stage respectively.

A similar expression called the coupled Saha-Boltzmann equation provides a mean to determine electron density, provided that two lines of the same element but with different ionization stages are known. This equation has the following form

$$n_e = \frac{2(2\pi m_e kT)^{3/2}}{h^3} \frac{I_{mn}^{\mathrm{I}} A_{ij}^{\mathrm{II}} g_i^{\mathrm{II}}}{I_{ij}^{\mathrm{II}} A_{mn}^{\mathrm{I}} g_m^{\mathrm{I}}} \; e^{-(\chi + E_i^{\mathrm{II}} - E_m^{\mathrm{I}})\,/\,kT} \tag{5.3}$$

$\chi$ stands for the ionization potential of the neutral species in its ground state, $E$ are the upper excitation energy of the level, $I$ is the integrated line intensity, $A$ is the line's transition probability, and $g_i$ is the statistical weight for the upper level. The superscripts I and II denote atomic and ionic parameters respectively|.

Equation 5.3 can be linearized; applying the logarithm and rearranging some terms leads to

$$\ln\left(\frac{I_{ij}^{II}\,A_{mn}^{I}\,g_m^{I}}{I_{mn}^{I}\,A_{ij}^{II}\,g_i^{II}}\right) = \ln\left(\frac{2(2\pi m_e kT)^{3/2}}{n_e\,h^3}\right) - \frac{\chi + E_i^{II} - E_m^{I}}{kT} \qquad (5.4)$$

The linear regression of this transformation for several ionic and atomic emission line combinations, results in a line whose slope is inversely proportional to the electron temperature, while its intercept is related to electron density.[121] As an added benefit, a linear regression is less sensitive to measurement noise.

In addition to LTE, another two conditions must be satisfied in order to obtain qualitative results from the plasma spectrum analysis. The first is that the plasma volume under observation must represent the sample composition; in other words, it must be *stoichiometric*. Chan and Russo[122] showed that laser power densities above 1 GW/cm$^2$, produces stoichiometric ablation on solid samples.

Crater depth also affects the plasma composition; Borizov et al.[123] found that when the depth/diameter ratio surpasses a value of 6, the composition of an ablated Pb/U sample was not stoichiometric.

The second condition is that plasma must be optically thin, meaning that light emitted must traverse and escape from the plasma without significant absorption or scattering. Cremers and Radziemski stated an easy way to check for the optical thickness of the plasma;[124] a flat-topped profile in the most intense lines is evidence of self-absorption. In severe cases, a dip appears at the line's central frequency; those lines are said to be self-reversed.

An experimental setup that accomplishes all the aforementioned conditions will produce a spectrum whose integrated line intensity for to the transition between the upper level $i$ and the lower level $j$ obeys:

$$I_{ij} = \frac{A_{ij}\,g_i}{U^s(T)}\,n^s\,e^{-E_i/kT} \qquad (5.5)$$

$U^s(T)$ represents the internal partition function of the species $s$ at temperature $T$, and $n^s$ is the total number density of species $s$ in the plasma; all other symbols have already been defined. Considering that the efficiency of the collection system acts as a scaling factor for the measured line intensity in equation 5.5, the expression shown in section 3.2, is obtained, which is reproduced here for the sake of clarity.

$$I_{ij} = FC_s\,\frac{A_{ij}\,g_i\,e^{-E_i/Tk_B}}{U^s(T)} \qquad (3.2)$$

where $F$ is a parameter that takes into account the optical efficiency of the collection system, the plasma volume and its density. $C_s$ is the concentration of the emitting species in the sample. Excepting $F$, $C_s$ and $T$, all the parameter in equation 3.2 can be consulted on literature or determined from experimental data; in particular, the value for $U^s(T)$ can be determined from known spectroscopic data once the plasma temperature has been calculated.

Taking the logarithm on both sides of equation 3.2 leads to

$$\ln \frac{I_{ij}}{A_{ij}\, g_i} = \frac{1}{k_B T}\, E_i + \ln \frac{C_s F}{U^s(T)}$$

$$y = \ln \frac{I_{ij}}{A_{ij}\, g_i}; \quad m = \frac{1}{kT}; \quad x = E_i; \quad q_s = \ln \frac{C_s F}{U^s(T)}$$

$$y = mx + q_s \tag{5.6}$$

Under this transformation it is possible to represent the measured intensity value for each line on a Boltzmann plane. Equation 5.6 suggest that both $m$ and $q_s$ can be determined applying linear regression to the set of $(x, y)$ points for the same emission species. The slope is inversely proportional to the plasma temperature while $q_s$ is related with the experimental factor $F$ and the species concentration.

With the aim for solving $F$, Ciucci et al. propose using an internal reference standard of known concentration when available, or assuming that the sum of the relative concentration for all the emission species in the plasma must equal unity. This leads to the so called *closure equation*

$$\sum_s C_s = \frac{1}{F} \sum_s U^s(T)\, e^{q_s} = 1 \tag{5.7}$$

Hahn and Omenetto[117] point to the closure equation as the main source of poor performance for determining trace elements' concentration. They argue that small uncertainties, which have a marginal effect in mayor components, will be translated in large uncertainties for minor and trace components. Reproducibility also affects the quantitative analysis precision; it is considered that reproducibility of LIBS measurements is not better than 3–5%, while $A_{ij}$ often have uncertainties between 10–20%.

In principle species' concentration could be measured with only just one line of known spectroscopic parameters, but better results will be obtained considering as many spectral lines as possible; a great number of lines average the uncertainties of $A_{ij}$ values. It is necessary to remark that all the plasma emission species must be considered, including ionic species. Ciucci's original paper suggests estimating ionic species by means of Saha's equation when ionic lines could not be measured.

Gomba et. al.[19] applied a distinct method for elemental quantification. They found species density ratios with equation 5.8, instead of computing $FC_s$ in a Boltzmann plane.

$$\frac{n_a^i}{n_b^j} = \frac{I_a^i\, U_a^i(T)}{I_b^j\, U_b^j(T)}\, e^{-(E_j - E_i)/kT} \quad \text{with} \quad I_s^k = \frac{I_{ij}}{A_{ij} g_i} \tag{5.8}$$

Exponents $i$ and $j$ represents the neutral or ionization stages, for emission species $a$ and $b$. By means of an iterative procedure, they compute values for $n_e$ and $n_s$ until their ratios match with experimental results.

## 5.1 Closure Equation Dismiss

Gomba et. al., propose an alternative closure equation based on electron density. An equivalent approach is used by Ciucci but only when ions cannot be detected. This closure is based on plasma neutrality condition, which states that plasma free electrons are due to ions, therefore

$$n_e = \sum_{i=1}^{k} n_i \tag{5.9}$$

where $n_i$ are the total density of ionic species for the element $i$. In turn, each $n_i$ is calculated with

$$
\begin{aligned}
n_i &= n_i^1 + 2n_i^2 + 3n_i^3 + \cdots \\
&= n_i^0 \left( \frac{n_i^1}{n_i^0} + 2\frac{n_i^2}{n_i^0} + 3\frac{n_i^3}{n_i^0} + \cdots \right) \\
n_i &= n_i^0 \sum_{z=1}^{M} z \frac{n_i^z}{n_i^0} = n_i{}^0 N_i
\end{aligned}
\tag{5.10}
$$

$n_i^z$ represents the density of the element $i$ in the ionization state $z$; $z = 0$ stands for neutral species, etc. $N_i$ is simply a shortcut for the sum. The first ion contributes with a single electron, while the second gives two electrons, and so forth.

The density ion ratio on equation 5.10 can be calculated by means of Saha's equation. Although equation 5.2 relates only successive ionization stages, the ratio between a high ionization stage and the neutral species is

$$\frac{n_i^z}{n_i^0} = \frac{n_i^1}{n_i^0} \times \frac{n_i^2}{n_i^1} \times \cdots \times \frac{n_i^z}{n_i^{z-1}} = \prod_{z=1}^{M} \frac{S_i^z}{n_e^z} \tag{5.11}$$

where $S_i^n$ represents the right side of equation 5.2.

At first sight, factor out $n_i^0$ in equation 5.10 seems unnecessary, but this form reveals that just $n_e$, $T$ and neutral species density $n_i^0$ are needed to compute the contribution of element $i$ to the total electron density. With similar arguments, it can be shown that when considering neutral and $M$ ionic species the plasma's total density of element $i$ is

$$m_i = n_i^0 \left( 1 + \sum_{z=1}^{M} \frac{n_i^z}{n_i^0} \right) = n_i{}^0 M_i \tag{5.12}$$

where $M_i$ represents the terms between parenthesis.

Now consider a simple example which only takes into account the first ionic state for a sample composed of $k$ elements. Using equation 5.10, plasma electron density can be estimated as

$$n_e = n_1 + n_2 + n_3 + \cdots + n_k$$

$$= n_1{}^0 \frac{n_1{}^1}{n_1{}^0} + n_2{}^0 \frac{n_2{}^1}{n_2{}^0} + n_3{}^0 \frac{n_3{}^1}{n_3{}^0} + \cdots + n_k{}^0 \frac{n_k{}^1}{n_k{}^0}$$

$$= n_1{}^0 \frac{S_1{}^1}{n_e} + n_2{}^0 \frac{S_2{}^1}{n_e} + n_3{}^0 \frac{S_3{}^1}{n_e} + \cdots + n_k{}^0 \frac{S_k{}^1}{n_e}$$

$$= \frac{n_1{}^0}{n_e} \left( S_1{}^1 + \frac{n_2{}^0}{n_1{}^0} S_2{}^1 + \frac{n_3{}^0}{n_1{}^0} S_3{}^1 + \cdots + \frac{n_k{}^0}{n_1{}^0} S_k{}^1 \right)$$

therefore

$$n_e^2 = n_1{}^0 \sum_{i=1}^{k} \frac{n_i{}^0}{n_1{}^0} S_i{}^1 \tag{5.13}$$

This expression can be generalized to cover ions higher than 1; in such case, a high order polynomial will result

$$\sum_{j=1}^{M} \left[ \frac{n_1{}^0}{(n_e)^{j+1}} \sum_{i=1}^{k} \left( \frac{n_i{}^0}{n_1{}^0} \prod_{z=1}^{j} z\, S_i^z \right) \right] = 1 \tag{5.14}$$

Ratios between neutral species can be obtained with CF procedure

$$\frac{n_2{}^0}{n_1{}^0} = \frac{FC_2{}^0}{FC_1{}^0} = \frac{U_2^0(T) e^{q_2{}^0}}{U_1^0(T) e^{q_1{}^0}} \tag{5.15}$$

Reader must note that the neutral species ratio can be computed from spectral data without knowing the $F$ factor. As temperature can be computed with the aid of both Boltzmann or Saha-Boltzmann plot, only $n_e$ and $n_1{}^0$ are unknown.

Ciucci proposes an iterative procedure for estimating $n_e$ when ion lines could not be measured, while Gomba iterates with approximate values of $n_i{}^0$. However, considering that only ratios among species can be derived from CF procedure, and as equation 5.13 implies, another equation is needed to solve the system, thus the electron density given by iterative procedures is not unique, instead heavily depends on initial estimation parameters.

The relevance of electron density in the CF method is best explained with an example. Figure 5.1, shows how $n_e$ affects the number of ions for each neutral atom measured, at a fixed temperature of 14000 °K. Clearly, a small variation of electron density below $10^{17}$ cm$^{-3}$ leads to very distinct total proportions, while a negligible variations are obtained near to $10^{18}$ cm$^{-3}$.

Fortunately, electron density can be determined within Saha-Boltzmann plane when ion lines are available, or by Stark broadening when not. Some cases may arise when $n_e$ cannot be reliably determined from Stark broadening due to the lack of spectral widths for a given emission line within a particular wavelength range or because the instrumental broadening is unknown. For those situations, only a confidence range can be given, based on an upper limit set by the McWhirter criterion and a lower limit established by the ratios of neutral species.

*Electron density effect on total density for a fixed temperature of 14000°K.*
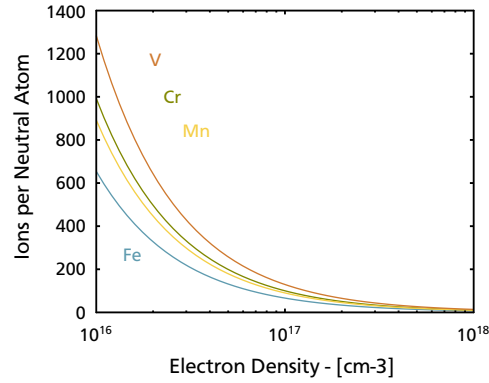


In order to be full closure equation independent, elemental ratios should be the preferred approach to report chemical content. When dealing with real samples, partial element detection is not uncommon; with the closure equation, the final proportion depends on the number of elements considered and errors will emerge if a species is left out. On the contrary, the ratio among elements will be constant. With the aim of demonstrating this claim, the general matrix solution of a CF system using the closure equation and equation 5.12 is shown

$$Am = a \qquad (5.16)$$

$$
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
M_2{}^0 C_2{}^0 & -M_1{}^0 C_1{}^0 & 0 & \cdots & 0 \\
M_3{}^0 C_3{}^0 & 0 & -M_1{}^0 C_1{}^0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
M_k{}^0 C_k{}^0 & 0 & 0 & \cdots & -M_1{}^0 C_1{}^0
\end{pmatrix}
\begin{pmatrix}
m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_k
\end{pmatrix}
=
\begin{pmatrix}
1 \\ 0 \\ 0 \\ \vdots \\ 0
\end{pmatrix}
$$

Solving by Cramer's rule leads to

$$m_i = M_i\, C_i{}^0\, (M_1\, C_1{}^0)^{k-2} / \det A \qquad (5.17)$$

Then, the total density ratio between two elements will be

$$\frac{m_i}{m_j} = \frac{M_i\, C_i{}^0}{M_j\, C_j{}^0} \qquad (5.18)$$

A similar expression holds for plasma neutrality condition and equation 5.10

$$Bn = b \qquad (5.19)$$

$$
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
N_2{}^0 C_2{}^0 & -N_1{}^0 C_1{}^0 & 0 & \cdots & 0 \\
N_3{}^0 C_3{}^0 & 0 & -N_1{}^0 C_1{}^0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
N_k{}^0 C_k{}^0 & 0 & 0 & \cdots & -N_1{}^0 C_1{}^0
\end{pmatrix}
\begin{pmatrix}
n_1 \\ n_2 \\ n_3 \\ \vdots \\ n_k
\end{pmatrix}
=
\begin{pmatrix}
n_e \\ 0 \\ 0 \\ \vdots \\ 0
\end{pmatrix}
$$

Solving by Cramer's rule leads to

$$n_i = n_e \ N_i \ C_i{}^0 \ (N_1 \ C_1{}^0)^{k-2} / \det B \qquad (5.20)$$

The electron density ratio between two elements will be

$$\frac{n_i}{n_j} = \frac{N_i \ C_i{}^0}{N_j \ C_j{}^0} \qquad (5.21)$$

Equations 5.18 and 5.21 could also be obtained by dividing equations 5.12 and 5.10 for two distinct elements; however the point here is to show that the ratio $n_i/n_j$ between distinct elements is independent of $\sum n_s$, but fully depends on $n_e$. Moreover, closure equation and plasma neutrality condition are fully equivalent. The first provides total density values as percentages, while the second provides the contribution of a given element to $n_e$.

The relevance of equation 5.15 will better understood with a simple example of a hypothetical sample which contains four elements whose proportion is $A$, $B$, $C$ and $D$ respectively. Under ideal conditions and providing that $n_e$ is determined, each $C_s{}^0$ will be accurately found and the CF procedure will provide accurate results.

Now suppose that due to instrumental noise, continuum signal or any other reason the element $C$ cannot be measured despite of having good estimates of the other elements. In this scenario, the closure equation leads to higher concentrations than those measured because proportions are compared against $A + B + D$ instead of $A + B + C + D$; in contrast, the ratio between elements would provide better results if it is compared against theoretical ratios $B/A$ and $D/A$.

The proposed method also saves time when only a few components of a given sample are of interest, because the analyst does not have to identify every element present in the sample, which is a time consuming task. Moreover, a sample with many trace elements may provide signals that cannot be properly identified and measured because they are confused with experimental noise, but as a whole they have a significant impact on closure equation; for example, steels usually have many trace elements that are difficult to measure as C, P, S and N.

Partial element measuring opens new opportunities. Now it is possible to make qualitative measurements with spectrometers of narrow spectral windows and high resolution, such as those based on a Czerny-Turner monochromator, without to worry about having the same laser energy in every spectral window.

The next sections will discuss how to extract single Lorentzian parameters from spectrum and data sources for partition function computing. Once all the required tools are gathered, the last section will show a full CF example.

## 5.2    LIBS Spectrum Fitting Procedure

Signals composed by the superposition of multiple Lorentzian profiles are acquired in a wide variety of experimental setups, pertaining to multiple science fields that includes atmospheric studies, chemistry,[125] electronics,[126] physics,[127] biology,[128] among others. In such experiments, the covered area of every single Lorentzian curve can be related to a specific physical property.

In particular, Calibration–Free (CF) method used in LIBS is a standardless quantitative elemental analysis, which uses the area of an emission band to measure the concentration of an atomic species.[71] When each Lorentzian curve present in a spectra—also named peak—is sufficiently isolated from others, the area can be estimated integrating the peak signal as was done by Takahashi et al.[129] However, when the signal presents overlapping with another peak this approach cannot be applied anymore. An analytical alternative is provided by a fitting process, which consists in giving an estimate of the individual parameters that defines every band present in the signal, from which, the individual peak's area is easily evaluated.

The non-linear nature of the Lorentzian profile, dictates the use of non-linear fitting methods, such as genetic algorithms[130] or the Levenberg-Marquardt (LM) algorithm, which certainly is the dominating approach to fit a signal composed of many non–linear profiles.[131] LM tends to be slow when a large number of peaks are analyzed, and its success is heavily dependent on an adequate estimation of the initial parameter conditions.

This section presents an automated alternative to deconvolve a signal composed of many Lorentzian curves, which can be easily adapted for Gaussian profiles. The proposed procedure estimates the peak's Lorentzian parameters from stationary points, then the peak's height is re-adjusted with a routine named Simultaneous Height Adjust (SHA), that takes into account peak's height distortion produced by neighboring signals. An initial estimation is used to calculate a single peak sample region free of the deviations introduced by its neighbors, which in turn, is used by a linear Least Square Fitting (LSF) algorithm in order to get the final peak parameters. The final result is a fast fitting algorithm, whose execution time is proportional to the number of fitted peaks that can be adapted to low computing power systems, such as portable devices.

### 5.2.1    Lorentzian Curve

The expression that represents a Lorentzian shape was covered before in section 3.2 equation 3.1; its area can be easily calculated applying the integral

$$S(\lambda) = \int_{-\infty}^{\infty} L(\lambda) \, d\lambda = I\pi\gamma \qquad (5.22)$$

Stationary points of a Lorentzian shape can be calculated by equaling to zero the corresponding high order derivatives of equation 3.1, their coordinates are shown on table 5.1. It must be noted that two roots of the Lorentzian curve aberrancy,[132] equals the left or right $\gamma$ coordinates, represented by $\gamma_-$ and $\gamma_+$ in this work; while

**Table 5.1**

*Lorentzian curve coordinates of some stationary points.*

| Derivative | Point | $x$ | $y$ |
|---|---|---|---|
| First | Maximum | $\lambda_0$ | $I$ |
| Second | Inflection | $\lambda_0 \pm \frac{\gamma}{2\sqrt{3}}$ | $\frac{3}{4}I$ |
| Third | Left ($\gamma_-$) or right ($\gamma_+$) $\gamma$ coord. | $\lambda_0 \pm \frac{\gamma}{2}$ | $\frac{1}{2}I$ |

the remaining root also leads to the central peak coordinates. The relevance of stationary points is that they provide the means to make a computational estimate of the non–overlapping Lorentzian peaks parameters, which will be used as the initial guess for the fitting procedure.

### 5.2.2   Simultaneous Height Adjust Fitting Procedure

The shape of an overlapped peak is distorted by its neighbor signals, in such a way that, the approach presented previously would produce $\gamma$ and $h$ parameters too different from their true values to be useful. The goal is to estimate the points that define a Lorentzian curve free of the deviations introduced by other bands, making possible to perform an individual fit in the least square sense.

As a first stage, the proposed procedure estimates the location of every band in the signal using the stationary points $\lambda_0$ and $h$, which are obtained from the first derivative and the $\gamma_-$, $\gamma_+$ values from third derivative, using the same equation exposed in section 3.4.

In order to reduce the number of points to include in the LSF procedure is necessary to associate each $\lambda_0$; $h$; $\gamma_-$ and $\gamma_+$ values to their corresponding band, as shown in figure 1, where the triangles establish the sampling interval. It can be noted that valleys central coordinate could be easily adopted as boundaries for the sampling interval; however according to Guo,[133] taking samples with less than 20 percent of peak height deteriorates the fitting quality. Additionally, when a suitable $\gamma_-$, $\gamma_+$ value cannot be identified because an overlapping of two bands occurs, the nearest valley is used instead. The $\gamma$ value is also determined from the sampling interval, see figure 1. In this case, the distance from $\lambda_0$ to the boundary value $\gamma_-$ or $\gamma_+$ that is closest to half height is taken as $\gamma/2$ value; $\gamma_+$ in this case.

**Figure 5.2**

*Lorentzian parameters estimation based on stationary points. Circular marks represent the $(\lambda_0, I)$ coordinates, while the point identified with a $\times$ is the calculated half height peak value.*
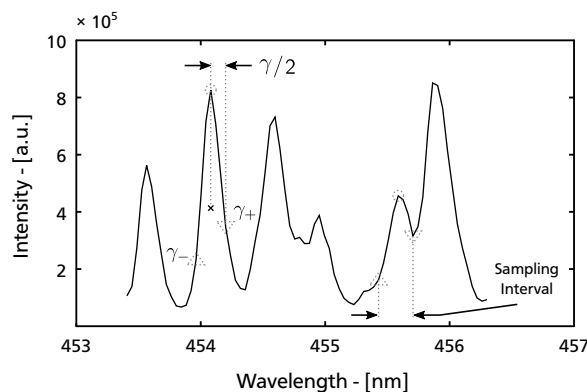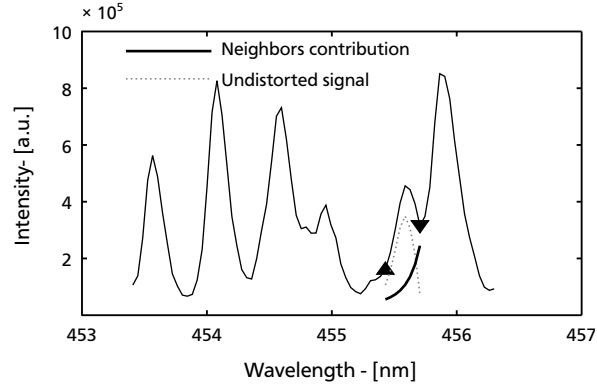
*Neighbors' height contribution represented with bold line is subtracted from the measurement signal to obtain an approximate set of fitting points free of distortion shown as a dashed line.*



At this stage the area of some bands is overestimated, due to band overlapping. For this reason a correction for the h parameter is introduced. First, a synthetic signal $R(x)$ is computed with the superposition of every detected peak $L_i(x)$

$$R(\lambda) = \sum_{i=1}^{n} L_i(\lambda) \tag{5.23}$$

$R(\lambda)$ only needs to be evaluated on the peak central locations $\lambda_{0_i}$ of the $i$-esim peak and compared against the measured spectrum $M(\lambda_{0_i})$ at the same point. This gives a set of adjustment factors $H$ for every peak according to:

$$H(\lambda_{0_i}) = \frac{R(\lambda_{0_i}) - M(\lambda_{0_i})}{2} \tag{5.24}$$

Adjusted factors $H(\lambda_{0_i})$ are subtracted recursively to the numerically estimated heights $I_i$, obtaining an improved height estimation with every iteration until a suitable stop condition is reached. The 2 in the denominator in equation 5.24 is used to avoid an excessive correction factor that could lead to negative values of h. From this point the iterated use of equation 5.24 will be referred as Simultaneous Height Adjust (SHA).

Based on this initial set of estimated Lorentzian parameters, given any peak in the signal, the contribution of other signals to their height can be estimated. This is a straightforward procedure that subtracts the synthetic signal evaluated on the peak's sample range from the measured spectrum, leading to a set of points that are an undistorted approximation of the Lorentzian peak, as figure 5.3 shows. The synthetic signal is calculated with every peak in the spectrum, except the evaluated peak.

The adjusted set of points can now be fitted in the least square sense to a Lorentzian curve. The classical approach is taking the inverse of equation 3.1, for obtaining a second order polynomial:[134]

$$\frac{1}{L(\lambda)} = \frac{4}{\gamma^2 I} \lambda^2 - \frac{8\lambda_0}{\gamma^2 I} \lambda + \left[ \frac{4\lambda_0^2}{\gamma^2 I} + \frac{1}{I} \right] \tag{5.25}$$

the coefficients of the polynomial $y = ax^2 + bx + c$, are related to the Lorentzian parameters with the following relationships

$$\lambda_0 = -\frac{b}{2a} \tag{5.26}$$

$$h = \frac{4a}{4ac - b^2} \tag{5.27}$$

$$\gamma = \frac{\sqrt{4ac - b^2}}{a} \tag{5.28}$$

Sometimes this model leads to overestimation of peak heights, while in others, the calculated coefficients leads to an imaginary value of $\gamma$. These drawbacks can be avoided employing a fixed height model fitting adjust

$$\frac{I}{L(x)} = \frac{4}{\gamma^2} x^2 - \frac{8\lambda_0}{\gamma^2} x + \left[ \frac{4\lambda_0^2}{\gamma^2} + 1 \right] \tag{5.29}$$

which is related to Lorentzian parameters by

$$\lambda_0 = -\frac{b}{2a} \tag{5.30}$$

$$\gamma = \frac{2}{\sqrt{a}} \tag{5.31}$$

After a peak has been processed, SHA is applied to the remaining unfitted peak heights, in order to take into account the contributions from the fitted peaks. At the end, with the aim to improve the spectrum parameter estimation, the whole spectra fitting procedure can be recursively applied employing the previous fitting values as the starting guess condition.

### 5.2.3   **Results and Discussion**

Performance of SHA is evaluated with a synthetically generated spectrum of Fe-I in the region between 290–370 nm, where the minor lines within an overlapped band are discarded. The routines were written in Matlab programming environment. As mentioned previously, two iterations were made in order to improve the fitting procedure, whose results were compared against LM algorithm.

The fitted signals quality was evaluated with the coefficient of determination $R^2$. Values closer to one indicate good correspondence between the model and the experimental signal. The performance test throws a $R^2$ equal to 0.9998, as shown in table 5.2. Peak's parameters calculated with the proposed method, shown a good agreement with the theoretical values used in the simulation. The error shown in table 5.2 is the average estimation error of spectrum's parameters bands. In particular, the width and central band position estimation error are two orders of magnitude lower than the resolution of many commercially available spectrometers, typically 0.02 nm. It is also worth to mention that the average height error is close to 0.2% when compared against the maximum height value in spectrum.

**Table 5.2**

*Average error in the Lorentzian curve parameters estimation for a Fe I synthetically generated spectrum without band overlapping.*

| Parameter | Metric | Value |
|---|---|---|
| Goodness of fit | $R^2$ | 0.9998 |
| Central wavelength | Average error | 0.0004 |
| $\lambda_0$ [nm] | $\sigma_{\lambda_0}$ | 0.0013 |
| Height | Average error | 0.1680 |
| $I$ [%] | $\sigma_I$ | 0.5542 |
| Width | Average error | 0.0007 |
| $\gamma$ [nm] | $\sigma_\gamma$ | 0.0038 |

On the other hand, Levenberg–Marquardt algorithm can perfectly recover all the theoretical values used to generate the synthetic spectrum. This formidable performance is a tradeoff between speed and accuracy, as can be seen in figure 5.4.

Execution time linear behavior for the proposed method contrasts with the time expended on LM algorithm, which grows exponentially as the number of peaks within the signal is increased. Figure 5.4 was calculated as the average of five runs on both procedures. The increase in the execution time in LM algorithm not only depends on the number of fitted peaks, but also on the number of points which comprises the evaluated signal, as shown in figure 5.5

**Figure 5.4**

*Mean elapsed time comparison between SHA and LM method in logarithmic scale, as a function of fitted peaks.*
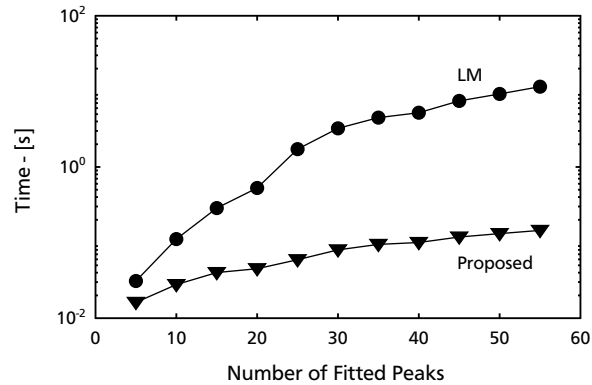


**Figure 5.5**

*Mean elapsed time comparison between SHA and LM method as function of the number samples in the signal.*
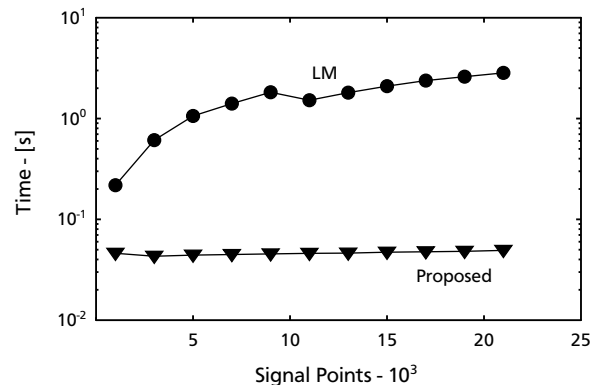
**Figure 5.6**  *Fitted results comparison between SHA and LM, applied to a D2 steel LIBS spectrum.*
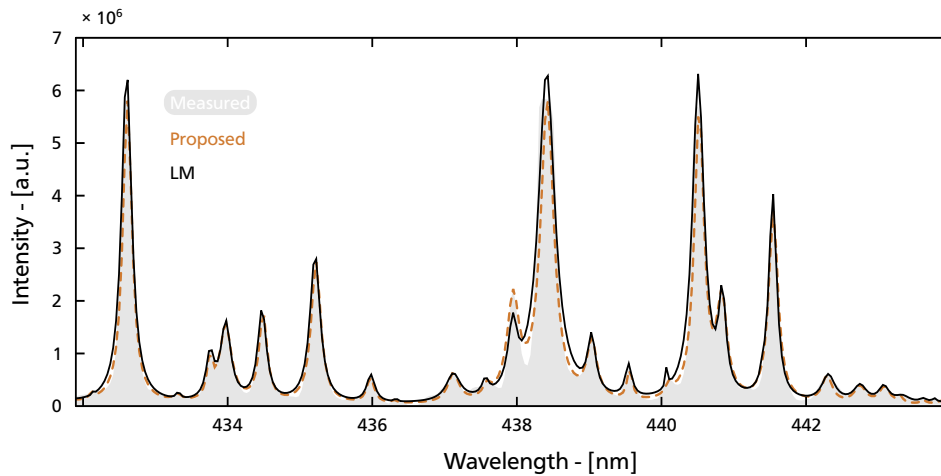


The full wavelength range is used to evaluate the predicted signal on each LM iteration. Besides, both target matrix values and derivatives grow with every added variable, therefore increasing the computing complexity. In contrast, SHA fitting procedure only evaluates the signal in a reduced set of points.

As a practical example, the fitting result for a LIBS spectrum is presented. This spectrum was generated by the atomic emission of Fe, V and Cr emission on a sample of D2 steel.

Both homogeneous and non-homogeneous broadenings, are present in LIBS spectra, thus, the most appropriate line fitting would be Voigt profiles, but as pointed by Fantoni et. al.,[135] Lorentzian line fits are suitable for most practical purposes in high-density plasma. According to Singh et. al.[49] LIBS plasmas have a very high electron density, $10^{15}$–$10^{18}$ cm$^{-3}$, for a considerable period of time in which Doppler and natural broadening are generally negligible. Moreover, Huang et. al.[136] have shown that when in a Voigt profile the Lorentzian width is greater than the Gaussian width in a 3:1 ratio both Voigt and Lorentzian profiles are nearly the same.

The initial estimate of the Lorentzian parameters were extracted from the signal stationary points, and used as initial guess for both LM and SHA giving a $R^2$ value of 0.967 and 0.928 respectively.

A LIBS spectrum band is affected by broadening mechanisms as collision and Doppler effects, which modifies the natural line shape. LM algorithm tends to compensate these broadened lines with an increment in the band height—as can be seen on figure 5.6—thus, predicting a line with a little more energy or area than those measured. The proposed algorithm is free of such effect and never overestimates height values.

### 5.2.4   Conclusion

This section presented an alternative approach to fit signals composed by multiple Lorentzian curves. SHA is a straightforward and fast procedure that overestimates the signal area in a minor proportion when compared against LM algorithm; this is an important feature for some applications. Execution time gain is accomplished at the expense of precision; fitted signals will have a slightly lower determination coefficient $R^2$ than those obtained with LM. However, for CF algorithm where theoretical values have a great uncertainty this feature is no longer a concern.

SHA is suitable for applications where a restriction on the computing power exists such as portable devices. Additionally, a low executing time is an important feature when fitting is applied to signals with both high resolution and large number of peaks, as is the case for those produced by Echelle spectrometers.

## 5.3   The Partition Function

The thermodynamic state of a system at equilibrium can be completely described by means of the *partition function*, from which, all the thermodynamic functions can be extracted.[137] The mathematical expression is

$$U(T) = \sum_i g_i \, e^{-E_i/kT} \tag{5.32}$$

where $T$ is the system´s temperature, $k$ represents the Boltzmann constant, $E_i$ is the microstate energy, and $g_i$ is the *degeneracy factor*; namely, the number of quantum states that have the same energy $E_i$. All the spectroscopic parameters required for computing $U(T)$ can be consulted on literature, as stated previously in section 3.1.

Equation 5.32 is a divergent series which must be bound in order to obtain useful values. As will be seen in the following paragraphs, there is not a unified stop criterion for computing $U(T)$. Based on this fact, three distinct information sources were processed, with the aim to select the most suitable for CF calculation.

As a first approach for determining the partition function, this work used the tables published by Irwin[138] who made a polynomial adjust for the neutral species and the first two ions for elements ranging from H to U, as well as some molecular species. The fitted polynomial is

$$\ln Q = \sum_{i=0}^{5} a_i \, (\ln T)^i \tag{5.33}$$

The provided coefficients for $a_i$ were saved in Matlab and verified against the $\ln Q(16000°K)$ field for checking type writing errors. Saved coefficients are then accessed by `FP_Irwin` function, which computes $U(T)$ and only needs as input the emission species and plasma temperature

Kurucz database also provides the species partition function values within a fixed set of temperatures; in this case information is organized in a tabular form within an `ASCII` file.

**Figure 5.7**

*Fitting results for partition function values obtained at Kurucz database for Fe I species.*
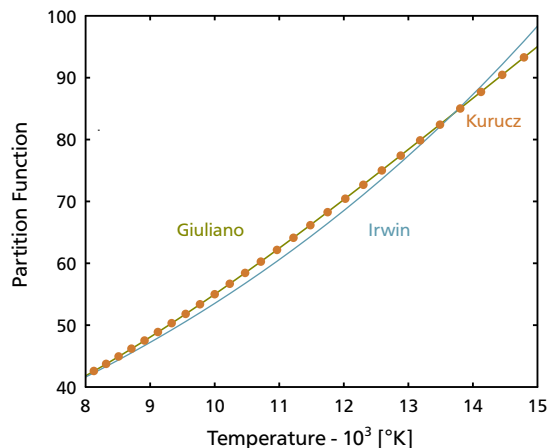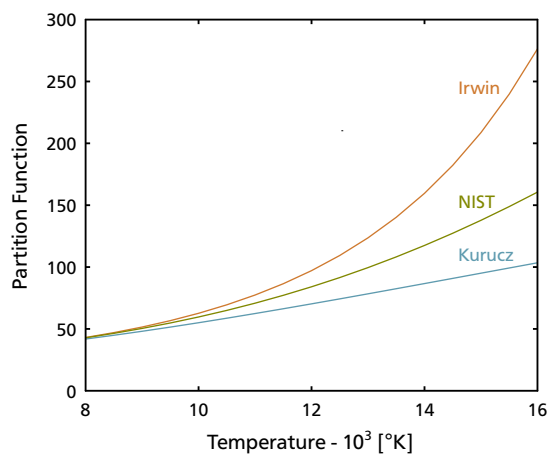


**Figure 5.8**

*NIST Atomic Spectra Database Levels Form.*



**Figure 5.9**

*Partition function values comparison for Kurucz, NIST and Irwin.*

Interpolation is necessary for those values of $T$ that are between the specified steps in Kurucz tables, and considering that the partition function does not follow a linear relationship this is far to be a straight forward task. For reaching of a simpler solution, Kurucz partition values were fitted to a polynomial to conform to the equation 5.33.

Considerable discrepancies for the $U(T)$ predictions at high temperatures were found while applying this model, as shown in figure 5.7. Giuliano et al.,[139] propose a simplified expression for $U(T)$ computing, whose mathematical expression is given in equation 5.34. Under this scheme a better prediction of the partition function values provided in Kurucz database was achieved, as shown in figure 5.7. In this example the resulting coefficient of determination for Irwin's model was 0.9975, while for Giuliano model was unity.

$$Q = \sum_{i=1}^{3} a_i \, e^{b_i/T} \tag{5.34}$$

NIST database can also calculate the partition function value for a given plasma temperature.[140] The NIST FrontEnd used in this task is depicted in figure 5.8. The resulting value is embedded inside a table along other information which can be selected on demand. Although there is free access to NIST page, a CF computing program will fully depend on Internet access; besides a complicated program is necessary for automatic fill form and extraction of $U(T)$. To avoid these drawbacks, the level information for every emission species was downloaded and translated to the `.MAT` format with a program similar to that described in section 3.1. Evaluating equation 5.32 with these data it is possible to duplicate partition function values provided at NIST web page.

To close this section, the behavior of the three computing methods for partition function are compared in figure 5.9, where high differences at plasma temperatures above 10,000 °K for Fe I species can be appreciated.

## 5.4   Calibration Free Implementation

This section presents the implementation details of the CF algorithm, employing a D2 steel sample to describe the process. In this case, the experimental setup uses a Nd:YAG Spectra-Physics laser whose emission is centered at 532 nm with 10 Hz pulse rate and 10 ns pulse duration; an average power of 115 mW was employed to interrogate the sample. Plasma emission was collected with an optical fiber and guided to an Acton research spectrograph with 1200 lines/mm grating. A spectrum comprises 10 accumulations with a delay time of 1.7 $\mu$s and 1 $\mu$s gate width.

Both amplitude and wavelength calibration are mandatory for quantification but the former must be applied as the last preprocessing step. Baseline signal was removed with a fourth order descendant SNIP with 47 point window. As seen on section 3.5, line height depends on plasma temperature values; this fact may alter line selection for trace elements, thus an alternative approach to get plasma temperature estimation is needed.

In this work, a reference synthetic spectrum of Fe I was computed using distinct temperatures, which was compared against the measured signal by means of the correlation coefficient. The temperature with better $R^2$ value was used for line selection, 14000 °K in this example, as shown in figure 5.10.

Tognoni et al.,[71] recommends the measuring of the integral line intensity over its wavelength profile; therefore, once the spectrum has been preprocessed, the individual Lorentzian shapes are calculated with the SHA procedure discussed on section 5.1, see figure 5.11. Then, Lorentzian peaks are matched with theoretical values. Only those measured profiles with a single theoretical emission line within its FWHM are selected, as shown in figure 5.12. Selected lines can be consulted in appendix C.

Reader may recall that thin plasma line emission is one of the requisites for applying CF method. Thin lines were detected using the ratio between two lines of the same species and the same upper level energy value. Under these circumstances the ratio between such lines by means of equation 5.5 results in

$$I_{ij} A_{mn} g_m = I_{mn} A_{ij} g_i \tag{5.35}$$

Equation 5.35 is independent of partition function, total number density or plasma temperature. Allowing lines with $\pm$ 5% of error, a pair of thin plasma lines was found, located at 419.143 and 421.034 nm. Two lines are insufficient to compute plasma temperature within a Boltzmann plot; however, relaxing the equal upper energy condition leads to the following expression

$$\frac{I_{ij} A_{mn} g_m e^{-(E_m - E_i)/kT}}{I_{mn} A_{ij} g_i} = 1 \tag{5.36}$$

Using the pair of thin plasma lines as reference for comparison, as well as the best correlation plasma temperature for background estimation and a maximum allowed error of $\pm$ 8%, five more Fe I lines were found. Among the lines suitable for CF calculation there is a Fe II line located at 417.885 nm. With this set of selected lines, the transformation given by equation 5.6 can be applied. A slope of -0.8682 was obtained, equivalent to a 13367 °K of temperature, along a $n_e$ equal to $9.9418 \times 10^{17}$ cm$^{-3}$, see figure 5.13

Before giving further details, a brief discussion about the other two conditions that must be satisfied for applying CF is in order. The first condition is plasma LTE, that is to say, at least the McWhirther criterion must be fulfilled. Using equation 5.1 with the electron density given by Saha-Boltmann plot, leads to $\Delta E = 17.5$ eV. The highest $\Delta E$ measured corresponds to a Fe I line that is located at 400.5242 nm or 3.09 eV, which is well below of the highest energy transition for which McWhirther criterion holds.

The second condition is stoichiometric plasma which is fulfilled with laser power densities above 1 GW/cm$^2$. The peak power for a laser pulse of 10 ns width and average power of 115 mW is 11.5 MW. On the other hand, the airy disk radius for a lens of 2.54 cm diameter and 7 cm focal length is 1.78 $\mu$m at 532 nm; hence, the system power density is $1.14 \times 10^{14}$ W/cm$^2$, so it is possible to conclude that

**Figure 5.10**  *Comparison of measured signal against synthetic spectrum at different plasma temperatures by means of correlation coefficient.*
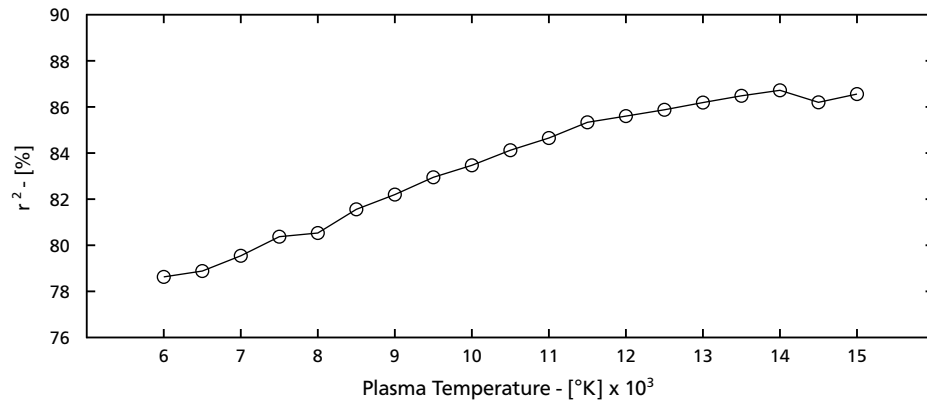


**Figure 5.11**  *D2 steel fitted LIBS spectrum.  The solid background represents the measured signal after preprocessing.*
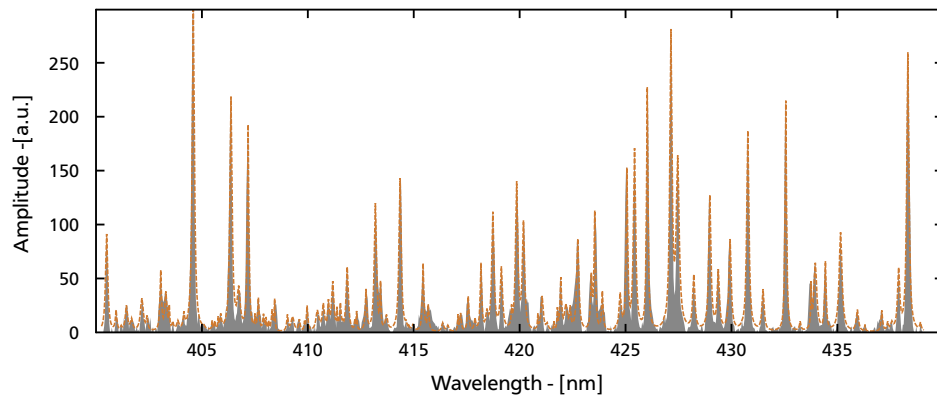


**Figure 5.12**  *Lines used in CF algorithm. Line's FWHM are marked with dots.*
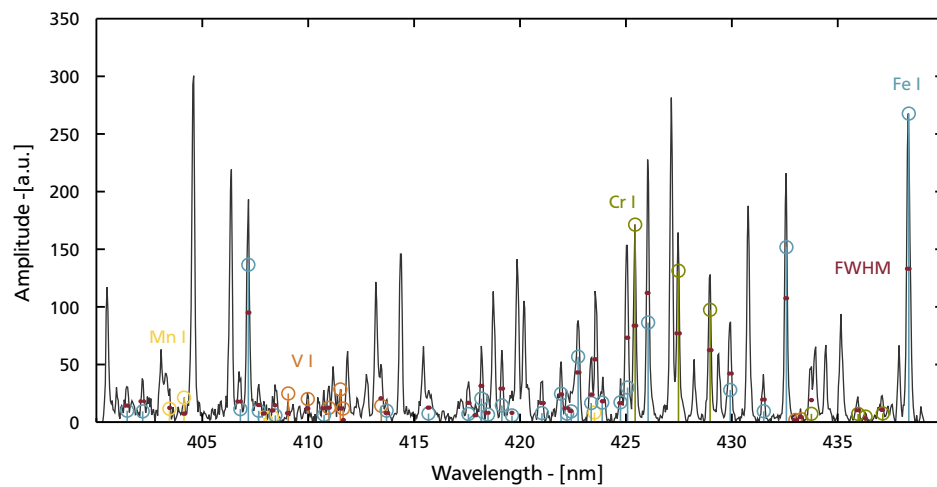
**Figure 5.13**  *Saha-Boltzmann plot for Fe I / Fe II selected lines.*
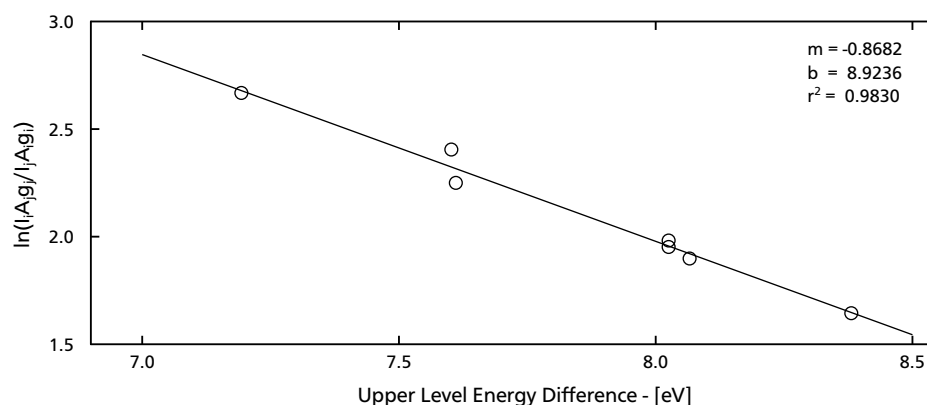


plasma is stoichiometric. With the certainty that the conditions for applying CF have been covered, it is time to address attention to the lines selection issue.

As plasma thin lines detection by means of equation 5.35 was able to locate only Fe I species, another line selection criterion must be used. For this purpose, the ratio between two lines of the same species eliminates $n_s$ dependency, as seen on equation 5.36, but contrary to Fe case, there is no pair of lines with the same upper energy level for V, Cr and Mn. Hence, every line was compared against each other computing the total number of lines that are between a tolerance of $\pm$ 10%. The set of lines with a maximum number of matches and lowest average error was selected. This criterion was applied for selecting Cr I lines.

It must be said that lines with values of $A_i g_i$ above $7 \times 10^8$ were discarded before performing comparisons for Fe I and Cr I, because lines with high emission probabilities are more prone to self-absorption effects.[141] Discarding lines by its $A_i g_i$ was not applied for trace elements because the most prominent lines were lost under this restriction. For trace elements, the ratio among all lines was also computed, but the first selection criterion changed. In this case, the line that leads to a minimum variance is selected, then, its standard deviation and mean are added in order to find a variation range; a line whose mean falls into this range is chosen.

Following, equation 5.6 is applied on the set of selected lines for each emission species, and the resulting coordinates are fitted to a fixed slope line, with the aim to establish the same temperature in every species; results are shown in figure 5.14.

Now, what remains to be done is quantify elemental content. As mentioned in section 3.5, the sample has C but it was not detected with the experimental setup employed because special cautions are required for this task; while Mo only has very weak lines in the measured region, besides, there is no information at NIST database for its neutral line's upper level energy.

**Figure 5.14**  *Boltzmann plot for selected emission lines of Fe I, Cr I, Mn I and V I.*
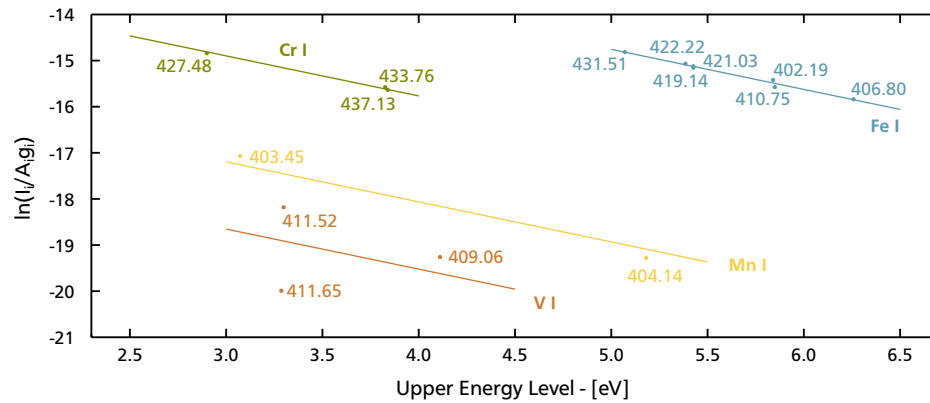


**Table 5.3**  *Elemental ratios normalized against Fe. $q_s$ represents linear fitting intercept.*

| Element | Literature | | CF-NIST | | CF-Ratio | |
|---------|-----------|-------|---------|----------|---------|--------|
|  | % Weight | Ratio | Ratio | $q_s$ | Irwin | Kurucz |
| Fe | 84.15 | 1.0000 | 1.0000 | -10.4161 | 1.0000 | 1.0000 |
| Cr | 12.00 | 0.1532 | 0.1535 | -12.2934 | 0.2054 | 0.1510 |
| V | 0.90 | 0.0117 | 0.0116 | -16.0496 | 0.0156 | 0.0138 |
| Mn | 0.35 | 0.0042 | 0.0064 | -14.5916 | 0.0090 | 0.0064 |
| C | 1.55 | | | | | |
| Si | 0.30 | | | | | |
| Mo | 0.75 | | | | | |

This implies that closure equation will lead to wrong concentrations; therefore, result evaluation must be done in terms of concentration ratios as discussed on section 5.1. Table 5.3 shows ratios reported in the literature [142][†] and CF computed ratios, considering just neutral and single ionized species.

Concentration was determined with equation 5.16, using partition functions values from NIST, Irwin and Kurucz. Total element density, including neutral and first ionic species, was computed with Saha's relationship and equation 5.12.

Best results were obtained with NIST partition function. Kuruckz partition function also has a good behavior but not as good as NIST. Finally Irwin's values leads to overestimate minor and trace elements concentration.

---

[†]Literature reports concentrations as weight proportions. Ratios were obtained transforming this quantities to number of atoms and then to ratios.

## 5.5   Conclusions

This section has shown the implementation of CF technique, suggesting computing procedures for solving traditionally considered cumbersome tasks, as elemental identification and spectral fitting with Lorentzian shapes. In particular, Lorentzian fitting makes preliminary line selection an easy task.

Based on plasma thin line definition, it was possible to automate a line selection routine for Fe and Cr. A similar procedure was applied for selecting lines for trace elements quantification. Trace elements processing still poses some challenges due to the very small line emissions which are difficult to identify and measure, thus, user intervention is needed to achieve good results.

It has also been introduced an alternative method to compare CF results based on elemental composition rations rather than concentrations. As the results show, elemental composition ratios are preserved no matter if closure equation or plasma neutrality conditions are used in the quantification procedure. Moreover, it has become evident that computing the experimental factor F does not play a key role when ratio composition is used as comparison basis; making unnecessary the use of a closure equation.

Best quantification results were obtained with both NIST and Kurucz partition function, while the reported values by Irwin leads to overestimate minor and trace concentrations. It must be added that knowing plasma electron density is crucial in CF experiments; the final determined composition heavily depends on its value. As discussed, little errors in $n_e$ are not translated in big quantification errors when the plasma has a high electron density value. This means that small delay windows should be preferred while measuring, as long as emission lines are not dimmed by the continuum.

# Chapter 6

# Conclusions

This thesis has shown the feasibility of measuring both low-frequency Raman and Laser-Induced Breakdown spectra in a single instrument. Not only a novel setup was developed, but also novel signal processing procedures were proposed, which takes a step forward to automatize both elemental identification and quantification.

Elemental identification is the basis of any chemical analysis but until now, the available methodologies required a great deal of effort to implement and not always provide adequate results. The proposed Wavelength Similarity Coefficient method could change this scenario because it is a straightforward method based in some often overlooked facts, as the temperature dependence of the spectral line relevance and the use of a simplified modeling for the plasma's line heights.

On the other hand, it was demonstrated that when a group of related samples is available, information beyond the elemental content of the sample can be gathered, avoiding the uncertainties inherent to the quantification process based on LIBS spectral analysis. Exploiting this fact, the classification of stainless steel according to its crystalline structure is a reality. In addition, the combination analysis of LIBS spectra with multivariate analysis, has shown that the manufacturing process of copper archaeological artifacts leaves behind a chemical signature, which could be used to identify the raw material source or assign a historic period based in the technologies that were available during their manufacturing.

Obtaining the individual emission line parameters form the measured spectrum is a fundamental step in the quantitative analysis. A new approach was introduced in this work, aimed to solve some drawbacks of the Levenberg-Marquardt method, as the convergence and line energy overestimates, which potentially leads to wrong elemental composition. Besides, it has been shown that partial element measuring is possible by means of LIBS analysis, as long as the elemental concentration is reported as elemental ratios, taking preferably the matrix element as reference.

Finally, it was also shown that electron temperature is a fundamental plasma parameter that is directly related with the precision of the quantification process, and should not be derived with iterative procedures, because it is directly related with the absolute species concentration, which is of course an unknown.

## 6.1   **Future Work**

The successful use of a low-power laser source for LIBS measurements suggest that a mobile RLS instrument could be implemented. Such device is of great interest for archaeologist because take out some pieces from its guard site is in some cases nearly impossible. This is not the only improvement that can be incorporated to the RLS instrument; a narrow laser line would allow lower frequencies to be detected. Another relevant feature to incorporate is the automatic identification of Raman spectra.

From a LIBS experimental point of view, an improvement on the measured intensity signal is beneficial. This could be done applying the double pulse laser technique or enhancing the signal with microwaves.

An Echelle spectrometer will help to reduce damage in samples during LIBS experiments. Unfortunately, the experience with this kind of devices has shown that they work better with high laser energies and single exposures; while this work obtained the best results taking several accumulations. The Echelle spectrometer software used for this test is very limited and must be improved in order to fully exploit all the advantages that can potentially provide.

Although a bunch of equations are the basis of quantitative elemental analysis, its application is far to be an easy task. There are many requirements that must be carefully fulfilled, as wavelength and amplitude calibration, baseline correction, line identification and deconvolution, as well as theoretical spectral parameters. In particular, baseline correction is a very sensitive procedure that can easily lead to wrong results. An alternative way to avoid this procedure is to measure at larger delay times, but as the line intensity also decreases with time, delay times with the best trade off between signal to noise ratio and baseline must be investigated.

# Appendix

# Appendix A

# Raman-LIBS Instrument

Making an instrument capable of measuring both LIBS and Raman spectra is an idea that has been around for many years; however, to the author's best knowledge, there is not yet a wide availability of commercial equipment. Instead, several setups are reported each year in laboratories around the world. Perhaps the most famous combined Raman-LIBS spectroscopy (RLS) instruments are those used in spacial missions, such as the instrument that will be part of the rover of European Space Agency's ExoMars mission that will be launched to Mars in 2018.[143]

Most of the reported applications are aimed to take advantages of RLS remote sensing capabilities in hostile environments such as the planetary exploration[144] and explosive detection.[145] In contrast, this chapter deals with a RLS instrument able to perform Low-Frequency Raman measurements.

Many materials have Low-Frequency Raman emissions, which can be used to:

1. Discriminate the crystalline structure of materials.[40]
2. Determine bond lengths in gases by measuring its rotational modes.[146]
3. Study advanced semiconductors supperlattice structures.[147]
4. Measuring the diameter of carbon nanotubes by means of the study of its radial breathing mode frequencies.[148]

Despite its flexibility, low-frequency Raman spectroscopy has been relegated to a secondary roll due to triple-stage spectrometers drawbacks, which is one of the most common methods used. Although triple-stage spectrometers can be used with a wide range of excitation wavelengths and have an excellent stray light rejection, they are expensive and have low transmission efficiencies.[149]

This scenario changed with the introduction of new improved filters. Renishaw company offers with their equipment a filter called Near-Excitation Tuneable filter that works as an edge filter capable of measuring Raman signals below 12 $cm^{-1}$ with a 512 nm wavelength excitation source.[150] On the other hand, Ondax Inc., offers a Volume Holographic Grating notch filter that have a high transmittance at both sides of the notch and allows to detect signals below 10 $cm^{-1}$ at 532 nm.[151] As Renishaw is supplied with vendor's equipment, this work uses Ondax filters to implement the RLS instrument. Both alternatives are cost effective solutions to triple-stage spectrometers with higher throughputs.

## A.1  Experimental Setup Description

The RLS instrument was designed to operate with two distinct excitation sources at 532 nm. The first is a high energy Nd:YAG laser that operates in the nanosecond regime made by Spectra-Physics, while the second is a low energy Nd:YAG laser made by TeemPhitonics, which operates in the picosecond regime, see section 2.3.1 for full specifications. Both lasers share a common optical path though the RLS instrument, as shown in figure A.1.

Despite its low power output, the narrow beam from the TeemPhotonics laser is able to infringe damage on mirrors surfaces, thus, a custom made beam expander was designed and placed at the laser output. Also, an USB controlled shutter was placed in the common path to both lasers with the aim of providing a safe way to handle samples and to reduce sample damage.

**Figure A.1**  *Laser input setup for the RLS system.*
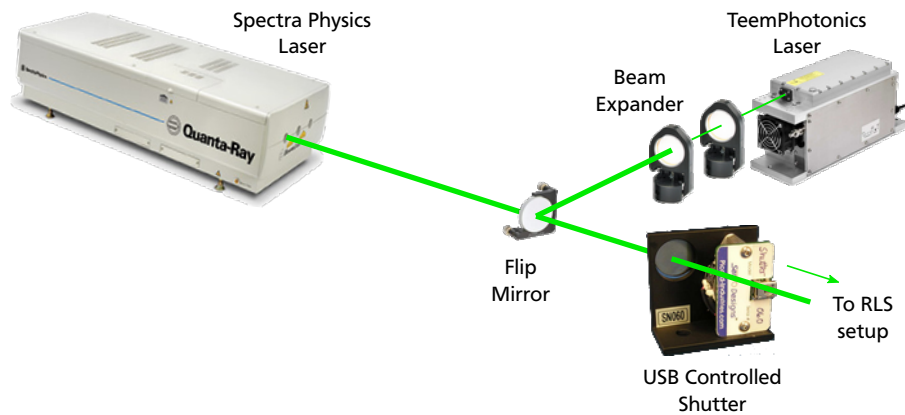


**Figure A.2**

*Focusing and collection optics for the RLS instrument.*
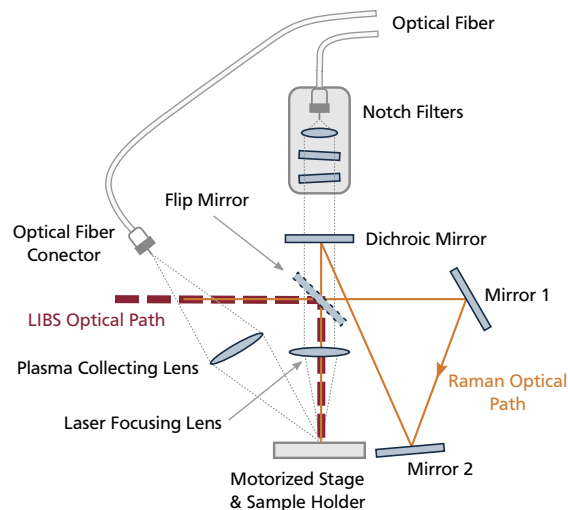
The main distinction between a time resolved Raman arrangement and a LIBS setup, is the filter that must be placed in the collection optics with the purpose of rejecting Raleigh scattering in Raman experiments. The proposed setup made use of two independent optical paths.

When LIBS experiments are done, a flip mirror guides the laser directly to a focusing lens of 7.5 cm focal length. The light emitted by the plasma is collected with a 2f system that employs a lens of 10 cm focal length, which focus the collected light in an optical fiber. A diagram of both focusing and collection optics is shown in figure A.2, where the dashed red line represents the focusing optical path for LIBS experiments, while the orange line is the path followed by the laser during Raman measurements.

When the flip mirror is moved away from the laser path, mirrors 1 and 2 guides the laser to a dichroic mirror. This last element, helps to reduce the noise in the resulting Raman signals and at the same time guides the light to the focusing lens, which is shared by Raman and LIBS excitation paths. A single filter is not enough to block the laser stray light at low-frequencies, therefore, two notch filters should be used in cascade. The proposed solution uses a commercial device that has two notch filters and optics to coupling the output to an optical fiber, as figure A.3 shows. Finally the optical fiber end is attached to any of the spectrometers listed in section 2.3.4.

**Figure A.3**

*Arrangement of the focusing and collection optics for the RLS system.*
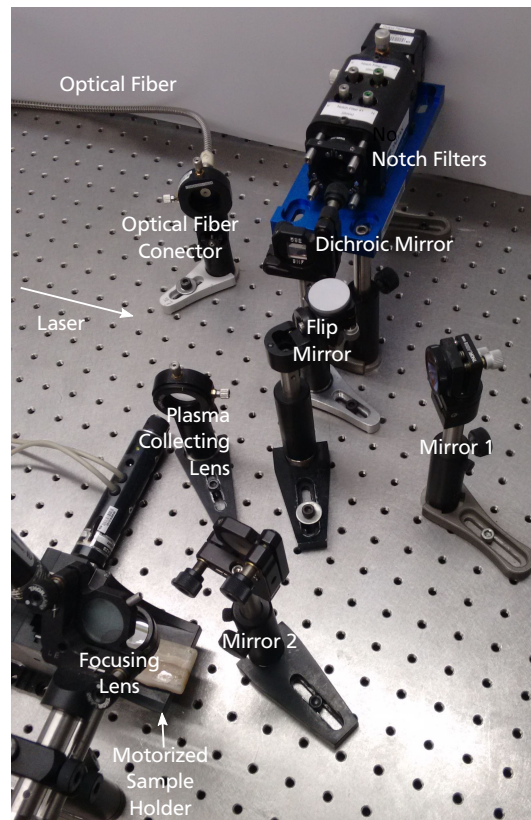
In its most common operation mode, the optical fiber is connected to LIBS port or to the Notch filter casing for measuring LIBS or Raman respectively, reflecting or transmitting the light with the flip mirror as described earlier. However, both Raman and LIBS could be measured using just the Raman optical path.

Ideally, the used cascade arrangement of notch filters have a width of 5.5 cm$^{-1}$ at either side of the central part of the notch at 50% transmission.[152] Therefore, less than 0.1 nm of spectral information is trimmed by the notch filters; in addition, the region between 531.8–532.2 nm doest not contain relevant lines for neutral emission atomic species.

As all the reported LIBS experiments in this thesis have been carried out with the independent focusing and collecting optical path, the following section presents the results of using the Raman path for low-frequency Raman spectra measurements in amino acids.

## A.2 Low-frequency Raman Measurements

The L-Cystine[153] is one of the most common substances to test the low-frequency response of a Raman instrument, because it has six easily distinguishable peaks below 100 cm$^{-1}$, as figure A.4 shows.

Each spectrum in this section was taken with an IsoPlane 160 spectrometer, while the signals were smoothed with a triangular filter of three points. Baseline was corrected with a fourth order descendant SNIP filter of 50 points.

The RLS instrument is able to measure bands above 30 cm$^{-1}$ in L-Cystine, as can be seen in figure A.5; this figure also shows the full Raman spectrum in the spectral region comprised between 30–3500 cm$^{-1}$. In this example, the 46 and 54 cm$^{-1}$ emission bands were mixed into a single band, while the remaining bands can be distinguished clearly.

**Figure A.4** *LCystine low-frequency Raman spectrum, taken from Ondax SureBlock XLF Notch Filter System brochure.* [152]
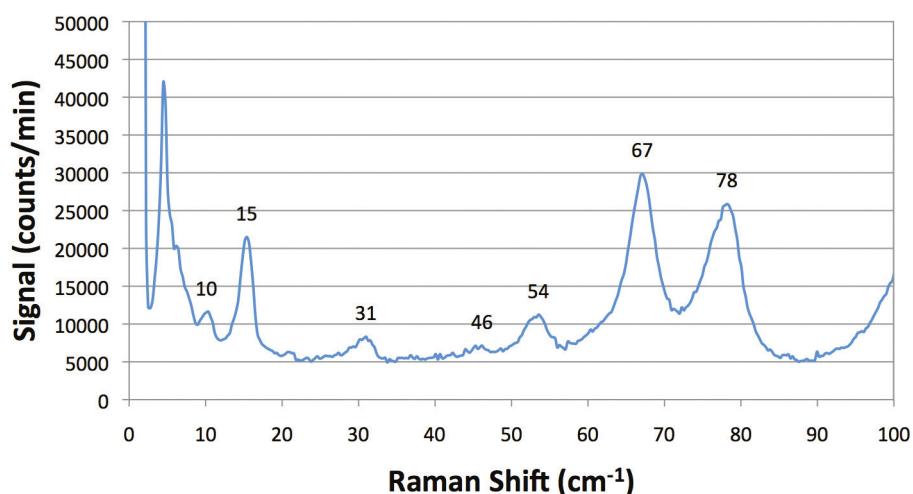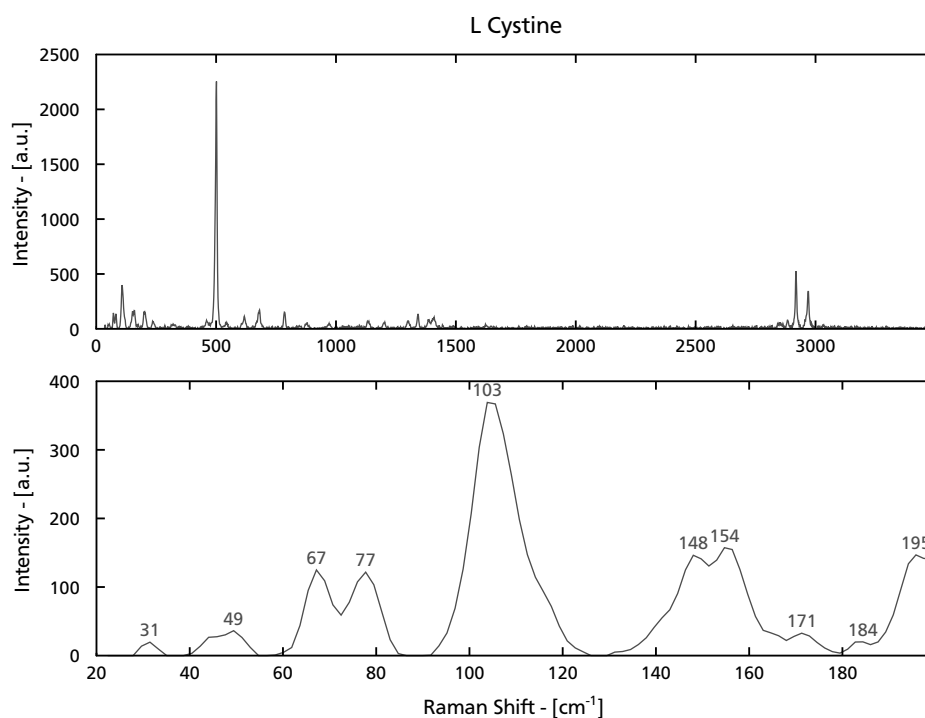
In order to measure emission bands closer to the laser line, a narrower laser in frequency is necessary. This kind of equipment is rather expensive, which limits a broader use. A more accessible solution are the laser-line filters that narrows the laser line with a reduction of power as a counterpart.

Besides L-Cystine, another 19 amino acids were measured. A list of the amino acids reported in this section can be found in table.A.1 Every amino acid with the exception of glycine can occur in two isomeric forms termed L and D, but only L-amino acids are manufactured by cells and incorporated into proteins.[154]

Amino acids have been broadly studied with standard Raman spectroscopy,[155] but few have been reported in the low-frequency region;[156] Therefore, this section only reports the low-frequency spectral portion.

**Table A.1** *Amino acids measured by means of RLS instrument.*

| | | | |
|---|---|---|---|
| L-4HProline | L-Alanine | L-Arginine | L-Asparagine |
| L-Aspartic Acid | L-Cysteine | L-Cystine | L-Glutamic Acid |
| L-Glutamine | Glycine | L-Isoleucine | L-Leucine |
| L-Methionine | L-Phenylalanine | L-Proline | L-Serine |
| L-Threonine | L-Thryptophan | L-Tyrosine | L-Valine |

**Figure A.6** *Low-frequency Raman spectra for L-4HProline, L-Alanine, L-Arginine and L-Asparagine amino acids.*
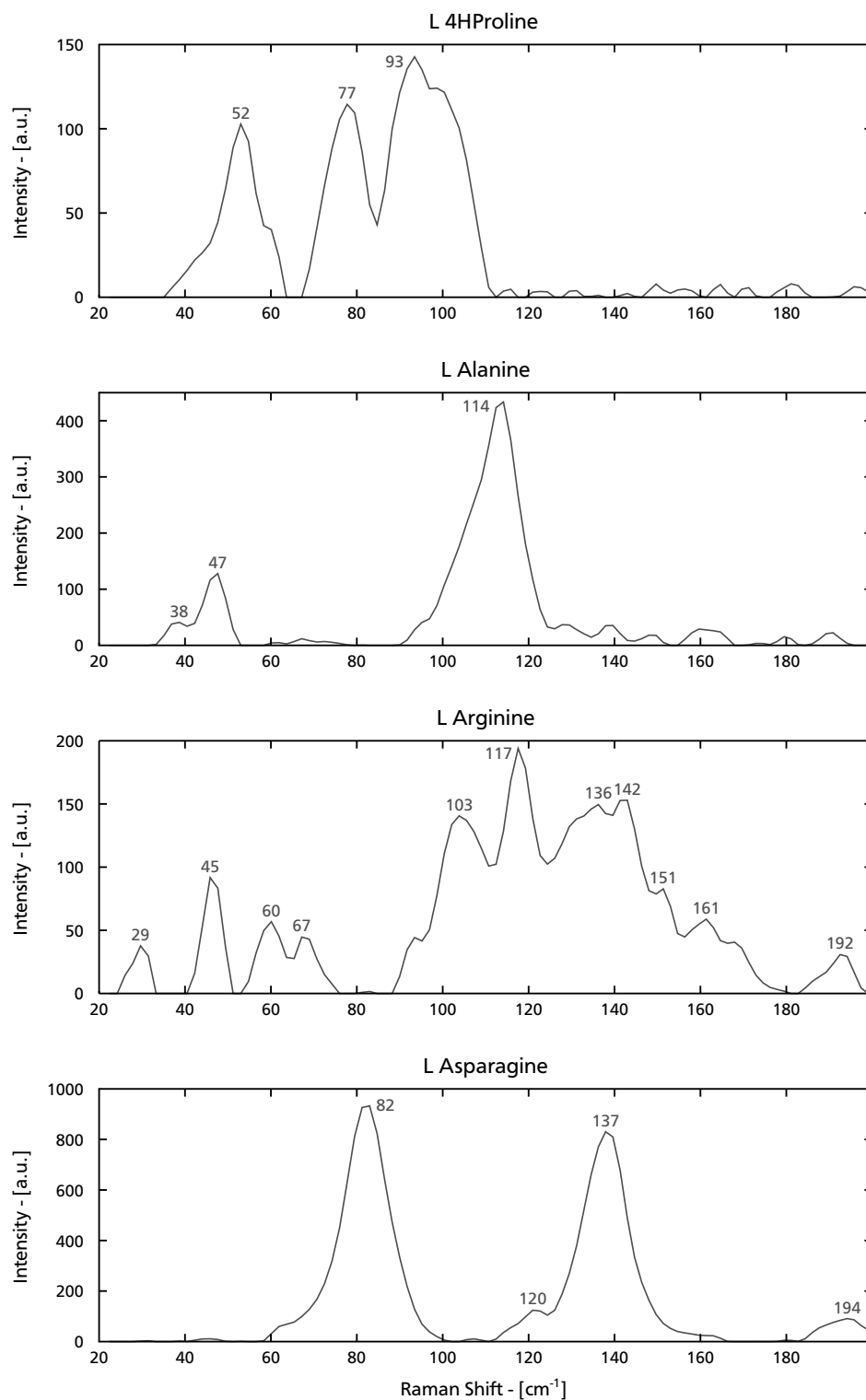
**Figure A.7**   *Low-frequency Raman spectra for L-Aspartic Acid, L-Cysteine, L-Glutamic Acid and L-Glutamine amino acids.*
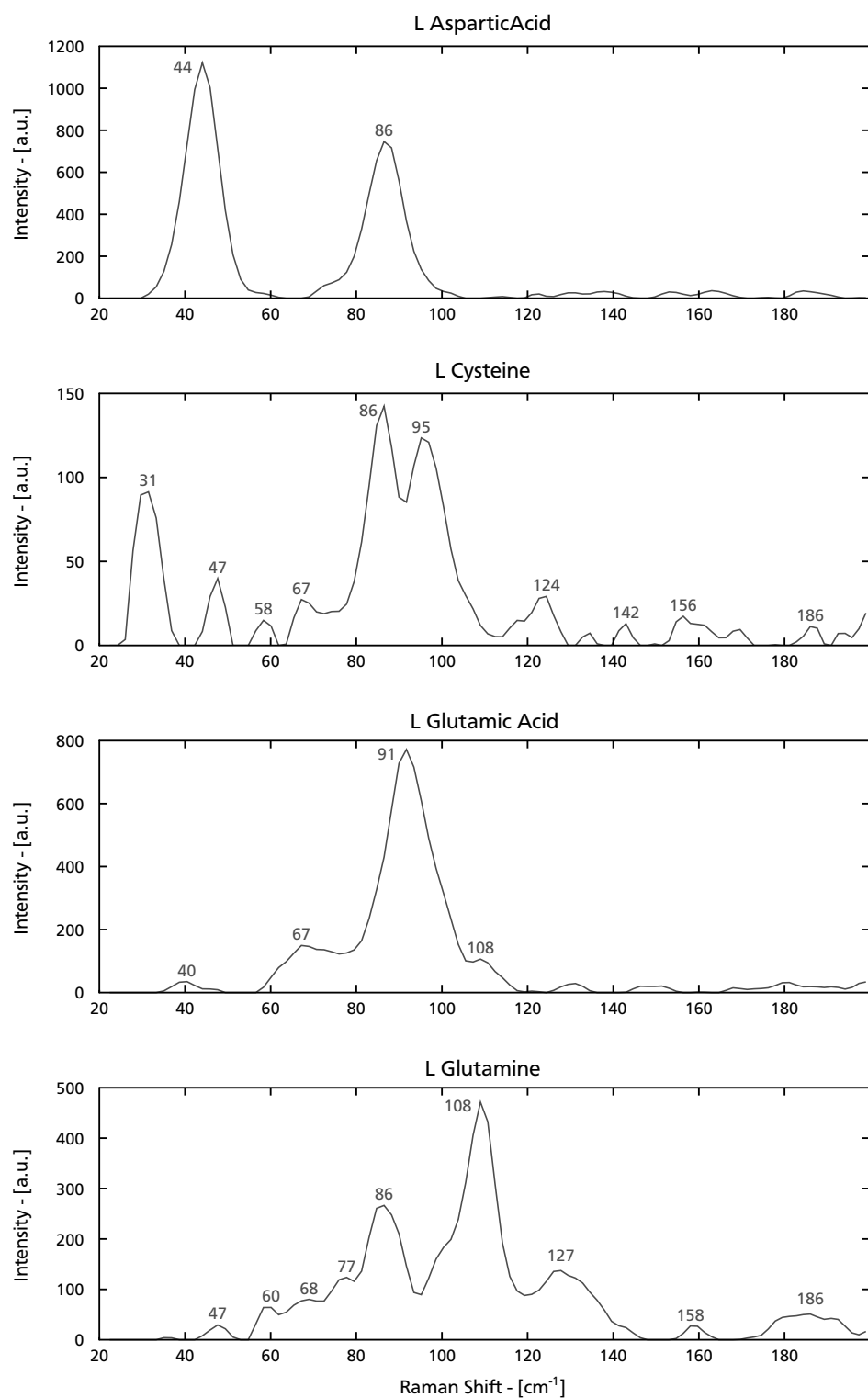
**Figure A.8** *Low-frequency Raman spectra for L-Glycine, L-Isoleucine, L-Leucine and L-Methionine amino acids.*
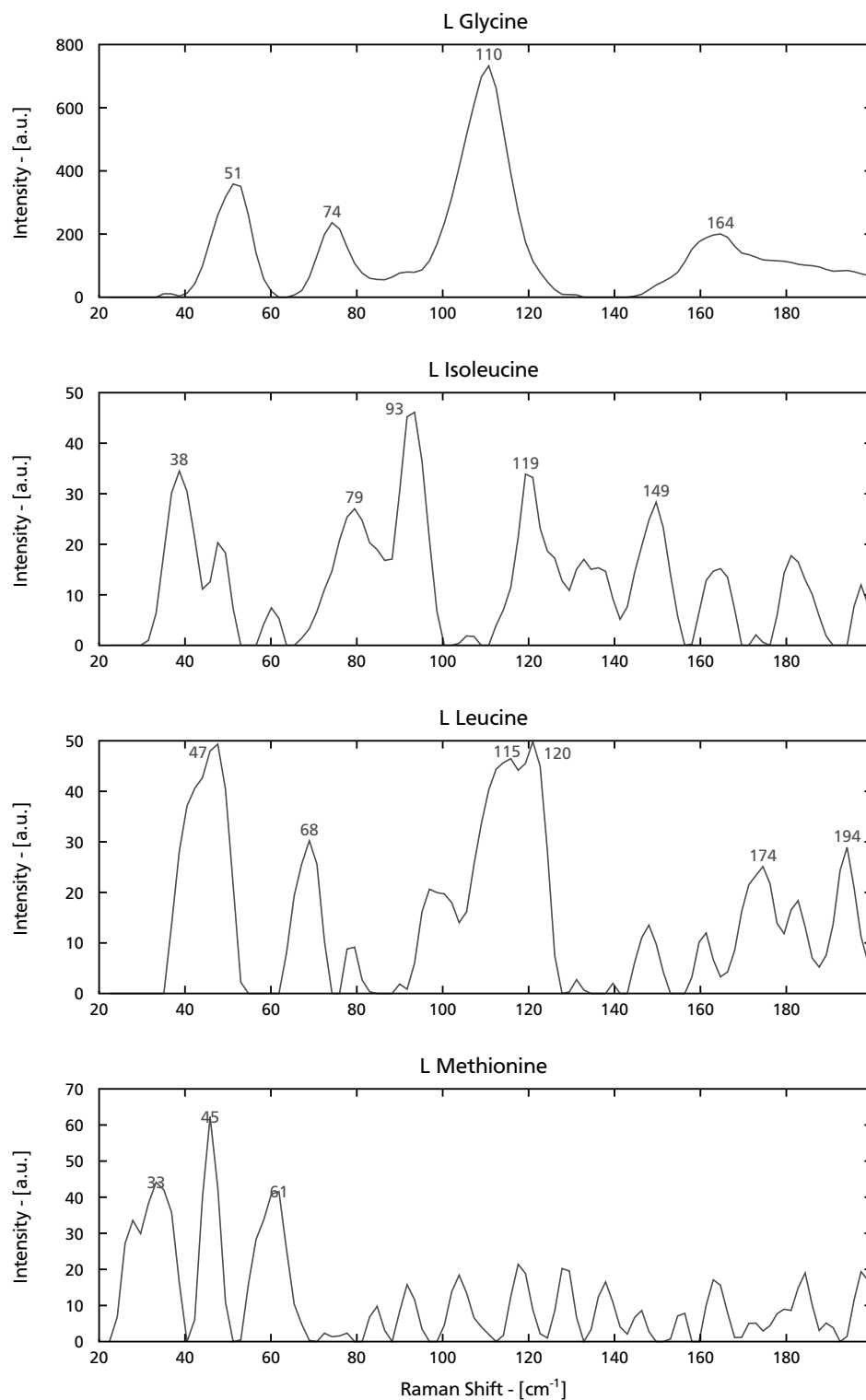
**Figure A.9** *Low-frequency Raman spectra for L-Phenylalanine, L-Proline, L-Serine and L-Threonine amino acids.*
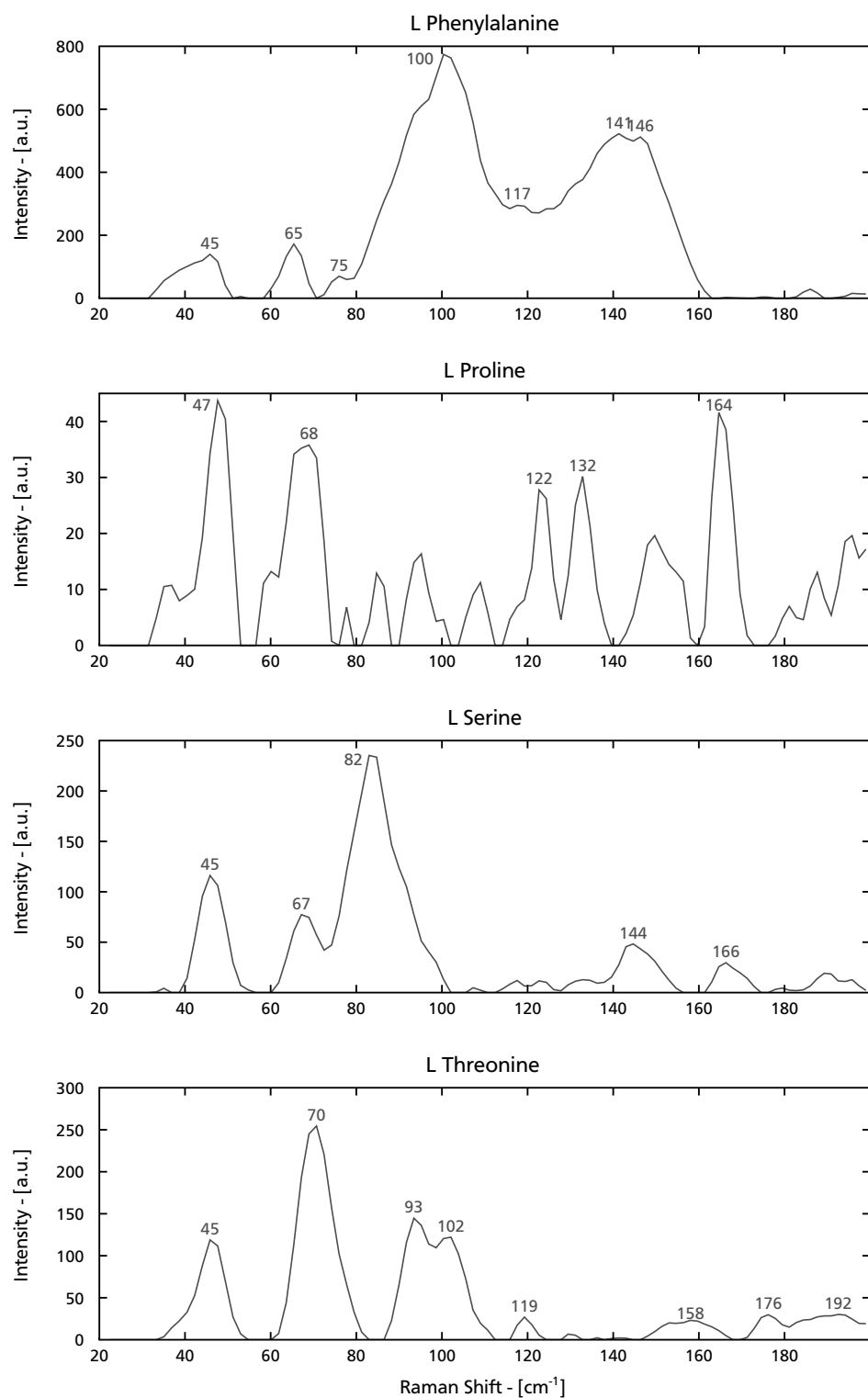
**Figure A.10** *Low-frequency Raman spectra for L-Thryptophan, L-Tyrosine and L-Valine amino acids.*
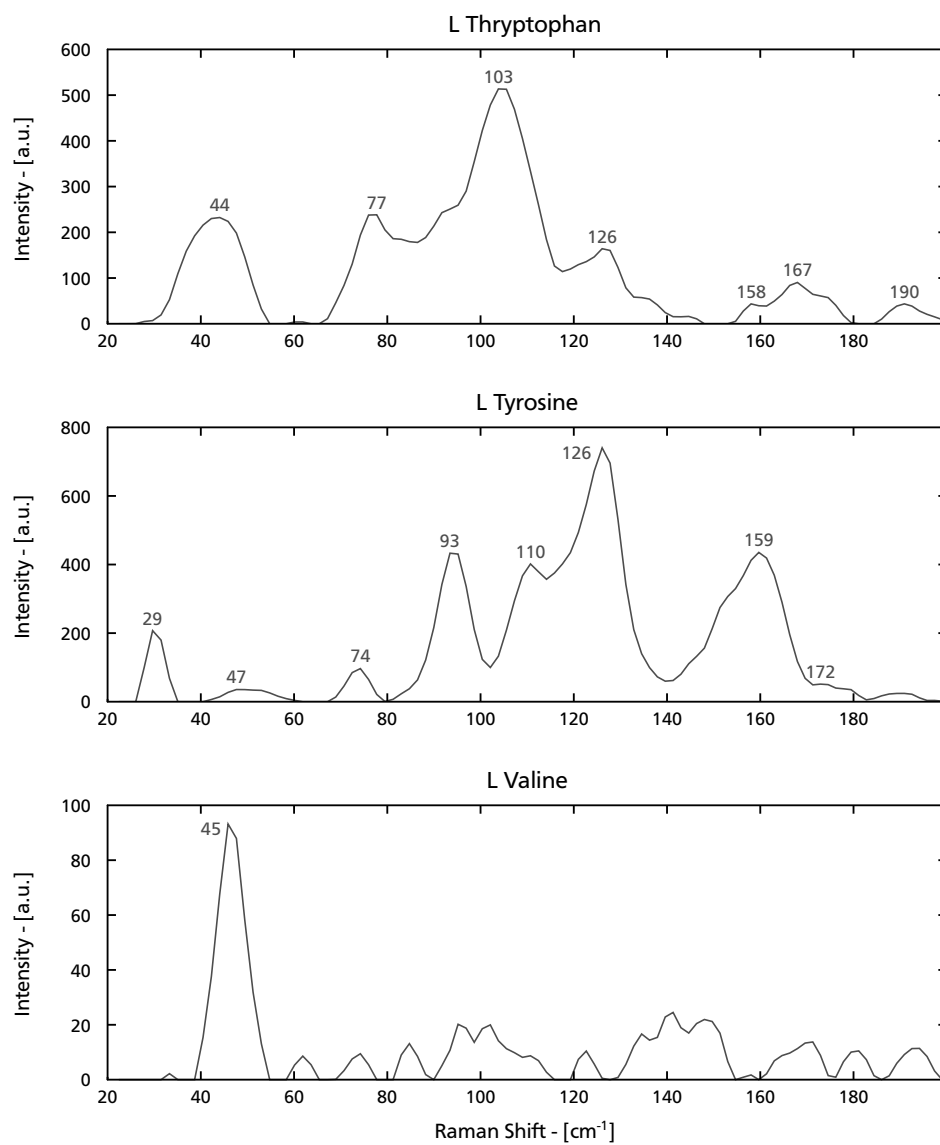
# Appendix B

# Routine Descriptions

DB_BaseTransition – Finds lines with transitions that involves base level. Input parameters:
    Path: Target directory for results.

DB_FileName – Creates a full file name based on the element's atomic number and temperature. When no path is provided, the function returns the element's atomic symbol.
    Element: Atomic number.
    Path: Target directory path.
    Temperature: Plasma temperature.

DB_FPIrwin – Verify scanned values from reference [138], and compares the eighth column against a calculated value with temperature of 16000 °K.
    None: This routine does not use parameters.

DB_FPKurucz – Converts the partition function values from Kurucz database, from text files to Matlab file format (MAT).
    None: This routine does not use parameters.

DB_FPNist – Downloads $g_i$ and $E_i$ values from NIST database, then converts results to Matlab format.
    None: This routine does not use parameters.

DB_Highest – Generates a list with the most intense lines of each element in Atom
    Atom: List with line emission information.
    Path: Target directory.
    Name: Target file name.

DB_Ionization – Extracts NIST database values of ionization potential energy. Results are stored in MAT format.
    None: This routine does not use parameters.

DB_MatchPeaks – Find emission lines based on its wavelength proximity. Returns the WSC value and a list with the index of matched lines.
    Atom: Reference line emission information list.
    Peaks: List of detected maximums in spectrum.
    Tolerance: Maximum distance between lines.

DB_Merge – Creates a list with the spectral information of every element in the source directory whose $\lambda \geq 270$ nm, sorted by wavelength ascending order. It also creates a report with the proportion of lines below a predefined threshold for each element.
    Path: Source directory.

Plot_Environment – Sets the background color and scale figure to occupy the whole screen. Also restricts zoom and pan to horizontal direction.
    Handle: Figure's handle.

Plot_GetColor – Creates a list of colors based on Jet color map.
    Elements: Number of desired colors.

Plot_Move – Moves and resizes the figure.
    Handle: Figure's handle
    X, Y: Lower left figure's position as fraction of full screen.
    Width: Figure width as fraction of full screen.
    Height: Figure height as fraction of full screen.

Plot_Tips – Sets data tips' precision
    handle: Figure's handle.

`Plot_TipsFormat` – Formating callback function for data tips. Sets three decimals for wavelength and four decimals for amplitude.

> `Event_obj:` System generated object.

`Spec_Peaks` – Find stationary peak's points. Returns a list of peak maximums, valleys and a list of peaks with the closest valley or inflection point.

> `X:` Spectrum wavelength vector.
> `Y:` Spectrum amplitude vector.
> `Threshold:` Lowest peak amplitude to consider.
> `Varargin:` Indicates if valleys and inflection point list must be generated.

`Spec_ReadCRV` – Reads spectrum measured with Catalina Instruments spectrometer. Returns both amplitude and wavelength as column vectors, as well as stored header information.

> `FileName:` Spectrum full filename.

`Spec_ReadSPC` – Reads spectrum saved in Thermoscientific file format. Return amplitude and wavelength as column vectors

> `FileName:` Spectrum full filename.

`Spec_ReadSPE` – Reads spectrum saved by both Acton and Princeton Instruments spectrometers. Returns amplitude and wavelength as column vectors, as well as stored header information.

> `FileName:` Spectrum full filename.

`CF_Do` – Calibration Free routine for D2 stainless steel. Returns: 1) a vector with plasma temperature in eV and electron density in $cm^{-3}$; 2) the intersection of Boltzmann plane adjusts; 3) The isolated lines; 4) Lines employed for quantification.

> `Atom:` List with the atomic number of elements.
> `Region:` Spectral region: 1–UV, 2–VIS, 3–IR.
> `SpecFilter:` Vector: a) Minimum wavelength, b) Maximum wavelength, c) Amplitude threshold for element loading.
> `AtomFilter:` Vector: a) Harmonic b) Ionization stage: 0–neutral, 1–first ion, etc. c) Plasma temperature in °K.
> `Lorz:` Lorentzian line information generated by `Spec_Fit` function.
> `varargin:` 1–Plots adjust results and lines used for CF, 2–Plots adjust results and isolated lines.

`CF_PSaha` – Computes Saha's equation employing the number of degenerated states.

> `Atom:` Atomic number
> `Lower:` Lower ionization stage 1–II/I, etc.
> `Temperature:` Plasma temperature in °K

`CF_Saha` – Computes Saha's equation using partition function.

> `Atom:` Atomic number
> `Lower:` Lower ionization stage 1–II/I, etc.
> `Temperature:` Plasma temperature in °K

`DB_Retrieve` – Downloads and stores NIST database in Matlab format.

> `Option:` 1=Download database, 2=File conversion, 3=Statistics extraction, 4=Converts and stores information for partition function.
> `Subdir:` Target directory

`Find` – Elemental search as discussed in chapter 3

> `X, Y:` Wavelength and amplitude column vectors.
> `Elements:` Elements to search, []–all elements.
> `Filter:` Vector: a) Spectral region 1–UV, 2–VIS, 3–IR, b) Ionic state c) Harmonic.
> `Display:` Vector: a) Minimum value of WSC b) Minimum value of $r^2$ c) Display options: 1–Plot, 2–List, 3–Both.
> `Tolerance:` Vector: a) Wavelength tolerance, b) Amplitude threshold for reference, c) Measured spectrum amplitude threshold.
> `varargin:` Plasma temperature in °K. A search by NIST reported amplitude values is done when no temperature is specified.

`Find_Band` – Displays all the emission species near to a specified wavelength and tolerance.

> `Lambda:` Central wavelength.
> `Tolerance:` Wavelength tolerance.

`Find_HighOrder` – Plots a spectrum and a wavelength scaled version.

> `Target:` Cell with instructions for loading and preprocessing a spectrum; see `Spec_Treatment`.
> `Source:` Same as Target, but this spectrum is used to obtain the high order harmonic.
> `Scale:` 2–first harmonic, etc.; 1/2–first low order harmonic, 1/3–second order harmonic etc.

`Find_LoadAtom` – Loads spectral line information from NIST database. Also computes emission coefficients normalized against the highest emission coefficient if a plasma temperature is specified.

    `AtomicNumber:` Element's atomic number.

    `Region:` Spectral region 1–UV, 2–VIS, 3–IR.

    `SpecFilter:` Vector: a) Minimum wavelength, b) Maximum wavelength, c) Amplitude threshold for element loading

    `AtomFilter:` Vector: a) Harmonic, b) Ionization stage c) Plasma temperature.

`Find_Multiple` – Performs elemental search under distinct conditions.

    `Spectra:` Cell with instructions for spectrum load and preprocessing; see `Spec_Treatment`.

    `FindOption:` Cell with parameters for searching. See `Find`.

`FP_Irwin` – Partition function calculation according to reference [138]

    `AtomicNumber:` Element's atomic number; 1–H.

    `Ion:` Ionic state; 0–Neutral, 1–First ion, etc.

    `Temperature:` Plasma temperature in °K.

`FP_Kurucz` – Partition function calculation reported by Kurucz database.

    `AtomicNumber:` Element's atomic number.

    `Ion:` Ionic state; 0–Neutral, 1–First ion, etc.

    `Temperature:` Plasma temperature in °K.

    `Column:` Partition function column [1–7].

`FP_Nist` – Partition function calculation with energy values provided by NIST database.

    `AtomicNumber:` Element's atomic number.

    `Ion:` Ionic state; 0–Neutral, 1–First ion, etc.

    `Temperature:` Plasma temperature in °K.

`PCA_Do` – Computes PCA.

    `varargin:` Can specify a X,Y vector spectrum, or a cell for loading and preprocessing a file. See `Spec_Treatment`.

`PCA_Ellipse` – Superimpose a confidence ellipse over a score or loading plot in 2D or 3D.

    `PCA:` Coordinates of points to evaluate.

    `Confidence:` Confidence range from 0–1.

    `Color:` Color of the ellipse.

`PCA_Load` – Loads in a single step several files.

    `Group:` File group cell generated with `PCA_Source`.

`PCA_PlotData` – Plots in a single graph many spectra.

    `x:` Wavelength column vector.

    `y:` Amplitude column vector.

    `Filter:` File group cell generated with `PCA_Source`.

`PCA_PlotPC` – Plots the Principal Components with highest variance in 2D or 3D.

    `X:` PC corresponding wavelength

    `V:` Principal components.

    `Range:` Elements with high variance to plot; e.g. 2:6.

    `PC:` Columns to include

    `Style:` 1–Plots points, 2–Plots reference lines 3–Plots both.

`PCA_PlotScores` – Score plot in 2D or 3D

    `V:` Scores

    `Group:` File group cell generated with `PCA_Source`.

    `Index:` Scores to plot; e.g. 1:5.

    `varargin:` Individual labels or group labels.

`PCA_Source` – Creates a cell for loading and labeling a group of files.

    `NameFilter:` Cell specifying name filters and paths.

`PCA_WaveToSymbol` – Replace wavelengths labels with its corresponding emission species based synthetic spectrum generation.

    `Rank:` Match statistics returned by `Find` function.

    `Synthetic:` Synthetic reference spectra returned by `Find` function.

    `X:` Spectrum Wavelength vector

    `Marks:` Cell with replacing labels

`Plot_LIBS` – Plots LIBS spectra with a unified style.

    `X:` Spectrum wavelength vector.

    `Y:` Spectrum amplitude vector

    `Style:` Plot style

    `varargin:` 0–Spectrum plot, 1–Spectrum and another spectrum plot, 2–Plot `Find` results.

`Plot_Tile` – Distributes figures in screen.
    `Handles:` Figure's handles
    `varargin:` Column vector–Vertical tile; Row vector–Horizontal tile.

`Plot_WindowName` – Changes figure's title.
    `Handle:` Figure's handle
    `Title:` Title string.

`Spec_CalibAmpl` – Spectrum amplitude calibration based on StellarNet calibration lamp spectrum.
    `X:` Spectrum wavelength vector.
    `Y:` Amplitude wavelength vector.

`Spec_CalibFile` – Generates a calibration polynomial based on the measured spectrum of StellarNet lamp and certified amplitudes.
    `None:` This function does not use parameters.

`Spec_CalibPoly` – Computes calibration polynomials based on a reference spectrum.
    `X:` Spectrum wavelength vector.
    `Y:` Amplitude wavelength vector.
    `Reference:` Reference spectrum.
    `Thres:` Amplitude threshold for line selection.
    `Tolerance:` Wavelength match tolerance.
    `Order:` Order of the polynomial.
    `varargin:` Generates a plot comparing reference and calibrated spectrum.

`Spec_CalibShift` – Calibration by shifting spectra.
    `X:` Spectrum wavelength vector.
    `Y:` Amplitude wavelength vector.
    `Element:` Atomic numbers.
    `Thres:` Amplitude threshold for reference.
    `AtomFilter:` Vector: a) Harmonic, b) Ionization stage c) Plasma temperature.
    `varargin:` Left shifts with any value.

`Spec_Correlate` – Computes correlation between two spectra; this function assumes that both reference and measured spectrum shares the wavelength vector.
    `Reference:` Amplitude wavelength vector.
    `Measured:` Amplitude wavelength vector.
    `Shift:` Number of points to shift.

`Spec_Fit` – Spectrum fit by SHA procedure.
    `X:` Spectrum wavelength vector.
    `Y:` Amplitude wavelength vector.
    `Thres:` Amplitude threshold for line detection.

`Spec_FitLM` – Spectrum fit by LM algorithm.
    `X:` Spectrum wavelength vector.
    `Y:` Amplitude wavelength vector.

`Spec_Glue` – Joins several files in a single spectrum.
    `Prefix:` Fixed portion of full file name.
    `Suffix:` Fixed portion of file name and extension.
    `Format:` Formating string options.
    `Name:` List with name's variable portion.

`Spec_Intersect` – Finds the wavelength portion of spectrum x that falls into a specified range.
    `Left:` Left wavelength range.
    `Right:` Right wavelength range.
    `X:` Spectrum wavelength vector.

`Spec_Read` – Wrapper function for reading `SPC`, `SPE` and `CRV` file formats.
    `Source:` Full file name.

`Spec_Smooth` – Spectrum smooth procedures.
    `Option:` 0–Triangular filter, 1–SNIP filter, 2–Boxcar filter
    `Spectra:` Spectrum amplitude vector.
    `Width:` Window length required by triangular and SNIP filters.

`Spec_Snip` – Background estimation by means of SNIP algorithm
    `Order:` SNIP order, [2, 4, 6 or 8].
    `Spectra:` Spectrum amplitude vector.
    `Iterations:` Window length.
    `Direction:` 1–ascending, 2–descending.

`Spec_Synthetic` – Synthetic spectrum calculation
    `Wavelength:` Spectrum wavelength vector.
    `Lorentzian:` Lorentzian parameters for the emission lines.

`Spec_Treatment` – Wrapper function for loading and preprocessing spectra with a single function. This function can be applied to several spectra at once. This function interprets pairs of commands.

`FilterName, []`: Loads the specified files. See `PCA_Source`.

`X,Y`: Provides a spectrum for processing.

`path, name`: Load the specified spectrum.

`'Trim', [min max]`: Trims the spectrum in the specified range.

`'SnipBase', [order window]`: Baseline removal by SNIP algorithm.

`'LinearBase', offset`: Subtracts offset from amplitude values.

`'ShiftCalib', shift`: Wavelength calibration by shifting. Negative values means left shifting.

`'PolyCalib', coef`: Wavelength calibration with a polynomial.

`'HeightCalib', []`: Amplitude calibration.

`'NoiseReduction', [filter window]`: Filter smooth. See `Spec_Smooth`.

`'Interpolate', X`: Spectrum interpolation to X wavelength vector.

`'Snv', []`: Applies SNV to a spectrum.

# Appendix C

# Isolated Lines for D2 Steel

A full description of the lines depicted in figure 5.12 of section 5.4 is presented. The lines used for quantification are marked in figure 5.14.

**Table C.1**  *Minor concentration elements' lines used in CF algorithm example of section 5.4*

| Wavelength | Element | Ion | $E_i$ | $A_{ij}g_{ij}$ | Peak Area |
|---|---|---|---|---|---|
| 405.0960 | 23 | 1 | 5.1901 | 14.0000 | 0.7107 |
| 409.0580 | 23 | 1 | 4.1107 | 8.5000 | 3.6665 |
| 409.9800 | 23 | 1 | 3.2986 | 3.2800 | 5.8327 |
| 410.9790 | 23 | 1 | 3.2778 | 2.0000 | 6.2027 |
| 411.5180 | 23 | 1 | 3.2986 | 4.6400 | 5.8898 |
| 411.6470 | 23 | 1 | 3.2863 | 1.9000 | 0.3952 |
| 413.4490 | 23 | 1 | 3.2986 | 2.3200 | 4.9594 |
| 433.2820 | 23 | 1 | 2.8777 | 0.2800 | 2.0602 |
| 437.9240 | 23 | 1 | 3.1310 | 13.0000 | 20.5700 |
| 403.3070 | 25 | 1 | 3.0733 | 0.9900 | 13.3050 |
| 403.4490 | 25 | 1 | 3.0722 | 0.6320 | 2.4418 |
| 404.1360 | 25 | 1 | 5.1812 | 7.8700 | 3.3338 |
| 407.9240 | 25 | 1 | 5.2251 | 1.5000 | 3.6689 |
| 408.2940 | 25 | 1 | 5.2140 | 1.7700 | 4.8729 |
| 423.5140 | 25 | 1 | 5.8150 | 5.5000 | 27.2573 |

**Table C.2** *Major concentration elements' lines used in CF algorithm example of section  5.4*

| Wavelength | Element | Ion | $E_i$ | $A_{ij}g_{ij}$ | Peak Area |
|---|---|---|---|---|---|
| 403.9091 | 24 | 1 | 6.9182 | 10.0000 | 2.9224 |
| 416.5515 | 24 | 1 | 7.4247 | 9.8000 | 2.0398 |
| 425.4352 | 24 | 1 | 2.9135 | 2.8400 | 62.8763 |
| 427.4812 | 24 | 1 | 2.8995 | 2.1500 | 77.1425 |
| 428.9731 | 24 | 1 | 2.8895 | 1.5800 | 46.8172 |
| 433.7557 | 24 | 1 | 3.8260 | 0.2740 | 4.7080 |
| 435.9625 | 24 | 1 | 3.8260 | 0.2700 | 7.3470 |
| 436.3129 | 24 | 1 | 5.8083 | 1.1000 | 0.7071 |
| 437.1275 | 24 | 1 | 3.8392 | 0.2900 | 4.7010 |
| 437.5330 | 24 | 1 | 5.8155 | 0.6500 | 2.4746 |
| 400.5242 | 26 | 1 | 4.6520 | 1.0200 | 34.6736 |
| 401.4531 | 26 | 1 | 6.1344 | 1.0700 | 8.8337 |
| 402.1866 | 26 | 1 | 5.8405 | 0.7700 | 15.4683 |
| 404.5812 | 26 | 1 | 4.5485 | 7.7600 | 113.8102 |
| 406.7978 | 26 | 1 | 6.2581 | 1.3600 | 17.9889 |
| 407.1738 | 26 | 1 | 4.6520 | 3.8200 | 48.1299 |
| 407.6629 | 26 | 1 | 6.2517 | 1.1900 | 7.4041 |
| 408.4491 | 26 | 1 | 6.3666 | 0.7790 | 3.5757 |
| 410.7488 | 26 | 1 | 5.8492 | 0.5220 | 8.9813 |
| 413.6997 | 26 | 1 | 6.4110 | 1.3800 | 4.9894 |
| 415.6798 | 26 | 1 | 5.8134 | 0.6000 | 5.9803 |
| 417.5636 | 26 | 1 | 5.8134 | 0.5700 | 8.0757 |
| 417.8855 | 26 | 2 | 5.5488 | 0.0140 | 2.6683 |
| 418.1754 | 26 | 1 | 5.7956 | 1.6200 | 15.5033 |
| 418.4892 | 26 | 1 | 5.7934 | 0.5150 | 5.6199 |
| 419.1430 | 26 | 1 | 5.4260 | 0.8190 | 21.5078 |
| 419.6208 | 26 | 1 | 6.3503 | 0.7630 | 1.8949 |
| 421.0343 | 26 | 1 | 5.4260 | 0.4440 | 12.0151 |
| 421.9360 | 26 | 1 | 6.5109 | 3.7400 | 11.8605 |
| 422.2213 | 26 | 1 | 5.3852 | 0.4030 | 11.5064 |
| 422.4171 | 26 | 1 | 6.3025 | 1.1700 | 4.6039 |
| 422.7426 | 26 | 1 | 6.2640 | 6.8800 | 32.0625 |
| 423.3603 | 26 | 1 | 5.4099 | 0.9250 | 11.8286 |
| 423.8810 | 26 | 1 | 6.3207 | 2.1700 | 8.9416 |
| 424.7425 | 26 | 1 | 6.2865 | 2.1300 | 12.2343 |
| 425.0787 | 26 | 1 | 4.4733 | 0.7140 | 36.6539 |
| 426.0474 | 26 | 1 | 5.3085 | 4.3900 | 56.2261 |
| 429.9235 | 26 | 1 | 5.3085 | 1.4200 | 31.5379 |
| 431.5084 | 26 | 1 | 5.0703 | 0.3880 | 14.2838 |
| 432.5762 | 26 | 1 | 4.4733 | 3.6100 | 53.5040 |
| 438.3545 | 26 | 1 | 4.3125 | 5.5000 | 96.5552 |

# References

[1] D. Harvey. *Modern Analytical Chemistry*. McGraw-Hill, 1999.

[2] Shiv K. Sharma, Anupam K. Misra, Paul G. Lucey, and Rachel C.F. Lentz. A combined remote raman and {LIBS} instrument for characterizing minerals with 532 nm laser excitation. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 73(3):468 – 476, 2009. Raman Spectroscopy Applied to the Earth Sciences - Sensu LatuSelected papers from GeoRaman 2008, 8th International Conference on Raman Spectroscopy Applied to the Earth Sciences - Sensu Latu, June 2008 Ghent, Belgium.

[3] Waldemar Hübert and Georg Ankerhold. Elemental misinterpretation in automated analysis of libs spectra. *Analytical and Bioanalytical Chemistry*, 400(10):3273–3278, 2011.

[4] M. Saritha and V.P.N. Nampoori. Identification of spectral lines of elements using artificial neural networks. *Microchemical Journal*, 91(2):170 – 175, 2009.

[5] Jeremy J. Hatch, Timothy R. McJunkin, Cynthia Hanson, and Jill R. Scott. Automated interpretation of libs spectra using a fuzzy logic inference engine. *Applied Optics*, 51(7):B155–B164, Mar 2012.

[6] J. Brian Loudermilk, David S. Himmelsbach, Franklin E. Barton, and James A. de Haseth. Novel search algorithms for a mid-infrared spectral library of cotton contaminants. *Applied Spectroscopy*, 62(6):661–670, Jun 2008.

[7] Stephen E. Stein and Donald R. Scott. Optimization and testing of mass spectral library search algorithms for compound identification. *Journal of the American Society for Mass Spectrometry*, 5(9):859–866, 1994.

[8] K. Baumann and J.T. Clerc. Computer-assisted {IR} spectra prediction — linked similarity searches for structures and spectra. *Analytica Chimica Acta*, 348(1–3):327 – 343, 1997.

[9] M.P. Mateo, G. Nicolás, V. Piñón, J.C. Alvarez, A. Ramil, and A. Yáñez. Versatile software for semiautomatic analysis and processing of laser-induced plasma spectra. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 60(7–8):1202–1210, 2005.

[10] G. Amato, G. Cristoforetti, S. Legnaioli, G. Lorenzetti, V. Palleschi, F. Sorrentino, and E. Tognoni. Progress towards an unassisted element identification from laser induced breakdown spectra with automatic ranking techniques inspired by text retrieval. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 65(8):664–670, 2010.

[11] I. Rosas-Román, M. Meneses-Nava, O. Barbosa-García, and Jose L. Maldonado. Semi-automatic elemental identification of laser-induced breakdown spectra using wavelength similarity coefficient. *Applied Spectroscopy*, 71(4):627–633, 2017.

[12] A Jurado-López and M.D Luque de Castro. Rank correlation of laser-induced breakdown spectroscopic data for the identification of alloys used in jewelry manufacture. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 58(7):1291 – 1299, 2003.

[13] A. Koujelev, M. Sabsabi, V. Motto-Ros, S. Laville, and S.L. Lui. Laser-induced breakdown spectroscopy with artificial neural network processing for material identification. *Planetary and Space Science*, 58(4):682 – 690, 2010. Exploring other worlds by exploring our own: The role of terrestrial analogue studies in planetary exploration.

[14] E. M. Rodriguez-Celis, I. B. Gornushkin, U. M. Heitmann, J. R. Almirall, B. W. Smith, J. D. Winefordner, and N. Omenetto. Laser induced breakdown spectroscopy as a tool for discrimination of glass for forensic applications. *Analytical and Bioanalytical Chemistry*, 391(5):1961, 2008.

[15] Jr. Frank C. De Lucia, Jennifer L. Gottfried, Chase A. Munson, and Andrzej W. Miziolek. Multivariate analysis of standoff laser-induced breakdown spectroscopy spectra for classification of explosive-containing residues. *Appl. Opt.*, 47(31):G112–G121, Nov 2008.

[16] Chase A. Munson, Frank C. De Lucia, Thuvan Piehler, Kevin L. McNesby, and Andrzej W. Miziolek. Investigation of statistics strategies for improving the discriminating power of laser-induced breakdown spectroscopy for chemical and biological warfare agent simulants. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 60(7):1217 – 1224, 2005. Laser Induced Plasma Spectroscopy and Applications (LIBS 2004) Third International Conference.

[17] U. Contreras, M.A. Meneses-Nava, D. Torres-Armenta, J. Robles-Camacho, and O. Barbosa-García. Identificación de aceros por espectroscopia de rompimiento inducido por láser(libs) y análisis de componentes principales. *Revista Mexicana de Física*, 58:184–190, 2012.

[18] E. Grifoni, S. Legnaioli, G. Lorenzetti, S. Pagnotta, F. Poggialini, and V. Palleschi. From calibration-free to fundamental parameters analysis: A comparison of three recently proposed approaches. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 124:40 – 46, 2016.

[19] J.M. Gomba, C. D'Angelo, D. Bertuccelli, and G. Bertuccelli. Spectroscopic characterization of laser induced breakdown in aluminium–lithium alloy samples for quantitative determination of traces. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 56(6):695 – 705, 2001. LIBS2000, 1st International Conference on Laser Induced Plasma Spectroscopy and Applications.

[20] Pavel Yaroshchyk, Doug Body, Richard J.S. Morrison, and Bruce L. Chadwick. A semi-quantitative standard-less analysis method for laser-induced breakdown spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 61(2):200 – 209, 2006.

[21] Eleonora D'Andrea, Stefano Pagnotta, Emanuela Grifoni, Stefano Legnaioli, Giulia Lorenzetti, Vincenzo Palleschi, and Beatrice Lazzerini. A hybrid calibration-free/artificial neural networks approach to the quantitative analysis of libs spectra. *Applied Physics B*, 118(3):353–360, 2015.

[22] G.H. Cavalcanti, D.V. Teixeira, S. Legnaioli, G. Lorenzetti, L. Pardini, and V. Palleschi. One-point calibration for calibration-free laser-induced breakdown spectroscopy quantitative analysis. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 87:51 – 56, 2013. Thematic Issue: 7th International Conference on Laser Induced Breakdown Spectroscopy (LIBS 2012), Luxor, Egypt, 29 September-4 October 2012.

[23] C. Aragón and J.A. Aguilera. Csigma graphs: A new approach for plasma characterization in laser-induced breakdown spectroscopy. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 149:90 – 102, 2014.

[24] F. Bredice, F. O. Borges, H. O. Di Rocco, R. S. Mercado, M. Villagrán-Muniz, and V. Palleschi. A procedure for estimating the electron temperature and the departure of the lte condition in a time-dependent, spatially homogeneous, optically thin plasma. *Brazilian Journal of Physics*, 43(4):239–246, 2013.

[25] Ümit Aydin, Peter Roth, Christoph Dominic Gehlen, and Reinhard Noll. Spectral line selection for time-resolved investigations of laser-induced plasmas by an iterative boltzmann plot method. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 63(10):1060 – 1065, 2008. A collection of papers presented at the Euro Mediterranean Symposium on Laser Induced Breakdown Spectroscopy (EMSLIBS 2007).

[26] Lanxiang Sun and Haibin Yu. Correction of self-absorption effect in calibration-free laser-induced breakdown spectroscopy by an internal reference method. *Talanta*, 79(2):388 – 395, 2009.

[27] Jia-Ming Li, Lian-Bo Guo, Chang-Mao Li, Nan Zhao, Xin-Yan Yang, Zhong-Qi Hao, Xiang-You Li, Xiao-Yan Zeng, and Yong-Feng Lu. Self-absorption reduction in laser-induced breakdown spectroscopy using laser-stimulated absorption. *Opt. Lett.*, 40(22):5224–5226, Nov 2015.

[28] J.A. Aguilera, C. Aragón, G. Cristoforetti, and E. Tognoni. Application of calibration-free laser-induced breakdown spectroscopy to radially resolved spectra from a copper-based alloy laser-induced plasma. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 64(7):685 – 689, 2009.

[29] I. Rosas-Román, M.A. Meneses-Nava, O. Barbosa-García, and J.L. Maldonado. Simultaneous height adjust fitting: An alternative automated fitting procedure for laser-induced plasma spectra composed by multiple lorentzian profiles. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 134:1 – 5, 2017.

[30] A. Giakoumaki, I. Osticioli, and D. Anglos. Spectroscopic analysis using a hybrid libs-raman system. *Applied Physics A*, 83(4):537–541, 2006.

[31] I. Osticioli, N.F.C. Mendes, A. Nevin, Francisco P.S.C. Gil, M. Becucci, and E. Castellucci. Analysis of natural and artificial ultramarine blue pigments using laser induced breakdown and pulsed raman spectroscopy, statistical analysis and light microscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 73(3):525 – 531, 2009. Raman Spectroscopy Applied to the Earth Sciences - Sensu LatuSelected papers from GeoRaman 2008, 8th International Conference on Raman Spectroscopy Applied to the Earth Sciences - Sensu Latu, June 2008 Ghent, Belgium.

[32] F. Matroodi and S. H. Tavassoli. Simultaneous raman and laser-induced breakdown spectroscopy by a single setup. *Applied Physics B*, 117(4):1081–1089, 2014.

[33] Romain Bruder, Vincent Detalle, and Claude Coupry. An example of the complementarity of laser-induced breakdown spectroscopy and raman microscopy for wall painting pigments analysis. *Journal of Raman Spectroscopy*, 38(7):909–915, 2007.

[34] R.C. Wiens, S.K. Sharma, J. Thompson, A. Misra, and P.G. Lucey. Joint analyses by laser-induced breakdown spectroscopy (libs) and raman spectroscopy at stand-off distances. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 61(10):2324–2334, 2005. cited By 70.

[35] Marek Hoehse, Igor Gornushkin, Sven Merk, and Ulrich Panne. Assessment of suitability of diode pumped solid state lasers for laser induced breakdown and raman spectroscopy. *J. Anal. At. Spectrom.*, 26:414–424, 2011.

[36] Qingyu Lin, Guanghui Niu, Qihui Wang, Qiaoling Yu, and Yixiang Duan. Combined Laser-Induced Breakdown with Raman Spectroscopy: Historical Technology Development and Recent Applications. *Applied Spectroscopy Reviews*, 48(6):487–508, 2013.

[37] M.J. Pelletier. Raman spectroscopy using an echelle spectrograph with ccd detection. *Applied Spectroscopy*, 44(10):1699–1705, 1990. cited By 0.

[38] Marek Hoehse, David Mory, Stefan Florek, Friederike Weritz, Igor Gornushkin, and Ulrich Panne. A combined laser-induced breakdown and raman spectroscopy echelle system for elemental and molecular microanalysis. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 64(11–12):1219 – 1227, 2009.

[39] Randy Heyler, James Carriere, and Brian Smith. Polymorph identification and analysis using ultralow-frequency raman spectroscopy. *Spectroscopy*, 28(6), 2013.

[40] David Tuschel. Why are the raman spectra of crystalline and amorphous solids different? *Spectroscopy*, 32(3):26–33, 2017.

[41] Geoffrey P.S. Smith, Gregory S. Huff, and Keith C. Gordon. Investigating crystallinity using low frequency raman spectroscopy: Applications in pharmaceutical analysis. *Spectroscopy*, 31(2):42–50, 2016.

[42] J.M. Hollas. *Modern Spectroscopy*. Wiley, 2004.

[43] D.W. Ball. *Field Guide to Spectroscopy*. Field Guide Series. International Society for Optics and Photonics, 2006.

[44] D.L. Andrews and A.A. Demidov. *An Introduction to Laser Spectroscopy: Second Edition*. Springer US, 2012.

[45] F. Rouessac and A. Rouessac. *Chemical Analysis: Modern Instrumentation Methods and Techniques*. Wiley, 2007.

[46] A.W. Miziolek, V. Palleschi, and I. Schechter. *Laser Induced Breakdown Spectroscopy*. Cambridge University Press, 2006.

[47] Blair Thornton, Tomoko Takahashi, Takumi Sato, Tetsuo Sakka, Ayaka Tamura, Ayumu Matsumoto, Tatsuo Nozaki, Toshihiko Ohki, and Koichi Ohki". Development of a deep-sea laser-induced breakdown spectrometer for in situ multi-element chemical analysis. *Deep Sea Research Part I: Oceanographic Research Papers*, 95:20 – 36, 2015.

[48] Nina L. Lanza, Roger C. Wiens, Samuel M. Clegg, Ann M. Ollila, Seth D. Humphries, Horton E. Newsom, and James E. Barefield. Calibrating the chemcam laser-induced breakdown spectroscopy instrument for carbonate minerals on mars. *Appl. Opt.*, 49(13):C211–C217, May 2010.

[49] J.P. Singh and S.N. Thakur. *Laser-Induced Breakdown Spectroscopy*. Elsevier Science, 2007.

[50] R. Noll. *Laser-Induced Breakdown Spectroscopy: Fundamentals and Applications*. Springer Berlin Heidelberg, 2012.

[51] B. Di Bartolo and O. Forte. *Advances in Spectroscopy for Lasers and Sensing*. Nato Science Series II:. Springer Netherlands, 2006.

[52] W. Demtröder. *Laser Spectroscopy: Basic Concepts and Instrumentation*. Advanced Texts in Physics. Springer Berlin Heidelberg, 2002.

[53] R.W. Solarz and J.A. Paisner. *Laser Spectroscopy and its Applications*. Optical Science and Engineering. Taylor & Francis, 1986.

[54] J.R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer US, 2007.

[55] *Synthetic Fused Silica*. [Technical Note]. Melles Griot, Rochester, New York 14620.

[56] *Optics: How to Build a Beam Expander*. [Brochure]. Newport Corporation, Irvine, California 92606.

[57] Princeton Instruments. *WinSpec/32*.

[58] Princeton Instruments. *Princeton Instruments SPE 3.0 File Format Specification*.

[59] *Format of the CRV Files used by KestrelSpec*. [Document]. Catalina Scientific, Tucson, Arizona 85705.

[60] *A Brief Guide to SPC File Format and Using GSPCIO*. [Document]. Thermo Galactic, Salem, New Hampshire 03079.

[61] *USB Shutter*. [Brochure]. Picard Industries, Albion, New York 14411.

[62] Zaber. *T-Series positioning products user's manual*.

[63] Charles Petzold. *Programming Windows 95*. Microsoft programming series. Microsoft Press, 1996.

[64] Joe Campbell. *C Programmer's Guide to Serial Communications*. Sams Publishing., 1987.

[65] Allen Denver. Serial communications in win32. [Online]. Available: http://msdn.microsoft.com/en-us/library/ms810467.aspx [1995, December 11].

[66] M. J. Pelletier. Quantitative analysis using raman spectrometry. *Appl. Spectrosc.*, 57(1):20A–42A, Jan 2003.

[67] S.W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Pub., 1997.

[68] Miroslav Morhác and Vladislav Matoušek. Peak clipping algorithms for background estimation in spectroscopic data. *Applied Spectroscopy*, 62(1):91–106, 2008.

[69] J. Gruber, J. Heitz, H. Strasser, D. Bäuerle, and N. Ramaseder. Rapid in-situ analysis of liquid steel by laser-induced breakdown spectroscopy. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 56(6):685 – 693, 2001. LIBS2000, 1st International Conference on Laser Induced Plasma Spectroscopy and Applications.

[70] Morihisa Saeki, Akio Iwanade, Chikara Ito, Ikuo Wakaida, Blair Thornton, Tetsuo Sakka, and Hironori Ohba. Development of a fiber-coupled laser-induced breakdown spectroscopy instrument for analysis of underwater debris in a nuclear reactor core. *Journal of Nuclear Science and Technology*, 51(7-8):930–938, 2014.

[71] E. Tognoni, G. Cristoforetti, S. Legnaioli, and V. Palleschi. Calibration-free laser-induced breakdown spectroscopy: State of the art. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 65(1):1 – 14, 2010.

[72] I. B. Gornushkin, M. Mueller, U. Panne, and J. D. Winefordner. Insights into linear and rank correlation for material identification in laser-induced breakdown spectroscopy and other spectral techniques. *Applied Spectroscopy*, 62(5):542–553, May 2008.

[73] I.B. Gornushkin, U. Panne, and J.D. Winefordner. Linear correlation for identification of materials by laser induced breakdown spectroscopy: Improvement via spectral filtering and masking. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 64(10):1040 – 1047, 2009.

[74] Peter R. Griffiths and Limin Shao. Self-weighted correlation coefficients and their application to measure spectral similarity. *Applied Spectroscopy*, 63(8):916–919, Aug 2009.

[75] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.2), [Online]. Available: http://physics.nist.gov/asd [2015, September 15]. National Institute of Standards and Technology, Gaithersburg, MD., 2014.

[76] R. Kurucz. Kurucz Database, [Online]. Available: http://kurucz.harvard.edu/ [2016, October 15]. Harvard-Smithsonian Center for Astrophysics, Cambridge, MA., 2016.

[77] J. E. Sansonetti and W. C. Martin. Handbook of basic atomic spectroscopic data. *Journal of Physical and Chemical Reference Data*, 34(4):1559–2259, 2005.

[78] Alex N. James. NIST Atomic Spectra Database import. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/34265-nist-atomic-spectra-database-import [2016, July 1]. Mathworks File Exchange.

[79] Phyton 2.7 programing language. [Online]. Available: https://www.python.org/download/releases/2.7/ [2016, July 1].

[80] Thomas Heller. Phyton extension for conversion of scripts into executable windows programs. [Online]. Available: http://www.py2exe.org/ [2016, July 1].

[81] The MathWorks Inc., 3 Apple Hill Drive, Natick, MA 01760-2098. *MATLAB Data Import and Export*, April 2011.

[82] A. Ciucci, M. Corsi, V. Palleschi, S. Rastelli, A. Salvetti, and E. Tognoni. New procedure for quantitative elemental analysis by laser-induced plasma spectroscopy. *Appl. Spectrosc.*, 53(8):960–964, Aug 1999.

[83] H.R. Griem. *Principles of Plasma Spectroscopy*. Cambridge Monographs on Plasma Physics. Cambridge University Press, 2005.

[84] M. Weeks. *Digital Signal Processing Using MATLAB and Wavelets*. Electrical engineering series. Infinity Science Press, 2007.

[85] E. Tzamali and D. Anglos. A parametric linear correlation method for the analysis of libs spectral data. In Johann Nimmrichter, Wolfgang Kautek, and Manfred Schreiner, editors, *Lasers in the Conservation of Artworks*, volume 116 of *Springer proceedings in physics*, pages 377–382. Springer Berlin Heidelberg, 2007.

[86] D.E. Knuth. *Art of Computer Programming, Volume 2: Seminumerical Algorithms, The*. Pearson Education, 2014.

[87] W.H. Press. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[88] G. Cristoforetti, A. De Giacomo, M. Dell'Aglio, S. Legnaioli, E. Tognoni, V. Palleschi, and N. Omenetto. Local thermodynamic equilibrium in laser-induced breakdown spectroscopy: Beyond the mcwhirter criterion. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 65(1):86 – 95, 2010.

[89] Nek M Shaikh, B Rashid, S Hafeez, Y Jamil, and M A Baig. Measurement of electron density and temperature of a laser-induced zinc plasma. *Journal of Physics D: Applied Physics*, 39(7):1384, 2006.

[90] Mohamed A. Khater. Laser-induced breakdown spectroscopy for light elements detection in steel: State of the art. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 81:1 – 10, 2013.

[91] J.F. Hair, W.C. Black, B.J. Babin, and R.E. Anderson. *Multivariate Data Analysis*. Always learning. Pearson Education Limited, 2013.

[92] Victor Ulises Contreras, Marco A. Meneses-Nava, Nancy Ornelas-Soto, Oracio Barbosa-García, Pedro L. López de Alba, José L. Maldonado, Gabriel Ramos-Ortiz, and Francisco J. Acevedo-Aguilarand Leticia López-Martínez. Fast and environmentally friendly quantitative analysis of active agents in anti-diabetic tablets by an alternative laser-induced breakdown spectroscopy (libs) method and comparison to a validated reversed-phase high-performance liquid chromatography (rp-hplc) method. *Applied Spectroscopy*, 66:1294–1301, 2012.

[93] J.E. Jackson. *A User's Guide to Principal Components*. Wiley Series in Probability and Statistics. Wiley, 2005.

[94] C. Croarkin and P. Tobias, NIST/SEMATECH e-Handbook of Statistical Methods, [Online]. Available: http://www.itl.nist.gov/div898/handbook/ [2017, February 9].

[95] Dan Kalman. A singularly valuable decomposition: The svd of a matrix. *The College Mathematics Journal*, 27(1):2–23, 1996.

[96] E. Johansson-J.Trygg C. Vikström L. Eriksson, T. Byrne. *Multi- and Megavariate Data Analysis, Part 1, Basic Principles and Applications*. Number p. 1. Umetrics Academy, 2006.

[97] S.I. Grossman. *Elementary Linear Algebra*. Saunders College Pub., 1994.

[98] R.J. Barnes, M.S. Dhanoa, and S.J. Lister. Standard normal variate transformation and detrending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.*, 43(5):772–777, May 1989.

[99] A.J. Johnson. Error Ellipse. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/4705-error-ellipse [2016, July 1]. Mathworks File Exchange.

[100] James Jones. Math 170 - Introduction to Statistics. [Online]. Available: https://people.richland.edu/james/lecture/m170/ [2016, July 1]. Richland College.

[101] Google Earth. [2016, July 1]. Google.

[102] Juan Morales, María del Sol Hernández-Bernal, Avto Goguitchaichvili, and José Luis Punzo-Díaz. An integrated magnetic, geochemical and archeointensity investigation of casting debris from ancient metallurgical sites of michoacán, western mesoamerica. *Studia Geophysica et Geodaetica*, pages 1–20, 2017.

[103] Robin Gill. *Igneous Rocks and Processes*. John Wiley & Sons Ltd, West Sussex, UK, 2010.

[104] Alfonso Martinez Vera, Edgar Rodriguez Medina, Jaime Núñez Espinal, Jorge E. Bustamante García, Jose E. Montiel Escobar, Mario Carranza Alvarado, , Javier López Avila, Alba Elena Perez Rodriguez, Carlos Francisco Yañez Mondragón, Miguel Angel Rivas Ayala, and Victor Manuel Enríquez Borja. *Geological-Mining Monograph of the State of Michoacan*. Consejo de Recursos Minerales, Pachuca, México, 1995.

[105] Panorama Minero del Estado de Michoacán. [Online]. Available: www.sgm.gob.mx/pdfs/MICHOACAN.pdf [December 2015]. Servicio Geológico Mexicano.

[106] Mineral Database. [Online]. Available: www.mindat.org/ [December 2015]. Hudson Institute of Mineralogy.

[107] Dorothy Hosler. Sound, color and meaning in the metallurgy of ancient west mexico. *World Archaeology*, 27(1):100–115, 1995.

[108] World Steel Association, Rue Colonel Bourg 120 - B-1140 Brussels - Belgium. *Steel Statistical Year Book 2016*, October 2016.

[109] H. Bhadeshia and R. Honeycombe. *Steels: Microstructure and Properties*. Elsevier Science, 2011.

[110] Ronald Lesley Plaut, Clara Herrera, Doris Maribel Escriba, Paulo Rangel Rios, and Angelo Fernando Padilha. A short review on wrought austenitic stainless steels at high temperatures: processing, microstructure, properties and performance. *Materials Research ibero-american Journal of Materials*, 10(4):453–460, 2007.

[111] Outokumpu, SE-774 22 Avesta, Sweden. *Handbook of Stainless Steel*, October 2013.

[112] Complete Stainless Steel table. [Online]. Available: http://industeel.arcelormittal.com/products/stainless-steels/ [February 2016]. Industeel - Arcelor Mittal.

[113] UR™ 2304 Data Sheet. [February 2016]. Industeel - Arcelor Mittal.

[114] UR™ 926 Data Sheet. [February 2016]. Industeel - Arcelor Mittal.

[115] UR™ 2507 - UR™ 2507Cu - UR™ 2507W Data Sheet. [February 2016]. Industeel - Arcelor Mittal.

[116] David Bergstrom. Characterization of ATI 2102 Lean Duplex Stainless Steel Pipe used for Water Distribution. [Online]. Available: https://www.atimetals.com/markets/oilandgas/Documents/ATI2102LeanDuplexforWater11.2011.pdf [December 2016]. ATI.

[117] David W. Hahn and Nicoló Omenetto. Laser-induced breakdown spectroscopy (libs), part ii: Review of instrumental and methodological approaches to material analysis and applications to different fields. *Appl. Spectrosc.*, 66(4):347–419, Apr 2012.

[118] Kathleen K. Herrera, Elisabetta Tognoni, Nicolo Omenetto, Benjamin W. Smith, and James D. Winefordner. Semi-quantitative analysis of metal alloys, brass and soil samples by calibration-free laser-induced breakdown spectroscopy: recent results and considerations. *J. Anal. At. Spectrom.*, 24:413–425, 2009.

[119] A. Fridman and L.A. Kennedy. *Plasma Physics and Engineering*. Taylor & Francis, 2004.

[120] M.I. Boulos, P. Fauchais, and E. Pfender. *Thermal Plasmas: Fundamentals and Applications*. Springer US, 2013.

[121] C. A. Bye and A. Scheeline. Saha-boltzmann statistics for determination of electron temperature and density in spark discharges using an echelle/ccd system. *Applied Spectroscopy*, 47:2022–2030, dec 1993.

[122] Wing Tat Chan and Richard E. Russo. Study of laser-material interactions using inductively coupled plasma-atomic emission spectrometry. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 46(11):1471 – 1486, 1991.

[123] Oleg V Borisov, Xianglei Mao, and Richard E Russo. Effects of crater development on fractionation and signal intensity during laser ablation inductively coupled plasma mass spectrometry. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 55(11):1693 – 1704, 2000.

[124] D.A. Cremers and L.J. Radziemski. *Handbook of Laser-Induced Breakdown Spectroscopy*. Wiley, 2006.

[125] K. S. Subrahmanyam, S. R. C. Vivekchand, A. Govindaraj, and C. N. R. Rao. A study of graphenes prepared by different methods: characterization, properties and solubilization. *J. Mater. Chem.*, 18:1517–1523, 2008.

[126] R. Haddad, L. Carmel, and D. Harel. A feature extraction algorithm for multi-peak signals in electronic noises. *Sensors and Actuators B: Chemical*, 120(2):467 – 472, 2007.

[127] Moritz Zaiß, Benjamin Schmitt, and Peter Bachert. Quantitative separation of {CEST} effect from magnetization transfer and spillover effects by lorentzian-line-fit analysis of z-spectra. *Journal of Magnetic Resonance*, 211(2):149 – 155, 2011.

[128] Susan E. Shadle, Douglas F. Allen, Hong Guo, Wendy K. Pogozelski, John S. Bashkin, and Thomas D. Tullius. Quantitative analysis of electrophoresis data: novel curve fitting methodology and its application to the determination of a protein—dna binding constant. *Nucleic Acids Research*, 25(4):850–860, 1997.

[129] Tomoko Takahashi, Blair Thornton, Koichi Ohki, and Tetsuo Sakka. Calibration-free analysis of immersed brass alloys using long-ns-duration pulse laser-induced breakdown spectroscopy with and without correction for non-stoichiometric ablation. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 111:8 – 14, 2015.

[130] Mustafa Karakaplan. Fitting lorentzian peaks with evolutionary genetic algorithm based on stochastic search procedure. *Analytica Chimica Acta*, 587(2):235 – 239, 2007.

[131] Jack G. Dodd and L.K. DeNoyer. *Curve-Fitting: Modeling Spectra*. John Wiley & Sons, Ltd, 2006.

[132] Steven H. Schot. Aberrancy: Geometry of the third derivative. *Mathematics Magazine*, 51(5):259–275, 1978.

[133] H. Guo. A simple algorithm for fitting a gaussian function [dsp tips and tricks]. *Signal Processing Magazine, IEEE*, 28(5):134–137, Sep 2011.

[134] R. De Levie. *Advanced Excel for Scientific Data Analysis*. Advanced Excel for Scientific Data Analysis. Oxford University Press, 2004.

[135] Di Bartolo and B. Ottavio Forte. *Frontiers Developments in Optics and Spectroscopy*. Ottavio Forte, 2007.

[136] Xianglei Huang and Yuk Ling Yung. A common misunderstanding about the voigt line profile. *Journal of the Atmospheric Sciences*, 61(13):1630–1632, 2004.

[137] M. Capitelli, G. Colonna, and A. D'Angola. *Fundamental Aspects of Plasma Chemical Physics: Thermodynamics*. Springer Series on Atomic,

Optical, and Plasma Physics. Springer New York, 2011.

[138] A. W Irwin. Polynomial partition function approximations of 344 atomic and molecular species. *Astrophysical Journal Supplement Series*, 45:621–633, April 1981.

[139] Giuliano D'Ammando, Gianpiero Colonna, and Mario Capitelli. A simplified approach to calculate atomic partition functions in plasmas. *Physics of Plasmas*, 20(3):032108, 2013.

[140] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team. NIST Atomic Spectra Database Levels Form, [Online]. Available: http://physics.nist.gov/PhysRefData/ASD/levels_form.html [2015, September 15]. National Institute of Standards and Technology, Gaithersburg, MD., 2014.

[141] Lanxiang Sun and Haibin Yu. Correction of self-absorption effect in calibration-free laser-induced breakdown spectroscopy by an internal reference method. *Talanta*, 79(2):388 – 395, 2009.

[142] Cold work tool steel – AISI D2, [Online]. Available: http://stainless.steelandtube.co.nz/wp-content/uploads/2014/06/ColdWorkToolSteel-D2.pdf [2016, September 15]. Steel & Tube, Auckland, New Zealand., 2016.

[143] G. Lopez-Reyes, R. Torío, C. Díaz, A. Sansano, and F. Rull. *RLS Instrument Algorithms and Tools for the Detection and Analysis of Traces of Astrobiological Interest*. Astrobiology Science Conference 2015 (Chicago, Illinois, USA, June 2015), 2015.

[144] P.J. Gasda, T.E. Acosta-Maeda, P. Lucey, A.K. Misra, S.K. Sharma, and G.J. Taylor. *A Compact Laser Induced Breakdown, Raman, and Fluorescence Spectroscopy Instrument for Mars Exploration*. 45th Lunar and Planetary Science Conference, 2014.

[145] Javier Moros, Juan Antonio Lorenzo, Patricia Lucena, Luciano Miguel Tobaria, and José Javier Laserna. *Simultaneous Raman Spectroscopy–Laser-Induced Breakdown Spectroscopy for Instant Standoff Analysis of Explosives Using a Mobile Integrated Sensor Platform*. *Analytical Chemistry*, 82(4):1389–1400, 2010.

[146] D.N. Sathyanarayana. *Vibrational Spectroscopy: Theory and Applications*. New Age International (P) Limited, 2015.

[147] P. H. Tan, D. Bougeard, G. Abstreiter, and K. Brunner. Raman scattering of folded acoustic phonons in self-assembled si/ge dot superlattices. *Applied Physics Letters*, 84(14):2632–2634, 2004.

[148] C. Fantini, A. Jorio, M. Souza, M. S. Strano, M. S. Dresselhaus, and M. A. Pimenta. Optical transition energies for carbon nanotubes from resonant raman spectroscopy: Environment and temperature effects. *Phys. Rev. Lett.*, 93:147406, Sep 2004.

[149] R.L. McCreery. *Raman Spectroscopy for Chemical Analysis*. Chemical Analysis: A Series of Monographs on Analytical Chemistry and Its Applications. Wiley, 2005.

[150] Near-Excitation Tuneable filter, low-frequency Raman spectroscopy, [Online]. Available: http://resources.renishaw.com/en/details/near-excitation-tuneable-filter-next--30950 [2016, September 15]. Renishaw. Gloucestershire, U.K., 2007.

[151] SureBlock™ ultra narrow-band notch filters, [Online]. Available: http://www.ondax.com/Downloads/SureBlock/SureBlock_NotchFilter_DS.pdf [2016, September 15]. Ondax. Monrovia CA, USA, 2015.

[152] SureBlock™ XLF Notch Filter System, [Online]. Available: www.ondax.com/Downloads/SureBlock/SureBlock_XLF.pdf [2016, September 15]. Ondax. Monrovia CA, USA, 2012.

[153] James Carriere. New opportunities in low-frequency Raman spectroscopy. *Photonics spectra*, 2011. Ondax Inc.

[154] J.M. Hornback. *Organic Chemistry*. Available Titles OWL Series. Cengage Learning, 2005.

[155] Joke De Gelder, Kris De Gussem, Peter Vandenabeele, and Luc Moens. Reference database of raman spectra of biological molecules. *Journal of Raman Spectroscopy*, 38(9):1133–1147, 2007.

[156] Brian C. Smith. Low-frequency raman spectra of amino acids measured with an astigmatism-free schmidt-czerny-turner spectrograph: Discovery of a second fingerprint region. Technical report, Princeton Instruments, Acton MA 01720 USA, 11 2013.