N° d'ordre : 41924

Université Lille 1 – Sciences et Technologies

Ecole Doctorale de Sciences de la Matière, du Rayonnement et de l'Environnement

Thèse de Doctorat

En vue d'obtenir la grade de

Docteur de l'Université de Lille 1

Discipline:

Optique, Lasers, Physico-Chimie et Atmosphère

Maya ABOU FADEL

Apports de la chimiométrie à la spectroscopie de Résonance Paramagnétique Electronique : nouvelles perspectives de traitement de données spectrales à dimensions spatiales (imagerie) et / ou temporelles.

Chemometric Contributions to Electron Paramagnetic Resonance Spectroscopy: New Perspectives for Processing of Spectral Data with Spatial Dimension (Imaging) and / or Time Evolution.

Soutenue le 3 décembre 2015 devant le jury :

Dr. Mario Chiesa	Professeur à l'Université de Turin
Dr. Anna De Juan	Professeur à l'Université de Barcelone
Dr. Sylvie Choua	MCF HDR à l'Université de Strasbourg
Dr. François Trompier	CR IRSN
Dr. Douglas Rutledge	Professeur à AgroParisTech
Dr. Ludovic Duponchel	Professeur à l'Université de Lille, Directeur de la thèse
Dr. Hervé Vezin	DR CNRS, Co-Directeur de la thèse

To my parents and my two brothers.

Acknowlegments

The work of this thesis was mainly done at the University of Lille 1 in the LASIR (UMR 8516) and three months at the University of Barcelona in Spain in the grup de quimometria (quimica analitica). I would like to thank all the people who contributed to the work described in the thesis and who made this period an unforgettable experience for me.

First and foremost, I would like to express my deepest sense of gratitude to my director of the thesis, Dr. Ludovic Duponchel, whose expertise, understanding, guidance and motivation added considerably to my graduate experience. I appreciate his leadership, his vast knowledge and skills in chemometrics as well as his suggestions and his demands for a high quality work in all my endeavors. I am sincerely grateful to my co-director of the thesis, Dr. Hervé Vezin, current director of the LASIR, for his insightful advices, his perseverance and his enthusiasm about this work. He gave me the opportunity to work on an innovative and rich spectroscopy.

Besides my director and co-director I would like to thank the remaining members of my thesis committee, Dr. Mario Chiesa, Dr. Anna De Juan, Dr. Sylvie Choua, Dr. Francois Trompier and Dr. Douglas Rutledge, for their time, for their interest in my work and for their positive comments.

A very special thank to Dr. Anna De Juan who welcomed me in her laboratory at Universitat de Barcelona, where I spent 3 months in Barcelona from October till December 2013. I was fortunate that I had the chance to work with her. I appreciate her time, her patience and her extensive knowledge about Multivariate Curve Resolution method. I gained a lot from her huge background and scientific curiosity.

I would like to thank Nadia Touati who initiated me in the more esoteric methods necessary to run the Electron Paramagnetic Resonance spectroscopy.

The members of the LASIR have contributed immensely on to my personal and professional time in Lille. The lab has been a source of friendships and good advice. I am thankful to my officemates, Silvère, Thomas, Remo, Adrien, Amélie, Xin, Gavyns for the simulating discussions, encouragement and all the fun we had. I give my sincere thanks to my fellow lab mates Mattéo, Romain, Siewert, Alexander, Patricia, Ya-Juan, Julia, Dima, Chaza, Ismail, Perrine, Vincent, Josselin, Cécilia, Samantha, Camille, Farah, Shino, Aleksandar,

Bruno, Franco, Fariza, Alzbeta, Bogdan, Fatima. They each helped make my time in the PhD program more fun and interesting. It was a pleasure to meet you all. I am genuinely grateful to Mateusz for his selfless support during the thesis. He was one of the first friendly faces to greet me when I began the doctoral program. I would like to thank also Isabelle, Nacer, Matthieu, Cyril, Myriam, Olivier, Michel, Vincent, Brigitte, Cécile, Gwendoline, Cyrille, Emilie, Maylis, Karima and Julien as well as all the LASIR colleagues. They created a great atmosphere for work.

My time in Barcelona was enriched with great friends: Victor, Flavia, Silvia, Sanae, Fiorella and especially Remo. Thank you for your continuous support, generous care and beautiful memories. I am indebted to thank, Laia, my flat mate in Barcelona for our long political discussions.

I acknowledge with tremendous and deep thanks to my friends Rita, Dayane, Jamile, Josette, Joseph, Nicolas, Assaad, Wafaa, Nazek, Vincent, Sabet, Imane, Samer, Nassim, Assala, Diana, Christine, Rouba, Samar, Christelle, Romain, Antoine, Maria, Wiola, Lena, Anne-Charlotte and Carole. I greatly value their friendship and I deeply appreciate their unwavering belief in me.

I would like to thank all my relatives including my grandma, my aunts, my uncles, my cousins as well as my friends in Lebanon, Canada and United States of America for their love and encouragement.

I gratefully acknowledge the funding sources that made my PhD work possible. I was funded by the University of Lille 1 and the region of Nord-Pas-De-Calais Picardie for the three years.

Most importantly, I would like to express my heartfelt gratitude to my beloved parents, my father Elias and my mother Hoda as well as to my siblings Majed and Firas, for their constant love, patience, prayers and strength in all my pursuits. They continuously supported and encouraged me to reach higher.

Thank you all
Maya Abou Fadel
University of Lille 1
December 2015

Contents

List of Abbreviations

EFA: Evolving Factor Analysis

EPR: Electron Paramagnetic Resonance

EPRI: Electron Paramagnetic Resonance Imaging

ESR: Electron Spin Resonance

FSIW-EFA: Fixed Size Image Window-Evolving Factor Analysis

FSMW-EFA: Fixed Size Moving Window-Evolving Factor Analysis

ICA: Independent Component Analysis

LOF: Lack of fit

MCR-ALS: Multivariate Curve Resolution-Alternating least squares

NMR: Nuclear Magnetic Resonance

PCA: Principal Component Analysis

SIMPLISMA: S*IMPLe* to use Interactive Self-modeling Algorithm

SVD: Singular Value Decomposition

List of notations

χ: magnetic susceptibility

M: magnetization of the material

B: applied external magnetic field

μ: magnetic dipoles

e: electronic charge

m: electronic mass

h: Planck constant

l: orbital quantum number

m_l: magnetic quantum number

H: spin Hamiltonian

E: energy level

Ψ_s: spin functions

g_e: free electron g-factor

S: spin operator

μ_B: magnetic moment

ν: microwave frequency

S: spin angular momentum

L: orbital angular momentum

n: spin population

k: Boltzmann constant

T: absolute temperature in K

ΔB: resonance line width

g_n: nuclear g factor

μ_n: nuclear magneton

γ_n nuclear magnetogyric ratio

I: nuclear spin angular momentum

a_{iso}: isotropic hyperfine coupling constant

Q: quality factor

E_1: electric field

B_1: magnetic field

G_x: linear magnetic field gradient

D: initial matrix

T: scores matrix

$\mathbf{L^T}$: loadings matrix

E: residual matrix

C: matrix that describes the concentration profile

$\mathbf{S^T}$: matrix that describes the spectral profile

T: transformation matrix

General Introduction

During the last few decades, spectacular advances in electronics and computer technology have led to the development of many analytical techniques in the experimental sciences which are capable of collecting large amount of data in a very short time. In particular, some modern spectroscopic techniques can provide during several seconds data set comprised of thousands of variables. Such richness of data, on one hand, provides a lot of information about the studied system, but on the other hand, requires often advanced data analysis method to explore and solve a specific analytical problem.

One of the spectroscopic techniques which are widely used in physical, chemical and biological sciences is Electron Paramagnetic Resonance (EPR) spectroscopy which was successfully applied for the first time in 1945 by E. Zavoisky. It utilizes the electron-spin resonance phenomenon and measures the resonant microwave power absorption spectra of unpaired electrons subjected to a varied magnetic field in an atom, a molecule, or a compound [1]. Therefore, this spectroscopy is now a platform of several irreplaceable tools to study for instance, free radicals, transition metals and other paramagnetic substances. However, the identification of various constituents in an EPR spectrum of a mixture of many paramagnetic components is often very difficult due to the high level of partial or complete superposition of spectral contribution of each species in a complex material. In the case of an entangled system, even the number of pure compounds can be quite difficult to determine. Furthermore, the absence of a rigid database containing the established EPR spectra of various paramagnetic species renders the interpretation of the spectral signature of each component much more challenging. This dilemma covers the entire range of EPR applications such as the analytical characterization of the composition of paramagnetic inorganic, organic and biological samples, kinetics investigations of radicals in the chemical reactions, as well as spectral-spatial information in images. Modern, fast and efficient data analysis methods are thus mandatory for EPR domain in order to resolve these complications.

A discipline which faces the analysis of complex data sets is chemometrics introduced by Svante Wold in the 1972 [2]. In general, it can be defined as mathematical methods that extract chemical information from complex experimental systems. There are a number of useful chemometric tools that can convert data into information. Univariate approach

considers one variable at a time and underutilizes the potential for discovering global information present in the data. Therefore, in case of complex data set, it supplies usually only partial answers to the problem under study which is a biased vision. On the other hand, multivariate tools allow more complete interpretation of the data structures. Among several multivariate approaches, Principal Component Analysis (PCA) and Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) are two unsupervised exploratory techniques especially suited for the study of unresolved mixtures in chemistry when no prior information is available about the nature and composition of these mixtures. PCA was formulated by Carl Pearson in 1901 and it is one of the oldest and most used algorithms [3]. It decomposes the matrix containing the raw data into two smaller matrices (scores and loadings) under very specific constraint of their orthogonality in such way that they contain linearly uncorrelated variables (Principal Components) and their product captures the essential data patterns of the initial matrix. It is an exploratory tool usually employed as a first approach to give an insight about the studied data and to estimate the number of components. On the other hand, MCR-ALS was developed by Romà Tauler in 1995 and is based on a bilinear decomposition [4]. Unlike PCA, it iteratively derives the scores and the loadings through the empty modeling approach and looks for a solution with more natural constraints to produce meaningful concentration and spectral profiles from the sole use of the raw data.

The general goal of this work is to demonstrate how the modern chemometric tools can solve the problems encountered in the analysis of various types of EPR data. In particular, the originality of this work lies in the first “study” or “exploration” of the chemometric method MCR-ALS in the analysis of EPR data sets. The part of the results presented in this thesis has been already valorized through two original publications in international journals: *Analytical Chemistry* and *Journal of Magnetic Resonance*. A third publication is in the process of writing and it is about the application of MCR-ALS on kinetic EPR data set. The obtained results were also exposed as oral communications and posters in national and international congresses. One oral communication presented some results of this thesis was selected as a runner-up in JEOL prize competition during *The 47th Annual International Meeting of the ESR Spectroscopy group of the Royal Society of Chemistry*, in United Kingdom.

This thesis is organized in five chapters where two main fields of study i.e. EPR spectroscopy and chemometrics are introduced in two first chapters and the results are presented and discussed in the others.

Chapter I describes the physical principles of electron paramagnetic resonance phenomenon and EPR spectroscopy. It provides information about the characteristics of EPR spectra such as fine and hyperfine structure, g-factor, or line-shape which serve as fingerprints of the paramagnetic species for unequivocal identification. The instrumentation of the classical EPR experiment as well as Electron Paramagnetic Resonance Imaging (EPRI) technique is also detailed with examples of applications.

In Chapter II, two multivariate and unsupervised exploratory chemometric techniques PCA and MCR-ALS are discussed. The principles of both methods are presented and compared. The multi-step procedures used in MCR-ALS to produce meaningful pure concentration and spectral profiles from the sole use of raw experimental data are detailed. In addition, the extension of this method to three dimensional hyperspectral images (data cube) is also described. Moreover, a state of art of chemometric tools that were already applied to EPR data sets is discussed.

Chapter III shows how MCR-ALS is able to solve the problem related to the superposition of the resonance lines in unresolved EPR spectra. Two experimental model systems containing mixed paramagnetic ions are presented and analyzed to validate the proposed methodology. The influence of the experimental artifacts, data pre-processing and application of different constraints on the resolution process is detailed. The results presented in this chapter are the first application of MCR-ALS method on EPR experimental spectra and demonstrate its benefits over the alternative methods.

Chapter IV demonstrates the applicability of MCR-ALS on EPR imaging experiments. It shows how to extend MCR-ALS approach to a three dimensional data set. The analysis is performed over a photo damaged CaF_2 plate possessing various paramagnetic defects and impurities. The extraction of information about the spatial distribution of such paramagnetic centers on the plate surface by the chemometric tools is demonstrated and discussed. The results presented in this chapter show the first use of MCR-ALS on EPR hyperspectral images.

Chapter V shows the combination of EPR spectroscopy and MCR-ALS in the study of kinetic reactions. It demonstrates how the simultaneous MCR-ALS analysis on data sets coming from two different experiments can help to resolve kinetic traces for reactions of an antioxidant, a hydroxycoumarin derivative, occurring when mixed with a base. The results shown in this chapter are the first application of MCR-ALS on EPR kinetic reactions.

This thesis was funded by the University of Lille 1 and the region of Nord-Pas-De-Calais Picardie for three years. The work was mainly done at the University of Lille 1 in the LASIR (UMR 8516) and three months at the University of Barcelona in Spain in the grup de quimometria (quimica Analitica) due to the historical collaborations.

Chapter I: Electron Paramagnetic Resonance (EPR) Spectroscopy

I.1 Introduction

Magnetic resonance phenomenon is attributed to the absorption of the energy coming from electromagnetic field which accompanies the change of the spin orientation of the electrons or the nuclei in the external magnetic field. In the case of the nuclei, such resonance is called Nuclear Magnetic Resonance (NMR), whereas for the electrons, Electron Paramagnetic Resonance (EPR) is a common name. A term Electron Spin Resonance (ESR) is also often used in the latter case to differ from the cyclotron resonance where the absorption is due to a periodic move of the electron on a circular path. In this work, only the resonance being result of the quantum nature of the electron spin is discussed, thus, the terms EPR and ESR are equivalent.

Since the interaction of the electromagnetic radiation with the magnetic materials provides inexhaustible amount of information, magnetic resonance spectroscopy techniques are important and widely used tools in the investigation of matter and the exploration of multitude of systems. In particular, EPR spectroscopy represents a technique that is based on the excitation of the electron spins in the magnetic field and the subsequent absorption of the electromagnetic radiation [1, 5, 6]. In consequence, EPR spectroscopy is addressed to the materials with unpaired electrons. Such electrons have magnetic moments which can align parallel or antiparallel to the external magnetic field. In terms of quantum mechanics, an energy level of such electron is split and the electron can move between the lower and upper state absorbing (or emitting) photon of energy *hν* which fulfill the resonance condition (see Figure I.1).

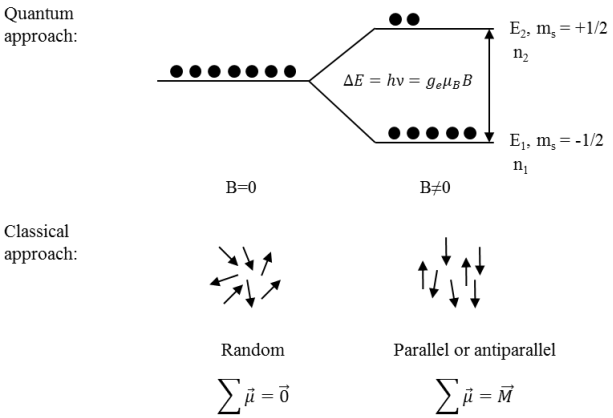


Figure I.1 : General concept of a paramagnetic system in the presence of an external magnetic field and an electromagnetic radiation.

In general, EPR spectroscopy is applicable to various systems and provides clear insights into many of the fundamental concepts of quantum mechanics. On the other hand, the extraction and the analysis of the information from EPR experiments is very often not a trivial task. Therefore, EPR experimental techniques as well as data analysis methods have been thoroughly developed for several decades.

Historically, EPR spectroscopy has its sources in the well-known and fundamental Stern-Gerlach experiment. In 1921, two German physicists Walter Gerlach and Otto Stern observed that a beam of silver atoms splits into two lines when passing through an inhomogenous magnetic field. Therefore, they showed that the electron magnetic moment in an atom can adopt only discrete orientations in a magnetic field, despite the sphericity of the atom [7]. Further on, Uhlenbeck and Goudsmit related the electron magnetic moment with the concept of the electron spin angular momentum [8]. Later, knowing that the hydrogen atom possesses an additional angular momentum arising from the nucleus proton, Bret and Rabi investigated the influence of the static magnetic field on the resultant energy levels. Subsequently, Rabi et al. studied the transition between the levels induced by oscillating fields. This experiment was the first observation of magnetic resonance [9, 10]. In 1944, I. I. Rabi was awarded for this pioneering work, a Nobel Prize in Physics.

A milestone in electron magnetic resonance was the work of Russian physicist E. Zavoisky who studied the paramagnetic salts. He was inspired by the advances of the emerged magnetic resonance and in 1945, he successively developed the first electron paramagnetic resonance experiment and observed the peak in the signal of a detected radiofrequency absorption line from a $CuCl_2.2H_2O$ sample [11, 12]. He found the resonance at a magnetic field of 4.76 mT and a frequency of 133 MHz. One year later, the first nuclear magnetic resonance in paraffin was recorded by Purcell et al. and independently in water by Bloch *et al.* [13, 14]. In 1952, the Nobel Prize in Physics was awarded jointly to both F. Bloch and E.M. Purcell.

After 1946, EPR spectroscopy was developing rapidly due to the widespread availability of complete and suitable microwave instrumentations following the World War II. For instance, equipment for the 9 GHz region were extensively used for radars and some components were easily available at low cost. In that time, Cummerow and Halliday were undertaking advanced EPR studies in United States as well as Baggulay and Griffiths in England [15, 16].

In 1958, Blume reported the first pulse EPR experiment [17]. In that time, the pulse EPR spectroscopy suffered expensive instrumentation, lack of microwave components, in addition to electronics and intrinsic problems of the limited microwave power. In the 1980s, the equipment needed became cheaper and manageable. Bruker released the first commercial pulse EPR spectrometer in the market followed by the first commercial high field spectrometer, ten years later [18, 19]. High frequency EPR has provided several advantages compared to low frequency techniques. It offers increased resolution and better absolute sensitivity that is particularly useful for studying systems whose number of paramagnetic species is inherently low. Since then, this EPR technique was subsequently refined and extended to become a work horse tool employed in a vast range area of fields, ranging from chemistry and biomedical research to solid state physics. Nowadays, EPR playground became wide and accessible to an ever growing research community. The most modern applications of EPR such as EPR imaging (see paragraph I.6) enable even monitoring of free radical reactions in vivo [20].

Where applicable, EPR spectroscopy yields accurate and detailed information that are unobtainable with other methods. This spectroscopy can furnish information about electronic and geometrical systems, dynamic ability, as well as mechanisms of reactions. Moreover, the influence of physical and chemical factors on the paramagnetic centers can be studied too. All the information mentioned above can be obtained from the analysis of the absorption signal of each corresponding center. EPR measurements reveal applications in a wide array of fields such as medicine, biology, pharmacy, material and environmental science [21-25].

I.2 Magnetic materials

All materials show magnetization when subjected to a sufficiently strong magnetic field. However, the interactions of the matter with magnetic field can vary depending on the intrinsic magnetic properties of a given material. In this context, the basic classification distinguish three categories: diamagnetic materials, paramagnetic materials, and ferromagnetic materials [26].

Diamagnetic materials:

All the electrons in an atoms or molecules of diamagnetic compounds are paired up and share their orbital with another electron (net spin=0). Diamagnetic materials have negative magnetic susceptibility χ which is defined as a dimensionless proportionality constant that indicates the degree of magnetization of a material in response to an applied magnetic field. It is expressed by equation (I.1)

$$\chi = \frac{M}{H} \tag{I.1}$$

Where M is the magnetization of the material (the magnetic dipole moment per unit volume) and H is the magnetic field strength. Thus, diamagnetic materials, where susceptibility is small and negative, create a magnetic moment in the direction opposite to the applied external magnetic field $\vec{B}$ resulting in a weak repulsion of the field. The physical background of diamagnetism is explained by the molecular or atomic magnetic dipole moments that are associated with electron orbital moments. In the presence of an external magnetic field, their orbitals are slightly deformed and their kinetic moment changes.

Paramagnetic materials:

In contrast to diamagnetism, paramagnetism is based on the existence of an intrinsic electronic magnetic moment. Electrons in paramagnetic compounds are not coupled in pairs of opposite spin in the same orbital causing a permanent dipole moment. In the case of more than one atom, these dipole moments do not interact and are randomly oriented due to the thermal movement. The magnetic field of the elementary dipole is compensated and the net resultant magnetic moment is zero. However, when the paramagnetic materials are placed in an external magnetic field, the spins tend to align giving a net magnetic moment in the same direction as the applied field (see Figure I.1). The whole material acts like a magnet that is slightly attracted by the magnetic field. As soon as the external field is removed, the material does not retain its magnetic properties. In the other words, paramagnetic materials have positive magnetic susceptibility χ, and react in an opposite way compared to diamagnetic materials.

Ferromagnetic materials:

Ferromagnetism is a property of existence of net resultant magnetic moment of the compound even in the absence of the external magnetic field. It indicates that the spontaneous magnetic moments of the electrons in such materials are ordered in a regular manner. In consequence, ferromagnetic materials possess the ability to magnetize strongly under an external magnetic field. However, the ferromagnetic properties of materials depend on the temperature. When it increases, the thermal motions compete with tendency for dipoles to align. In consequence, above a certain temperature (called the Curie temperature) the spontaneous magnetization disappears and the material becomes paramagnetic.

I.3 Which substances are paramagnetic?

As mentioned above, paramagnetic materials exhibit permanent magnetic dipoles. In the simple case of individual atom, such a dipole is results of the motion of the electron around the nucleus and can be classically expressed as:

$$\mu = -\frac{e}{2m} L \tag{I.2}$$

where e and m are electronic charge and mass, respectively, and L is the orbital angular momentum of the electron. Taking into account the quantum mechanics principles, it is known that both orbital momentum as well as its projection along the specified axis are quantized i.e. they can take only discrete values.

$$L = \frac{h}{2m} \sqrt{l(l+1)} \tag{I.3}$$

And

$$L = \frac{h}{2\pi} m_l \tag{I.4}$$

where h is Planck constant, l is the orbital quantum number ($l = 0,1,2 \dots$) and m_l is the magnetic quantum number ($m_l = -1,0\ or\ 1$). The quantum numbers describe the set of allowed states that electrons may occupy in the atom or molecule. For a full description of the

electronic configuration, the spin of electron and its projection with quantum number ($m_s = \pm\frac{1}{2}$) must be included.

In the context of these considerations, paramagnetic material i.e. substance with permanent magnetic dipoles has to possess the odd number of electrons. However, most of the chemical substances are on the contrary diamagnetic what can be illustrated by the electron shell of salt NaCl, compared to the shells of its free atoms.

Na^+:$1s^2\ 2s^2\ 2p^6\ 3s^0$ Cl^-:$1s^2\ 2s^2\ 2p^2\ 3p^6$
Na: $1s^2\ 2s^2\ 2p^6\ 3s^1$ Cl:$1s^2\ 2s^2\ 2p^2\ 3p^5$

Ions have completely filled electron shells, for instance, *s*-orbital ($l = 0$) contains two electrons with different spins $m_s = \pm\frac{1}{2}$ (along with Pauli exclusion principle) three *p*-orbitals ($l = 1$) with two electrons in each with different spins having $m_l = \pm 1,0$. Full shelled ions and atoms do possess neither spin nor orbital angular momentum. Nevertheless, there are some classes of compounds which exhibit the unpaired electrons and paramagnetic properties in consequence.

I.3.1 Free radicals

Free radicals are atoms or group of atoms which have at least one unpaired electron. Thus, free radicals are paramagnetic. They can be formed from atoms by losing or gaining the electron(s) spontaneously as well as by exposure to heat, light or other factors. Radical chemistry began about 1900 when the first persistent free radical, triphenylmethyl, was obtained from hexaphenylethane by Moses Gomberg [27].

$(C_6H_5)_3C\text{-}C(C_6H_5)_3 \longleftrightarrow 2(C_6H_5)_3C\cdot$
hexaphenylethane triphenylmethyl

However, most of the free radicals are very unstable and appear as short living intermediate products in various reactions. When the outermost shell of an atom is not full, it is unstable and will try to stabilize itself by either gaining or losing an electron to either fill or empty the shell. Therefore, free radicals are highly reactive. They can be formed during

photolysis, by the influence of ionization radiation and in certain enzymes. One of the driving force in radical chemistry was the investigation of X or γ-irradiation damage of molecules giving also radicals, in addition to monitoring of living tissues [28, 29]. It is noteworthy that free radicals play an important role in many biological processes. In spite of their inherent instability under suitable conditions, free radicals can have sufficient lifetime to be observed even in the liquid. In 1960, the appearance of free radicals in liquid hydrocarbons during radiolysis was demonstrated by means of EPR spectroscopy [30]. Moreover, in the solid state free radicals can become stabilized almost indefinitely in many materials. In terms of spin multiplicity $(2S + 1)$, the simple free radicals (one unpaired electron with $S = \frac{1}{2}$) are doublets.

I.3.2 Triplet state molecules

Some molecules with even numbers of electrons, two of which, however, are not paired, are also paramagnetic. The oxygen molecule, O_2, is the simplest and most popular example of the kind. The spins of the two electrons ($S = \frac{1}{2}$) in degenerate orbitals give rise to four independent spin states in total. One of them is a singlet state (total spin $S = 0$) and others are triplet states ($S = 1$). According to Hund's rules, the triplet states are energetically more favorable, hence the ground state of the oxygen molecule is a triplet. Much more common species are excited triplet states, with finite lifetime, formed by many compounds under visible or ultraviolet light excitation and giving an excited energy system with a total spin quantum number ($S = 1$). Since the triplet states possess two non-paired electrons which behave as a pair of radical centers, they are often called triplet diradicals. Triplet state molecules have been widely studied by EPR spectroscopy in liquids, single crystals or amorphous matrices (glasses) [31].

I.3.3 Metals with partially filled inner shells

Since paramagnetism requires the presence of the unpaired electrons, paramagnetic properties can be also observed in the compounds of some elements with partially filled sub-

shells. The most popular group of such species is compounds of transition metals such as Cr, Fe, Co, Cu, Mn, V, and Mo. In transition metals d sub-shell is incomplete, thus, this class of metals contains iron group $3d$, palladium group $4d$ and platinum group $5d$ in the periodic table. In addition, the elements with partially filled f sub-shell i.e. lanthanide group $4f$ and actinide group $5f$ exhibit similar paramagnetic properties. EPR spectra of transition metal ions are affected by the environment more than in the case of a free ions or triplet state molecules. This is especially visible in the condensed phase where the complexity of the EPR signal is attributed to the interaction with the surrounding ions and electric dipole that causes the splitting of energy levels for example, the five degenerate 3d orbitals of the iron group. Due to abundant applications, transition metal complexes or rare earth ions have been the subject of numerous experimental and theoretical EPR studies [32-35].

I.4 Basic principles of continuous wave-EPR spectroscopy

There are several methods to record an EPR spectrum. They can be briefly divided into continuous wave (CW) and pulsed EPR spectroscopic techniques. CW EPR is a traditional and the most widely used method which is based on low intensity microwave radiation that continuously irradiates the studied sample. On the other hand, the pulsed EPR spectroscopy relies on the short pulses of high power microwave radiation sent to the sample and the response is recorded in the absence of radiation. For the purposes of this work, this chapter is focused on the CW EPR method.

I.4.1 Spin Hamiltonian

Any system which has discrete energy levels and is described by defined quantum numbers can be represented by an eigenvalue equation, called Schrodinger equation that is appropriate to the property being observed. In such equation, the Hamiltonian *H* is the operator corresponding to the total energy of the system. Since EPR is concerned with quantization of spin, the concept of *Spin Hamiltonian* is used. This concept describes the system with a simplified form of Shrödinger wave equation describing only the lowest

electronic state of the molecule in addition to magnetic interactions. In this description, the simplified operator, H, is the spin Hamiltonian, the simplified wave functions, Ψ_s, are the spin functions, and the eigenvalues E are the energy values of the ground state spin manifold.

$$H\Psi_s = E\Psi_s \tag{I.5}$$

The spin Hamiltonian can be derived by the Hamiltonian of the whole system by separating the energetic contributions involving the spin from all the other contributions [36]. It contains all the interactions of the electron spin with the external magnetic field and the internal magnetic moments. Among the different interactions, only three interactions are discussed in this chapter i.e. H_{EZ} electron Zeeman interaction, H_{NZ} nuclear Zeeman interaction and H_{HF} hyperfine interaction.

I.4.2 The electron Zeeman interaction

A free electron has an intrinsic angular momentum called spin S being present even without external forces. Consequently, every free electron induces magnetic field due to the moving charge and it acts as a small bar magnet or magnetic dipole having a magnetic moment (see Figure I.2).

Figure I.2 : Free unpaired electron with a spin angular momentum (curved arrow) and magnetic moment (vertical arrow).

This moment can be expressed as:

$$\mu_e = -g_e\mu_B S \tag{I.6}$$

where g_e is the free electron g-factor, μ_B is Bohr magneton and S is spin operator. In the case of a free electron, the exact value of the g-factor is 2.0023193043617. This value is predicted by the quantum electrodynamics and is the most accurately determined fundamental constant by both theory and experiment [37]. Bohr magneton is expressed as

$\mu_B = \frac{e.\hbar}{2m_e}$ and it is the magnetic moment for one unit of quantum mechanical angular momentum with the electron of charge e and mass m_e ($\hbar = \frac{h}{2\pi}$ is the reduced Planck constant) .

When the electron is subjected to an external magnetic field B, its energy is

$$E = -\mu \cdot B \tag{I.7}$$

In quantum mechanics, the interaction between the electron spin and the external magnetic field B is expressed following the simple spin Hamiltonian (a dominant term)

$$H_{EZ} = -\mu \cdot B = g\mu_B S \cdot B \tag{I.8}$$

As a result of this interaction, the two energy levels $m_s = \pm\frac{1}{2}$ of the degenerate spin states split. As the magnetic field increases, the degeneracy level lifts linearly between the level $m_s = -\frac{1}{2}$ where the spin is parallel to the magnetic field (more stable level) and the level $m_s = \frac{1}{2}$ where the spin is anti-parallel to the magnetic field, as illustrated in Figure I.3. This effect is known as Zeeman splitting. The energy of each electronic level becomes

$$E = g_e\mu_B B m_s = \pm\frac{1}{2}\, g_e\mu_B B \tag{I.9}$$

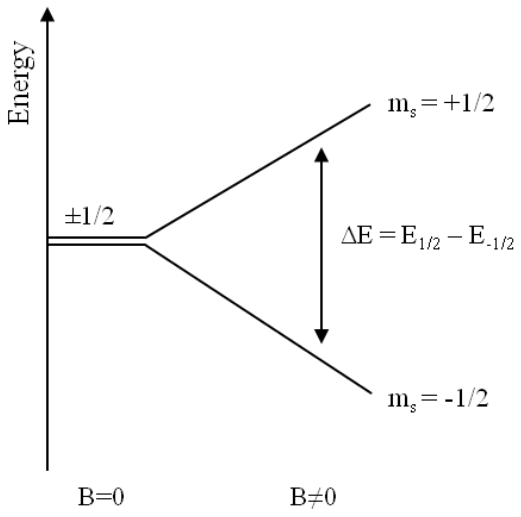


Figure I.3: Zeeman effect.

Resonance condition

According to the selection rule $\Delta m_s = \pm 1$ (see Figure I.4), EPR consists of inducing transitions from a lower energy level to a higher energy level by applying an electromagnetic radiation with a magnetic field B_1 perpendicular to the magnetic field B oscillating with a frequency ν and verifying the resonance condition

$$\Delta E = E_{1/2} - E_{-1/2} = h\upsilon = g_e \mu_B B_{res} \quad \text{(I.10)}$$

Therefore, the basis of EPR spectroscopy consists of inducing transitions between the existing Zeeman levels subjected to a varied magnetic field B and the absorption of a constant frequency υ to give an absorption spectrum (typically in the first derivative form) plotted against the magnetic field. These transitions are often referred to as the spin-flip effect and the ΔE is the photon energy needed to flip the spin population in a given magnetic field. The applied electromagnetic frequency υ belongs to the microwave region and can range between 1 and 500 GHz.

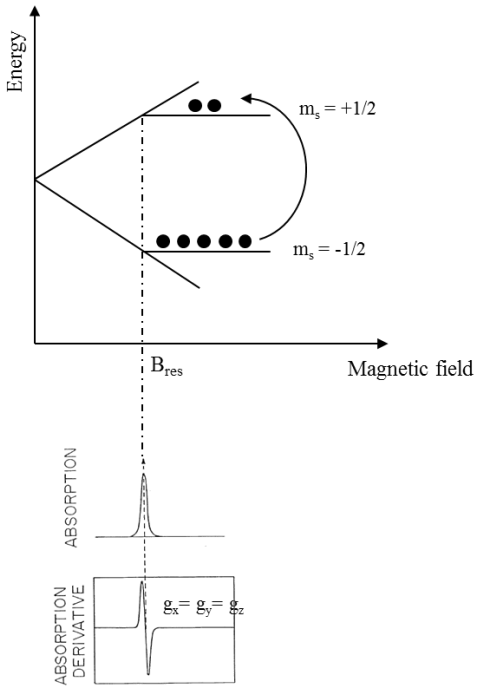


Figure I.4: Resonance.

I.4.3 Relaxation processes

In order to observe a net microwave absorption, the population of the electrons at the ground state must be greater than the population at the excited state. At thermal equilibrium and under external magnetic field, the spin population is split between Zeeman's levels according to Maxwell-Boltzmann law:

$$\frac{n_1}{n_2} = e^{\frac{-\Delta E}{kT}} = e^{\frac{-g_e \mu_e B}{kT}} \quad \text{(I.11)}$$

where n_1 and n_2 are the spin population characterized by m_s values $-\frac{1}{2}$ and $+\frac{1}{2}$ respectively, k is the Boltzmann constant and T is the absolute temperature in K.

At a high microwave power, the population of the two levels n_1 and n_2 tends to equalize, and therefore, no transfer of energy occurs. Consequently, the intensity of the signal decreases and becomes no longer proportional to the number of spins. This phenomenon is called saturation. Subsequently, the odd electrons from the upper level give up their $h\nu$ quantum and return back to their equilibrium position. The release of the energy may occur via a spin relaxation process either spin-lattice relaxation, defined by the time T_1 or spin-spin relaxation, defined by the time T_2 (Figure I.5).

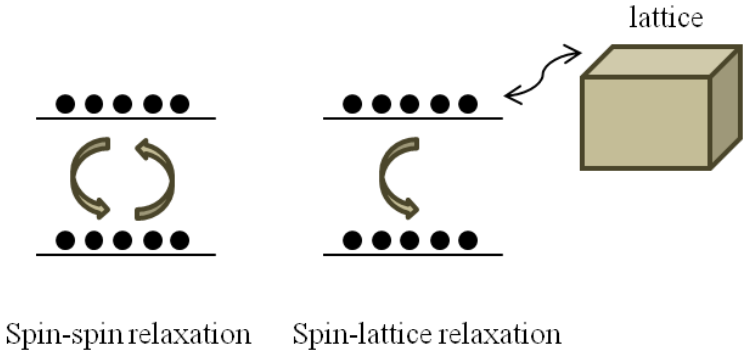


Figure I.5: Two types of spin relaxation.

Spin-lattice relaxation is the mechanism of energy dissipation within the lattice where the spin belongs. The spins give back the energy already gained by collision with other atoms in the form of vibrational, rotational and translational energy. The relaxation process can be also reached by spin-spin relaxation where the energy exchange occurs between the spins and not to the lattice keeping the energy constant among them.

Both spin-spin relaxation and spin-lattice relaxation contribute to the resonance line width (ΔB)

$$\Delta B \propto \frac{1}{T_1} + \frac{1}{T_2} \quad \text{(I.12)}$$

In general, the line width depends on the spin-spin relaxation, therefore $T_1 > T_2$. T_2 increases when the spin concentration decreases. When T_1 becomes very short (below 10^{-7} seconds), it broadens the EPR lines due to its effect on the lifetime of the species.

I.4.4 Spin-orbital coupling

The electron present in a molecule has a different total angular momentum compared to an electron placed in space since, in addition to the intrinsic spin angular momentum *S*, it also possesses orbital angular momentum *L* (Figure I.6).

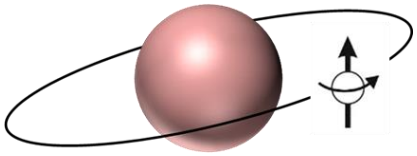


Figure I.6 : orbital angular momentum.

Hence, two magnetic moments interact with each other exist and the energy of the resulting spin-orbit interaction depends on their relative orientation. It can be assumed that the spin-orbit coupling term is proportional to *S*. In such a case, the g-factor of the free electron g_e is replaced by g in resonance condition.

$$\Delta E = h\upsilon = g\mu_B B \quad \text{(I.13)}$$

The magnitude of the spin-orbital coupling depends on the size of the nucleus associated with the unpaired electron. For instance, organic free radicals that contain only H, C, O, and N atoms has a small contribution from spin-orbit coupling and gives a g factor very close to $g_e \approx 2$. Whereas, much heavier elements, such as metals, may give g-factor considerably different from g_e. The g-factor is specific to each paramagnetic center, and can be calculated from the position of the resonance line in EPR spectrum

$$g = \frac{h\upsilon}{\mu_B B} = 0.7145 \frac{\upsilon\ (MHz)}{B\ (Gauss)} \quad \text{(I.14)}$$

with $h = 6.626 \times 10^{-34}\ J \cdot s$ and $\mu_B = 9.274 \times 10^{-28}\ J \cdot G^{-1}$

It is worth to note that the position of the resonance peak expressed in field units (Gauss, or Tesla (1T = 10^4 G)) is a convenient way to characterize compounds. In this context, the parameter which we actually wish to obtain from EPR experiment is the g-factor.

I.4.5 Anisotropy of g-factor

In the previous section, only one orientation of a molecule placed in the magnetic field is discussed. The g-factor in the molecule depends on the spin-orbital coupling between the ground state and the excited state. Knowing that orbitals are oriented in the molecule, the magnitude of this mixing is direction dependent i.e. anisotropic in general case. As the molecule is rotated in the spectrometer, the EPR spectrum changes due to the g-factor anisotropy. Every paramagnetic molecule has a unique axis system called principal axis system. The g-factors measured along these axes are called the principal g-factors that are labelled as g_x, g_y, and g_z. In a single crystal, the paramagnetic molecules are in a fixed orientation (unique g). In liquid solution, all the anisotropy is averaged out (unique g), whereas frozen samples give powder spectrum that represents the summation of all possible orientations (g_x, g_y, and g_z). Three different symmetries can be obtained (see Figure I.7).

In a metal ion that has a symmetric environment, the three principal axes are identical ($x = y = z$) because electrons in the different d-orbitals have equal interactions in all directions. Therefore, the orbital moment is equal in all directions as well as the magnetic moment. When the molecule is placed in the external field in different positions, the magnetic momentum is always the same giving one g-value (Figure I.7a). Resonance occurs at one value of the external magnetic field and one corresponding absorption peak is observed. This occurs in octahedral and tetrahedral geometries.

$$g_x = g_y = g_z \tag{I.15}$$

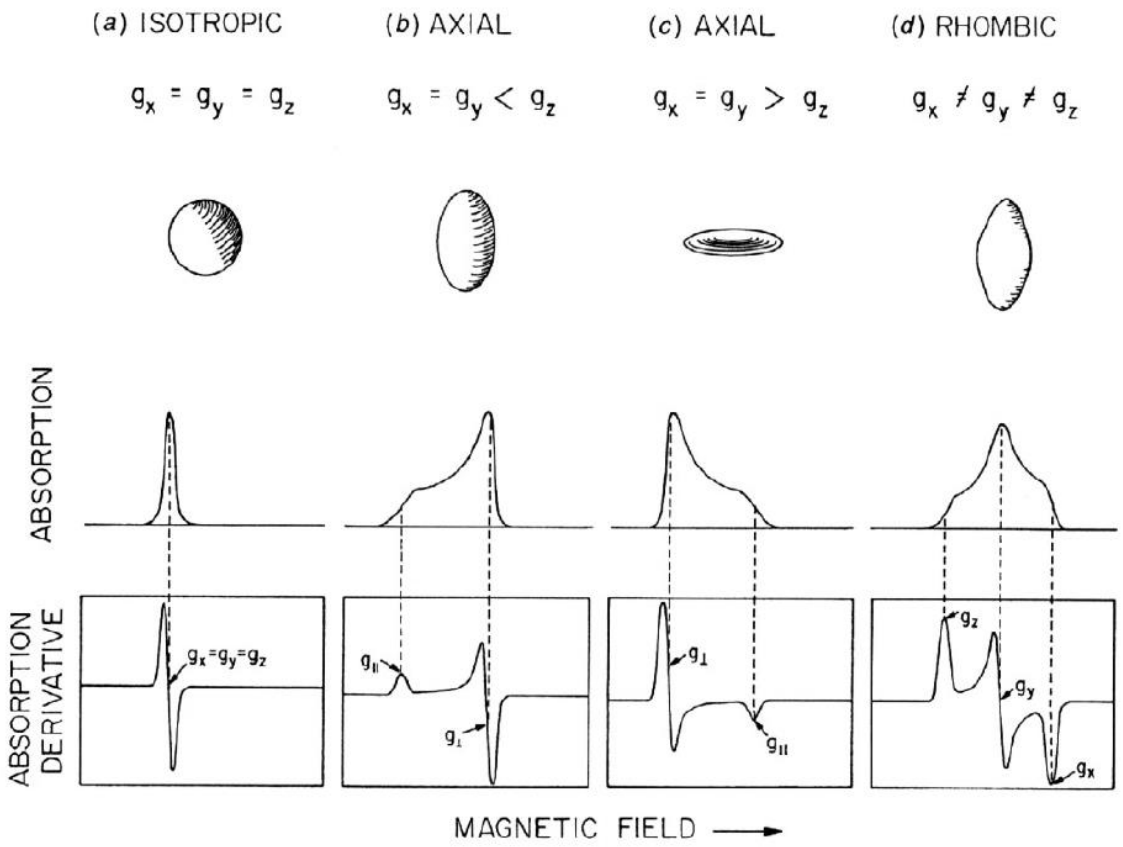


Figure I.7 : EPR spectra (absorption and first-derivative spectra) for the three different classes of anisotropy.

Axial symmetry is when a unique axis in a molecule differs from two others. The position of the molecule placed in the external magnetic field does matter. A metal that is coordinated by two equal ligands in the z-direction and four different but equal ligands in both x and y directions can be considered as an example. The resulting g-factor is different for the situation where *B* is parallel to the z axis or parallel to either xy plane (Figure I.8).

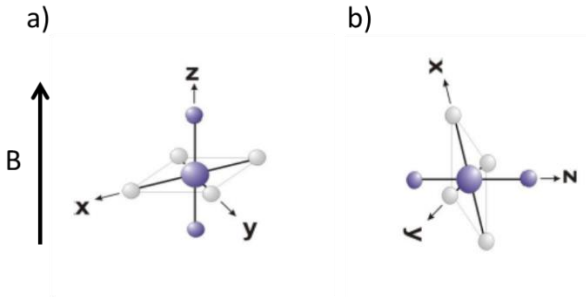


Figure I.8 : An example of metal complex where central atom is coordinated by two equal ligands in the z-direction and four different but equal ligands in xy plane. a) ***B*** is parallel to the z-direction and b) ***B*** is parallel to xy plane.

When the molecule is placed in the external field B parallel to the z-direction of the molecule, it has a large magnetic moment μ_z. In order to attain resonance, a small magnetic field is required. The rotation of the molecule from the position, where the magnetic field is parallel to the z-direction, to the position where magnetic field is parallel to the x,y plane (magnetic field perpendicular to x-axis) decreases the total magnetic moment from μ_z to $\mu_{x,y}$. All the orientations occur in the frozen samples, consequently, giving a large number of overlapping absorption lines starting at B_z and ending at $B_{x,y}$. EPR spectrum is the sum of all these lines representing a total absorption in the x,y direction much larger than in the z-direction because the ions occur more frequently with its x- or y- direction parallel to the magnetic field.

The g-factor along the unique axis is said to be parallel with it, $g_z = g_{\|}$ while the remaining two axes are perpendicular to it $g_{x,y} = g_{\perp}$. This situation is characteristic for trigonal bipyramid, square-based pyramid, square plane and elongated octahedral geometries (Figure I.7b and c).

$$g_x = g_y \neq g_z \tag{I.16}$$

When all the principal axis are different from each other, rhombic symmetry is obtained which is characterized by three distinct g-values

$$g_x \neq g_y \neq g_z \tag{I.17}$$

Three special cases can, thus, occur for the resonant field positions corresponding to the orientation of the applied field along with x, y and z axis, respectively. Consequently, the resultant EPR spectrum exhibits three distinct features (see Figure I.7d).

I.4.6 Nuclear Zeeman interaction

In analogy to the interaction between the unpaired electron and the magnetic field, there also exists interaction between the magnetic nuclei of the paramagnetic atom with the

magnetic field when the nuclear spin I ≠ 0 (Table I.1). Therefore, the nuclear magnetic moment can be expressed as:

$$\mu_n = g_n \mu_n I \tag{I.18}$$

where g_n is the nuclear g factor,

$\mu_n = \frac{e.\hbar}{2m_p}$ is the nuclear magneton that is equal to $5.050 \times 10^{-29} JG^{-1}$,($\gamma_n = g_n \mu_n$ is also often used and called the nuclear magnetogyric ratio), and $\vec{I}$ is the nuclear spin angular momentum.

Atom	Nuclear spin I
^{1}H	$\frac{1}{2}$
^{13}C	$\frac{1}{2}$
^{14}N	1
^{19}F	$\frac{1}{2}$
^{49}V	$\frac{7}{2}$
^{55}Mn	$\frac{5}{2}$
^{63}Cu	$\frac{3}{2}$

Table I.1 : Nuclear spin of some atoms.

Thus, the spin Hamiltonian for the nuclear Zeeman interaction and its energy in the presence of magnetic field can be expressed as:

$$H_{NZ} = \mu \cdot B = -\gamma_n I \cdot B \tag{I.19}$$

$$E = -\gamma_n B m_I \tag{I.20}$$

In the case of a proton ($m_I = \pm\frac{1}{2}$), the two degenerated spin states are split in the presence of magnetic field similarly to the electron.

I.4.7 Isotropic Hyperfine interaction

The energy of the unpaired electron depends on the interaction between the unpaired electron (Electron Zeeman effect) and the nucleus (Nuclear Zeeman effect) with the magnetic field as well as on the interaction between the unpaired electron and the magnetic nuclei called hyperfine interaction. The hyperfine interaction is one of the most important sources of information in EPR spectroscopy. It characterizes the interaction between the electron spin and the nuclear spin and therefore provides information about the direct magnetic environment of the spin. Its Hamiltonian spin, also called Fermi contact, can be described by

$$H_{HF} = I\ a_{iso}\ S \tag{I.21}$$

The parameter a_{iso} is the isotropic hyperfine coupling constant that represents the energy of interaction between the electronic and the nuclear spin.

I.4.7.1 Interaction of the electron with one nucleus

Considering a simple two-spin system with an unpaired electron $S = \frac{1}{2}$ and a single nucleus $I = \frac{1}{2}$ in an external magnetic field, the electronic energy levels are perturbed and, therefore, each decomposes into 2I+1 sub-energy levels (see Figure I.9). The Hamiltonian spin of this interaction is

$$H = H_{EZ} + H_{NZ} + H_{HF} \tag{I.22}$$

$$H = g_e \mu_B S \cdot B - \gamma_N I \cdot B + I\ a_{iso}\ S \tag{I.23}$$

where μ_B is much larger than μ_N, and nuclear Zeeman effect can be considered negligible in the Hamiltonian

$$H = g_e \mu_B S \cdot B + I\ a_{iso}\ S \tag{I.24}$$

The possible energy levels of the system can be described by formula:

$$E\ (m_s, m_I) = g_e \mu_B B m_s + m_I\ A\ m_s \tag{I.25}$$

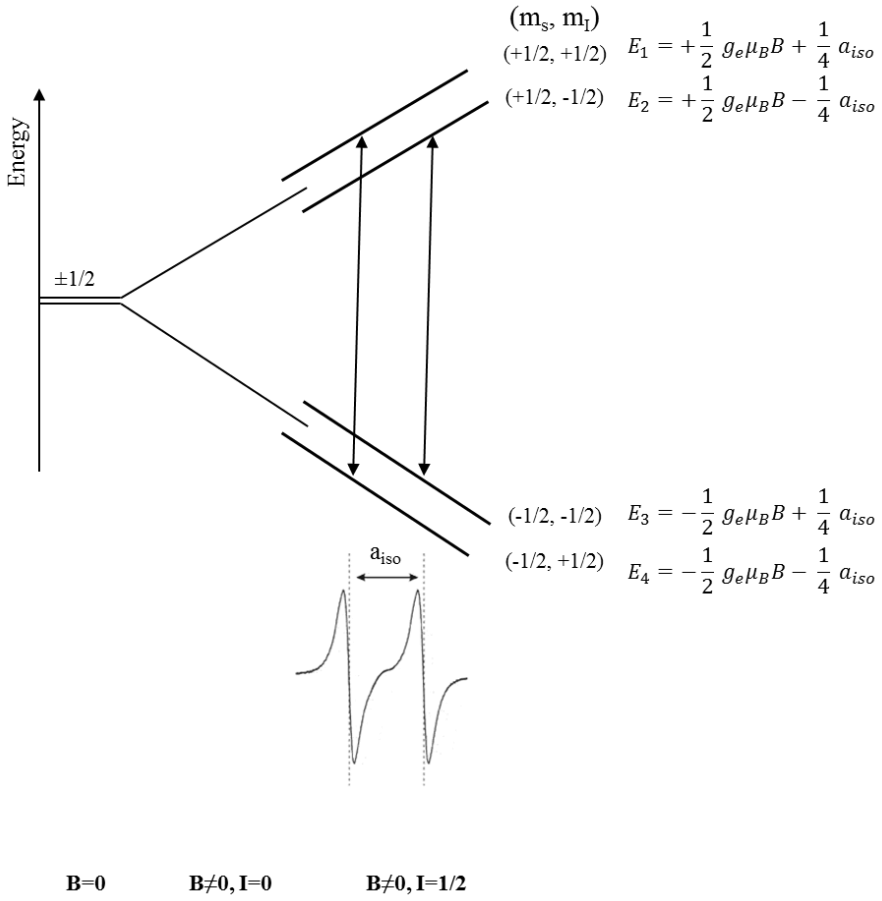


Figure I.9 : Magnetic energy level diagram and the allowed transitions for a two spin system.

According to the selection rules ($\Delta m_S = \pm 1, \Delta m_I = 0$), two allowed resonance transitions

$$\Delta E_{1-4} = E_1 - E_4 = g_e \mu_B B - \frac{1}{2} a_{iso} \tag{I.26}$$

$$\Delta E_{2-3} = E_2 - E_3 = g_e \mu_B B + \frac{1}{2} a_{iso} \tag{I.27}$$

can be experimentally observed.

These two transitions give rise to two absorption peaks in EPR spectrum at different magnetic field positions, separated by a_{iso} the isotropic hyperfine coupling (see Figure I.9).

I.4.7.2 Interaction of the electron with several nuclei

Interaction with equivalent nuclei

In a system that contains more than one equivalent nuclei, the interaction of the unpaired electron with n equivalent nuclei of spin I gives an EPR spectrum with $2nI + 1$ equally spaced peaks. The relative intensities of the absorption peaks are represented by the binomial expansion of $(1 + x)^n$ and the successive sets of coefficients for an increasing n are given by Pascal's triangle (Table I.2).

n	Pattern	Coefficients
0		1
1	Doublet	1 1
2	Triplet	1 2 1
3	Quartet	1 3 3 1
4	Pentet	1 4 6 4 1
5	Sextet	1 5 10 10 5 1
6	Septet	1 6 15 20 15 6 1
7	Octet	1 7 21 35 35 21 7 1
8	Nonet	1 8 28 56 70 56 28 8 1

Table I.2 : Coefficients for the binomial expansion $(\mathbf{1} + \boldsymbol{x})^{\boldsymbol{n}}$.

Considering a paramagnetic system including two equivalent hydrogen nuclei $\left(I = \frac{1}{2}\right)$, each equivalent nuclei splits the energy levels into a doublet having the same splitting constant. Therefore, the splitting of the other equivalent nuclei gives an overlap of energy levels. For 2 equivalent protons, the number of lines is given $2nI + 1$, then $2 \times 2 \times \frac{1}{2} + 1 = 3$ lines, producing a triplet hyperfine pattern with an intensity ratio 1:2:1 (Table I.2).

Interaction with non-equivalent nuclei

The unpaired electron can also interact with several sets of inequivalent nuclei, with different hyperfine coupling. The resulting EPR spectrum is complex in such case and the total number of lines can be calculated for m sets of equivalent nuclei, each counting n number of equivalent nuclei

$$N = (2n_1I_1 + 1)(2n_2I_2 + 1)\cdots(2n_nI_n + 1) \quad \text{(I.28)}$$

In the case of an unpaired electron interacting with two inequivalent protons, the energy levels are split by the first proton with a_1 coupling and then, the second proton splits the energy levels with a coupling constant of a_2. The EPR spectrum obtained by the four possible transitions consists of 'Doublet of doublets' where each doublet possesses an intensity ratio 1:1 (Figure I.10).

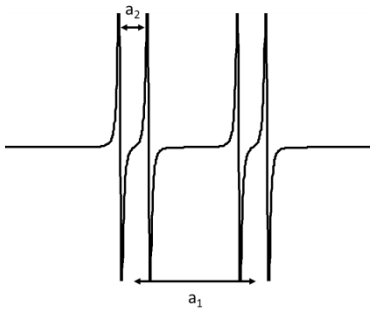


Figure I.10: EPR spectrum of a system containing an unpaired electron and two inequivalent protons.

I.5 Instrumentation

EPR spectrometers are present in different sizes and models depending on the year of production and on the frequency and field of operation (Figure I.11). In continuous wave EPR spectroscopy, the sample is placed in a magnetic field that varies in a linear field sweep exposing it to a fixed frequency of microwave radiation. This technique gives much information about the paramagnetic species studied such as oxidation degree, geometry, nuclearity, and the coordination number making it an indispensable characterization technique in diverse domains of chemistry as well as in physics.

Figure I.11 : E500 CW EPR spectrometer in X-band (Bruker).

Every EPR spectrometer consists of several components which are drawn in Figure I.12 and briefly described in the following paragraphs.

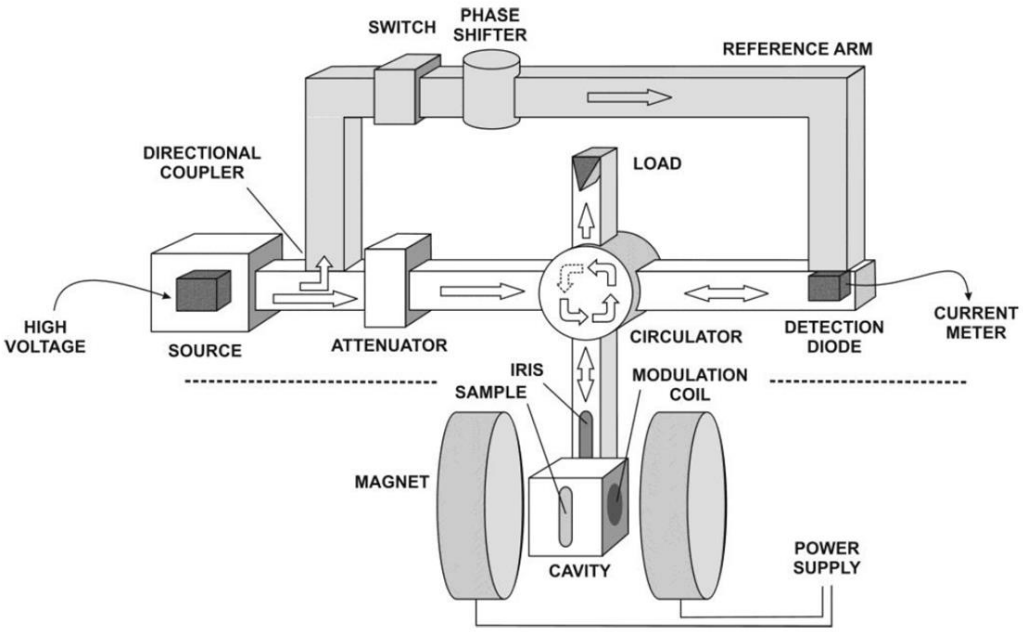


Figure I.12 : Microwaves flow in EPR spectrometer [38].

I.5.1 Microwave bridge

The microwave bridge starts from the microwave source and ends at the detector, including attenuator, reference arm, circulator, and insulator.

The source includes all the components that generate the microwaves and the transmission line. The klystron is the most often used source of energy, and it is a vacuum tube that produces microwave oscillations in a small range of frequencies. Gunn diode is another popular source that is used in modern spectrometers. The frequency ν of the produced electromagnetic wave is often equal to 9.5 GHz known as X-band. Other widely used frequencies are noted in Table I.3.

Band	Frequency (GHz)	Wavelength (cm)	Magnetic field (T) for g=2
L	1.2	24.99	0.043
S	3.2	9.37	0.114
X	9.4	3.19	0.336
K	25	1.20	0.892
Q	35	0.88	1.249
W	95	0.32	3.394
D	135	0.22	4.800

Table I.3 : Properties of different frequencies used in EPR spectroscopy.

Some modern spectrometers offer several bands available in the single instruments. Such multifrequency EPR spectrometers provide an experimental route to study versatile interactions in paramagnetic materials. In terms of frequency used in the source, the spectrometers can be divided into high and low frequency types. The advantage of using high frequency EPR spectrometer ($\nu > 40GHz$) is the possibility of the study of small samples and the discrimination between similar species in addition to the study of both very fast (nanosecond range) and very slow motions. On the other hand, low frequency EPR spectrometers have the capacity to obtain spectra from living animals and organs, including the concentration of oxygen in tissues, redox metabolism, and distribution of spin-labeled molecules.

The frequency of the source can vary in a small interval and can be adjusted in the beginning of the experiment. Three important components are present between the source and the cavity (where the sample is placed): isolator, attenuator and circulator. The output power of the microwave source can be regulated by the attenuator that can precisely and accurately control the power that the sample sees. Circulator directs the microwave power from the source to the cavity and simultaneously redirects the reflected power to the detector. Isolator protects the source from the reflected microwave energy since EPR spectrometers are

reflection spectrometers that measure the amount of radiation reflected back from the microwave cavity containing the sample. Thus, reflected microwaves return to the circulator, where they are directed to a detector. The diode detector converts the microwave power to electric current since it is a silicon crystal diode in contact with a tungsten wire. At a low power level, less than 1 microwatt, the diode current is proportional to the microwave power and the detector is called square law detector. At higher power level, greater than 1 milliwatt, the diode current is proportional to the square root of the microwave power and the detector is called linear detector. The diode should operate in the linear region in order to obtain optimal sensitivity and quantitative signal intensity measurements. A reference arm, an additional element in the microwave bridge, supplies the detector with additional microwave power controlled by a second attenuator for optimal performance of the detector. A phase shifter is present to insure that the reference arm microwaves are in phase with the reflected microwaves when the two signals meet at the detector. The detector diodes are sensitive and can be damaged in the case of excess current. A circuitry protection in the bridge monitors the current and when the current exceeds 400 microamperes, the bridge lowers the microwave power level reducing the damages and improper operating procedures.

I.5.2 Cavity

The wave-guides or coaxial cables transmit the microwaves to the cavity where the sample is placed and the microwaves are concentrated. Automatic frequency control (AFC) locks the source at the cavity's resonant frequency. Microwave cavity is a metal box with rectangular or cylindrical shape that resonates with the microwaves same as organ pipe resonates with sound waves. At resonance, the frequency of the cavity microwaves is not reflected back but remains in the cavity. Microwaves generate electrical currents in the walls of the cavity, hence, energy is lost in the walls in the form of heat. The cavity, also called resonator, is characterized by the quality factor Q that indicates the efficiency of the cavity in microwave energy storage. The sensitivity of the spectrometer increases as Q increases. Q factor can be expressed as

$$Q = \frac{\nu_{res}}{\Delta\nu} \tag{I.29}$$

where ν_{res} is the resonant frequency of the cavity and $\Delta\nu$ is the width of the line at half height of the resonance (see Figure I.13).

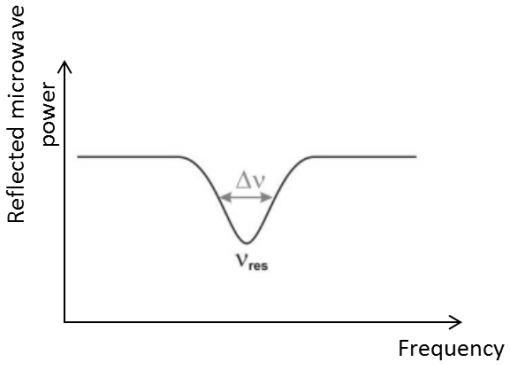


Figure I.13 : Reflected microwave power as a function of the frequency produced by the microwave source. Based on [39].

Standing wave is created in the cavity as a consequence of the resonance. These standing electromagnetic waves have electric field E_1 and magnetic field B_1 out of phase (magnetic field is maximum, electric field is minimum and vice versa). The spectrometer is designed for an optimal placement of the sample in the cavity, where the electric field is minimum and the magnetic field is maximum in order to obtain the best signal with the highest sensitivity. The microwaves are coupled in the cavity via an opening called iris. The size of the iris controls the amount of microwaves that enter and are reflected back from the cavity by matching the resistance of the waves in the cavity and the waveguide. The iris screw in front of the iris allows the adjustment of the wave matching by moving the screw up and down and changing the size of the iris (Figure I.14).

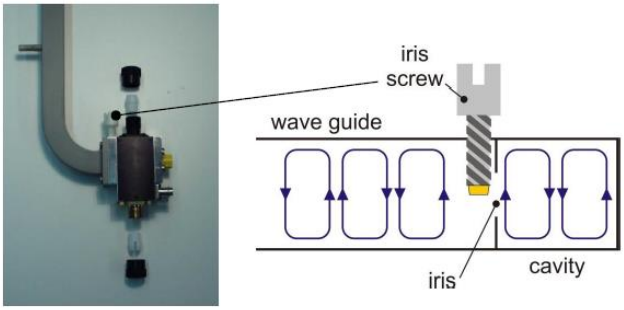


Figure I.14 : The waveguide and the cavity. Based on [39].

Prior to the beginning of the measurements, the system is tuned and the cavity is coupled. During the measurements, Q is lowered when the sample absorbs the microwave

energy and therefore, the coupling changes and the losses increase. The cavity is not coupled and the microwaves will be reflected back to the bridge giving an EPR signal.

I.5.3 Magnet

During EPR measurements, the sample and the cavity are subjected to a magnetic field (in Gauss) generated by the magnet. In practice, this function is provided by an electromagnet which is supplied by stabilized voltage. Since the key role of the electromagnet is ensuring the sweep capability combined with high resolution of magnetic field, high-class and matched power supplies with the superior linear current output are required. Moreover, the magnet should exhibit high field homogeneity over a large volume and be adjusted to the microwave source band to meet the resonance condition.

I.5.4 Signal channel

After passing through the detector, the amplifier amplifies the registered signal without changing the signal-to-noise ratio. Phase sensitive detection enhances the sensitivity of the spectrometer, decreases the noise from the detection diode and eliminates the baseline instabilities due to the drift in DC electronics. This is achieved by modulating the magnetic field coming from the magnet by the modulation Helmholtz coils that are placed along the direction of the field, at each side of the cavity. The field is modulated sinusoidally at a set modulation frequency. The field modulation quickly sweeps through part of the signal in the case of an EPR signal and the microwaves reflected back from the cavity are amplitude modulated at the same frequency. The signals that are amplitude modulated are only detected; however, other signals that do not fulfill these requirements are suppressed. For an EPR signal that is approximately linear over an interval as wide as the modulation amplitude, the EPR signal is transformed into a sine wave with an amplitude proportional to the slope of the signal at the resonance conditions (Figure I.15). A time constant is used to filter out the noise in order to improve the sensitivity (see also paragraph I.5.6). The processed signal is digitized and a computer is used for recording and analyzing data. As a result, a first derivative of the signal is measured.

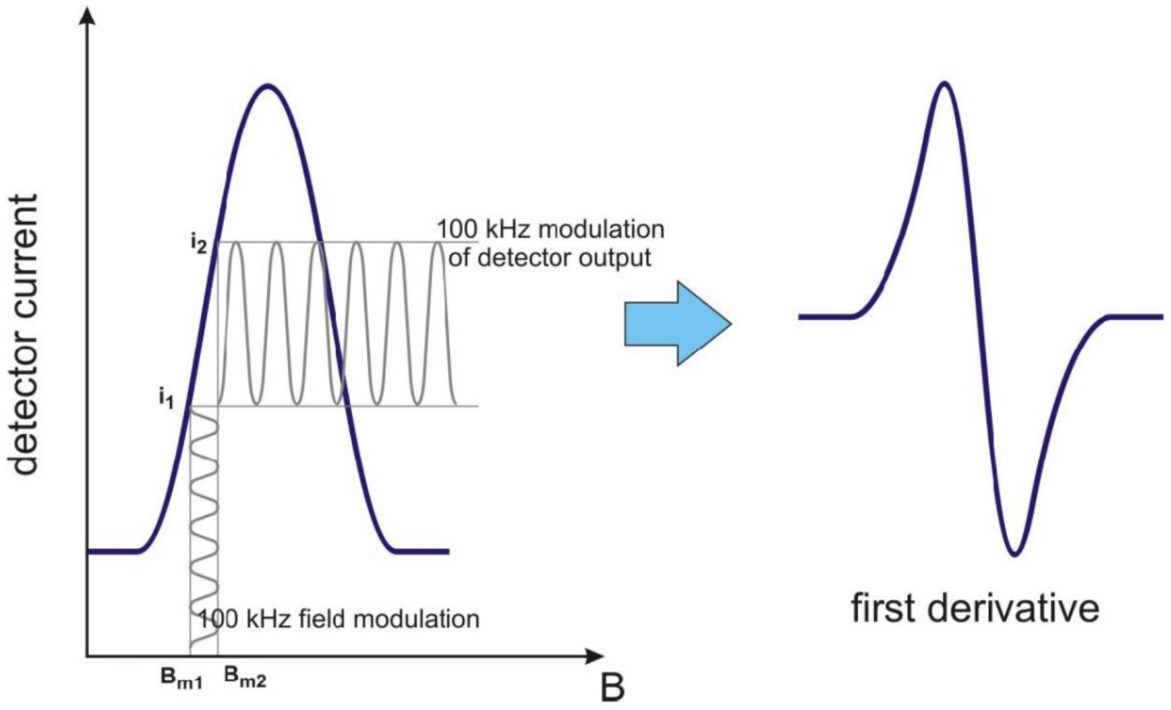


Figure I.15 : Schematic representation of phase sensitive detection. As the main field $\boldsymbol{B}$ is scanned through the EPR line, a small additional oscillating magnetic field $\boldsymbol{B_m}$ (commonly at 100 kHz) is applied in the same direction. As $\boldsymbol{B_m}$ increases between $\boldsymbol{B_{m1}}$ and $\boldsymbol{B_{m2}}$ the detector output increases from $\boldsymbol{i_1}$ to $\boldsymbol{i_2}$. When the magnitude $\boldsymbol{B_m}$ is small relative to the line width, the detector oscillating at 100 kHz has a peak-to-peak that approximates the slope of the absorption curve. Therefore the output is the derivative of the absorption curve [40].

I.5.5 Sample

The materials adapted to be measured in the EPR spectrometer can be solid-powder or liquid solution. Many samples can be measured at room temperature, but in some cases lower temperatures are needed. In such a case, a cryostat is typically fitted to the spectrometer and liquid and gaseous helium flows through the system in response to the working of the flow pump. Various types of cuvettes for solid, liquid and gaseous sample are in use. These cuvettes are made up of quartz that is transparent in EPR spectroscopy. Capillaries and flat cells are often used for liquid samples, whereas sample tubes can be used for both solid and liquid samples.

I.5.6 Basic working parameters

To get valuable signal in real experiment, many instrumental factors need to be optimized and carefully adjusted. Some of them are usually optimized and fixed by the producent of spectrometer, whereas others must be adjusted by the user to match the desired experimental conditions. The majority of such parameters govern the experiment and are easily accessible in the modern spectrometers via dedicated software installed on the computer connected to the spectrometer. Such software allows choosing the region of the field of spectra which is determined by two parameters: the *center field* and *sweep width*. The former is the chosen magnetic field in the middle of the spectra range, whereas the latter is as wide as the width of the signal. These values are different for various types of samples. The wrong setting of some parameters may lead to a deformation or complete/total loss of the signal. *Receiver gain* is a parameter that controls the gain of the digitizing detected signal and depends on its strength. The receiver gain must not be too low, as in Figure I.16b, and neither too high in Figure I.16c.

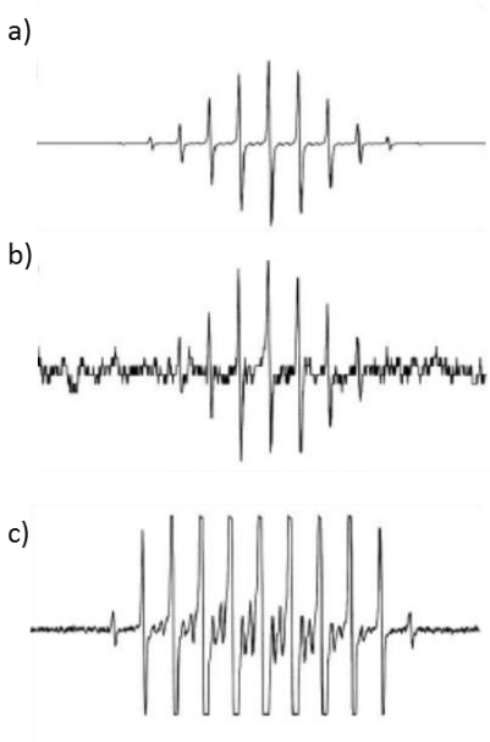


Figure I.16 : Effect of the receiver gain on EPR spectrum. (a) optimal, (b) too low, and (c) too high. Based on [39].

Modulation amplitude and *modulation frequency* are parameters associated to the phase sensitive detection. In general, increasing the modulation amplitude increases the EPR signal till a certain point. When the modulation amplitude is too large, larger than the line width of the EPR signal, the detected signal broadens and distorts the resonance line. On the other hand, low modulation amplitude can result in poor signal-to-noise ratio. A good compromise between the signal intensity and signal distortion is when the modulation amplitude is equal to the width of an EPR spectrum. In addition, if the modulation amplitude is greater than the splitting between two peaks, the two signals cannot be resolved (Figure I.17). Modulation frequency is normally 100 kHz.

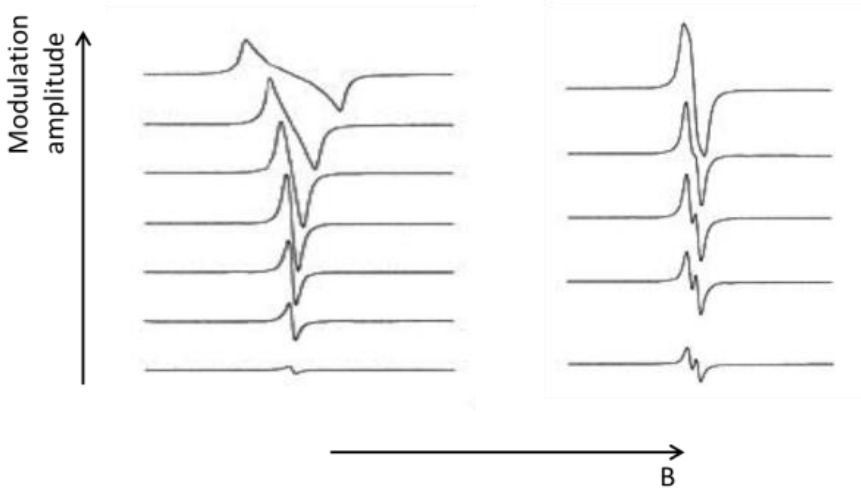


Figure I.17 : Effect of increasing modulation amplitude on signal intensity and signal shape. Based on [39].

Time constant is a parameter that controls the sensitivity of the signal. It filters out the noise by slowing down the response time of the spectrometer. In general, as the time constant is increased, the noise level decreases (Figure I.8). However, the chosen time constant must be suitable for the rate at which the magnetic field is scanned. Some distortion and even unwanted filtration of the main signal can occur as a consequence of a too long time constant with respect to the rate of the magnetic field scan. In experimental practice, the time needed to scan a single EPR signal should be ten times greater than the length of the time constant.

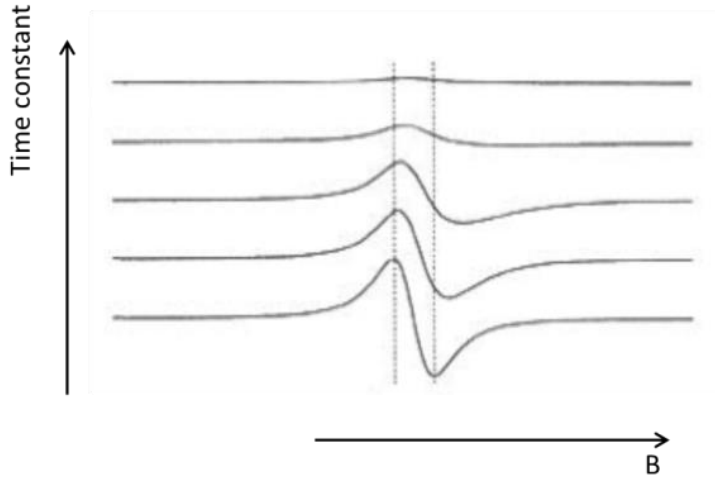


Figure I.18 : Effect of increasing the time constant on the signal intensity and the signal shape. Based on [39].

Conversion time is a parameter related to the time constant. The transfer of the EPR signal to the digital data acquisition system takes place in the signal channel and it incorporates an integrating ADC (analog to digital converter). During this conversion, noise is removed from the signal. In order not to obtain a distorted or filtered out signal, the time constant is set equal to the conversion time.

Signal averaging is equivalent to acquiring a spectrum with a long scan time and longtime constant. It is used to increase the signal to noise ratio.

EPR signal intensity is proportional to the square root of the power. As the power increases, the intensity of the signal increases until saturation where the signal broadens and becomes weaker. Therefore, for an optimal EPR intensity signal, several microwave power must be tried (*Microwave power level*).

I.6 EPR imaging

Besides the pure spectral information obtained for studied materials by EPR spectroscopy the spatial information is often desired. In spite of advances in EPR technology, it seems that the microscopic objects cannot be resolved by this method because EPR wavelength is too big (tens of micrometers or bigger). The situation changes when electron

paramagnetic resonance imaging (EPRI) of paramagnetic atoms in objects can be performed using magnetic field gradients [41].

The imaging of spin distribution by means of magnetic resonance methods has developed rapidly since 1973 when Paul Lauterbur described a method for generating magnetic resonance (MR) images using frequency encoding by application of the magnetic field gradient [42]. A few years later, in 1979, Hoch and Day first demonstrated the feasibility of EPRI when they described the distribution of color centers in natural diamonds [43]. Since that time, a great deal of effort was made to develop this technique due to numerous potential applications [44, 45]. EPRI is a tool for investigating aspects for many applications in material science including polymers, ceramics, semiconductors, inorganic materials as well as their defects and impurities [46-48]. However, the strongest driving force in the development of EPRI has been the possibility to apply this technique in medicine and biological studies for *in vivo* imaging [49-54].

I.6.1 Principles of EPR imaging

As mentioned above, the key concept of EPRI is based on the use of magnetic field gradient. Classical imaging by spatial scanning of the investigated volume with the simultaneous detection of the spectral response is not practical in case of EPR due to the long wavelengths used in this technique. Instead, the linear gradient of magnetic field can be applied to observe spatial density of the spins. The concept of the gradual field in EPR is presented in Figure I.19. If we consider four identical paramagnetic centers distributed in 2D space under homogeneous magnetic field B_0, the EPR signal coming from the entire space consists only of one peak corresponding to resonance condition $B_0 = B_{res}$ (Figure I.19). In this case, the location of individual centers is not distinguishable and the EPR signal does not contain any spatial information.

On the other hand, when linear gradient of magnetic field is applied, the differences between some centers can be observed. This is clearly visible in Figure I.19B when the linear gradient of the magnetic field alongside *x* axis is introduced. In such a case, the paramagnetic centers 1 and 3 are located in different magnetic field than spins 2 and 4. Consequently, the resonance condition is different for two types of location ($B_{res} = B_0 + x \; or \; B_0 -$

x) according to x axis and two peaks appears in the EPR signal. Thus, it can be said that this signal contains some spatial information of the system.

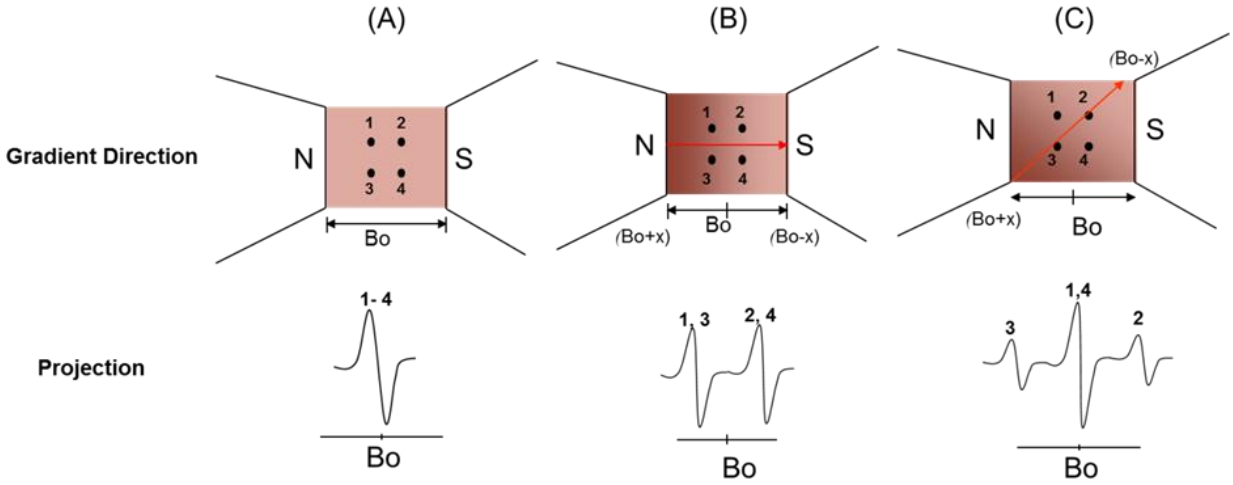


Figure I.19 : The concept of gradient of magnetic field in EPR imaging.

Namely, it shows that some paramagnetic centers are present in two different locations and equal amplitudes of the individual peaks indicates that spin density in these locations is equal. If the orientation of the gradient changes, additional information can be obtained. It is shown in Figure I.19C where the gradient is applied in diagonal direction of *xy* plane. Then, the centers 1 and 4 are in the same field, giving single peak in the spectrum, whereas the spins 2 and 3 meet the resonance conditions at different values of the magnetic field. As a result, three different peaks appear in the EPR spectrum.

If we apply high number of projections for different orientations of the gradual magnetic field, the useful spatial information can be encoded in the set of EPR spectra. Standard spatial EPR imaging in CW mode uses the magnetic field gradient at constant magnitude. After registration of the spectrum, a gradient orientation rotates by a fixed angle, depending on the number of projections e.g. within 0-180 degree range for 2D imaging. In order to extract the spatial information, some back-projection algorithms have to be applied to re-construct the image (Figure I.20) [55].

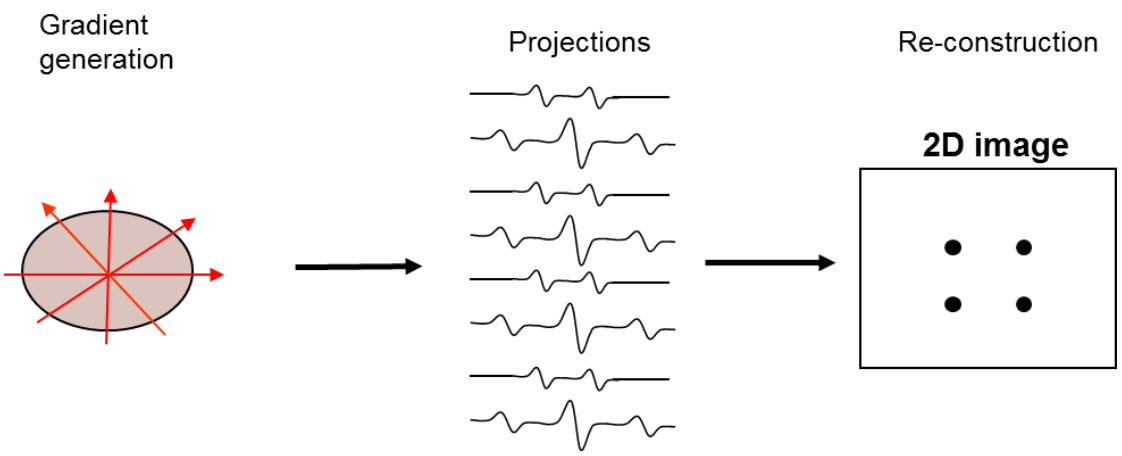


Figure I.20 : The main steps of EPR spatial imaging.

In other words, in EPR imaging experiments, the microwave power is absorbed by the unpaired electron located at point *x* when the resonance condition is fulfilled

$$\nu = (\frac{g\mu_e}{h})(B_{res} + xG_x)$$

ν is the microwave frequency, B_{res} is the resonant magnetic field, G_x is the linear magnetic field gradient (in $Gauss\ cm^{-1}$). Thus, the field gradients produce a correspondence between spin location and B_{res}.

I.6.2 Spectral-spatial imaging

As shown above, EPRI in the spatial domain (1D, 2D or 3D) provides distribution of the paramagnetic centers in the studied objects. However, it does not give any information regarding spectral shape because it is assumed to be invariant throughout the object during the spatial reconstruction procedure. Thus, for samples having complex hyperfine structure or multiple radical species, it is not possible to obtain an accurate map of spin distribution because the information obtained is limited only to the spin content and not the nature of the spins in the spatial volume element. To acquire more complete information of the system, an additional spectral dimension is, thus, required in EPRI. The spectral dimension can be added to spatial images by collecting projections along a given direction of the magnetic field gradient at various gradient magnitudes [56, 57]. Thus, sampling gradient space is realized by the use of not only different gradients orientations (as in pure spatial EPRI) but also different amplitudes. We can also say that the change of the gradient amplitude in spectral-spatial

image is mathematically equivalent to the rotation of the projection direction in spatial image (Figure I.21). In the most general case, combination of a 2D spatial and spectral acquisition gives a 2D spectral-spatial image in which EPR line shape is measured in every voxel.

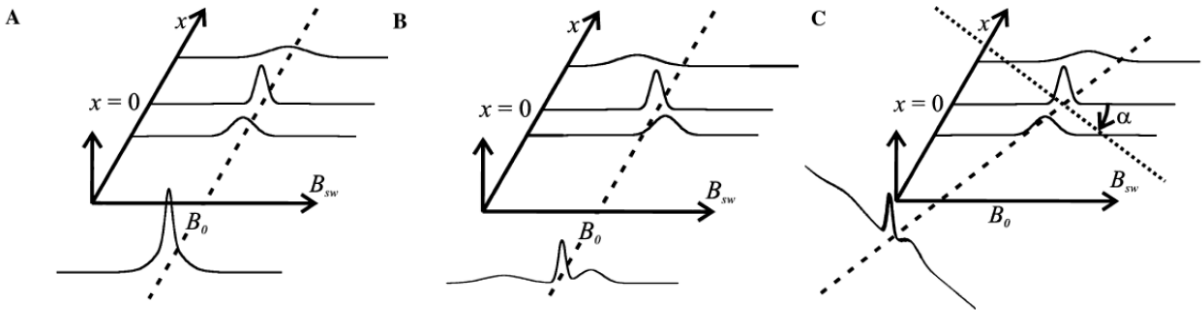


Figure I.21 : The effect of a linear magnetic field gradient on spectra recorded for a 2D spectral–spatial object. (A) Without an applied magnetic field gradient, the recorded spectrum is the sum of the individual spatial components or a projection of the object onto the spectral axis.(B) When a small positive gradient is applied, the individual spatial components pass through resonance at values of B that depend on their spatial positions and the components are resolved in the recorded spectrum. (C) With proper scaling, the spectrum shown in (B) as a projection of the spectral–spatial object with projection angle α [58].

I.6.3 Instrumentation and data analysis in EPRI

EPR imaging system is a continuous wave EPR spectrometer with small modifications, one of which is the gradient coils that are fixed on the poles of the magnet. EPR spectrum is measured by scanning the Zeeman magnetic field in the z direction and the gradient coils supply gradients along x $(\delta B_z/\delta x)$, along y $(\delta B_z/\delta y)$, or along z $(\delta B_z/\delta z)$. The coils supply a linear field gradient in the direction parallel to the external magnetic field (z-axis). Typical EPRI systems exhibit the gradient fields of up to 150 G/cm [53].

The important challenge in EPRI is the need for a resonator design of minimum thickness, which makes it possible to achieve higher magnetic field gradients for a given coil's driving power. Thinner resonators also enable multidimensional gradient coils to be placed in the gap. Loop–gap resonators (LGR) provide straightforward design and high filling factors [53].

The frequency used in EPRI system depends on the type and size of the studied objects. For *in vivo* applications, typically, low frequencies are used because they provide deeper penetration (see Table I.4). However, lower frequencies causes simultaneously worse spectral resolution and lower sensitivity. Moreover, the use of some bands is limited in case of samples containing a lot of water. For example, X-band is not practical to study biological samples because most of the microwave field is absorbed in a non-resonance way and is transformed to the heat. Thus, the choice of the working band is always a compromise between different factors. Therefore, most EPRI experiments in material science are performed at X-band (9 GHz), whereas ESRI is applied *in vivo* studies where experiments are performed at lower frequencies (below 2 GHz) that can accommodate large samples with high water content.

Frequency	300 MHz	750 MHz	1GHz	3GHz	9GHz
Penetration depth	> 10 cm	6-8 cm	1-2 cm	~ 3 mm	~ 1 mm

Table I.4 : Penetration depth in function of source frequency in EPR.

One of the biggest challenges in the software part of EPRI is fast and effective image reconstruction. In the simplest case, the image can be reconstructed by back-projection method. The idea of this technique is schematically presented in Figure I.22. After performing a number of projections under different gradient orientation, the back-projection procedure is applied and the total intensity of the signal in every pixel of 2D image is summarized. In practice, much more complicated filtered back-projection algorithms are needed to remove the hyperfine-based image artifacts and to enhance the image resolution. Several methods for these purposes have been developed during the last few decades [59, 60].

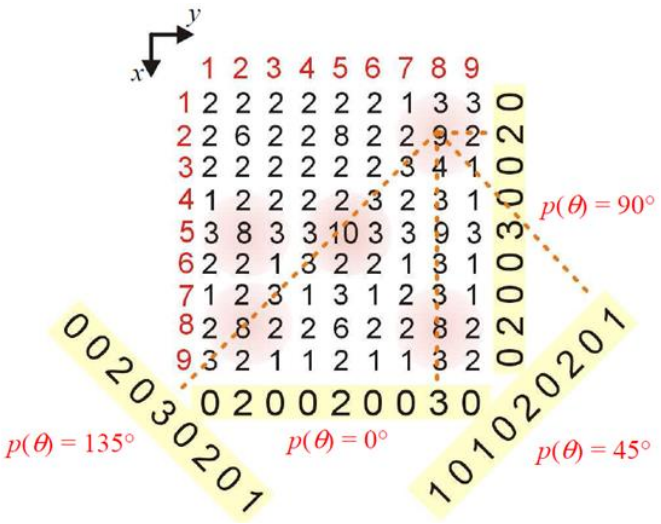


Figure I.22 : Reconstruction of the 2D image by back-projection method.

I.7 Conclusion

Electron paramagnetic resonance (EPR) spectroscopy is employed to study atoms or molecules possessing one unpaired electron in their valence e.g. free radicals, transition metals and other paramagnetic substances most commonly in liquids and solids. EPR is a characterization method used in several fields for instance geology, material science, environmental science, biology, and many others. EPR data provide a wealth of information of the samples not available by other techniques. They are, for instance, fine and hyperfine; g-factor; line-shape or saturation. All these kinds of information can serve as fingerprints of the paramagnetic species for unequivocal identification. Moreover, EPR have been extended to its imaging form (EPRI) what dynamically multiplied the number of possible applications. The development of low frequency EPR instrumentation at L-band 1–2GHz, or lower frequencies, and lumped circuit resonators has made it possible to perform EPR measurements on lossy biological samples.

The amazing richness of experimental data obtained by EPR is often associated with their complexity. Many EPR spectra and images exhibit very complicated structure due to the overlap of numerous lines of different shapes. Thus, fast and efficient analysis of such data is nowadays a real challenge for spectroscopists. For this reason, the modern effective data analysis methods are mandatory for the EPR domain. The next chapter is dedicated to this issue.

Chapter II: Chemometrics

II.1 Introduction

Analytical chemistry provides a vast number of samples that are multicomponent systems. These samples are far from simple and contain more than one component to be interpreted simultaneously. Therefore, more complex instrumentation is needed to cope with these systems in an efficient way as well as tools are needed to interpret and identify the information obtained. Advances in technology and the increasing availability of potent instrumentation offer now the possibility to obtain high amounts of data on each sample analyzed in a reasonable, often negligible time frame. In the case of spectroscopic techniques, rapid analysis of a sample gives a spectrum that can be considered as a data vector in which the order of the variable e.g. absorbance at consecutive wavelengths or magnetic fields has a physical meaning. In other cases, a set of samples can be described by a number of chemical and physical parameters contributing a spectra for each sample. Despite the differences in compositional variation, complexity and origin, the raw measurements recorded in a multicomponent chemical system can be often described with a simple model consisting of the composition-weighted sum of signals of their pure components. A set of tools for extracting information from chemical system by data-driven procedures makes up a discipline called chemometrics [61].

A number of steps are required to extract and properly interpret the potential information embodied within the data. A deep understanding of the nature of the analytical data is the basic step for any proper data treatment because different data types usually require different processing strategies that closely depend on their nature. For this reason, the data analyst should be aware of the problem under study i.e. sampling to instrumental analysis. Such knowledge makes the difference between the mathematician and the chemometrician. Chemometrician is a chemist who is acquainted with experimental data and utilizes mathematical methods for the conversion of numerical records into relevant chemical information [62].

The analytical food chemist William Sealy Gosset (1876-1937) is considered as one of the fathers of chemometrics. In fact, he worked in a company where he studied a number of statistical tools and adapted them to better solve actual chemical problems. His studies were presented using a pseudonym since his company did not permit him to publish any data. As a modest contributor in the field, he adopted a pen name Student. His most famous work is on

the definition of the probability distribution that is usually referred to as the Student's *t* distribution [63].

In 1972, the notation of chemometrics was introduced by Svante Wold, for the first time, for identifying the discipline that performs the extraction of useful chemical information from complex experimental systems [2]. Two years later, the International Chemometrics Society was founded and led to the first description of the discipline. Chemometrics is a chemical discipline that uses mathematics, statistics, and formal logic in order 1) to select optimal experimental procedures 2) to provide maximum relevant chemical information by analyzing chemical data and 3) to obtain knowledge about chemical systems [61, 64]. This discipline is strongly related to the use of computers in chemistry. The approach is what all analytical chemists try to do i.e. to design optimal analytical procedures and to try to obtain as much as possible information from the results with the difference that chemometrics does it with the help of mathematical methods. Chemometrics can be useful to scientists including biochemists, organic chemists, pharmaceutical chemists, medical doctors, pharmacists as well as biologists that apply a chemical discipline and consequently perform a chemical analysis.

In general, starting with a certain chemical knowledge, scientists define a hypothesis. They need experimental data to be able to test the hypothesis and to verify its validity. Therefore, they need to decide which experiment to carry out that is the first point of the definition of chemometrics. Chemometrician's approach is the use of mathematical and statistical techniques i.e. experimental design methodology. Second point of the definition is that chemometricians use the data generated by the experiments to extract information, deriving a model by computing a regression equation that describes how the result of the measurement is related to the experimental variables. Chemists use information and their scientific background to generate more knowledge about the system studied which is third point in the definition [65].

II.2 Measurements and data

Depending on the situation, the system studied and the instrument used, a measurement can give a scalar x (1×1), a vector $\mathbf{d}$ $(1 \times n)$, a matrix $\mathbf{D}$ $(m \times n)$, or a cube $\mathbf{X}$ $(x \times y \times n)$ (Figure II.1). The single spectroscopic measurement gives a vector $\mathbf{d}$ with n elements that

corresponds to *n* different wavelengths or magnetic fields. Such a vector is traditionally called spectrum. In this chapter, we will consider spectra obtained by EPR spectroscopy, therefore *n* is magnetic field. The matrix **D** can be easily obtained by regrouping the different *m* spectra registered in the experiment. The series of complete spectra can be acquired in an inherent order (time, pH,…) e.g. spectroscopic data evolved in time (kinetics) where each spectrum is registered at a different time *t*. On the other hand, series of independent samples can be obtained where the order of the data does not play a role. An image sample is often displayed by spectroscopies as a cube of three dimensional structure containing two spatial axes (*x*- and *y*-) and one spectral axis (*n*-). A single scalar data element in the cube is called pixel and for each pixel a spectrum is provided along the different *n* magnetic fields.

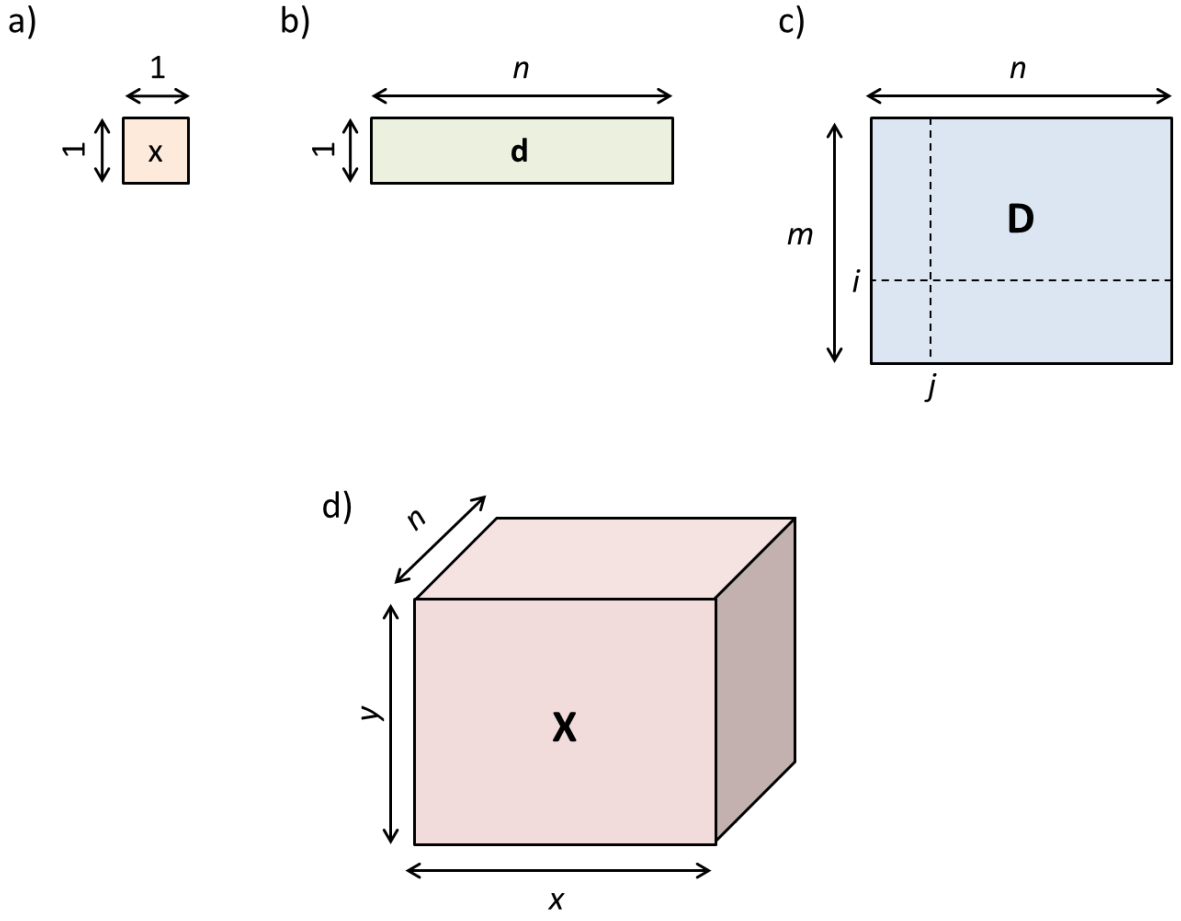


Figure II.1 : Schematic representation of a measurement dataset acquired under the form of a) scalar, b) vector, c) matrix and d) cube.

II.3 Chemometric methods

Statistics offer a number of useful tools that can convert data into information. They make up the chemometric methods among which some are univariate and others are multivariate. Univariate methods, as the name indicates, consider one variable at a time. Nevertheless, these methods usually supply only partial answers to the problem under study since they underutilize the potential for discovering the global information present in the data. For example, univariate methods are not able to take into account the inter-correlation between variables that is an informative feature if it was recognized and properly interpreted. On the other hand, multivariate strategies allow more complete interpretation of the data structures [5-9]. Multivariate methods can be principally divided into two main groups one of which is based on linear model, whereas the other is based on non-linear model. Furthermore, each main group can be further divided into supervised and unsupervised. The first ones look for determined features within data, explicitly oriented to address particular issues. Conversely, unsupervised methods show the data as they are and require the sole use of the spectral data. Some methods give qualitative information others give quantitative. Several examples of each group of multivariate chemometric methods are shown in the diagram of Figure II.2. *Positive Matrix Factorization* (PMF) [66], *Independent Component Analysis* (ICA) [67], *Principal Component Analysis* (PCA) [68] and *Multivariate Curve Resolution-Alternating Least Squares* (MCR-ALS) [4] are good examples of linear and unsupervised analysis. *Partial Least Squares* (PLS) [69], *Principal Component Regression* (PCR) [70], *Soft Independent Modeling of Class Analogy* (SIMCA) [71], *Linear Discriminant Analysis* (LDA) [72] and *K Nearest Neighbors* (K-NN) [73] are examples of linear supervised tools. Similarly, chemometric methods based on non-linear model are divided into unsupervised analysis i.e. *Kohonen Artificial Neural Network* [74] and supervised tools *Multilayer Artificial Neural Network* [75].

In the next paragraphs, two multivariate and unsupervised exploratory techniques based on linear model for data analysis that were used in this thesis i.e. PCA and MCR-ALS, will be discussed in details. It is the first approach between the tools cited previously and EPR spectroscopy.

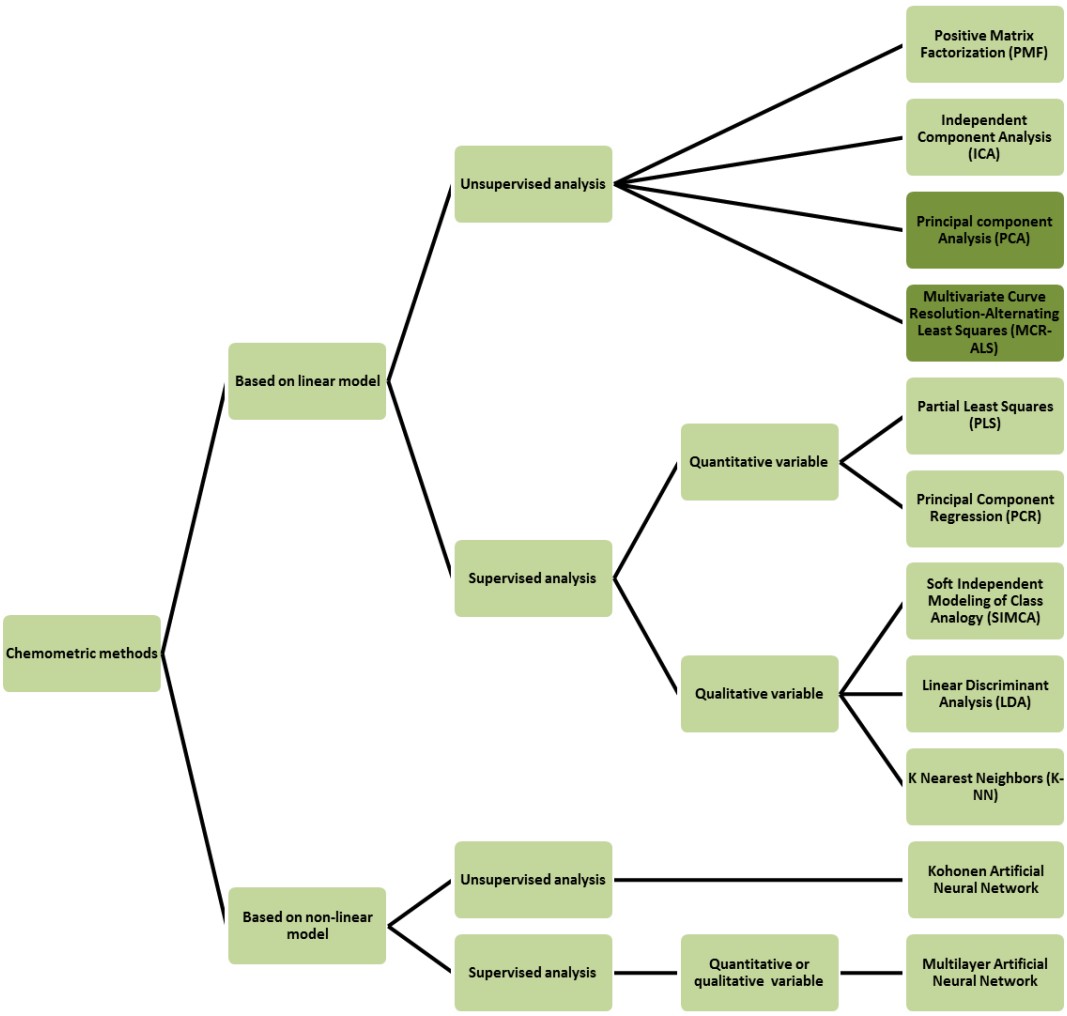


Figure II.2 : Diagram representing different chemometric methods.

II.3.1 Principal Component Analysis (PCA)

Principal Component Analysis is a statistical technique that linearly transforms an original data set of possibly correlated variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables. The goal of this technique is to reduce the dimensionality of the original data set in which the uncorrelated variables are much easier to understand and to use in further analysis [68].

II.3.1.1 History

Principal Component Analysis (PCA) is the oldest and mostly used algorithm in the unsupervised classification techniques. The concept was first formulated in 1901 by the British mathematician Carl Pearson who was concerned with finding lines and planes that best fit a set of points in a *p*-dimensional space [3]. 50 years before the widespread of computers, Pearson's optimistic comments regarding computation were interesting. He confirms that his methods can be easily applied to numerical problems and calculations become complicated (although possible to perform) with four and more variables. However, the entire potential of the technique was not fully exploited until the advent of computers. The dynamic expansion of its use coincided with the widespread introduction of microelectronics.

Over twenty years after Pearson's work, Fisher and MacKenzie mentioned briefly PCA as more suitable method for the modelling of response than analysis of variance and they performed the two-way analysis of agricultural trail [76]. Further in their work, they outlined the Nonlinear Iterative Partial Least Squares Algorithm (NIPALS) that was adapted to chemometric field by Wold in 80s [77]. In 1933, Hotelling made the major impact on the method and developed PCA to its present stage [78]. In the same period, the development of Factor Analysis (FA) was initiated by Thurstone and other psychologists that is closely related to PCA [79]. Since then, the utility of PCA has been applied in many diverse scientific fields. In chemistry, this method under the name of Principal Component Analysis was introduced by Malinowski around 1960 and after 1970 a large number of chemical applications have been published by Malinowski et al. [80] and Ramos et al. [81]. Interestingly, in some other fields of science, PCA has lived a secluded life in the beginning and for example in geology was overshadowed by its twin brother Factor Analysis (FA) [82, 83]. It is worth adding that PCA concept goes under several names in different disciplines. For example *Singular Value Decomposition* (SVD) is used in numerical analysis [84, 85] and *Karhunen-Loéve expansion* in electrical engineering [86]. Besides, *Eigenvector Analysis* and *Characteristics Vector Analysis* are often used in physical sciences. In addition, the term *Hotelling transformation* is often used in image analysis and *Correspondence Analysis* is a name of special doubled-scaled variant of PCA that is favored in French-speaking countries and Canada.

II.3.1.2 Principle

Principal Component Analysis (PCA) is a way that identifies pattern in a data and highlights the similarities and differences among them [87]. More precisely, PCA is a mathematical manipulation of data matrix that compresses the data consisting of a large number of interrelated variables by reducing its dimensions, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, 'factors' or 'Principal Components (PC)', that are uncorrelated and ordered so that the first few retain most variation present in all of the original variables. The starting point of all multivariate data analyses is a data matrix (or data table) that possesses rows and columns. The former are also called 'objects' that represents the samples and the latter are also termed 'variables' that compromise the measurements made on the objects. PCA is a bilinear decomposition of the data matrix into smaller matrices called scores and loadings (Figure II.3).

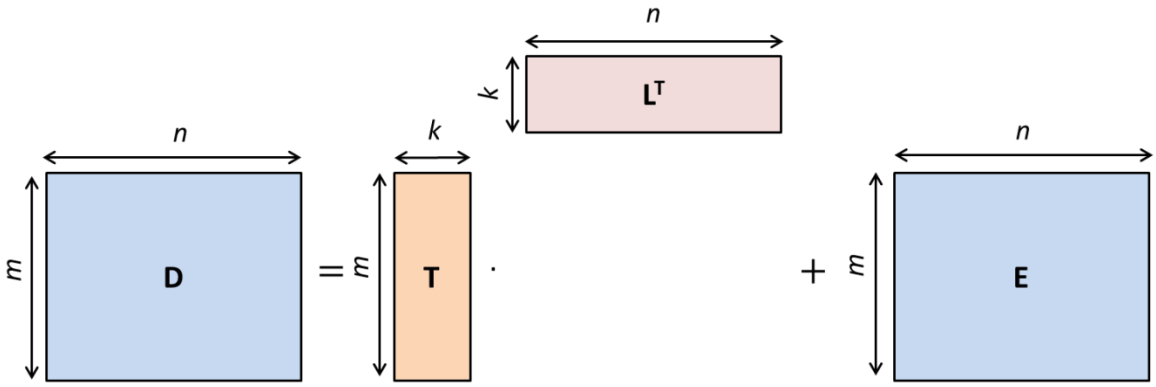


Figure II.3 : PCA decomposition the data (**D**) into scores (**T**), loadings (**L**) and residuals (**E**).

Therefore, new axes using factors instead of original variables are predefined and called as Principal Components. They are aligned with the maximum signal variance that is high amount of information. In other words, the eigenvector with the largest eigenvalue is the direction of greatest variation. The following equation shows the mathematical formula of Principal Component Analysis

$$\boldsymbol{D}_{(m\times n)} = \boldsymbol{T}_{(m\times k)}\, \boldsymbol{L}^{\boldsymbol{T}}_{(k\times n)} + \boldsymbol{E}_{(m\times n)} \qquad \text{(II.1)}$$

The matrix **D** represents the original data with *m* rows and *n* columns. **T** is the scores matrix with *m* rows and *k* columns and **L** is the loadings matrix with *k* rows and *n* columns. It is worthy to note that the number of *k* columns corresponds to the number of Principal Components. The latter can be selected by the user or can be proposed by the algorithm. The matrix **E** represents residuals often noise not included in the **T** and **L** decomposition or sometimes non-significant Principal Components.

PCA concept is schematically shown in Figure II.4. Considering a model plane spanning the space of the two variable system, the new bi-plot is formed of two orthogonal PCs. Samples are presented by a red cross in the space. The relationship between the samples is defined by the distance between them. If they are near, they are more similar than those that are further. Typically, in analytical chemistry, more than two varieties are studied, then, PCA describes the spread or variation of the distances in a few axes.

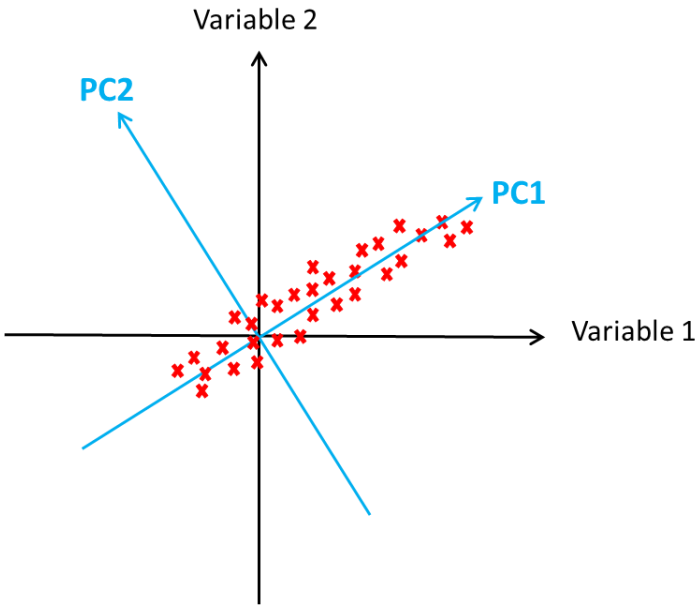


Figure II.4 : The first two Principal Components of a raw plot data in a two-measurement system.

In general, the first Principal Component (PC1) explains the maximum amount of variation in the dataset in one direction. The second Principal Component (PC2) accounts to the next most significant variation not explained by the first one and so on. All the Principal Components are perpendicular (orthogonal) to each other. In the context of a series of spectroscopic data set, even though spectra can have hundreds or thousands of magnetic fields (n), the relevant information is contained in a very small number of Principal Components. This is due to the redundant information in the signal for some magnetic fields. Therefore, although the number

of Principal Components that can be calculated is often large (because it is the smaller of the number of samples or variables), in practice only the first few ones describe spectral variations linked to the chemical composition of the spectra and the rest account for noise-related signal contributions.

Samples have coordinates in the original space as well as with respect to the new Principal Components. The coordinates of the samples relative to the Principal Components are the *scores*. Each PC is constructed from the linear combinations of the original measurement variables. If α is an angle between the variable original axes and the Principal Components (Figure II.5), the cosine α values are the *loadings* and can range between -1 and 1 describing the contribution of each axis to the Principal Component PC. The larger the cosine absolute value, the closer the two directions, therefore the larger the contribution of the original variable to PC.

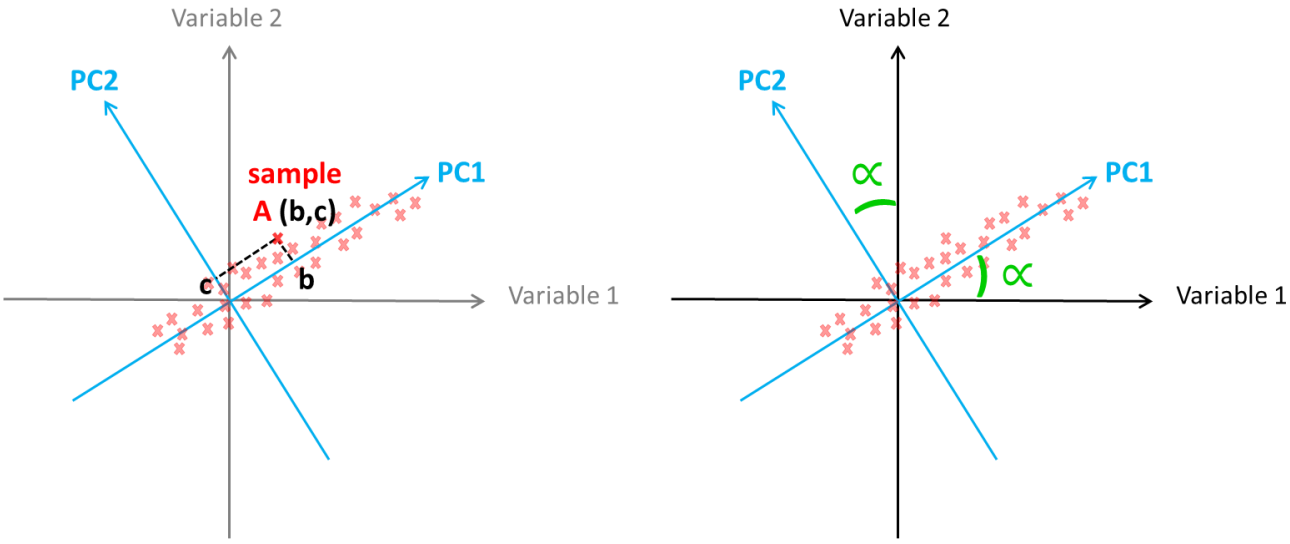


Figure II.5 : Graphical representation of scores (left) and loadings (right).

PCA can be practically realized by means of different mathematical operations in the computation. The simplest method is the Nonlinear Iterative Partial Least Squares (NIPALS) but in practice more powerful algorithms that are based on the matrix diagonalization are used [88]. One of the most popular method is the *Singular Value Decomposition* (SVD). In this chapter only this method among all used for PCA will be discussed.

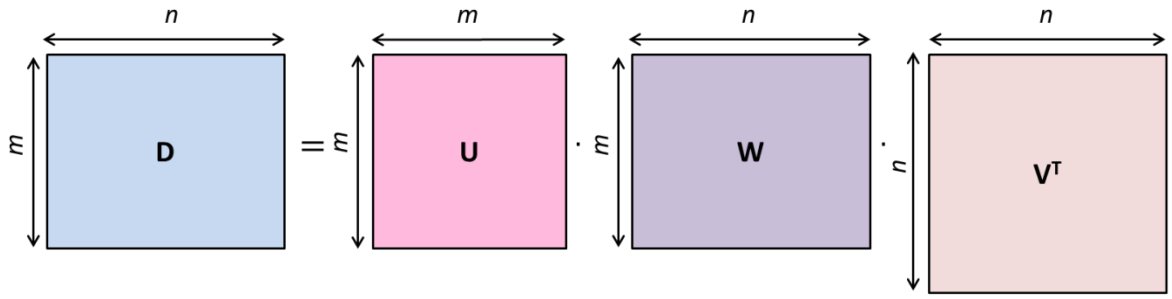


Figure II.6 : Schematic representation of PCA-SVD: decomposition of matrix **D**.

Singular value decomposition algorithm was first introduced by Beltrami (1873) and Jordan (1874) and was described in a modern way by Golub and Reinsch in 1970 [89]. SVD decomposes the data matrix, **D**, into the matrices coded as **U**, **W**, and **V**, where **U** and **V** are orthonormalized matrices such as $\mathbf{UU^T}=\mathbf{I}$ and $\mathbf{VV^T}=\mathbf{I}$. The eigenvectors of $\mathbf{DD^T}$ are the columns of the matrix **U** $(m \times m)$, whereas the eigenvectors of $\mathbf{D^TD}$ are the columns of matrix **V** $(n \times n)$. The matrix **W** $(m \times n)$ is called the diagonal matrix in which its diagonal entries are the non-negative square roots of eigenvalues also called singular values.

In the case where only k significant factors are considered, the product **U** and **W** from SVD correspond to the score matrix **T** in PCA. And the matrix **V** is equivalent to the loadings $\mathbf{L^T}$. Note that *k* can have a maximum value that is equal to *m* or *n* depending which one possesses the minimum value. Thus, data matrix **D** can be expressed as:

$$\boldsymbol{D}_{(m\times n)} = \boldsymbol{U}_{(m\times k)}\boldsymbol{W}_{(k\times k)}\,\boldsymbol{V}^{\boldsymbol{T}}_{(k\times n)} + \boldsymbol{E}_{(m\times n)} \qquad \text{(II.2)}$$

where **U** contains *k* orthonormalized eigenvectors, **W** is a diagonal matrix containing *k* singular values of **D**, and **V** contains *k* eigenvectors of **D**. **U** and **V** contain orthonormal column vectors and are ordered to reflect a decreasing sequence of singular values of **W**.

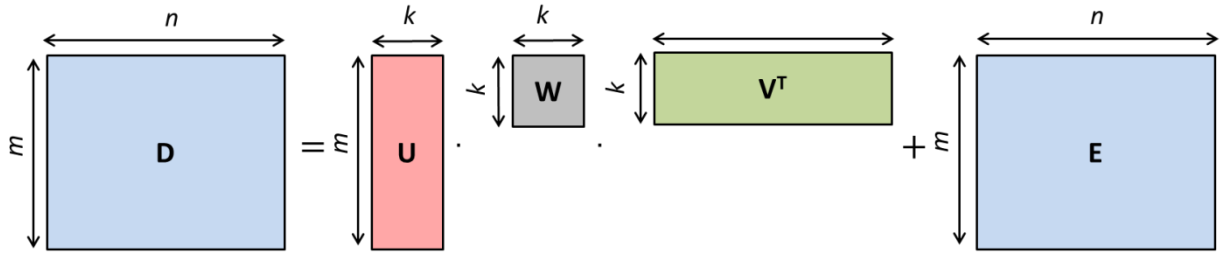


Figure II.7 : Schematic representation of PCA-SVD.

PCA performs the decomposition of the matrix containing the raw data under the very specific constraint of factor orthogonality on both scores and loadings, normalization on the loadings and maximum explained variance for successive extracted components. Under these conditions, PCA gives favorable mathematical property of unique solutions (one possible solution). Principal Components obtained do not reveal the profiles of pure components since PCs are calculated to keep a zero correlation between them (the loadings are completely uncorrelated), however, real spectra constituents often do not obey this condition. For this reason, Principal Components are never associated with real chemical pure components because the mathematical properties and chemical meaning of both of them are different. Nevertheless, PCA provides useful information about the studied system. It is an exploratory tool that is employed as a first approach to give an excellent overview of the system studied in smaller dimensions without losing the richness of information in the original measurements. It is used for the determination of the number of pure components, for feature and noise reduction purposes and constitutes the basis for other more complex pattern recognition that deals with more natural constraints.

Similarly to PCA, other methods exist that represent the raw data in a space of smaller dimensionality trying to keep all the relevant information. For instance, we can mention Independent Component Analysis (ICA) that relies on a bilinear decomposition and projects the data in a small space according to mathematical criteria [67, 90]. The concept of ICA is an extension of PCA where the former finds the linear representation of nongaussian data so that the components are statistically independent. In addition, it accounts for higher order statistics. However, the straightforward association of independent components with real compounds is not efficient unless the real spectra of the compounds obey the same mathematical conditions.

Unlike PCA and ICA, the unsupervised computational tool MCR-ALS iteratively derives the scores and the loadings through the empty modelling approach. The latter, described in more details in this chapter, look for solutions in agreement with more natural constraints such as non-negativity instead of orthogonality. MCR-ALS produces meaningful pure concentration and spectral profiles, however, the mathematical property of uniqueness is lost [91].

II.3.2 Multivariate curve resolution-alternating least squares (MCR-ALS)

Soft modelling approaches, such as multivariate curve resolution also known as Self-Modelling Curve Resolution (SMCR), are used more frequently because of their flexibility and describe processes without using explicitly the underlying chemical model linked to them [92]. In contrast with hard modelling, soft modelling avoids errors caused by the assumption of a wrong model and does not allow the modelling of chemical components that do not participate in the process under study [93]. There exists many soft modelling approaches such as *Positive Matrix Factorization* (PMF) [66], *Iterative Target Transformation Factor Analysis* (ITTFA) [94, 95], *Resolving Factor Analysis* (RFA) [96], *Principal Component Analysis* (PCA) [68], *Independent Component Analysis* (ICA) [67, 90] and *Multivariate Curve Resolution-Alternating Least Squares* (MCR-ALS) [4], the tool used in our thesis.

Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) belongs to the class of soft-modeling algorithms developed by Romà Tauler in 1995 [4]. It is a widespread chemometric tool that is designed to unravel pure compound information from unresolved and unknown experimental multicomponent systems. The method is based on describing a bilinear factor analysis model (table or data matrix) from multicomponent system [97]. Such a bilinear model can include many kinds of processes and mixtures such as chemical reactions, industrial processes, chromatographic elutions, spectroscopic images or environmental data. The purpose of MCR-ALS is the characterization of a multicomponent system basing on a relevant and sufficiently discriminating experimental output. The algorithm resolves the pure contributions of all species present in a mixed spectroscopic dataset providing the concentration profiles and related pure spectra without requiring any chemical information about the underlying physicochemical system such as number, evolution and identification of the species [98, 99]. This powerful approach uses solely the experimental measurements i.e. low demanding needs, no one of the pure components in the system has to be known beforehand and any information about the system may be used but is not essential.

MCR-ALS has been successfully used to analyze data sets from various spectroscopic techniques such as Raman spectroscopy [100-102], Raman imaging [101], near-infrared spectroscopy [103-105], infrared imaging [106] FT-IR [107-109], UV-Vis spectroscopy [110-113], NMR spectroscopy [114-117], high performance liquid chromatography [118-120], gas

chromatography [121-123], voltammetry [124], DNA microarrays [125], mass spectrometry [126] and many others.

II.3.2.1 Principle

Supposing that the data follow a bilinear model, a two-way data sets acquired by the spectroscopic techniques can be typically organized in a matrix **D** that is composed of *m* rows and *n* columns containing raw information about all the components present in the system. Considering a raw data that is composed of only two pure components e.g. 1 and 2 represented in matrices $\mathbf{D_1}$ and $\mathbf{D_2}$, respectively (Figure II.8).

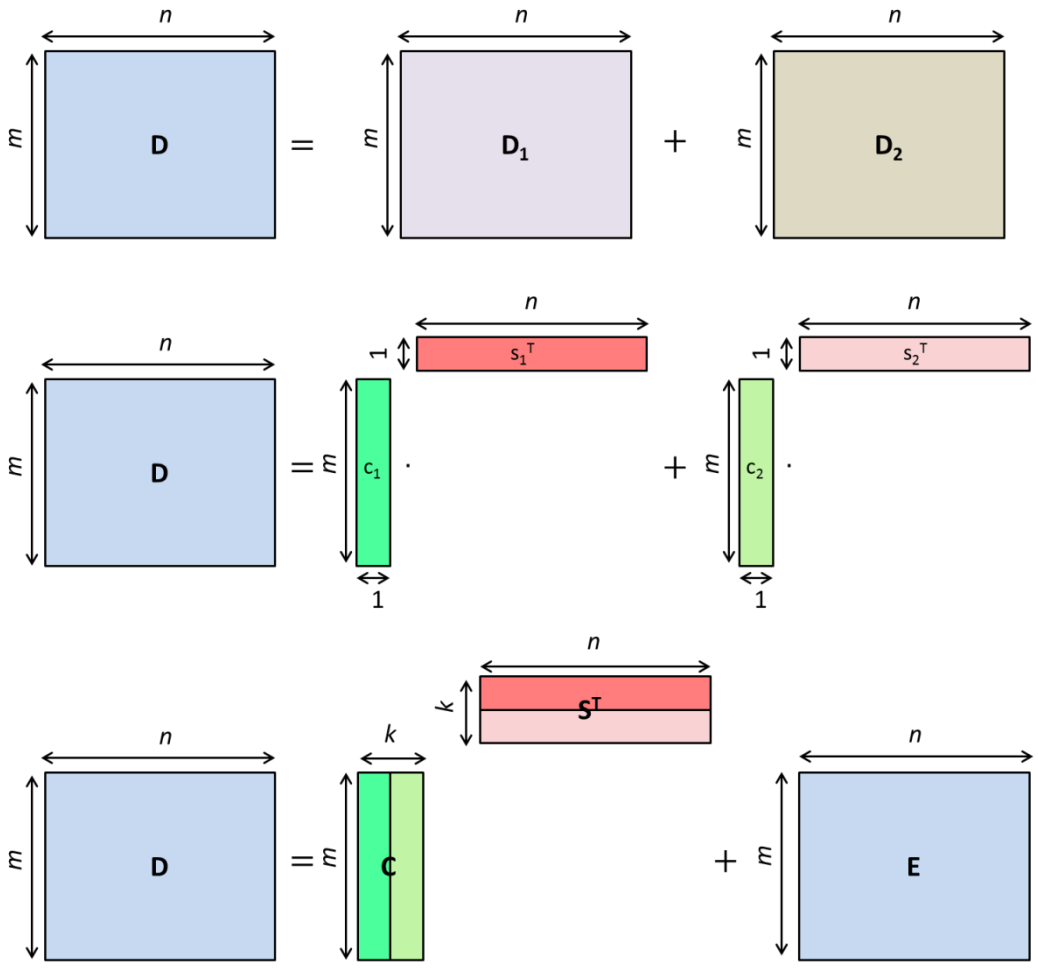


Figure II.8 : Bilinear decomposition of the matrix **D**.

In general, for a system containing *i* components each matrix is obtained from the dyad of

profiles i.e. concentrations c_i and spectra s_i^T.

$$\boldsymbol{D} = \sum_i \boldsymbol{D}_i + \boldsymbol{E} = \boldsymbol{D}_1 + \boldsymbol{D}_2 + \boldsymbol{E} \quad \text{(II.3)}$$

$$\boldsymbol{D} = \sum_i \boldsymbol{c}_i \boldsymbol{s}_i^{\boldsymbol{T}} + E = \sum_i \boldsymbol{c}_2 \boldsymbol{s}_2^{\boldsymbol{T}} + \boldsymbol{c}_2 \boldsymbol{s}_2^{\boldsymbol{T}} + \boldsymbol{E} \quad \text{(II.4)}$$

The additive model in the above equations can be expressed in a compact way grouping all the concentration and the spectral profiles in the matrices **C** and $\mathbf{S}^{\mathbf{T}}$, respectively.

$$\boldsymbol{D} = \boldsymbol{C}\boldsymbol{S}^{\boldsymbol{T}} + \boldsymbol{E} \quad \text{(II.5)}$$

The additive bilinear model is of analogous structure to the generalized law of Lambert–Beer in matrix form and is by extension the bilinear MCR model [127]. Therefore, the aim of a multivariate curve resolution (MCR) algorithm is providing a mathematical bilinear decomposition of a multivariate multicomponent experimental data matrix **D**, sized ($m \times n$), which represents the total instrumental response of the mixture samples in a system, into the product of two simpler matrices **C**, sized ($m \times k$) and $\mathbf{S}^{\mathbf{T}}$, sized ($k \times n$). **C** is the matrix that describes the individual contribution (column concentration profiles) of the k pure components involved in each one of the mixture samples or processes, $\mathbf{S}^{\mathbf{T}}$ is the matrix of the k related spectral contributions (row pure spectral profiles), and **E** is the error matrix, the residual variation of the data set that is not related to any chemical contribution i.e. mainly noise.

Essentially, MCR-ALS is least-squares based iterative resolution method that alternatingly optimizes both **C** and $\mathbf{S}^{\mathbf{T}}$ matrices which are subject to some constraints. The individual steps followed to carry out the MCR-ALS procedure are represented in Figure II.9 and explained in detail below.

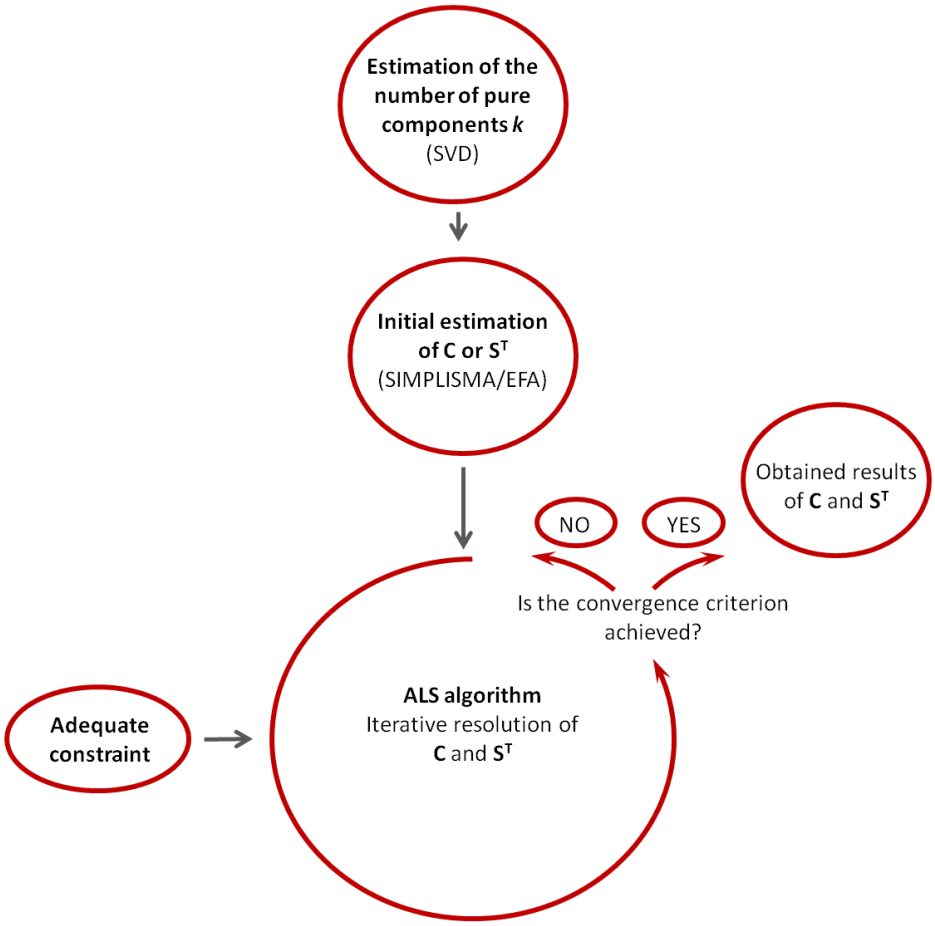


Figure II.9 : Scheme of different steps of the resolution process in MCR-ALS.

The sequence of the application of the algorithm involves three main steps. The first step is the determination of the number of pure components in the raw data **D**. The next step is the generation of an initial estimate (e.g. **C**-type or $\mathbf{S}^T$-type matrices). The final step is the constrained alternating least squares optimization of **C** and $\mathbf{S}^T$ until convergence is achieved.

II.3.2.2 Estimation of the number of pure components

Multivariate curve resolution methods do not need prerequisites or prior information about the number, evolution and chemical identity of the components present in the mixtures. The preliminary information that can be obtained from an exploratory analysis of the measurements can influence positively the resolution of the system. Nevertheless, an estimation of the number of pure components is a necessary step in order to apply MCR methods. This information will define the dimension of the pure **C** and $\mathbf{S}^T$ matrices. The

determination of an incorrect number of components will hinder the correct resolution of the respective data matrix (Nested method). Thus, the method starts by evaluating *k*, the number of pure contributions present in the system, for instance, by Singular Value Decomposition (SVD), the algorithm used for PCA. Singular values associated with relevant chemical contributions are large, whereas experimental error and noise are associated with low and non-relevant singular values. The number of significant singular values corresponds to the mathematical rank *k* of the matrix **D**. The choice of *k* is the starting point of MCR-ALS methodology. Figure II.10 shows a result of SVD applied on a series of 18 samples. The plot of log (Eigenvalues) as a function of eigenvalue number indicates that 4 different constituents are significant in the samples.

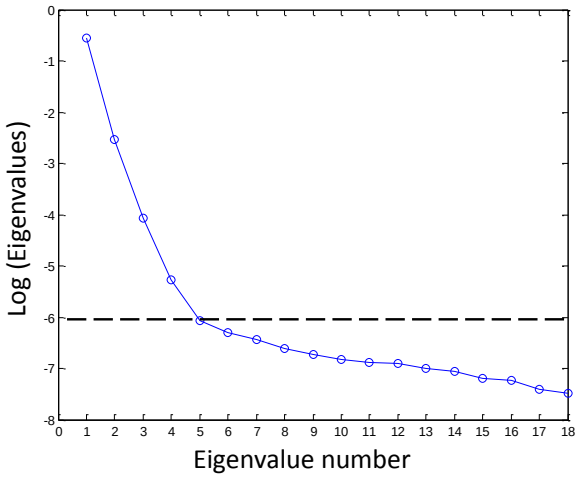


Figure II.10 : Plot of log (Eigenvalues) as a function of eigenvalue number obtained by SVD.

In some cases, the number of significant components estimated by SVD is lower than the true number of the pure species in the studied system, this phenomenon is known as rank deficiency and will be detailed in section II.3.2.7 [128-130].

In addition to SVD, there exist in the literature other methods that aim at choosing the appropriate dimensionality of the data such as Bayesian Information Criterion (BIC) [131], Akaike Information Criterion (AIC) [132, 133] and Integrated Completed Likelihood Criterion (ICL) [134]. The three Bayesian methods are based on the

$$IC_M = -2\log L + p_M \qquad \text{(II.6)}$$

Where $p_M > 0$ is the penalty applied to the likelihood L of the model M. The common aim for the three criteria is to find the best balance between the fit of the model to the data and its

complexity. The model for which the IC_M is the minimum, balance will be achieved. However, the difference between each criterion is the value of the penalty [135]. AIC provides an estimate of the distance between an appropriate model and the truth, for which

$$p_M = 2k \tag{II.7}$$

where k is the number of parameters in the model M.

BIC was designed to find the most probable model given the data as an estimate of the Bayes factor for two competing models. The penalty in this case is

$$p_M = k \ln(n) \tag{II.8}$$

where n is the sample size.

ICL is a criterion designed to select the model leading to the greatest evidence for clustering the data by maximizing the integrated likelihood.

For ICL,

$$p_M = k \ln(n) - 2ENT \tag{II.9}$$

where the so-called entropy term

$$ENT = \sum_{i=1}^{n} \sum_{s=1}^{s} p_{is} \ln p_{is} \tag{II.10}$$

that quantifies the ability of a mixture model to provide well-separated class with p_{is} being the estimated posterior probability that the individual i belongs to component s of the model M. When $p_{is} > 0.5$, classification is achieved by assigning individuals to a component *a posteriori* such as the individual i to component s of the model. If the components are well separated, ENT is almost zero and the classification is almost perfect. If not, ENT is large and positive and the rate of error of classification of individuals increases.

AIC overestimates the true number of mixture because, by construction, it tends to select too complex model [136]. BIC has been shown to select the actual model if it is in the set of the candidate models. However, when dealing with real data, for which there is no true model, BIC may overestimate the number of components as it does not account for the separation of the mixture components. ICL was derived from BIC to overcome its limitations by including an additional term (entropy) that quantifies the degree of separation of the mixture

components, therefore favoring the well-separated components. By accounting for the quality of the classification ICL should avoid overestimating the number of components however it may underestimate the number if the components are poorly separated. Many studies show that the performance of these indices is very dependent on the structure of the data. Our chemometric strategy is, thus, centered on the use of the SVD.

Indeed, Singular values possesses two main characteristics that makes the SVD better than the Bayesian methods and more convenient to be used for rank determination [137]. First, singular values are stable, in the presence of a little disturbance in the matrix, the change of the singular value is not great. The inherent feature of matrix of singular value can efficiently reveal essential property of the matrix.

II.3.2.3 Determination of initial estimates $\mathbf{C_{ini}}$ or $\mathbf{S^T_{ini}}$

Next step in MCR-ALS is the determination of the initial estimates. Initial estimates of the pure concentration profiles $\mathbf{C}_{ini}$ (C-type) or pure spectral profiles $\mathbf{S}^T_{ini}$ (S^T-type) are needed to start the iterative alternating least-squares (ALS) optimization. These initial estimates can be either set manually using chemical knowledge or be found by other methodologies from the matrix **D**. For this task there is a wealth of methods which can be briefly divided into two groups. The first one is applied when no structure in the concentration direction is available. The most commonly used method of this group is SIMPLISMA (see below). The second type methods are designed to obtain initial estimates for process-like data where the concentration profile evolves according to a given pattern. The most popular representative of this kind of method is EFA (see below). Both methods are non-iterative.

SIMPLe-to-use Interactive Self Modeling Analysis (SIMPLISMA) is a method based on the purest variable selection by Windig *et al.* and it can work irrespectively of the presence or absence of a structured concentration direction in the data set [138, 139]. This method works by searching the most dissimilar rows or columns in the raw measurements data set **D** (see Figure II.11) and gives the pure spectra $\mathbf{S}^T$ or pure concentration profiles **C**, respectively. The main drawback of this method is that it is limited to work with positive data. In a dataset that contains negative values, the use of pre-processed spectra (e.g. squared or integrated spectra) is an alternative to perform the selection of the purest rows or columns.

Then, considering a classical least squares procedure, initial estimates can be built afterwards from the original data set related to the purest rows or columns selected.

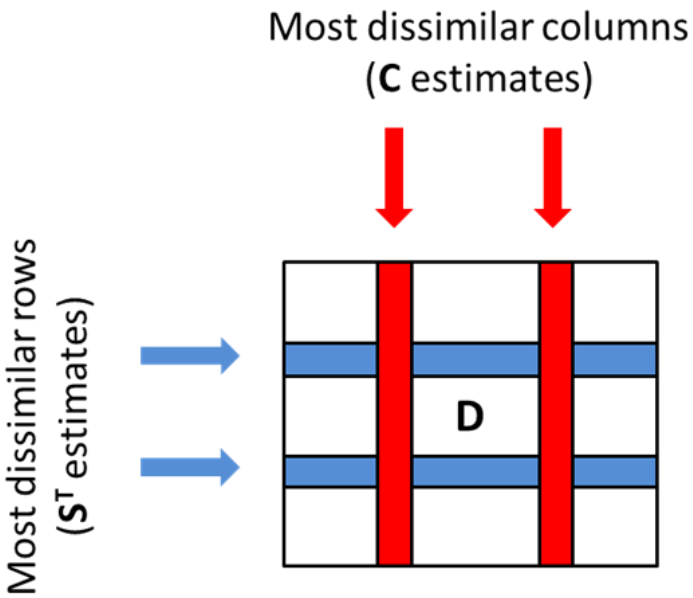


Figure II.11 : Graphical representation of SIMPLISMA for two component system.

Considering the application of SIMPLISMA on the spectral direction of the bilinear matrix $\mathbf{D}(m \times n)$ to give the initial estimate **S**, the purity p_i of the variable i can be calculated by the ratio

$$p_i = \frac{\sigma_i}{\mu_i + \alpha} \tag{II.11}$$

where σ_i is the standard deviation and μ_i is the variable mean. α is a constant offset (percent of the mean) designed to prevent an artificially high purity value in cases where the mean value μ_i approaches zero in noise-related spectra.

The purity function compares the variance between measurements of different absolute magnitude. The higher the purity variable value, the higher is the relative component purity for this variable. The plot of the purity value at each variable i gives a purity spectrum. The maximum intensity peak in this plot corresponds to the first pure spectrum in the initial estimate **S**. To find the next pure variable, the effect of the chosen component is mathematically eliminated from the purity spectrum that reflects the residual variance in the data not modeled by the previously chosen components. Subsequent maximum in the new

purity spectrum corresponds to the second initial spectrum and so on until the number of rank is achieved.

Evolving factor analysis (EFA) is a method particularly applicable in the study of processes where the concentration profiles of the different components evolve smoothly following a sequence pattern. It was introduced by Keller *et al.* and is used to follow the change of the rank of the data matrix **D** as a function of the number of line considered (time, pH, temperature,...) which is obtained by PCA on an increasing data matrix [140]. The idea of the method is schematically presented in Figure II.12.

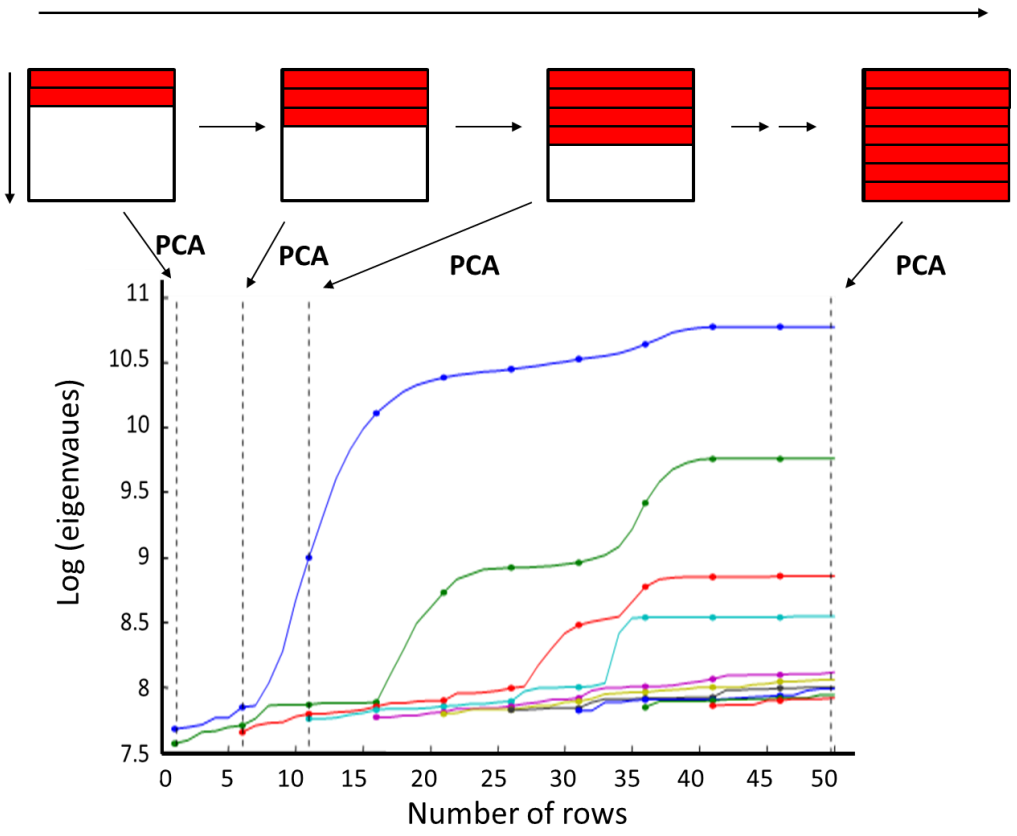


Figure II.12: Principle of forward EFA.

For the first submatrix $\mathbf{D_i}$ that is built from the first two rows of the data set, the eigenvalues are calculated. The third row is then added and the eigenvalues are calculated again and so on. In each step, a row is added and PCA is computed till the total number of spectra in the matrix **D**. The process of the factor analysis evolves forward direction on the rows and it is known as the forward EFA. The logarithm of the eigenvalues is plotted as a function of the line considered. All the points calculated for the first eigenvalues are connected as well as the points calculated for the second eigenvalues and so on. The eigenvalue that rises out of the

noise level indicates the appearance of a new component and therefore, an increase of the rank by one is noted. The number of eigenvalues above the noise is the number of pure compounds in the system and corresponds to the rank.

Analogue to the forward EFA is the backward EFA which is performed from the end of the ordered rows to obtain the disappearance of the components (see Figure II.13). However, the plot of the log eigenvalues is read from right to left and the increase of the eigenvalues above the noise level indicates the substance leaving the system. In order to associate the appearance of a compound with its disappearance, the first component emerging is the first decaying and so on [141, 142].

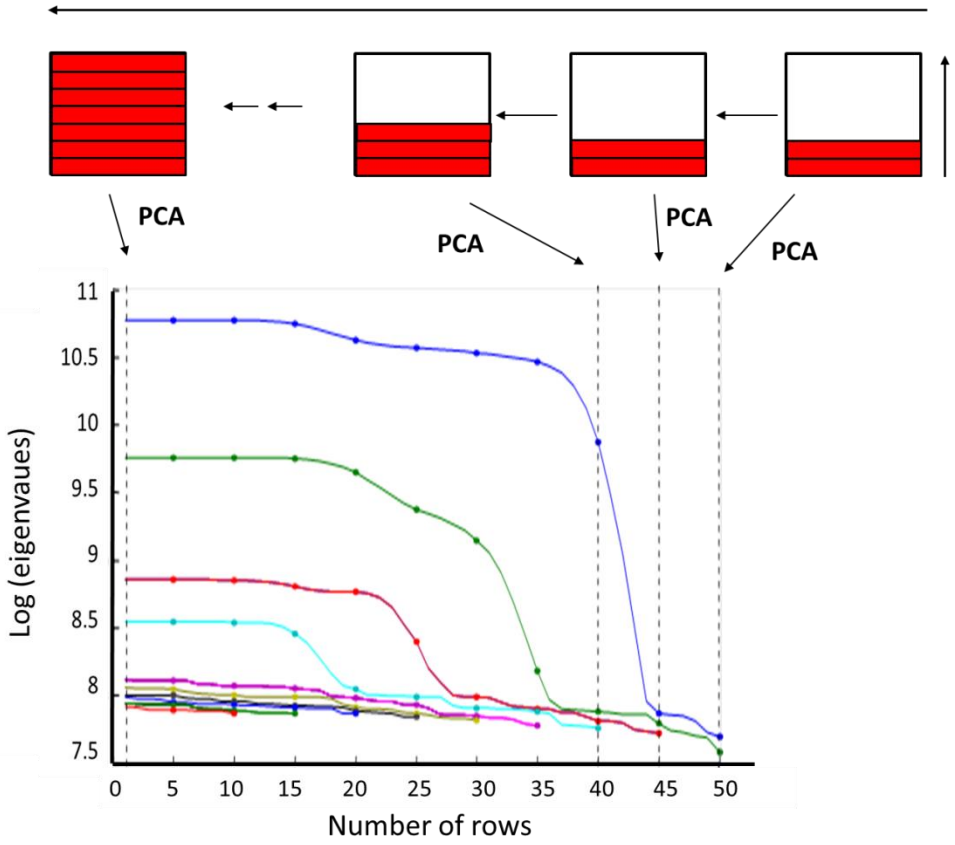


Figure II.13: Principle of backward EFA.

The final result obtained by the EFA is the plot of the eigenvalues as a function of the number of line which is connected with evolving variable such as time, pH or temperature depending on the data set studied (Figure II.14). Such a plot used for initial estimation of the concentration profile estimates the evolution of each pure component in the system [143, 144]. It is worthy to add that EFA gives profile of concentration as initial estimate but does not involve any assumptions about the shape of the signal. Since there are four significant

Principal Components, therefore, there are four real factors (four compounds) in the sample. Figure II.14 shows the initial estimation of the concentration profile of the four components.

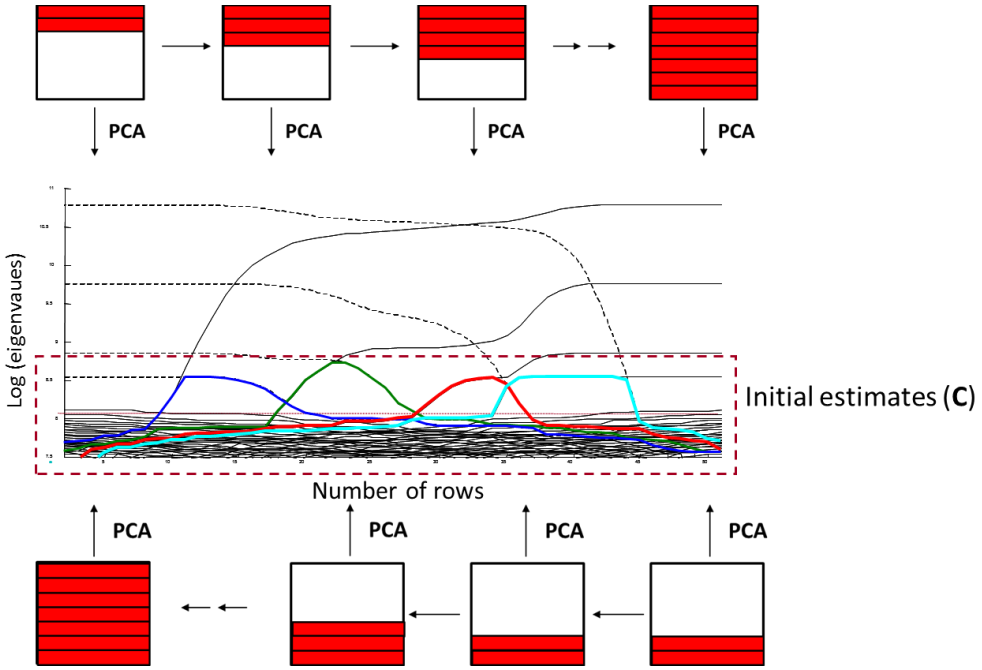


Figure II.14 : EFA graph obtained for forward (black line) and backward (dashed line) calculations. Initial estimates **C** obtained as a result of EFA.

Similarly to many other methods, EFA was modified and adapted to systems with different requirements. One of the variant of the original EFA is so called Fixed Size Moving Window-EFA (FSMW-EFA) [145]. This is a method where PCA analyses are performed with a fixed size window that is moved along the series of the spectra. The resulting local rank map obtained describes the variation of the component overlap in the process direction. Another very useful alternative is Fixed Size Image Window-EFA (FSIW-EFA) [146]. This is a local rank analysis method used in the spatial structure of the images where order is not important in the data set. This method that is derived from the FSMW-EFA will be detailed later in this chapter.

II.3.2.4 Alternating least squares (ALS)

Alternating least-squares optimization of equation (II.5) is started by using the original raw measurements of matrix **D** and the initial guess of either **C** or $\mathbf{S}^T$ matrices. This involves

in each iterative cycle the calculation of both **C** and $\mathbf{S}^{\mathbf{T}}$ matrices. The first step is calculation of the concentration profile **C** starting with the matrix **D** and the initial estimate $\mathbf{S}^{\mathbf{T}}$

$$\boldsymbol{C} = \boldsymbol{D}\,\boldsymbol{S}\,(\boldsymbol{S}^{\boldsymbol{T}}\boldsymbol{S})^{-1} \tag{II.12}$$

Knowing the matrix **D** and the concentration profile **C** that is obtained from the previous step, the next step is calculation of the spectral profile $\mathbf{S}^{\mathbf{T}}$

$$\boldsymbol{S}^{\boldsymbol{T}} = (\boldsymbol{C}^{\boldsymbol{T}}\boldsymbol{C})^{-1}\boldsymbol{C}^{\boldsymbol{T}}\boldsymbol{D} \tag{II.13}$$

Next, the matrix which is the product of concentration and spectral profiles $\mathbf{CS}^{\mathrm{T}}$ is compared with the experimental data matrix **D** and the difference can be expressed as matrix **E** (see section II.3.1.2). After that, the next iteration starts. "Least squares" means that the overall solution minimizes the sum of the squares of the errors i.e. the entries of the matrix **E**. The iterative process of alternating least squares stops when the algorithm converges. The convergence criterion in MCR-ALS is achieved when the relative difference between two consecutive iterations, which is called *lack of fit*, is below a threshold or when a predefined number of iteration is reached. The lack of fit, defined in equation (II.14), is evaluated from the elements of matrix **E** i.e. difference between the experimental data matrix **D** and the resolved data matrix, which is the product of concentration and spectral profiles $\mathbf{CS}^{\mathrm{T}}$. Another parameter called *explained variance* (r^2), is also used to evaluate the quality of MCR-ALS model and is defined in equation (II.15)

$$lack\ of\ fit\ (\%) = 100\sqrt{\frac{\sum_{i,j} e_{ij}^2}{\sum_{i,j} d_{ij}^2}} \tag{II.14}$$

$$r^2 = 100\left(1 - \frac{\sum i,j\ e_{i,j}^2}{\sum i,j\ d_{i,j}^2}\right) \tag{II.15}$$

where d_{ij} is an element of the input data matrix **D** (sample *i*, magnetic field *j*), e_{ij} is the related residue obtained from the difference between the input element and the MCR-ALS reproduction.

II.3.2.5 Uncertainty in MCR solutions

Although MCR-ALS gives good results from the sole use of the raw experimental measurements, the main theoretical issue affecting the resolution is the non-unicity of the solution (a band of feasible solutions) so called ambiguity [147, 148]. Noise-related uncertainties in MCR are attributed to the instrumental measurements errors, whereas ambiguity is a problem related to the modeling procedure itself. The concept of ambiguity can be defined as the different combinations of sets of concentration profiles and spectra that can reproduce the original data set with the same fit quality. Three kinds of ambiguity can be observed:

Permutation ambiguity. The resulting MCR-ALS components are not sorted in a specific order i.e. they can be shuffled within the concentration and the spectral matrix giving the same results but keeping the right correspondence of the dyads.

Instensity ambiguity. Dyads of profiles having the same shape but different relative scales between the concentration and spectral profile regive the original data set. For this reason, concentration values and pure response intensities in $\mathbf{c_i}$ and $\mathbf{s_i^T}$ profiles are always in arbitrary units unless the reference information on real intensities is available and actively used in the resolution process. In order to suppress the intensity ambiguity, normalization of concentration profiles or spectra is often used.

$$\boldsymbol{D} = \sum_i \boldsymbol{c_i s_i^T} + \boldsymbol{E} = \sum_i (\boldsymbol{c_i} k_i)\left(\boldsymbol{s_i^T}\frac{1}{k_i}\right) + \boldsymbol{E} \qquad \text{(II.16)}$$

Rotational ambiguity. Sets of concentration profiles and spectra with different shapes can reproduce the original data set with the same residuals. Going back to the MCR model equation (II.5), it can also be written as:

$$\boldsymbol{D} = (\boldsymbol{CT})\left(\boldsymbol{T^{-1}S^T}\right) + \boldsymbol{E} \qquad \text{(II.17)}$$

Where **T** is any transformation matrix and this gives :

$$\boldsymbol{D'} = \boldsymbol{C'S'^T} \qquad \text{(II.18)}$$

$$\boldsymbol{C'} = \boldsymbol{CT} \tag{II.19}$$

$$\boldsymbol{S'^{T}} = \boldsymbol{T^{-1}S^{T}} \tag{II.20}$$

where the matrices $\mathbf{C'}$ and $\mathbf{S'^{T}}$ are rotationally equivalent to the real $\mathbf{C}$ and $\mathbf{S^{T}}$. Thus, the decomposition of $\mathbf{D}$ based on the $\mathbf{C'}$ and $\mathbf{S'^{T}}$ leads to the same fit quality as the one based on $\mathbf{C}$ and $\mathbf{S^{T}}$.

The main solution for decreasing and suppressing ambiguities in MCR solutions is the application of constraints and the work with multisets that will be both detailed in the next sections.

For each component profile, a set of feasible solutions under constraints defines a range (band) of feasible solutions and this band may be delimited by maximum and minimum band boundaries. These boundaries are related to specific rotation matrices $\mathbf{T}$ for each component *k* that are called $\mathbf{T}_{\max,k}$ and $\mathbf{T}_{\min,k}$. Considering a particular set of solutions obtained by equation (II.5) fulfilling the constraints defined by the problem $\mathbf{C_{inic}}$ and $\mathbf{S^{T}_{inic}}$. The maximum band boundaries for the concentration and spectral profiles are $\mathbf{C_{max,k}}$ and $\mathbf{S^{T}_{max,k}}$, respectively. The minimum band boundaries for the concentration and spectral profiles are $\mathbf{C_{min,k}}$ and $\mathbf{S^{T}_{min,k}}$, respectively. The equation can be expressed as follows :

$$\begin{aligned} \boldsymbol{D} = \boldsymbol{C_{inic}S^{T}_{inic}} = \boldsymbol{C_{inic}T_{min}T^{-1}_{min}S^{T}_{inic}} = \boldsymbol{C_{min,k}S^{T}_{min,k}} \\ = \boldsymbol{C_{inic}T_{max,k}T^{-1}_{max,k}S^{T}_{inic}} = \boldsymbol{C_{max,k}S^{T}_{max,k}} \end{aligned} \tag{II.21}$$

When calculating the feasible bands, the main aim of the method is to obtain the values of $\mathbf{T}_{\max,k}$ and $\mathbf{T}_{\min,k}$. MCR-bands is an algorithm that finds the minimum and maximum boundaries of feasible solutions obtained by MCR-ALS [149-151]. It is based on constrained optimization of an objective function that is defined by the ratio of the Frobenious norm of the signal of the considered component *k* that depends on the considered $\mathbf{T}$ matrix over the Frobenious norm of the total signal with dyads of profiles obeying the constraints.

$$f_k(\boldsymbol{T}) = \frac{\|\boldsymbol{c_k s_k^T}\|}{\|\boldsymbol{CS^T}\|} \tag{II.22}$$

The optimization of the objective function gives the maximum and minimum boundaries when it is maximized or minimized respectively i.e. two extreme MCR dyads. The method is

applicable to systems with unlimited number of components and can describe the extent and the location of the ambiguity.

The extent of the ambiguity. If the difference between the boundaries is zero, then the component is resolved in a unique way. The more the difference between the maximum and the minimum, the larger is the ambiguity associated with the resolution of a particular component.

The location of the ambiguity. When we plot the dyads of profiles for a particular component in MCR-bands, it can be seen whether ambiguity affects the concentration profiles or spectra for a particular component.

Ambiguous solution can be satisfactory if the components in the studied system are uniquely resolved or affected by a very low degree of ambiguity.

II.3.2.6 Constraints

One of the attractive features of resolution methods is that no prior knowledge, either chemical or mathematical, is needed to analyze the data set of interest. Additional knowledge, when existing, can be useful to orient the resolution process and to improve the final results obtained (**C** and $\mathbf{S}^{\mathbf{T}}$). The ALS starts with the initial estimates of **C** or $\mathbf{S}^{\mathbf{T}}$ and works by optimizing iteratively the concentration and the response profiles using the available information about the system to reduce the so-called rotational ambiguity inherent to any data decomposition with MCR [152, 153]. Such information can be introduced through the implementation of the constraints. A constraint is defined as a particular characteristic of chemical or mathematical nature that the profiles of the pure species must obey to ensure a final solution with chemical meaning (Figure II.15). Constraints, translated into mathematical language, force the iterative optimization process to model the profiles respecting the conditions desired. Within the family of constraints, selectivity is essential to resolve any kind of data set. However, the misuse or the wrong choice of the constraint may play a negative role in the resolution process such as distort in the related profile leading to a damaged convergence of the optimization process. A variety of different types of constraints is possible and thus, can be applied in a very flexible manner. Most of the constraints are directly linked to the chemical properties fulfilled by the pure concentration or spectral profile such as non-

negativity, closure, unimodality, hard-modelling, and equality. Other constraints are related to mathematical features and therefore can be applied to all data sets regardless of their chemical nature e.g. local rank/selectivity. In general, constraints are driving forces of the iterative process to the right solution, if they are well implemented and fulfilled by the data set.

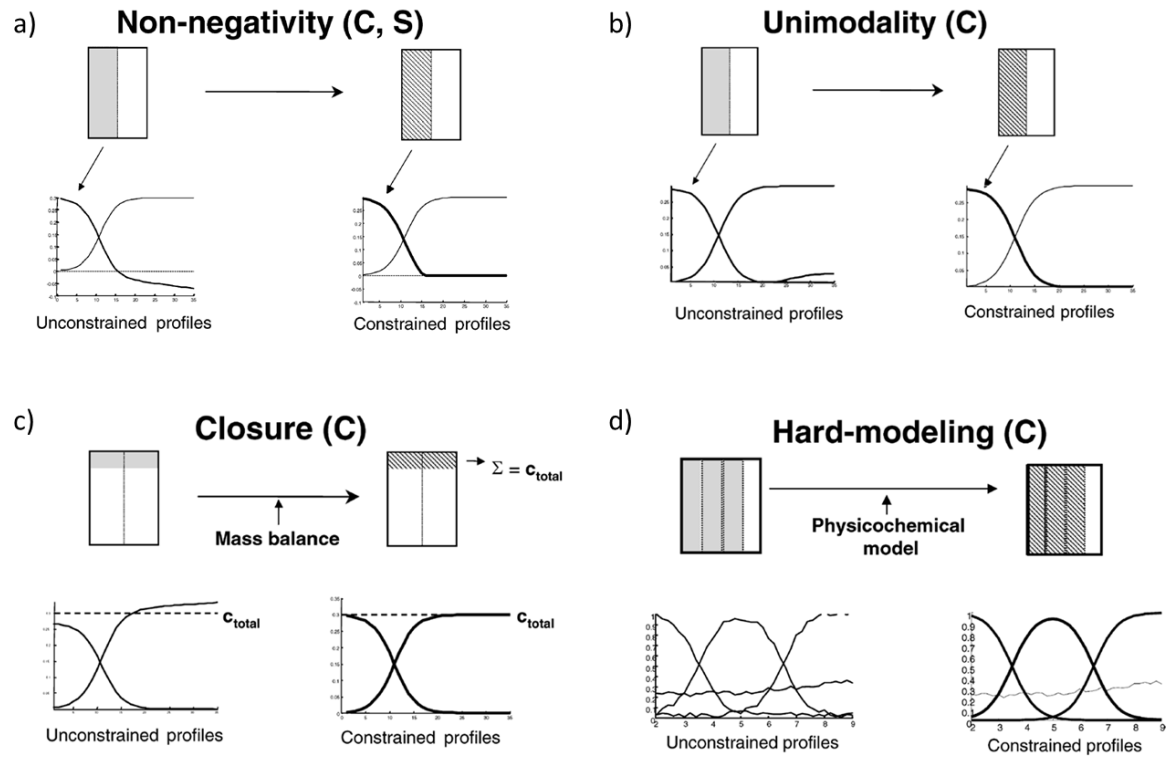


Figure II.15 : Common constraints used in ALS iterations [154].

Non-negativity. Non-negativity is the most popular constraint applied to concentrations and to many instrumental signals. It forces the profiles to be formed by positive values and can be implemented by replacing negative values with zeros (Figure II.15a).

Unimodality. Unimodality is a constraint applicable in the presence of only one maximum in the profile, thus, it imposes the monotonic character of it (Figure II.15b). In consequence, it is very useful in the concentration profiles related to chemical processes such as reactions.

Closure. Closure is also called mass balance valid in many systems (Figure II.15c). It is typically applied to the concentration profiles when the total concentration of all the components in the reaction system is conserved.

Hard-modeling. This constraint mainly forces the concentration profiles to be fitted by a parametric physicochemical model i.e. equilibrium and kinetic constants. Therefore, it implies performing a model-fitting task during the iterative optimization process (Figure II.15d).

Equality constraint. It imposes a given (previously known) shape or even a value of the pure spectra and/or concentration profiles to be invariant in the iterative process.

Local rank/selectivity. This constraint introduces information about certain pixels in images where one species is absent (selectivity), or where some species are missing (local rank) in the concentration profile.

II.3.2.7 Rank deficiency

If the number of significant components detected by the rank analysis of the data matrix is lower than the number of chemical contributions truly present in the system this so called rank deficiency [128, 155]. Such situation may occur when, for instance, at the beginning of a chemical reaction or process, more than one component already exists or parallel reactions take place. This problem may also occur when the reaction of one of the component immediately takes place and produces a constant signal during the time range of the experiments. In addition, it may occur when the concentration or spectral profiles of different chemical components in a mixture are linearly dependent or present identical shapes in the most trivial task. The relation between the chemical rank of a data matrix **D** and its bilinear decomposition is

$$\text{Rank}\,(\mathbf{D}) \;=\; \min\,(\text{rank}\,(\mathbf{C}), \text{rank}\,(\boldsymbol{S^T})) \qquad \text{(II.23)}$$

Linear dependencies can exist in either **C** or $\mathbf{S}^{\mathbf{T}}$, causing a decrease in the mathematical rank and as a consequence the chemical rank of **D** decreases too. In general, linear dependencies are usually encountered for the concentration profiles in a kinetic reaction. $A + B \rightarrow C$ is an example where the number of components present is three, however, the number of independent concentration profiles is only two since A and B have the same concentration profiles. The mathematical rank of matrix **C** in the above example is two and therefore, the chemical rank of **D** will be the same no matter how different the spectra of A, B and C are. Analogue situation of rank-deficiency can be encountered in the spectral matrix when for

example two chemical components have identical spectra or when one species has a spectrum that is a linear combination of the spectra of two or more components in the system. In this case, the rank deficiency will occur in the spectral matrix and will result in the overall diminution of the chemical rank of the matrix **D** no matter how different the concentration profiles of the component in the system can be.

MCR analysis resolves the concentration and spectral profile of each component according to the chemical rank of the data matrix **D**. In the case of rank-deficient data matrix, MCR can only resolve a lower number of dyads that does not represent the correct pure profiles of the components but to the linear combination of them. Therefore, rank-deficiency poses problems in the analysis of MCR and the use of appropriate matrix augmentation strategy that will be detailed in the next section is the best way to solve this problem [156, 157]. For the kinetic example tackled in the previous paragraph, the simultaneous analysis of the system under conditions where the initial relative amounts of *A* and *B* are different will create a new augmented data matrix with the rank increased in one unit and therefore, the rank deficiency will be eliminated.

II.3.2.8 Multiset

In spite of the great potential of MCR-ALS method, sometimes the information included in a single data set is insufficient to provide a satisfactory quantitative description. In such case, the multiset analysis i.e. simultaneous analysis of data coming from several experiments is necessary (Figure II.16) [156]. The major advantages of extended MCR of multiple data matrices allow the elimination of rank-deficiency as well as to decrease the rotational ambiguity.

Two independent and complementary approaches are possible in multiset analysis. The first one relies on the use of several different spectroscopic techniques applied to the same multicomponent system (multi-technique analysis). In this case, the row-wise augmented matrix is built (see Figure II.16a). Bilinear decomposition gives the concentration matrix **C** which is common and spectral matrix $\mathbf{S}^{\mathbf{T}}$ which is composed from independent spectral matrices corresponding to the respective spectroscopic techniques.

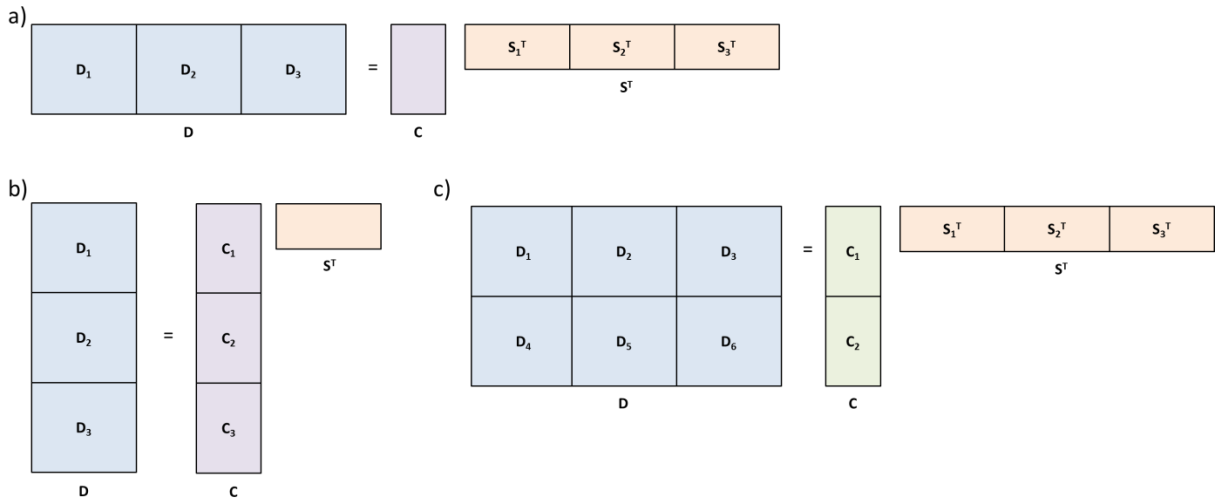


Figure II.16 : Augmented bilinear models. a) Row-wise augmented matrix, b) Column-wise augmented matrix and c) Row- and column-wise augmented matrix.

On the other hand, a column-wise augmented matrix is built (see Figure II.16b) when the experiments of the same technique are performed for several different mixtures containing some common species (multi-experiment analysis). This approach results in one spectral matrix $\mathbf{S^T}$ which is common for all the experiments, whereas the concentration matrix **C** is composed from independent concentration submatrices attributed to different mixtures. In addition, the row- and column- wise augmented matrix can be applied on many data sets coming from different mixtures monitored with several techniques (see Figure II.16c).

II.3.3 Data pre-processing

Experimental practice shows that the recorded signals can be often affected by a number of different variations from sources other than the analytical system under study such as undesirable physical phenomena, electric instrument components or surroundings. Unwanted signals can be random or systematic. Random unwanted signals are due to sporadic interferences or associated with random phenomena that often follow a standard normal or a Poisson probability distribution. This type of noise is called white noise and it can be sometimes substantial part of the original useful spectrum. On the other hand, systematic unwanted variations are due to the instrumental trends or to external influences. They may affect the signal by intense and irregularly shaped baseline shifts that have a low contribution in the signal.

The undesirable signal contributions can be removed or at least substantially decreased by *signal pre-processing*. This procedure is also called row pre-processing since the mathematical transformations act on each signal independently corresponding to a row in the data matrix. It aims not only at reducing unwanted signals (random and systematic) but also reducing the size of the data and enhancing the quality of the spectral dataset. There exist several pre-processing procedures such as filtering, normalization, baseline correction and scaling and centering. Nevertheless, only a couple of pre-processing techniques will be detailed below.

De-noising. There exist many methods designed to de-noise spectroscopic datasets. Among the methods some rely on classical smoothing procedures i.e. by averaging or fitting neighboring fragments of spectra to a polynomial function as Savitzky-Golay approach [158]. Other classical smoothing procedures use mathematical signal filtering e.g. Fourier transform or wavelet-based filters [159, 160]. The application of the above methodologies should be precise so that no relevant features of the dataset are removed together with the noise. PCA is another method used to de-noise the spectra since it decomposes the original data into a bilinear model of latent variables. Among these variables, few can be related to chemical variation and the rest to noise. Reproducing the data set taking only the model with the relevant components is a good way to separate the noise from the spectra.

Baseline correction. It is a common pre-processing step applied after spectral de-noising.

Derivatization and integration. The former can be carried out to stress subtle differences in spectral features among spectra, improve spectral resolution and correct the baseline. The first derivative of a signal $y = f(x)$ is the rate of change of y in x i.e. $y' = \frac{dy}{dx}$ that can be explained as the slope of the line tangent to the signal. The second derivative can be considered as further derivation of the first derivative i.e. $y'' = \frac{d^2y}{dx^2}$ and it represents a measure of the curvature of the original signal (rate of change of the slope). Unfortunately, derivatives should be used carefully used because they may magnify the noise characterized by the high slope variations and increase the complexity of the spectrum. The inverse transformation of the derivatization is the integration i.e. $A = \int f(x)dx$.

Asymmetric least-squares (AsLS). It is based on a local fitting of the whole spectrum with a baseline obtained by Whittaker smoother [161]. In this pre-processing technique, the parameters are tuned according to the shape of the baseline. The first parameter (λ) is linked

to the smoothness of the fit and the second one (p) is linked to the penalty imposed to the points giving positive residuals in the fit. This correction is particularly applied in order to correct smooth baselines with a significantly lower frequency than the useful signal.

Mean Centering.

Mean-centering calculates the mean of each column and then, deletes systematic location differences between the variables by subtracting each variable by the mean column.

II.4 Data cube

Besides two dimensional data sets obtained by most of the classical spectroscopic techniques, the so called *data cube* can be also the subject of chemometrics. The most common types of such structures are images containing two spatial dimensions and the third one which can have spectroscopic nature. Image analysis is a wide denomination that encloses several kinds of images. Typical examples are simple classical gray scale images, multispectral images that are collected using low number of spectral channels or hyperspectral images that are exploited in their full extension with several hundreds or thousands of spectral channels. To understand well all the challenges of image analysis in chemometrics it is worthy to follow through the features and development of image acquisition.

II.4.1 Imaging overview

Spectroscopic techniques, which are used for imaging purposes, are based on integrating principles to improve signal-to-noise ratio. It means that a detector records the signal in a given sample area, a measurement is repeated several times and the spectral intensities throughout the sample areas are added up. However, in case of non-homogeneous samples, the measurement has to be repeated many times in a systematic way to reveal the non-homogeneity of the investigated area. A single (or insufficiently repeated) measurement may be erroneously used to represent the bulk value of the entire sample. On the other hand, the measurement which is repeated many times extends the acquisition time, increases the

cost and can lead to the damage of the sample. Instead of increasing repetition of detection, the additional information, revealing inhomogeneity of the sample, can be obtained by spectroscopic characteristic of the studied area. Recent advances in spectroscopic detector technologies that enable the detection of hyperspectral images have provided such advantage.

Figure II.17 and Figure II.18 represent the historical development of hyperspectral images from an imaging perspective. Black and white photography was first used to create two dimensional images by compressing a range of colors measured throughout the image area into a set of gray values representing the average spectral intensity at each location in the image. Gray scale image reflects the light intensity over the electromagnetic spectrum in a single band. By adding the color filters to the cameras or films, the acquired images could represent the narrower bands of the color spectrum. Since that time, a color image reflects the intensity over the red, green and blue bands of the spectrum. This is presented today in the typical color digital photographs, compromised of a set of three separate grayscale images representing the three main color bands: red, green and blue (RGB). This is shown in Figure II.17 where a digital color image possesses three channels of information.

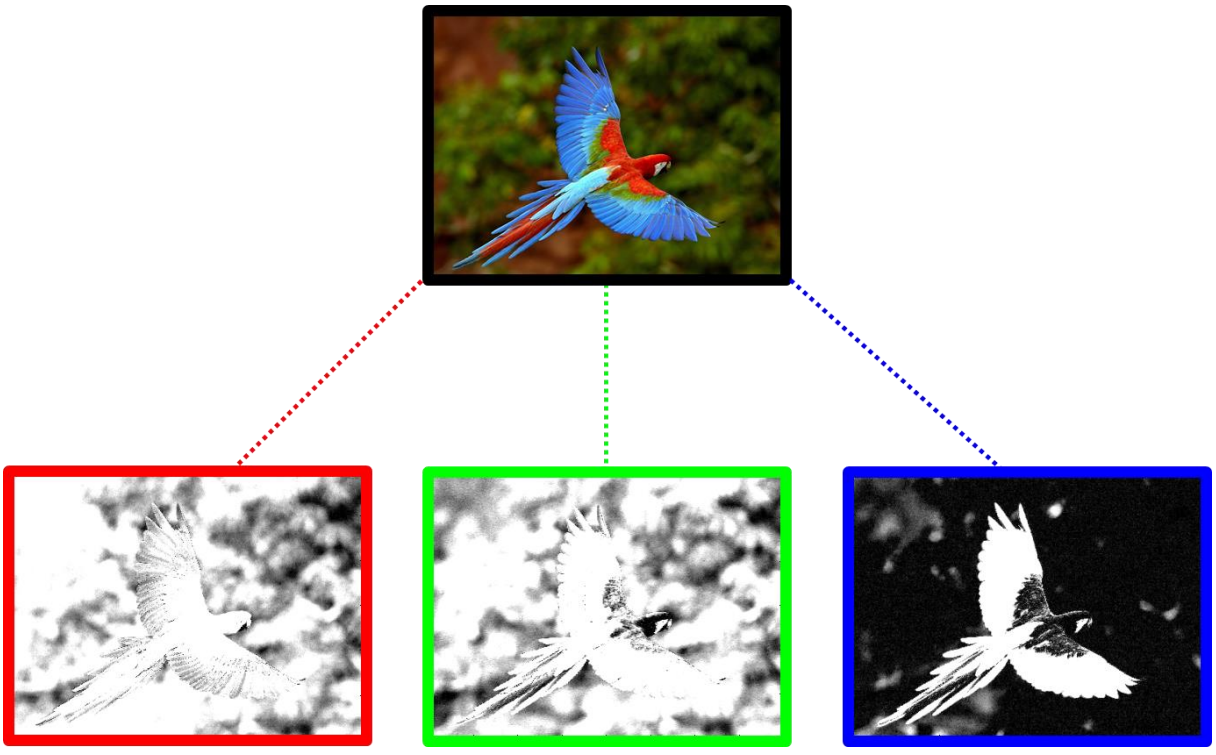


Figure II.17 : Typical digital color picture composed from the three channels red, green, and blue intensities. In grayscale color bands white color represents maximal intensity and black minimal intensity.

To increase the spectral resolution, the additional channels can be introduced by using 4 or more additional sets of color filters giving a multi-spectral image shown in Figure II.18. Higher number of spectral channels can greatly increase the amount of information included in the image. Historically, this approach was first popularized in the field of remote sensing where images of earth's surface were acquired from spacecraft, satellites and airplanes. In 1968, the Apollo 9 mission included 4 Hasselblad cameras aimed at the same target, while the Skylab missions (1973-1974) included a six camera multi-spectral system. In 1972, the Landsat I satellite launched containing a multispectral scanner which permitted the acquisition of red, green and two infrared channels [162]. Later, similar multichannel images have been used in microscopy where a filter wheel was used to acquire sets of images representing different color bands [163]. Nowadays, most of the spectroscopic techniques are used to acquire the images containing substantial spectral information.

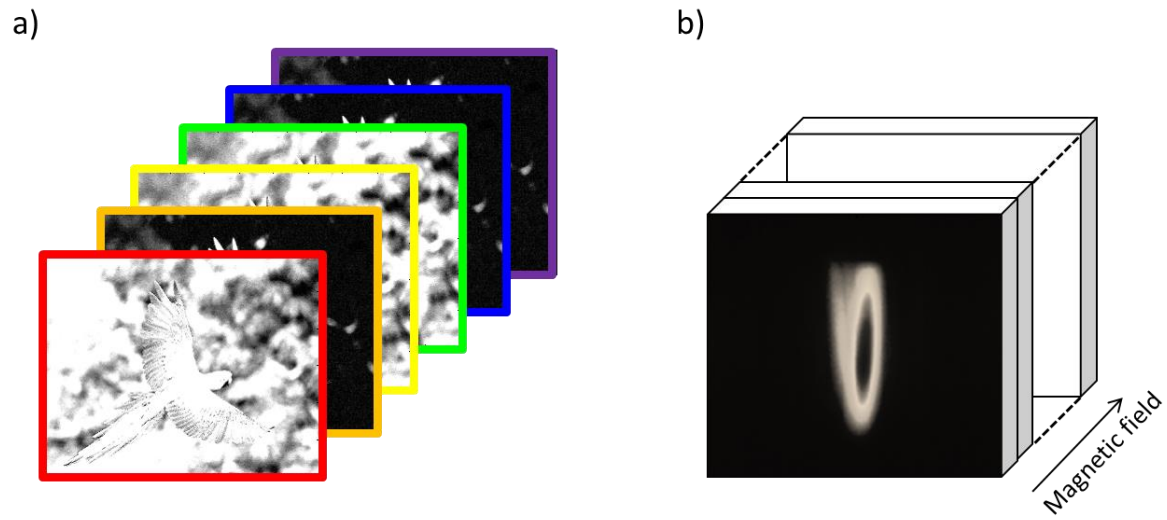


Figure II.18 : a) Multi-spectral image containing six channels (color bands) obtained by the use of the color filters. b) Hyperspectral EPR image.

When many different spectral channels (bands) are obtained, these images are called hyperspectral images. Although there is no clear defined number of channels that differentiate between multispectral and hyperspectral images, in general, it is commonly accepted that hyperspectral images typically contain over 100 channels. Figure II.18a shows the example of multispectral image where each pixel contains six spectral channels, whereas Figure II.18b represents a typical hyperspectral image acquired by EPR spectroscopy where each pixel contains a spectrum composed of hundreds of spectral channels i.e. magnetic fields in this case.

Hyperspectral imaging is beyond the conventional imaging and spectroscopy because it acquires, simultaneously, both spectral and spatial information about the sample [164-168]. Imaging technique is the science of acquiring spatial and temporal data information from the objects using a digital camera. On the other hand, spectroscopy is the science of acquiring the spectral characteristics of an object to describe light intensities emerging from the molecules at different wavelengths or magnetic fields and therefore, provides a fingerprint of that object. Since typical images are considered two-dimensional data, an addition of a new dimension i.e. 'spectrum' results in hyperspectral image data perceived as a three dimensional data cube (or hypercube). In hyperspectral imaging techniques, the sample is physically preserved and compartmented into small surface or volume areas called pixels or voxels, respectively, on which the spectra are recorded. Hyperspectral images are displayed as cubes in 3 dimensions (3D) among which the two dimensions (x- and y-) are pixel coordinates and the third (z-) is the spectral one. In the case of multilayer image, a hypercube (4D) is formed with three spatial directions (x-, y- and z-voxel coordinates) and the spectral direction. The main impetus for developing a hyperspectral imaging system is to enable the identification of chemical composition of different components extracted from the spectral features and their spatial distribution from the pixel-to-pixel variation in the studied sample. Some features of imaging, hyperspectral imaging and spectroscopy are compared in Table II.1.

Features	**Imaging**	**Spectroscopy**	**Hyperspectral imaging**
Spatial information	✓	×	✓
Spectral information	×	✓	✓
Multi-constituent information	×	✓	✓
Building chemical images	×	×	✓
Flexibility of spectral information extraction	×	×	✓

Table II.1 : Differences between imaging, spectroscopy and hyperspectral imaging techniques.

Hyperspectral images can be acquired by several techniques such as mid infrared spectroscopy [169], near infrared spectroscopy [170], Raman spectroscopy [171], fluorescence spectroscopy [172] or electron paramagnetic resonance [41]. Hyperspectral

imaging provides very rich information content and high excellent feature identification capabilities making it highly suitable for numerous applications such as biomedical diagnostic, material science, analytical chemistry, environmental analysis, pharmaceutical and food industry [173-180].

The advantages of hyperspectral imaging over traditional methods include minimal sample preparation and nondestructive nature due to fast acquisition time and visualizing spatial distribution of numerous chemical compositions simultaneously. Hyperspectral imaging records a spectral volume that contains a complete spectrum for every spot (pixel) in the sample rather than collecting a single spectrum at one spot in spectroscopy. In addition, modern hyperspectral imaging provides both quantitative and qualitative measurements due to its high spectral resolution. It is a chemical-free assessment method that enables safety and environmental protection by eliminating pollutant, dangerous solvents and chemical reagents. It is a flexible method in choosing the region of interest even after image acquisition.

It seems that the enhancement in instrumental developments, the availability of high-speed computers and the progress in appropriate chemometric procedures may allow hyperspectral imaging to become a dominant experimental technique in the near future.

Hyperspectral images are measurements of high quality and quantity of information. Suitable tools are thus also required to handle the massive spectral output from the image to extract the interpretable knowledge. Hence, the use of chemometric tools in image analysis is often mandatory in the exploration of the full measurements opposite to classical approaches, where only a partial visualization of the raw data is obtained. Image analysis is a field of constant development and progress in chemometrics. Some solid knowledge has been already built and new work and ideas are still under study.

II.4.2 Multivariate image analysis

In 1986, Geladi *et al*. first presented the theoretical aspects of extracting chemical information from hyperspectral images [181]. They showed that hyperspectral image, i.e. three dimensional data structure, can be unfolded to produce a large matrix of spectra. The obtained matrix can be then treated by ordinary chemometric tools used in the analysis of spectroscopic data. Later, the same authors published the fundamental works in the

application of chemometric tools starting first with Multivariate Image Analysis (MIA) that was associated to the Principal Component Analysis (PCA) [182, 183]. The following essential theoretical works in Multivariate Image Regression (MIR) included the application of Principal Component Regression (PCR) [184, 185] as well as Partial Least Squares (PLS) algorithms [186]. In addition, in 1995 Geladi introduced the concept of image sampling and local modeling specific to hyperspectral images. In recent years, the image analysis toolboxes have grown to include the application of Independent Component Analysis (ICA) [187] and Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) [188].

As with the exploration of spectroscopic data, multivariate resolution methods are among the most recent tools implemented in the field of image analysis. Resolution methods provide the distribution map and the corresponding spectral signature of the pure image constituents from the sole raw image measurement.

The spectrum of each pixel is represented by the concentration-weighted sum of the contributions of the pure spectra of the image constituents. The pure spectra of the constituents are the same along the image but the concentration weights vary from pixel to pixel depending on the pixel composition. The spectra and the pixel position in the sample provide a spatial-related compositional description of the sample [189, 190]. Indeed, all the pixels in an image obey the Beer-Lambert law, which translates them into a bilinear model [93].

In order to make the 3D image data suitable for MCR-ALS analysis, unfolding the image cube ($x \times y \times magnetic\ field$) into a 2D data table by the classical line by line unfolding method is required. The *x* first spectra corresponding to the *x* first pixels of the cube are placed along the *x* first lines in the matrix **D** and so on. This process is shown in Figure II.19 [188]. As a consequence of transformation of the *x*- and *y*- pixel's coordinates into a single dimension in the matrix **D**, the concept of pixel neighborhood is partially lost and the pixel is not surrounded by all its neighbors in the original image.

The original mixed measurement is described as a sum of the signal contribution of the different components i.e. $\mathbf{D_1}$ and $\mathbf{D_2}$ for a two-component system. Each one of the pure signal contribution can be expressed by a bilinear model i.e. the product of concentration weight of the particular *i*th component along the pixel direction $\boldsymbol{c_i}$ and its corresponding dyad $\mathbf{s_i}$ the pure unit spectrum. When the analysis is finished, the stretched concentration profile $\boldsymbol{c_i}$ should be

refolded to recover the 2D structure of the original image giving the distribution map of the particular component. The pure spectra are used for identification purposes and the consequent association with functional groups or matching with a spectra library. **E** contains the experimental error due to signal variation not associated with chemical information.

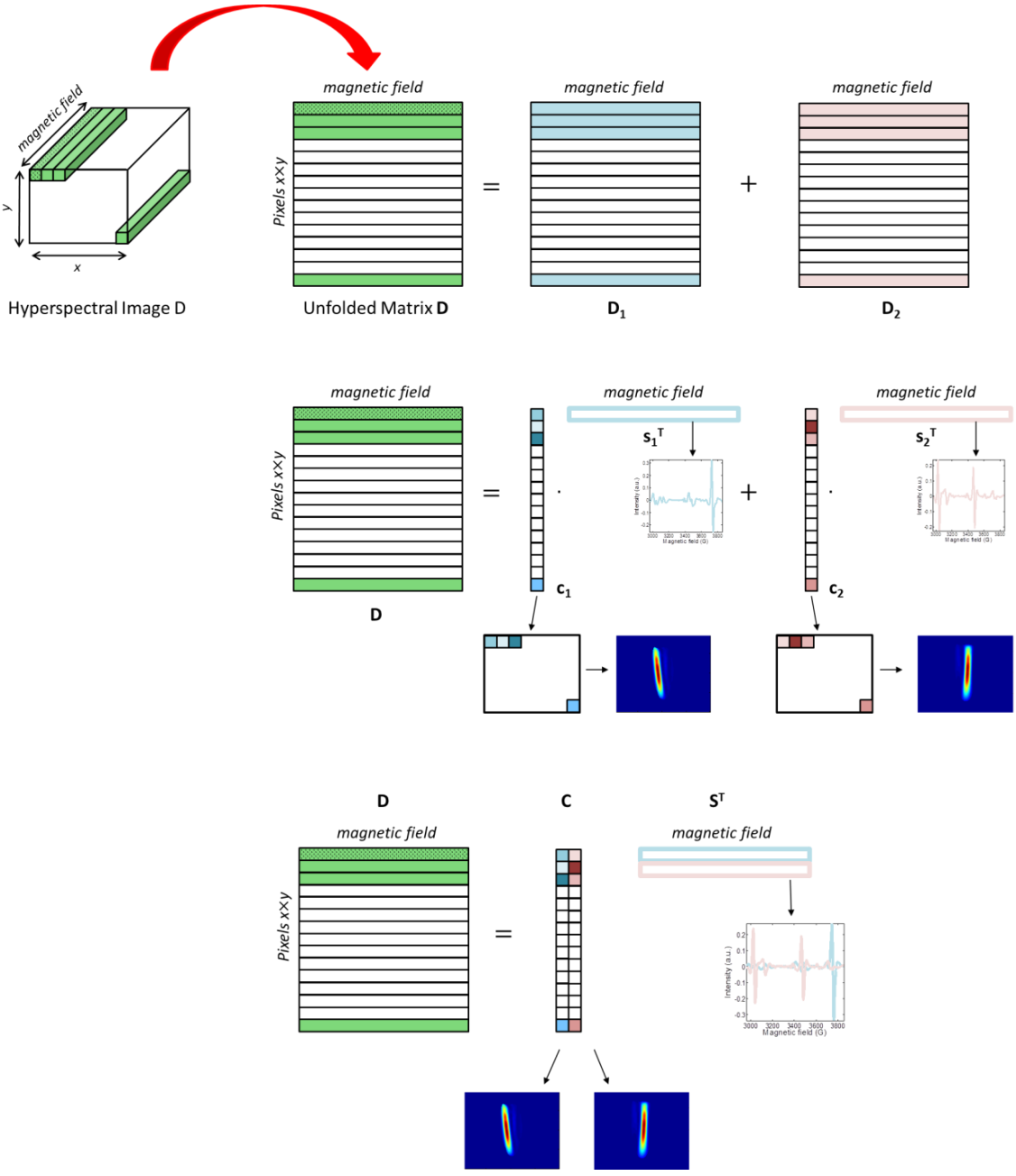


Figure II.19 : Bilinear model of hyperspectral image.

As noted in section (II.3.2.1), MCR-ALS aims at solving a bilinear factorization of the spectral data matrix **D** ($m \times n$) into the product of two simpler matrices, **C** ($k \times n$) and $\mathbf{S}^{\mathbf{T}}$($n \times m$), having a physical sense. In case of image analysis, **D** is the unfolded spectral data matrix that possesses m rows corresponding to m different spectra acquired at different pixels ($x \times y$) and n

columns corresponding to *n* different magnetic fields. *k* is the number of pure compounds present in the system. **C** is the pure concentration matrix, the profiles of which are individually folded back to recover the distribution maps for each pure compound in the analyzed sample. **S** is the matrix spectra of pure constituents used for molecular identification. **E** is the error matrix, which expresses unmodelled variations i.e. noise.

$$\boldsymbol{D} = \boldsymbol{C}\boldsymbol{S}^{\boldsymbol{T}} + \boldsymbol{E} \quad \text{(II.5)}$$

MCR-ALS procedure for images is similar to that described for spectroscopic data (see paragraph II.3.2.1) and is performed according to the following steps:

1. Determination of the rank of the unfolded data matrix **D**. (SVD)
2. Generation of initial estimates $\mathbf{C_{ini}}$ or $\mathbf{S^T_{ini}}$. (EFA or SIMPLISMA)
3. Optimization of **C** and $\mathbf{S^T}$ by iterative least squares algorithm with imposed constraints (if applicable).
4. Reproduction of **D** from the calculated $\mathbf{CS^T}$ when the lack of fit, is satisfactory.

The typical constraints used in the analysis of images are the non-negativity of the concentration direction (pixels) and non-negativity in the spectral direction (if applicable). Process-related constraints such as the unimodality, closure or hard-modelling cannot be used due to the lack of global continuous smooth evolution in the concentration profile. In addition, unfolding the cube destroys the 2D image pattern and therefore alternative constraints (see below) can be applied to decrease the rotational ambiguity.

II.4.3 Local rank constraint application on spatial information

Local rank constraints in image analysis introduce information about the presence and absence of some species in profile windows (pixels). In the simplest cases, for some pixels all the species can be present or only one of them (selectivity). However, in general, on profile windows some species are missing and local rank information is needed [191]. In order to set the local rank constraints in the image, Fixed Size Image Window-Evolving Factor Analysis (FSIW-EFA) is typically applied as a first step to obtain necessary pixel-to-pixel local rank information.

What is FSIW-EFA and how does it work?

As discussed in paragraph II.3.2.3, EFA-based methods are applied to inherent order of datasets. They are, thus, not adequate to be used in unfolded images since the concept of pixel neighborhood is lost in the image data sets, for instance the pixel is no longer surrounded by all its neighbors. In 2005, de Juan *et al*. established the combination of Fixed Size Image Window-Evolving Factor Analysis (FSIW-EFA) that is a local rank analysis method designed specifically for spectroscopic image analysis [146]. It is an algorithm that scans locally all the pixel areas in the analyzed sample. It is the first step to set local rank constraints for image resolution and its results is to provide valuable information about the local rank complexity in different pixel areas of the image.

FSIW-EFA performs local Principal Component Analyses in the whole image. It is based on a small fixed window containing a pixel and its neighboring pixels in the two spatial dimensions of the image and then moves one pixel at a time until the whole image is covered (Figure II.20). The PCA analyses performed on each one of the pixel areas provide information about the local complexity of the image. The window is fixed by the user at a given size depending on the spatial resolution required for the analysis. In general, the number of pixels in the window should be equal or greater than the estimation of the total number of constituents in the image in order to allow rank values representing all possible situations of compounds overlap in the image. Meanwhile, the number of pixels should be as small as possible to preserve the spatial resolution of the image and to discriminate better the pixel-to-pixel rank differences.

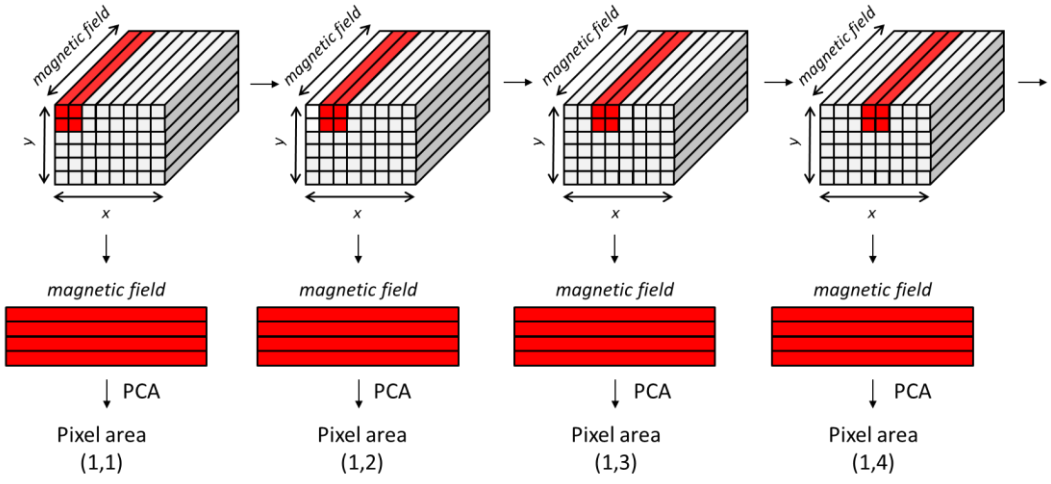


Figure II.20 : Window design and subsequent PCA analyses of FISW-EFA algorithm.

For example, in an image that contains 4 constituents, a window size of (2×2) is the most convenient dimension to be used. Then, PCA analyses are done on the unfolded data matrices containing the spectra in the related windows. As a result, singular value plots are reconstructed to recover the 2-dimensional image. There will be as many flat plots as singular values present in the image window i.e. *x* and *y* axes corresponding to the *x* and *y* pixels and a color coded magnitude of the singular values at appropriate pixel location. Warm colors indicate zones of the image where the related singular value has a high intensity. For example, singular values of high magnitude can be detected in the 2nd singular value plot locating image zones where two compounds overlap. High intensity zones in the 3rd singular plot correspond to three compounds overlap and so on. These singular value plots represent the degree of overlap among the different constituents.

Then, singular values are represented versus the unfolded pixels in an image. The plot has as many lines as the singular values calculated in each PCA window analyses. Large values represent significant contributions to the signal related to the presence of chemical constituents, whereas small values describe the experimental noise. A threshold value or band is selected by the user to distinguish the significant singular values from the noise-related singular values. Depending on the threshold selection, complete local rank maps and partial local rank map are obtained.

Complete local rank maps are obtained by the threshold value. It recovers the 2D structure of the original surface image and gives information about the complexity in the pixels (zones) of the image and how many constituents overlap each pixel area.

Partial local rank maps are obtained by the threshold band that marks the lowest and highest boundaries for sensible threshold values. The partial local rank presents pixels whose rank is invariant within the threshold boundaries. The pixels in the partial local rank are those with the most reliable rank estimation. From these pixels only the ones with a rank lower than the total rank may potentially include local rank information.

The final local maps obtained by FSIW-EFA are an overview of the image overlap and are basis to set the local rank constraints for the image. Incorporation of the local rank constraint starts by calculating the number of missing components in the pixel i that can be constrained.

$$number\ of\ missing\ components\ i = image\ total\ rank - local\ rank\ pixel\ i$$

In order to identify this component, correlation coefficient is calculated between the raw pixel spectrum and each one of the reference spectrum (usually obtained from SIMPLISMA or an $\mathbf{S^T}$ matrix obtained from previous MCR-ALS application on the same data set). The component with the lowest correlation coefficient with the spectrum will likely be the one missing in this pixel. In pixels where more than one component is absent, the missing components will have the smallest correlation coefficients with the pixel spectrum. Absence of particular component in a pixel is not confirmed unless the correlation coefficient between the pixel spectrum and the reference spectrum of the component is equal or smaller than the largest element in the correlation matrix for the particular component. Therefore, reliable identification of the missing components in a pixel is achieved by the combination of local rank and reference spectral information.

To end up, the information accepted by the rank and compound identification are translated in a matrix called **csel** matrix that has the same size as the concentration matrix **C** ($number\ of\ pixel \times number\ of\ components$) and will be implanted in the resolution process.

II.5 Data analysis in EPR spectroscopy: Sate of Art

The spectrum of a mixture of many constituents acquired by Electron Paramagnetic Resonance is often difficult to interpret due to the high level of partial or complete superposition of spectral contribution of various species present in the complex materials. In case of an entangled system, even the number of pure compounds can be quite difficult to determine. Moreover, characterization of the spectral signature of each component causes further problems due to the absence of a rigid database containing the established EPR spectra of various paramagnetic species. This dilemma covers the entire range of EPR applications such as the analytical characterization of the composition of paramagnetic inorganic, organic and biological samples, kinetics investigations of radicals in the chemical reactions, as well as spectral-spatial information in images. In practice, an experimental trade-off is often demanded to resolve these complications. Typically, an increase of magnetic fields is used to improve the spectral resolution [192]. When magnetic fields will increase, the separation of pure EPR signals will increase since the absorption maxima at B_0 shift according to the

compound specific g-values. However, such approach, if available, increases significantly the cost of apparatus and can be also undesired due to the stability of some samples. Therefore, EPR specialists typically use some simulations and advanced fitting procedures to interpret their data sets. In the first case, they benefit from the fact that each paramagnetic species produces a characteristic EPR signature according to well-defined spin Hamiltonian parameters. Simulation relies thus on the theoretical calculation of the effective spin Hamiltonian that gives the relation between experimental data and the predictions of quantum mechanics. There exists a wealth of simulation software such as EasySpin and WinSim that can be applicable to liquid and solid state EPR spectra [193].

The simulations are often combined with simultaneous fitting of the experimental data by means of different algorithms. For instant, Duling's method consists of rule-based perturbations with trial and error calculations. It is used for fitting EPR data with multiple free radicals as formed in chemical and biochemical spin-trapping systems [194, 195]. The proposed algorithm aims at refining hypothetical simulation parameters to give optimal values that will allow the researcher to test various possible free-radical formation pathways in an unbiased manner. Nevertheless, such calculations are very dependent on the initial parameters and on the correctness of the postulated models. Hence, the interpretation of the simulation is complicated in case of absence of any initial information about the studied data.

To avoid imposing the rigid models, the chemometric methods could be used to resolve successfully a multicomponent EPR dataset. However, in the related literature only very sporadic examples of application of PCA and multivariate techniques in EPR spectroscopy have been presented up to date.

In 1997, Steinbock *et al*. introduced PCA to EPR spectroscopy [196]. This chemometric tool is mainly used to explore EPR data sets and to determine the number of pure contributions needed [197, 198]. The studied samples were two and three component systems consisting of aqueous solutions of paramagnetic ions that exhibit distinct but overlapping spectra. In their work, Steinbock *et al*. applied PCA on the integrated EPR spectra since they supposed that PCA requires non-negative values, which is absolutely wrong. Novotny *et al.* applied also PCA on EPR data set [199]. Their aim was to isolate the paramagnetic ions that can affect the results obtained by other spectroscopic techniques in humic substances in order to elucidate structures of these humic substances. Nagai *et al.* applied PCA on EPR spectra of sodium ion-implemented silicon dioxide film [200]. Although this method is able to observe

relevant sources of variations in the data set, it cannot extract the pure spectra of each constituent since Principal Component profiles are imposed to be orthogonal (and in consequence uncorrelated), whereas real EPR spectra of different species overlap and reveal correlation among them. Thus, the practical applicability of PCA can be often limited to the determination of the number of pure components.

Improved chemometric approach is the application of Maximum-likelihood common factor analysis (MLCFA) proposed by Moens *et al.* and Vanhaelewyn *et al.* [201, 202]. It is a multivariate statistical technique that aims to detect the total number of constituents and to decompose a set of multicomponent EPR powder spectra e.g. sucrose EPR spectra into their individual pure components by using minimization procedure or target transformation. MLCFA assumes that each row (observed spectrum) is affected linearly by a small number of hypothetical constructs e.g. spectrum of specific radical (so called common factors). The latter is common if it affects at least two variables. This chemometric method assumes that the correlation between the observed variables is only due to a small number of common factors. The values of the different parameters in MLCFA model are estimated by the maximization of the so called likelihood function. The determination of the correct number of common factors is based on sequential procedure with a stepwise increasing number of factors. The hypothesis assuming the existence of k common factors is tested by the $\chi^2 - tests$. Once the hypothesis is validated, the common factor spectra are predicted using linear regression method. Afterwards, the spectra are transformed into an equal number of real component spectra by the use of minimization procedure i.e. minimize the discrepancies between the theoretical line profile (spin Hamiltonian and line width parameters) and its optimum spectrum production by means of a linear combination of common factor spectra. The final step is the simulation of the spectra. In spite of many advantages, MLCFA possesses also some important limitations. Indeed, it retains only the components which express fairly high percentage of the total variance, the remaining ones being considered as non-significant. In addition, this method has difficulty in identifying the real-component spectrum whose line shape are not described by means of analytical functions (i.e Gaussian or Lorentzian character). It also requires sometimes the availability of a reference spectral library.

Another noteworthy attempt in use of chemometric tools in EPR data analysis is Second Order Blind Identification with the Robust Orthogonalisation algorithm (SOBI-RO) followed by Constrained Independent Component Analysis (CICA) [203]. The cumulative usage of

both algorithms was applied on EPR spin-trapping in order to characterize EPR spectra of adducts so that different trapped radicals can be easily distinguished. SOBI-RO algorithm was used to roughly estimate the number of spectral components, afterwards fine determination was performed to increase the reliability for the relative contribution by the use of CICA. The important limitation of the application of these two algorithms is that CICA needs reference spectra for fitting the estimated signal obtained by SOBI-RO.

A statistical approach was proposed by Ren *et al.*, the Blind Source Separation (BSS) method implementing the Independent Component Analysis (ICA) in order to extract the different radicals from standard chemical system, cultured cell suspense and ex vivo rat kidneys by spin trapping EPR techniques [204]. Blind Source Separation (BSS) is a technique used to extract EPR spectra from their mixtures. It supposes that the mixture is a linear decomposition of a mixing matrix and the source signal i.e. the EPR spectrum is assumed to be a linear superposition of the source components according to the principal superposition. The aim is to extract the source signals from the mixtures that are the only known data. Further statistical assumptions on the source signals are needed to make the blind source separation problem solvable. Estimation of the source vector with preserved waveforms but possibly undetermined scales and orders are needed. The algorithm FastICA is used to implement ICA since the statistical distribution of the sources are supposed to be non-Gaussian and mutually independent. This algorithm is used to construct a measure of non-Gaussianity using an objective function for ICA estimation. This method can estimate the source spectra and their percentages in the total signal from the set of the observed mixture spectra. However, EPR spectra are neither completely independent nor completely uncorrelated, the decomposition result of the ICA-based blind source separation approach was not satisfying. Therefore, alternative statistical assumptions methods are required.

Later on, an alternative attempt was proposed by the same group of researchers, the novel sparse component analysis method for blind source separation that exploits the sparsity of the EPR spectra [205]. This method has been applied to the analysis of superoxide, hydroxyl and nitric oxide free radicals. A signal is considered sparse if it has some peaks and relatively flat area in between the peaks (noise). A robust sparsity function has some desired properties that can be used to extract, simultaneously, all the source signals from their linear mixtures on condition that the source signals are sparse and have some degree of non-

overlapping among each other. However, sparse component analysis can only retrieve sparse signals.

An interesting example of multivariate approach was applied by a predecessor of MCR-ALS algorithm at the end of 1980s to explore EPR data sets (called SPFAC) [206, 207]. However, some problems may arise in systems that contain minor species with low absorbance as well as in the study of complicated systems.

All the mentioned above methods used in analysis of multicomponent EPR data sets demand a number of presumptions and approximations on the studied system. Thus, the new approach, basing on the chemometric tools is proposed in the next chapter to overcome the limitations of the previous methods.

II.6 Conclusions

Advances in technology and the increasing availability of spectroscopic instrumentations offer a huge amount of data delivered in a negligible time frame. However, the large and complex data sets require the advanced analysis methods to be processed and interpreted in reasonable time. Chemometric methods are successful solutions to treat fastly the data and extract the desired information. Among different chemometric techniques, two unsupervised methods have been significantly developed and deserve the particular attention. They are PCA and MCR-ALS.

PCA is an exploratory method that offers an overview of the problem studied. It is often employed for noise reduction purposes and constitutes the basis for other more complex pattern recognition techniques. One of its mathematical representations is the SVD that is used to determine the rank of data matrix and thus, to estimate the number of species in the investigated system. In turn, MCR-ALS aims to finally resolve the chemical constituents of non-selective mixed original experimental output. This methodology is, with no doubt, a powerful and flexible algorithm, adaptable to many different kinds of instrumental techniques and chemical scenarios i.e. different kinds of data structures (single data sets and multiset structures) and meanings (diverse instrumental responses, environmental, omics data…). The main advantages of MCR-ALS are flexibility and the fact that the user receives the results

which can be interpreted straightforwardly i.e. the pure concentration profile and corresponding pure spectral signature of each constituent. Moreover, the possibility of using the suitable constraints during the iterative process allows decreasing enormously the ambiguity of the solution. In addition, the adaptation of MCR-ALS to imaging data treatments, by means of unfolding/refolding 3 dimensional spectral/spatial data sets, opened the perspective for fast and efficient resolving the complex hyperspectral images.

Chapter III:Chemometric Methods to Unmix Multicomponent EPR Data Sets

III.1 Introduction

In the literature, there exists a shy application of chemometrics on EPR data sets in order to resolve the fact of spectral superimposition of various species. The main aim of this chapter is to demonstrate how chemometric techniques, described in Chapter II i.e. PCA and MCR-ALS, can be applied to solve the problem related to the superimposition of peaks in unresolved EPR spectroscopic data sets. The concept of resolving the complex EPR spectra is presented in Figure III.1. The spectrum AB is an EPR signal of a mixture of only two components (A and B) but even a brief inspection of the experimental spectrum indicates that, without any prior knowledge about the identity of the mixture, resolving such a system is not a trivial task.

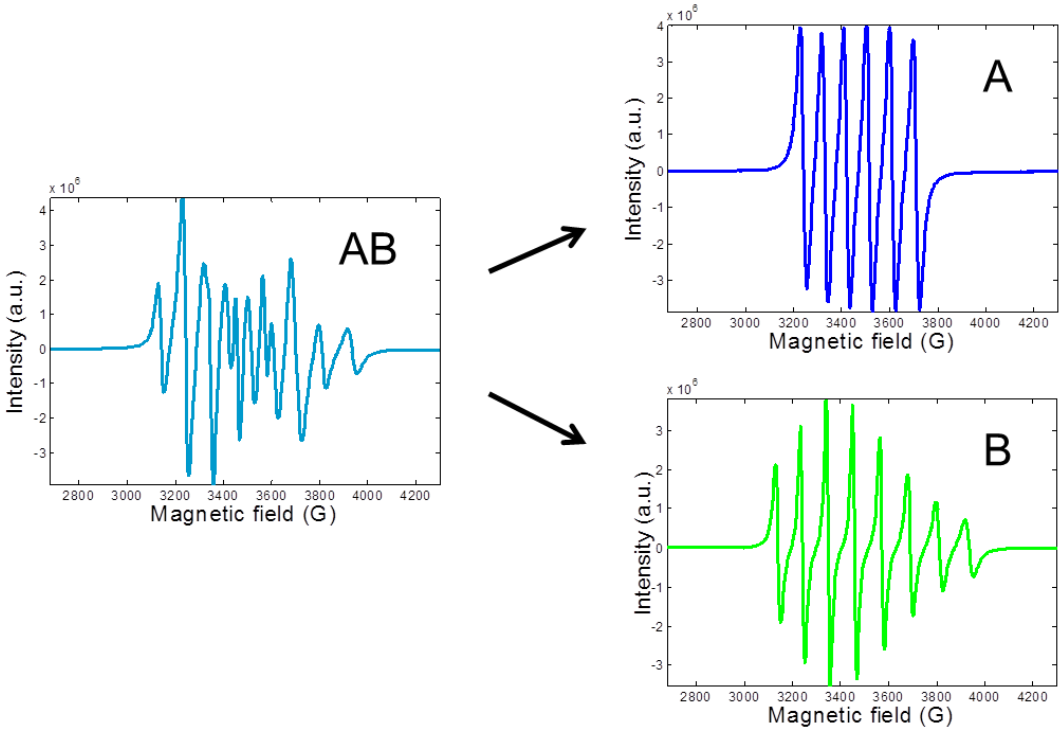


Figure III.1 The concept of resolving complex EPR spectrum composed of two multipeak and strongly overlapped constituents.

In the presented methodology, PCA provides valuable preliminary information about the most important variables involved in the dataset as well as the identification of the number of pure components. Thus, it constitutes the basis for further analysis. MCR-ALS is proposed here as the most suitable chemometric method for the final resolution of EPR dataset. It is

applied to obtain the concentration and the spectral profile of each pure component from the sole raw data table.

In order to validate the proposed methodology, two experimental model systems were chosen and carefully prepared. The first one consists of mixtures of two well-characterized and non-interacting paramagnetic ions VO^{2+} and Mn^{2+}. It means that the EPR spectral contribution of a compound is not distorted by the presence of the second species. Both ions exhibit multipeak EPR spectra which overlap each other. The second model system was similar to the first one but supplemented with Cu^{2+} ions to increase the complexity of the mixture. In both cases, the set of samples was prepared as mixtures with a different ratio of concentrations of the individual components. Then, EPR spectra were acquired and explored first by PCA, then, treated by MCR-ALS algorithm to retrieve the concentration and the spectral profile of each pure constituent.

The results presented in this chapter represent the first use of MCR-ALS method on the EPR experimental spectra and they demonstrate the benefits of applying both PCA and MCR-ALS on the complex EPR data set. These results have been partially published in [208].

III.2 Sample preparation

Three compounds containing paramagnetic ions were chosen to prepare binary and ternary mixtures in distilled water by varying the concentration of the chemical species. The first compound is cupric sulfate anhydrous ($CuSO_4$), the second compound is manganese II acetate tetrahydrate ($(CH_3CO_2)_2Mn.4H_2O$) and the third compound is vanadyl sulfate ($VOSO_4.5H_2O$). Consequently, Cu^{2+}, Mn^{2+} and VO^{2+} are the three paramagnetic ions which can be present in the experimental solutions, respectively. Distilled water was selected for this work because all the chosen compounds are well soluble in this solvent, whereas in other considered ones i.e. methanol and ethanol at least one of the three compounds was not soluble enough to get the desired concentrations.

Two-component system. Twelve different aqueous mixtures were prepared containing the two paramagnetic centers VO^{2+} and Mn^{2+} with concentration ranges of 10^{-1} to 10^{-4} M and 10^{-2} to 10^{-4} M, respectively (see Table III.1).

Sample number	Concentration of Mn^{2+} ions (in M)	Concentration of VO^{2+} ions (in M)
1	10^{-2}	10^{-1}
2	10^{-2}	10^{-2}
3	10^{-2}	10^{-3}
4	10^{-2}	10^{-4}
5	10^{-3}	10^{-1}
6	10^{-3}	10^{-2}
7	10^{-3}	10^{-3}
8	10^{-3}	10^{-4}
9	10^{-4}	10^{-1}
10	10^{-4}	10^{-2}
11	10^{-4}	10^{-3}
12	10^{-4}	10^{-4}

Table III.1: Concentrations of Mn^{2+} and VO^{2+} ions in each aqueous mixture prepared in the two-component system.

Figure III.2 shows the plot of the concentration profile of manganese and vanadium ions.

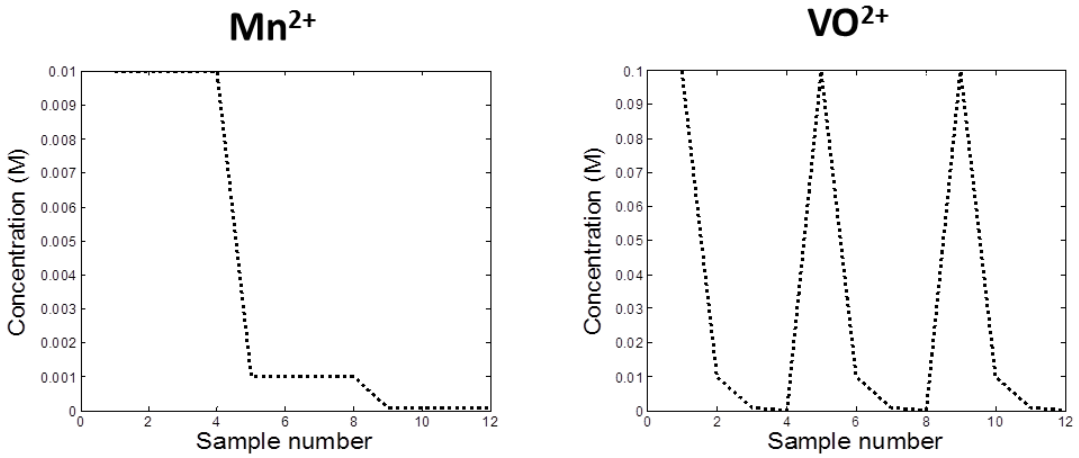


Figure III.2 : Plot of the concentration profiles of manganese and vanadium ions.

Three-component system. In addition to the previous two-component system, a three-component system was prepared from the three paramagnetic centers Cu^{2+}, VO^{2+} and Mn^{2+} ions by varying the concentration of each between 10^{-1}M and 10^{-4}M to obtain seventeen different aqueous mixtures (see Table III.2). Systematic linear correlation among the concentration of the ions in the consecutive samples was avoided.

Number of sample	Concentration of Mn^{2+} ions (in M)	Concentration of Cu^{2+} ions (in M)	Concentration of VO^{2+} ions (in M)
1	10^{-3}	10^{-1}	10^{-2}
2	10^{-4}	10^{-1}	10^{-2}
3	10^{-2}	10^{-2}	10^{-2}
4	10^{-3}	10^{-2}	10^{-2}
5	10^{-4}	10^{-2}	10^{-2}
6	10^{-2}	10^{-1}	10^{-3}
7	10^{-3}	10^{-1}	10^{-3}
8	10^{-4}	10^{-1}	10^{-3}
9	10^{-2}	10^{-2}	10^{-3}
10	10^{-3}	10^{-2}	10^{-3}
11	10^{-4}	10^{-2}	10^{-3}
12	10^{-2}	10^{-3}	10^{-2}
13	10^{-3}	10^{-3}	10^{-2}
14	10^{-4}	10^{-3}	10^{-2}
15	10^{-2}	10^{-3}	10^{-3}
16	10^{-3}	10^{-3}	10^{-3}
17	10^{-4}	10^{-3}	10^{-3}

Table III.2: Concentrations of Mn^{2+}, Cu^{2+} and VO^{2+} ions in each aqueous mixture prepared in the three-component system.

Figure III.3 shows the plot of the concentration profile of manganese, copper and vanadium ions.

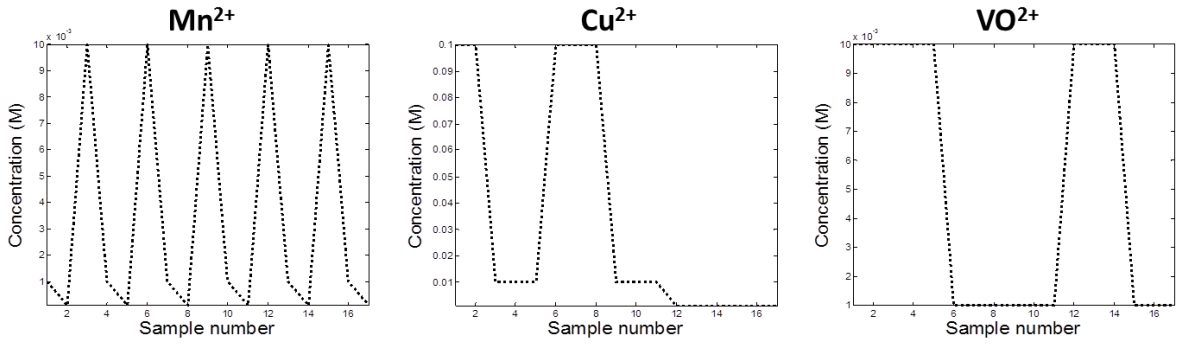


Figure III.3 : Plot of the concentration profiles of manganese, copper and vanadium ions.

III.3 Spectral acquisition

Samples were analyzed with a continuous wave-electron paramagnetic resonance spectrometer (CW-EPR). Once conditions were optimized for the measurement of the signal of interest, a spectrum required few seconds to be recorded. EPR spectra were recorded at a constant room temperature of 20°C, using a Bruker ELEXSYS 500 spectrometer operating at the X-Band. All the spectra were recorded at a modulation field frequency of 100 kHz and a microwave frequency of 9.80 GHz, with an amplitude modulation of 0.4 mT and a microwave power of 10 mW corresponding to non-saturation conditions. A flat cell was initially used since it is often used for liquid samples and represents, relatively, high precision and accuracy. After preliminary experiments, this cell was replaced with micro capillary tube due to reproducibility problems (see below). The weak pitch from Bruker was used as standard reference and contained a known concentration of spin/mass (1.29×10^{13} spins/g). The following spectrometer parameters were used during the spectral acquisition: receiver gain 40 dB, time constant 40.96 ms, conversion time 39.25 ms. The spin concentration is given by the double integration of the first derivative of the EPR signal. Spectra were measured over a 2048 point range.

For this study, the stability of the investigated solutions during *in situ* EPR analysis was first verified. The replicate samples were prepared several times and replicates of the EPR measurements were carried out into the spectrometer under the same conditions. For the first impression, EPR spectral data did not suffer significant variation after a couple of hours from the preparation i.e. the general shape was conserved. However, preliminary MCR-ALS analysis applied on the series of the spectra, revealed that a good reproducibility of the spectra was not achieved.

One of the key aspects of resolution of complicated multicomponent spectra is reproducibility of the experiments. During the preliminary study of this work, it was noticed that some peaks in the obtained EPR spectra suffered the spectral shifts rendering the data treatment very difficult (Figure III.4). The observed effect is especially undesired in MCR-ALS analysis because it does not favor such perturbations. Figure III.4 shows a series of spectra of a two-component. For example, a peak recorded at 3200 Gauss in one spectrum and the analogue peak in another spectrum (with different concentration ratio) recorded at 3220 Gauss due to some experimental shifts can represent the same component. However, MCR-

ALS procedure considers these two peaks as two different species while EPR specialists consider it as one peak at about 3200 Gauss. In this case, the data set loses its bilinearity character.

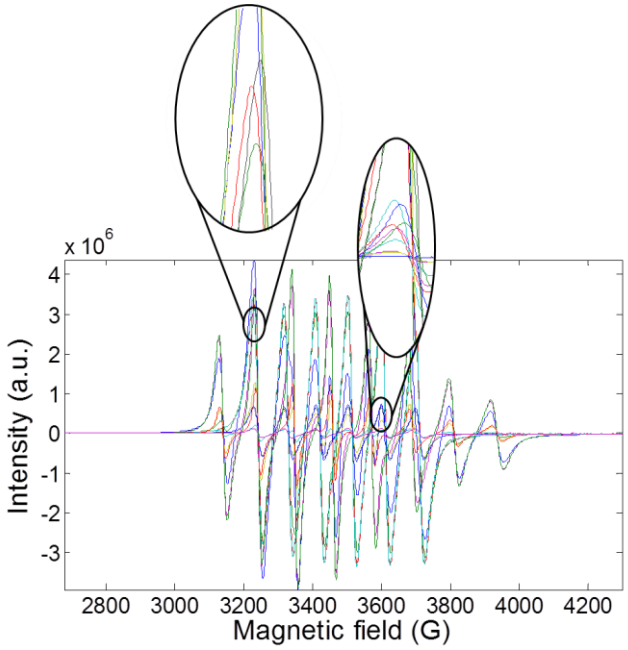


Figure III.4 : Shifts observed in the raw EPR dataset.

Moreover, Singular Value Decomposition is applied on the shifted matrix **D** and it gives a biased rank. Figure III.5 shows the plot of the first ten eigenvalues and three components are related to the chemical information instead of two components.

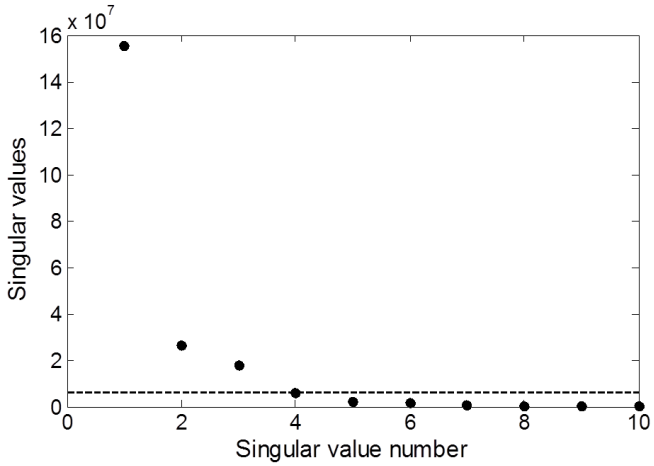


Figure III.5 : Eigenvalues obtained by Singular Value Decomposition of the shifted spectral data matrix **D**.

To avoid such misinterpretation, several pretreatments of the data set were done on Matlab to suppress and to solve the problem of such shifts. In the first approach, one of the

spectra among the series was fixed as a reference spectrum, then the position (in Gauss) of each peak in a spectra was compared with its analog peak in the reference spectrum. As a result, a given spectrum was shifted by several gauss in relation to the reference one that is a sort of alignment procedure. All the spectra were corrected in this way and MCR-ALS analysis could be applied. Unfortunately, it was noticed that sometimes the shifts were not homogeneous in all spectra for the corresponding peaks. In addition, the broadening and narrowing of some bands occurring from one spectrum to another were also detected. Even though the problem of the broadening/narrowing of the peaks is a subject of modern chemometric methods, the established solution of this issue has not been achieved yet. For this reason, an effort was made to eliminate or, at least, to minimize such unwanted effects on the instrumental level. It is commonly known that a multitude of error sources arising from instrumental and sample-associated factors may affect the accuracy and reproducibility of EPR spectral acquisition. The most popular examples are sample material (dielectric constant), sample size and shape, sample cell wall thickness as well as sample orientation and positioning in the microwave cavity [209]. The majority of these errors can occur simultaneously and synergistically and can be controlled by the EPR spectrometer operators and users. The most effective way to minimize the influence of such errors in multisampling experiments is the standardization of procedures for all the measurements and post-recording spectral manipulations. In this work, the sample material, sample size and shape as well as the sample cells were kept the same through all the different experiments. In addition, the same quantity $(40\mu L)$ of each sample was added in the flat cell for each acquisition. However, the very careful inspection of the experimental procedures revealed that the reason of the peak shifts and broadenings originated from small differences in the sample position and orientation that may induce relatively large changes in the resonator quality factor (Q value). In consequence, the spectrometer sensitivity varied between different measurements with the subsequent loss of accuracy and precision [210]. Even though the flat cell used in this study was later accurately positioned with the sample holders and carefully aligned in the same orientation for each sample, i.e. the plane part facing the radiations, the broadening of the peaks was still observed although with a slight amelioration. Nevertheless, more precision was still required for the chemometric treatments. Satisfactory results were finally obtained by the use of a micro capillary tube instead of a flat cell (see Figure III.6). The crucial advantage of micro capillary tubes $(50\mu L)$ is higher quality and more homogeneous wall thickness that keeps the quality factor constant over all the measurements.

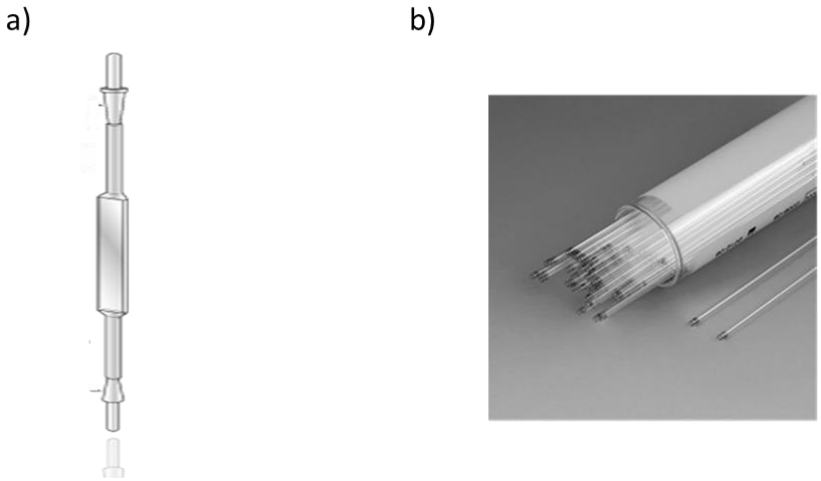


Figure III.6 : Experimental cells used in EPR measurements: a) flat cell b) micro capillary tubes.

After the replacement of the cell, all the EPR spectral acquisitions were repeated and the satisfying reproducibility as well as stability of the spectra for further chemometrics analysis were achieved.

III.4 EPR spectral signature of Cu^{2+}, Mn^{2+} and VO^{2+} single ions

Prior to the study of the mixtures, three different aqueous solutions containing the pure paramagnetic ions Cu^{2+}, Mn^{2+} and VO^{2+} were prepared, respectively. Their corresponding spectra acquired by EPR spectroscopy are shown in Figure III.7. These spectra were acquired for the sake of comparison and they serve as reference spectra to be compared with the pure spectra resolved from multicomponent systems that will be obtained by chemometric methods.

Copper ions (Cu^{2+}) possesses an electron spin $S = \frac{1}{2}$ and a nuclear spin $I = \frac{3}{2}$, therefore according to the equation $2I + 1$, four absorption maxima are expected. However, working in the conditions of aqueous media makes the four closely lying lines unresolved and merged in one broad peak (singlet) which appears around 3200 G (Figure III.7a).

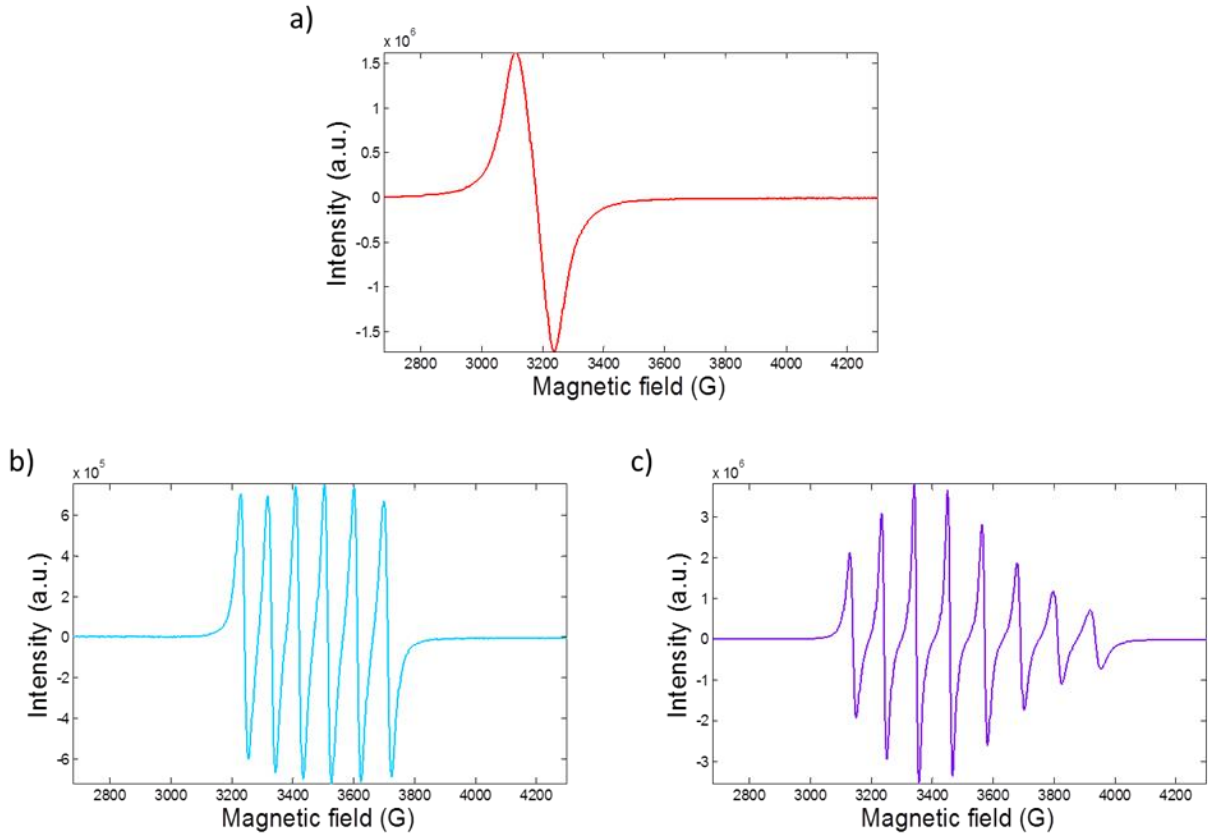


Figure III.7 : Raw EPR spectral signals of a) Copper b) Manganese and c) Vanadium ions.

Manganese ion (Mn^{2+}) exhibits an electron spin $S = \frac{1}{2}$ and a nuclear spin $I = \frac{5}{2}$, thus, six well-defined resonance lines (sextet) can be obtained in EPR spectrum (Figure III.7b).

Vanadium ion (VO^{2+}) yields an eight-line hyperfine spectrum (octet) in isotropic media since it is a $S = \frac{1}{2}$ and $I = \frac{7}{2}$ system (Figure III.7c). The g-values of $CuSO_4$, $(CH_3CO_2)_2Mn.4H_2O$ and $VOSO_4$.5H2O in aqueous solution are 2.21, 2.06 and 1.96, respectively [211-213].

Since the integrated spectra are sometimes more convenient for chemometric analysis due to their non-negativity (see below) such spectrum of each paramagnetic species was calculated and is presented in Figure III.8.

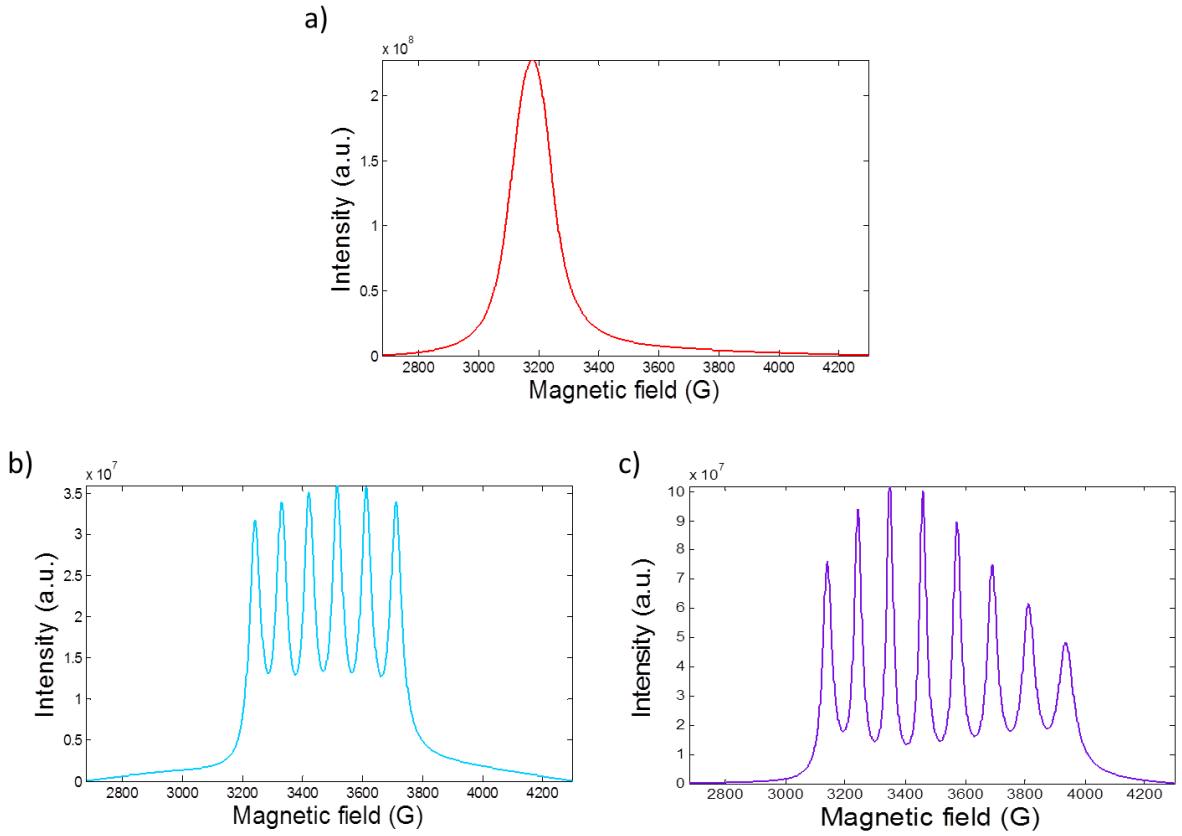


Figure III.8 : Integrated EPR spectral signals of a) Copper b) Manganese and c) Vanadium ions.

III.5 Results 2-component system

The first studied system was a two-component mixture composed of Mn^{2+} and VO^{2+} cations. The raw spectra of all twelve prepared mixtures are presented in Figure III.9 and are kept in their first derivative form. Each spectrum corresponds to the aqueous solution with various concentration of both compounds (Mn^{2+} and VO^{2+}) acquired at 2048 different points (from 2678 Gauss till 4300 Gauss). It is worth noting that no pure spectrum, corresponding to one of the components, is present among all the spectra. Moreover, one can notice the appearance of new peaks as a result of the spectral overlapping between Mn^{2+} and VO^{2+} ions whereas some other peaks are hidden. It indicates that the straightforward identification of the number of components (assuming no *a priori* knowledge) as well as the extraction of individual spectral components is not possible.

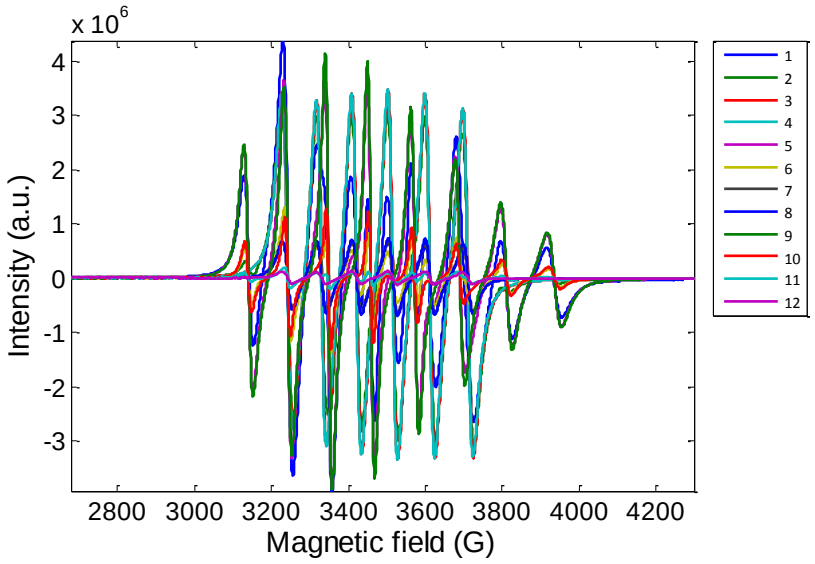


Figure III.9 : EPR spectra of twelve aqueous binary mixtures containing various concentrations of Mn^{2+} and VO^{2+} ions. The numbers in the legend correspond to the samples described in Table III.1.

For chemometric analysis, the series of twelve spectra is represented in a matrix **D** i.e. 12 (objects) different rows where each row corresponds to an EPR spectrum of a given binary mixture, and 2048 (variables) columns where each column corresponds to a magnetic field (see Figure III.10). For the sake of initial exploration and estimation of the number of components, PCA was first applied on the data followed by MCR-ALS application for the extraction of the pure concentration and the corresponding spectral profiles.

III.5.1 Principal Component Analysis (PCA)

From the statistics point of view, the considered data matrix can be described as 12 spectra which are placed in a 2048 (variables) axis-system. PCA is, thus, used to reduce the dimensions of the original set of interrelated variables (2048) into sets of uncorrelated *k* variables and retains, as much as possible, the variation present in the original data set. It linearly decomposes the matrix **D** into two smaller matrices i.e. scores **T** and loadings $\mathbf{L}^{\mathbf{T}}$ (see Figure III.10).

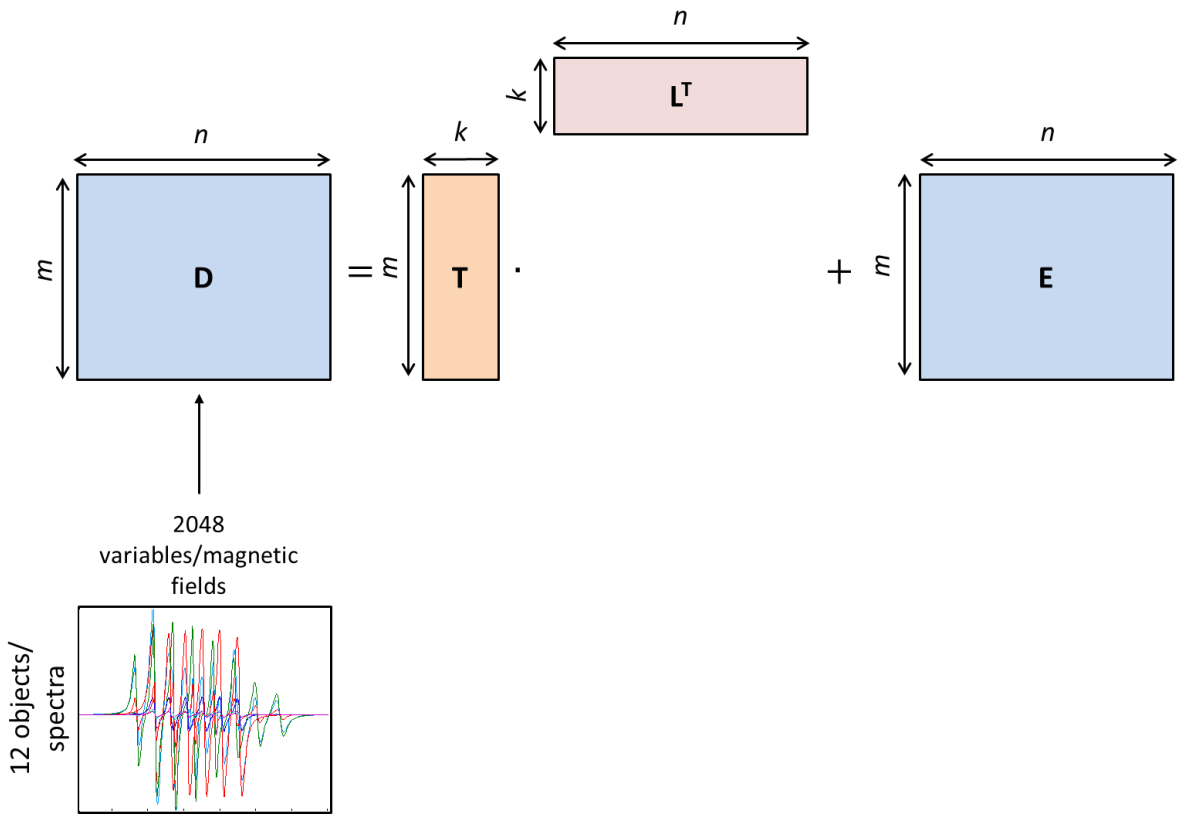


Figure III.10 : General concept of PCA.

The *k* derived variables which maximize the variance accounted for in the original variables are the Principal Components (PC) and establish new set of variables. In this case, two Principal Components PC1 and PC2 (among 12 possible PCs) represent the most variation linked to the chemical composition of the spectra and the rest are considered as nonsignificant PCs that contribute to the noise related signal. The first Principal Component (PC1) explains the 65.25% of the total variance (maximum variation), whereas the second Principal Component (PC2) shows the second maximum variation (34.23%) and the next PCs are not mentioned here due to their negligible variation. Therefore, the first two PCs present almost 99.5% of the overall variation and consequently they can, plausibly, express the total system. This result confirms that from the sole raw data matrix, PCA can correctly predict the presence of two (k=2) significant components in the binary mixtures. Thus, the dimensionality of the data set was reduced from 2048 correlated dimensions to 2 uncorrelated dimensions that retain most of the information from the original variables. In other words, the 2048 axis-system (original coordinate system) was reduced into a 2 axis-system (new coordinate) where the PCs are orthogonal. This new system is presented in Figure III.11 where the samples are presented by numbered dots (from 1 to 12) corresponding to the samples gathered in Table III.1.

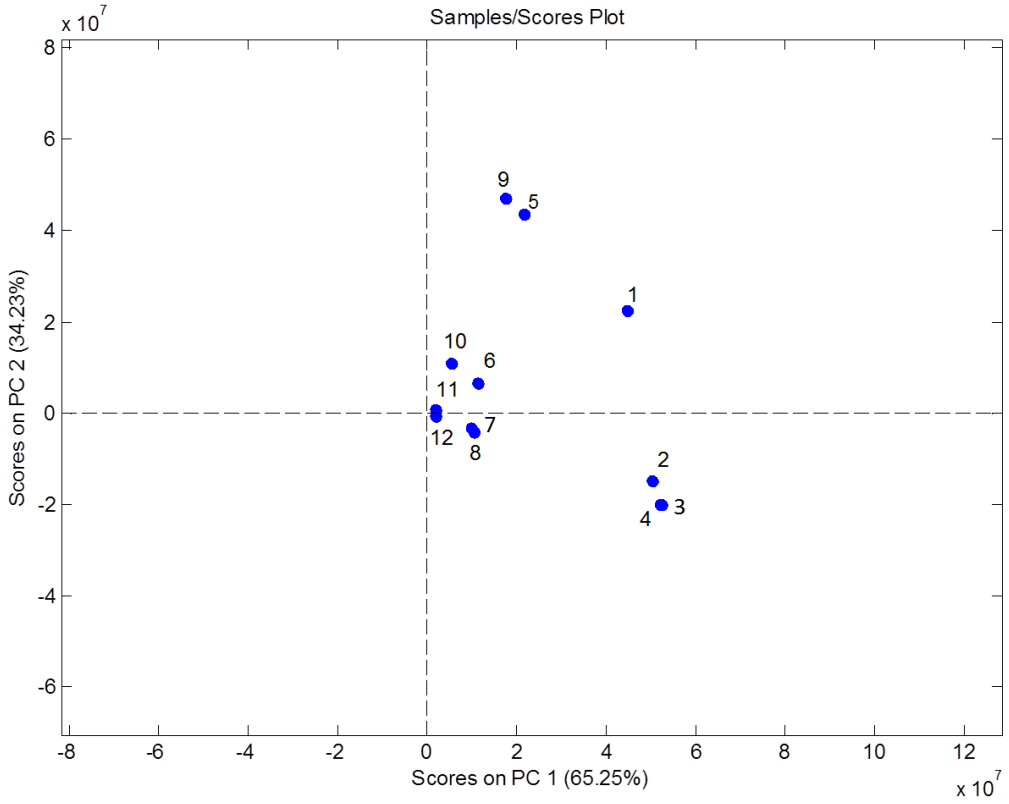


Figure III.11 : Score plots of PC1 and PC2.

As aforementioned, the bilinear decomposition of the matrix **D** gives two substantially smaller sets of uncorrelated variables i.e. scores **T** and loadings $\mathbf{L}^{\mathbf{T}}$. Figure III.12 presents both scores and loadings on PC1 and PC2. They represent the projection of 12 samples on the first Principal Component (PC1) in blue and on the second Principal Component (PC2) in magenta. The corresponding loadings $\mathbf{L}^{\mathbf{T}}$ are presented as well by the same colors.

Careful inspection of the results presented in Figure III.11 and Figure III.12 provides a series of interesting observations. According to the PC1-PC2 scores plot in Figure III.11, it can be noticed that there exist four groups in which the spectra (rows in data matrix) are similar. Precisely, the first group contains spectrum 5 and 9, the second group includes the spectra number 2, 3 and 4, the third group is represented only by spectrum number 1 that is half way between the first and second group. Finally, the biggest group (fourth group) contains spectra 6, 7, 8, 10, 11, and 12. To see better the similarity of the grouped spectra, the plot of the second group (2, 3 and 4) is presented as an example in Figure III.13.

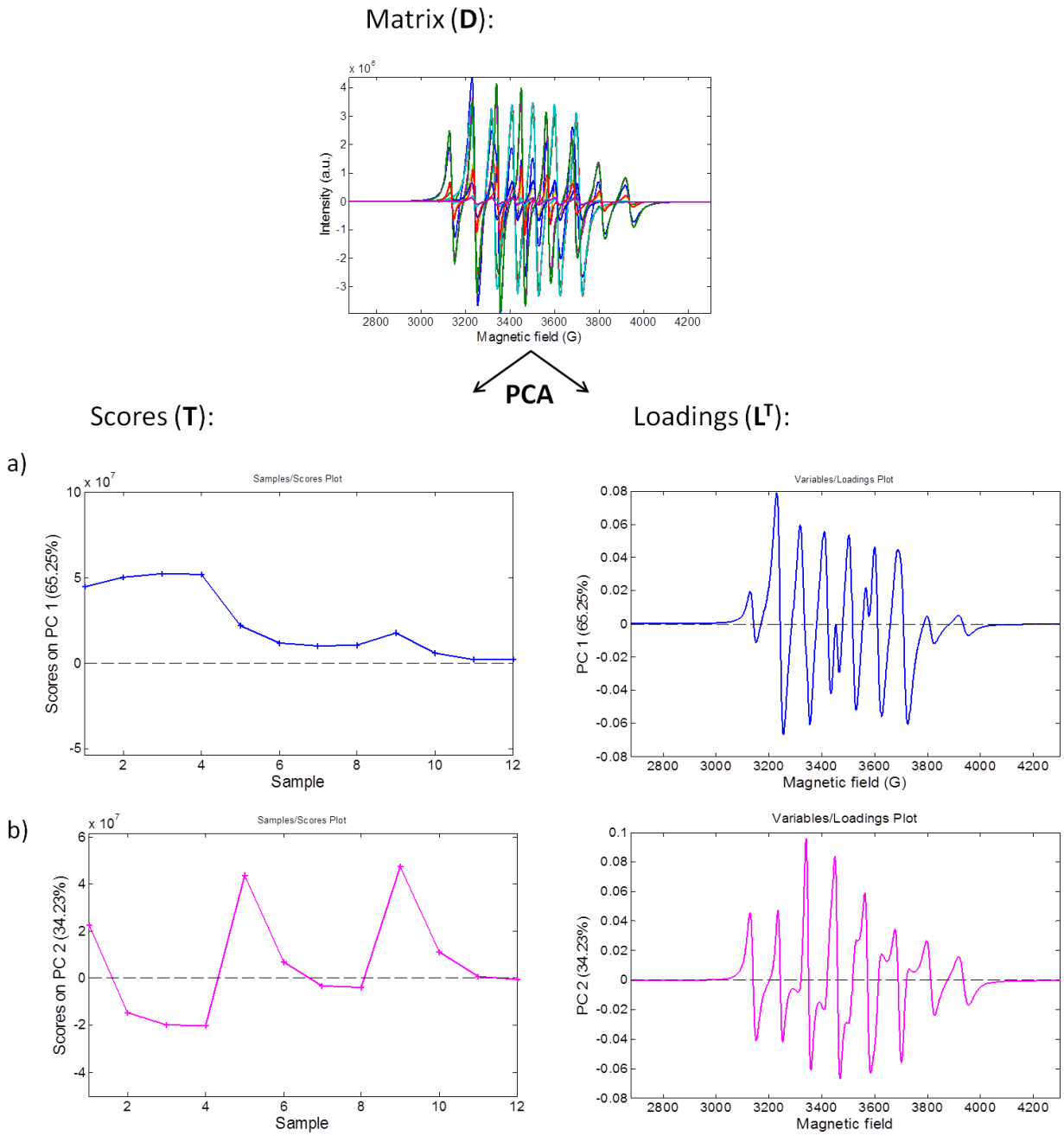


Figure III.12 : The plots of scores **T** and corresponding loadings $\mathbf{L}^{\mathbf{T}}$ obtained on PC1 (a) and PC2 (b) by bilinear decomposition of matrix **D**.

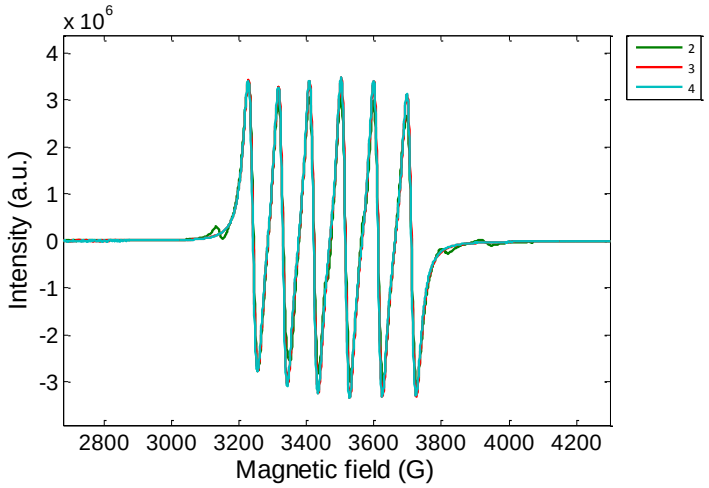


Figure III.13 : Plot of the spectra recorded for the samples 2,3 and 4.

On the other hand, the comparison of sample 1 with the spectra belonging to the second group i.e. 5 and 9 (see Figure III.14) confirms the fact that they are very different.

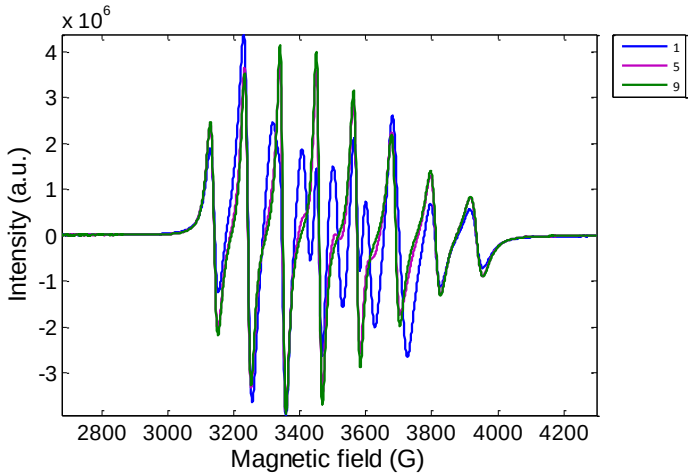


Figure III.14 : Plot of the spectra recorded for the sample 1, 5 and 9.

From another point of view, the inspection of Figure III.12 reveals that all the samples (1-12) exhibit positive scores on PC1. On the other hand, half of the samples (1, 5, 6, 9, 10 and 11) have a positive score on PC2, whereas the rest shows a negative score on PC2 (2, 3, 4, 7, 8 and 12). Figure III.11 shows that samples can be also grouped in another way. Namely, samples 1, 2, 3 and 4 have the same scores on PC1 but different values of PC2. These samples have the highest scores on PC1 compared to the other samples.

Moreover, the analysis of the loadings can indicate which variables in the dataset are important and which ones may be of little influence. Six main peaks in the loading plot of PC1 at 3229, 3318, 3408, 3502, 3600, 3687 Gauss (see Figure III.12) are peaks that

characterize the manganese ions (compare with Figure III.7). The other smaller peaks e.g. 3129, 3796, 3917 Gauss are peaks related to the vanadium that are not well separated. On the other hand eight peaks in the loading plot of PC2 at 3129, 3234, 3340, 3449, 3562, 3677, 3796, 3917 Gauss are the bands that characterize the vanadium ion (see Figure III.7). For instance, the sets of variables from 2678 till 3020 and from 4100 till 4300 do not have any importance on the two Principal Components (PC1 and PC2) and this indicates that these variables have no contribution to the variation in the dataset (silent regions).

Combining the scores and loadings, it is visible that the samples that have high scores on PC1 contribute strongly in the spectral signature of PC1 loadings, whereas the samples that have high scores on PC2 predominate in the spectral signature of PC2 loadings. The scores on the Principal Components are similar to the concentration profile of the species i.e manganese and vanadium, respectively (Figure III.2). Therefore, based on the PCA interpretations, more information than only the number of pure species present in the mixture can be deduced.

The above analysis proves that information obtained by PCA application with no prior knowledge can be useful. The scores give an insight into the concentration profile of the species present in the mixture, whereas the related loadings include the spectral information of the species. It is worth stressing that PCA decrease the dimensionality of the spectra preserving all the relevant information enclosed in the Principal Components. Despite the unquestionable benefits PCA possess, however, it has important drawbacks. On one hand, the loadings on PC1 and PC2 still seem mixed and not pure. On the other hand, the negative scores in PC2 are unacceptable for real concentrations. These negative values are due to the fact that PCs are calculated to keep the null correlation among them (loadings are uncorrelated). However, the real spectra do not obey this condition. Therefore, scores and loadings establish an excellent first approach to visualize the series of spectra in small dimensions without losing the richness of the information of the original measurements.

III.5.2 MCR-ALS

After the exploratory analysis of PCA, MCR-ALS was applied on the same set of twelve binary mixtures. Similar to PCA, this method is also based on the bilinear decomposition of the matrix **D** that contains the raw data (12 spectra) to obtain 2 matrices i.e.

the concentration profiles **C** and the spectral signature $\mathbf{S}^T$ of each pure component (Figure III.15). The matrix **D** represents (*m*) 12 rows where each row corresponds to a spectrum acquired at a certain concentration of manganese and vanadium ions, and (*n*) 2048 columns where each column corresponds to a given magnetic field.

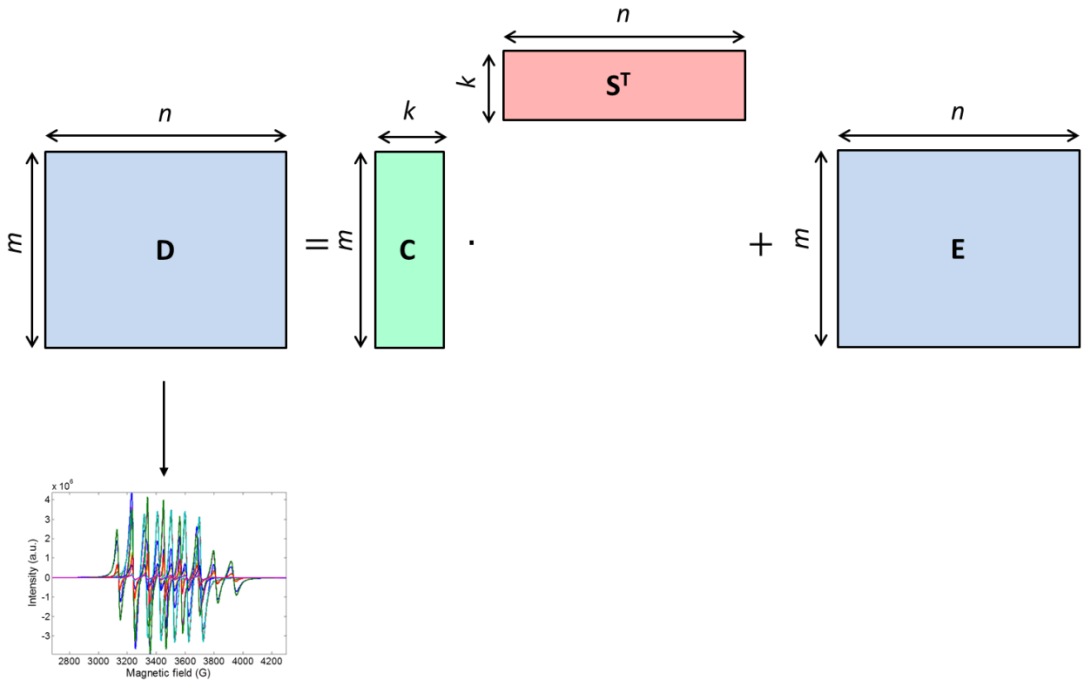


Figure III.15 : General concept of MCR-ALS.

As described in Chapter II, MCR-ALS method is divided into the consecutive steps: rank estimation, initial estimation and constrained ALS procedure to obtain the final results.

III.5.2.1 Rank estimation

The first step performed in MCR-ALS was rank estimation of the initial matrix **D**. The aim of this step was to determine the number of independent components (assuming no *a priori* knowledge) present in the experimental mixture. Actually, such information was already obtained by the application of PCA (see above) which correctly predicted two Principal Components (PC1 and PC2), hence, two species in the experimental mixtures. In practical use of MCR-ALS, one of the PCA based methods, i.e. Singular Value Decomposition (SVD), is typically used to determine the mathematical rank *k* of the data matrix. SVD analysis performed on the studied data set confirmed the presence of two distinct

pure components in the investigated mixtures. The plot of the first ten eigenvalues obtained from the matrix **D** is shown in Figure III.16.

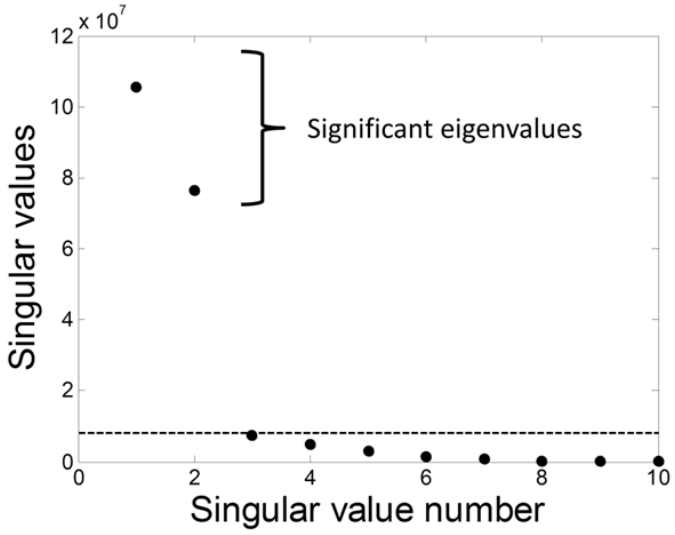


Figure III.16 : Eigenvalues obtained by Singular Value Decomposition of spectral data matrix **D**.

Considering the large gap observed between eigenvalues, there is a clear difference among values related to the components with the chemical information and to the noise. The dotted line gives an indication of the noise level in the eigenvalue scale. Above this line, two considerably higher eigenvalues are observed and, therefore, the rank of the matrix is two and two chemically meaningful contributions are deduced even without prior knowledge of the chemical system.

III.5.2.2 Initial estimation

The next step in the MCR-ALS process was the initial estimation of the concentration or spectral profiles. In the studied case, SIMPLISMA was chosen as a method (see Chapter II for details) among the wealth of present methods since the data in the initial matrix do not follow any monotonic order in concentration direction. In this context, initial estimate may be finding either the purest magnetic fields or the purest sample number in the raw data. The rows related to the purest sample number give an approximation of the pure spectra of the constituent in the mixture, and the column related to the purest spectral channel allow the

estimation of the concentration profile. The purest spectra and concentration profiles obtained by SIMPLISMA give an exploratory overview about the characteristics of the various compounds in the sample. It cannot be associated to the pure components since selective sample numbers and spectral channels in the measurements are not ensured. The sample number direction was chosen, the purest spectra $\mathbf{S}^{\mathbf{T}}_{\mathbf{ini}}$ of the binary mixtures are obtained. In general, SIMPLISMA is a method designed to work with positive data. Therefore, initial estimates of the EPR spectral profiles were obtained by applying this method on the matrix containing the absolute values of the derivative spectra in **D**. In this way, the rows of the most dissimilar spectra could be obtained and, then, original EPR spectra related to these rows were used as initial estimates (Figure III.17). Spectral profiles obtained by this method were used as initial estimates for further ALS analysis.

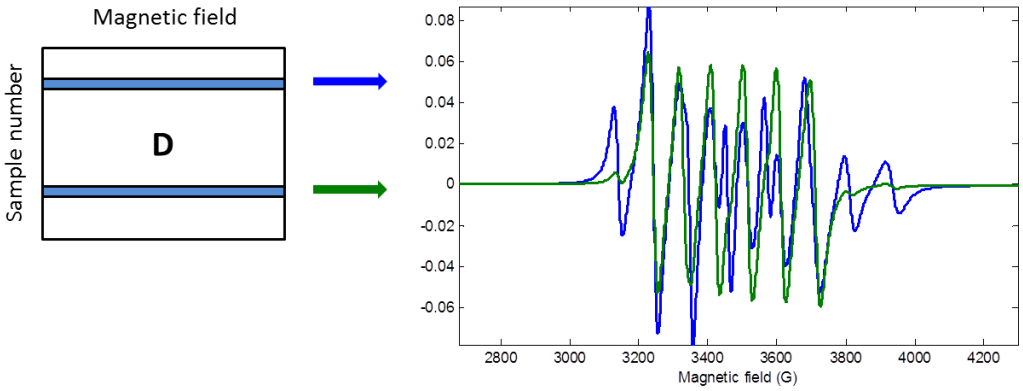


Figure III.17 : SIMPLISMA analysis on EPR data set. Representation of the spectra of the purest samples $\mathbf{S}^{\mathbf{T}}_{\mathbf{ini}}$.

III.5.2.3 Constrained ALS

After the estimation of the matrix rank $k = 2$ as well as the initial spectra, alternating least-squares optimization process according to the equation $\boldsymbol{D} = \boldsymbol{C} \cdot \boldsymbol{S}^{\boldsymbol{T}}$ was started by using the original raw measurements of the matrix **D** and the initial guess of spectra $\mathbf{S}^{\mathbf{T}}_{\mathbf{ini}}$ matrix. The constraints implemented during the optimization step were non-negativity for the concentration profiles and normalization for the spectral profiles. The iterative process of alternating least squares was stopped when the convergence criteria were achieved. The results obtained from overall MCR-ALS analysis of **D** are presented in Figure III.18. The

final results represent the pure spectra matrix $\mathbf{S}^T$ and the related pure concentration matrix **C** of studied manganese/vanadium ions mixture.

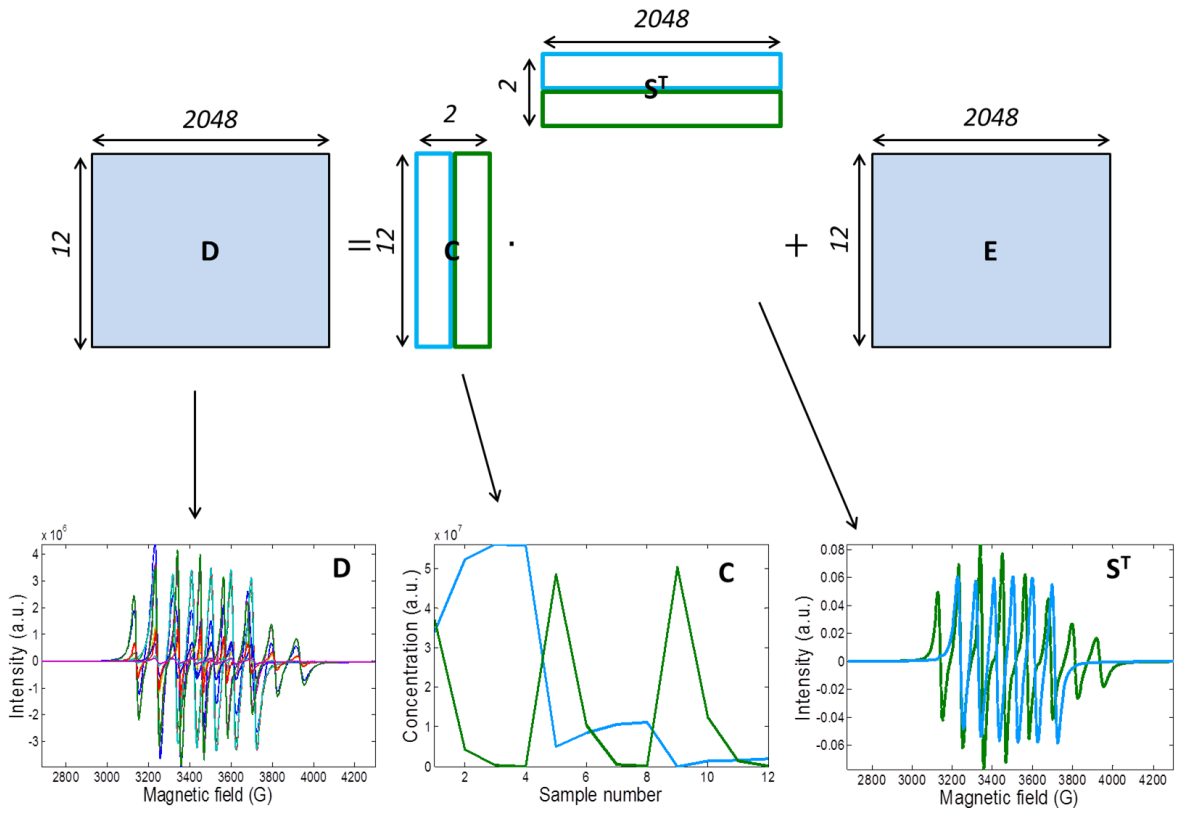


Figure III.18 : Application of MCR-ALS on raw data set **D** to obtain the concentration profile **C** and the spectral signature $\mathbf{S}^T$ of the paramagnetic ions VO^{2+} and Mn^{2+}. The first column (blue) in matrix **C** is the concentration profile of manganese ions and corresponds to the first row in the spectral profile $\mathbf{S}^T$ that is the spectral signature of this species. The second column in the concentration profile **C** (green) represents the concentration profile of vanadium ions and its corresponding spectral signature in matrix $\mathbf{S}^T$ is the second row.

The first column (blue) in matrix **C** corresponds to the concentration profile of Mn^{2+} and the first row in the spectral profile $\mathbf{S}^T$ (blue) is attributed to the spectral signature of this ion. The second column in the matrix **C** (green) represents the concentration profile of VO^{2+} and its corresponding spectrum is shown as second row in matrix $\mathbf{S}^T$ (green).

III.5.2.4 Reliability of MCR-ALS solution

The quality of the obtained solution was first evaluated by inspection of residuals and calculation of statistical parameters. The lack of fit (see eq. II.14 in Chapter II) below 7% was

determined, whereas explained variance (see eq. II.15 in Chapter II) was above 99%. Such values are satisfying results because they correspond with the experimental noise limits. The residual matrix **E** (see Figure III.19) has also been checked qualitatively and quantitatively in order to assess the reliability of the resolution. The intensity of the residuals is negligible (10^{-5}) compared to the initial dataset (10^{6}). However, residuals have a spectral structure that can be due to small shifts still present in the initial data set or due to impurities.

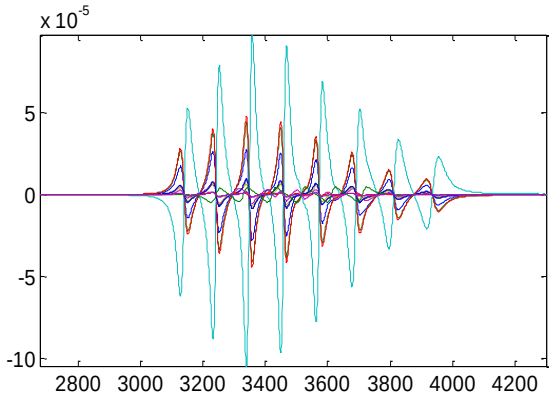


Figure III.19 : Residuals obtained at the end of MCR procedure applied on the binary mixtures.

As discussed in Chapter II, every MCR-ALS solution is not unique and ambiguity problem should also be verified. For this reason, the feasible band boundaries, i.e. MCR bands, can be calculated (see paragraph II.3.2.5 in Chapter II). They represent the minimum and maximum boundaries of the calculated profiles and, therefore, can monitor the extent and the location of the ambiguity associated to the results. MCR-bands were calculated for the two components in their spectral and concentration profiles. The difference between the minimum and maximum boundaries is zero. Therefore, the two paramagnetic ions are resolved in a unique way and there is no rotational ambiguity.

In order to definitively check the correctness of the obtained results and in consequence to validate the used approach, the calculated profiles were compared with the real concentrations and spectra. Figure III.20 shows the comparison between the pure spectra extracted by MCR-ALS (solid red line) and the corresponding experimentally measured spectra of individual manganese and vanadium ions (dotted black line obtained from Figure III.7). Impressive results can be noticed since it is visible that the perfect retrieval of the pure spectra for each paramagnetic ion was achieved.

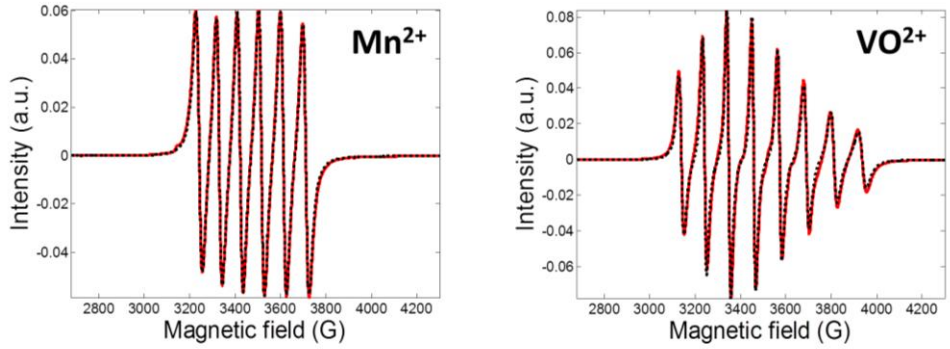


Figure III.20 : Pure spectra extracted by MCR-ALS (solid red line) and the corresponding experimentally measured spectra of pure manganese and vanadium ions (dotted black line).

In this part, Figure III.21 presents the comparison between the pure concentration profiles extracted by MCR-ALS (solid red line) and the reference data (dotted black line) of manganese and vanadium ions, respectively. It is necessary to mention in this place that the relative concentrations between different species in a given mixture cannot be directly compared because they are represented in arbitrary units scaling during the optimization process. However, the concentration of a given species between the different samples is comparable and the profile shape expresses the actual concentration evolution from sample to sample. In this case, the profiles are almost retrieved with only slight differences.

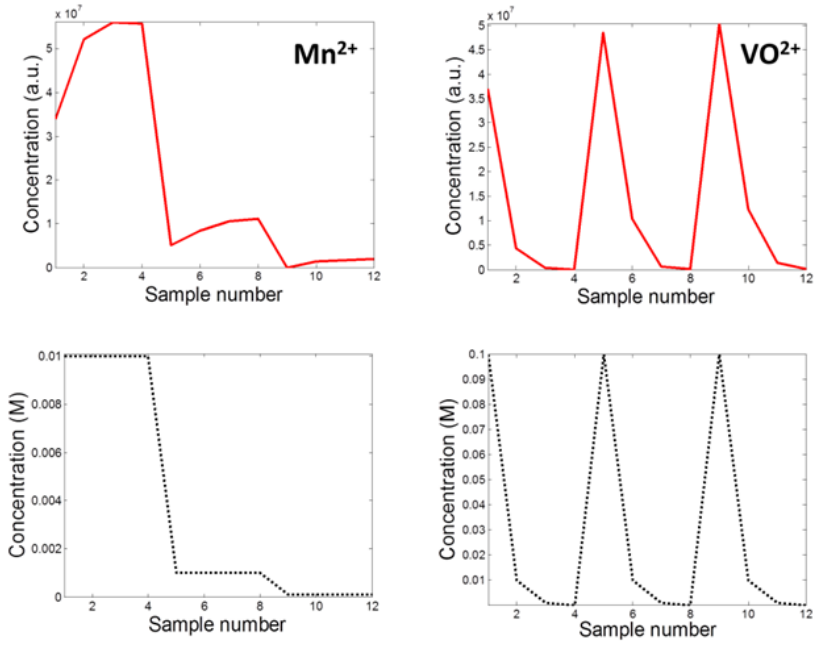


Figure III.21 : Pure concentration profiles extracted by MCR-ALS (solid red line) and the corresponding reference concentrations of manganese and vanadium ions (dotted black line).

Thus, it can be concluded that MCR-ALS algorithm is able to recover the true EPR features of the experimental data sets for binary mixtures with strongly overlapped multipeak spectral signatures. Applying MCR-ALS algorithm on this first dataset confirmed the feasibility of the study with no prior knowledge about the mixture and prompted us to investigate more complicated systems.

III.6 Results 3-component system

After the analysis of binary mixtures, the new three-component system was prepared (see Table III.2) to check the applicability of MCR-ALS on the more complicated EPR data set. Addition of Cu^{2+} ion to the previously studied two-component system entangle significantly a spectral blend because Cu^{2+} ion exhibits a broad spectrum extending over the signal of the two other ions (Mn^{2+} and VO^{2+}). Such a situation is much more demanding from the analytical point of view. It was confirmed during the first attempts when PCA and MCR-ALS were performed directly on the raw data, in their first derivative form (Figure III.22a). This approach was not efficient especially in the results obtained by MCR-ALS because of the presence of more components with the ubiquitous spectral overlapping. Due to the complexity of the system and only two constraints applied (i.e. non-negativity on the concentration direction and normalization on spectral profile), the uncertainty linked to the final solutions was too high. To overcame this difficulty and to reduce the rotational ambiguity, the integrated spectra (Figure III.22b), generated by a home-made code, were used instead of the original data in the analysis. Such an approach made it possible to use an additional constraint i.e. non-negativity in the spectral direction. Finally, MCR-ALS was carried out on the integrated data, under the constraints of non-negativity on both, concentration and spectral directions in addition to the normalization in the spectral direction.

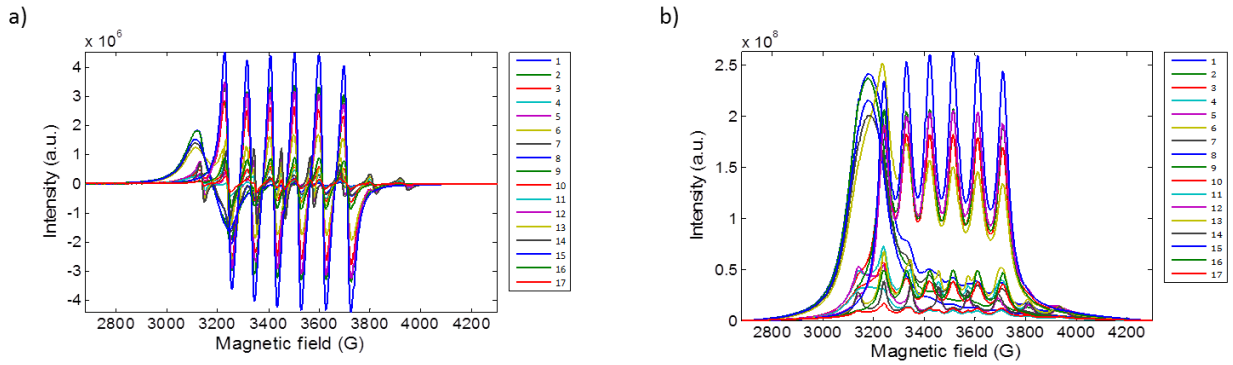


Figure III.22 : Seventeen EPR spectra of the aqueous ternary mixtures containing manganese, vanadium and copper ions with various concentrations presented a) in their first derivative form b) in their integrated form.

All spectra of seventeen mixtures (see Table III.2) were recorded in the same spectral range as in case of binary mixtures (2678 – 4300 Gauss). Some selected spectra, in their integrated form, are presented in Figure III.23. It can be clearly noticed that they reveal a high overlap degree. In particular, in Figure III.23a, the major broad band arises from the copper ions in addition to some weak peaks that appear related to the Mn^{2+} and VO^{2+} ions. On the other hand, Figure III.23b, shows the situation when the peaks originating from Mn^{2+} and VO^{2+} become dominant. In both cases, it is evident that visual extraction of the number and the shape of the pure spectra present in the mixtures (in the case of no prior knowledge) are practically unfeasible. In addition, it is also very difficult to detect and identify low concentration species with conventional analysis methods. Thus, the data matrix consisting all integrated seventeen spectra (ordered in rows) was built and analyzed by PCA and MCR-ALS methods.

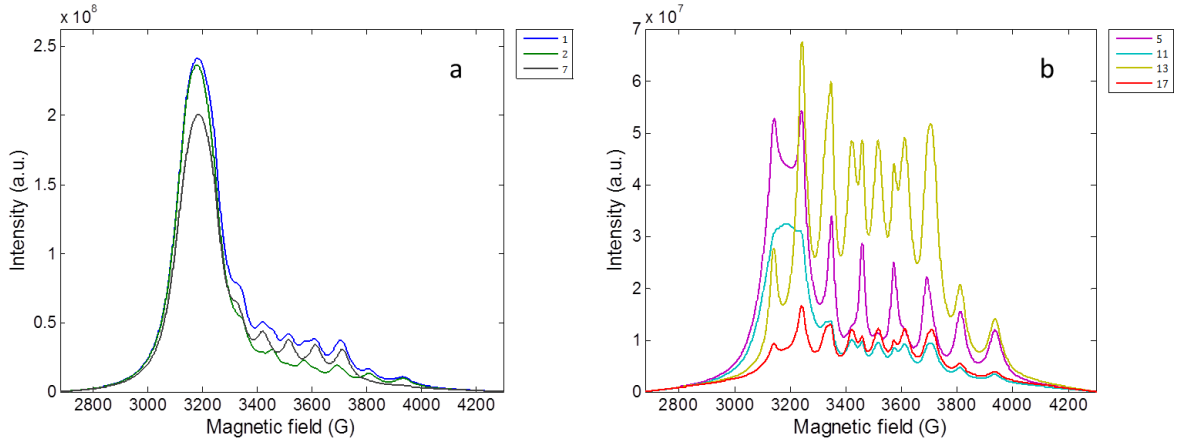


Figure III.23 : Selected integrated EPR spectra of ternary mixtures showing the high level of complexity and signal overlap: a) spectra with Cu^{2+} dominant signal b) spectra with Mn^{2+} and VO^{2+} dominant signal.

III.6.1 PCA

Principal Components were calculated from the integrated data. As a result, the first Principal Component (PC1) explains the maximum variance of 81.93% of the total variance, whereas the second Principal Component PC2 expresses the variance of 17.87%. The third Principal Component PC3 represents the much lower value of 0.16%. This results show that the presence of the third component in the mixture can be a subject of speculations. Thus, in contrast to 2-component system, in case of complicated tertiary mixture, the PCA method does not clearly answer what the number of components is. It is also visible in the plot of scores presented in Figure III.24 where the samples are presented by numbered dots (from 1 to 17). The large scatter of the data does not allow grouping the samples as easily as in the two-component system.

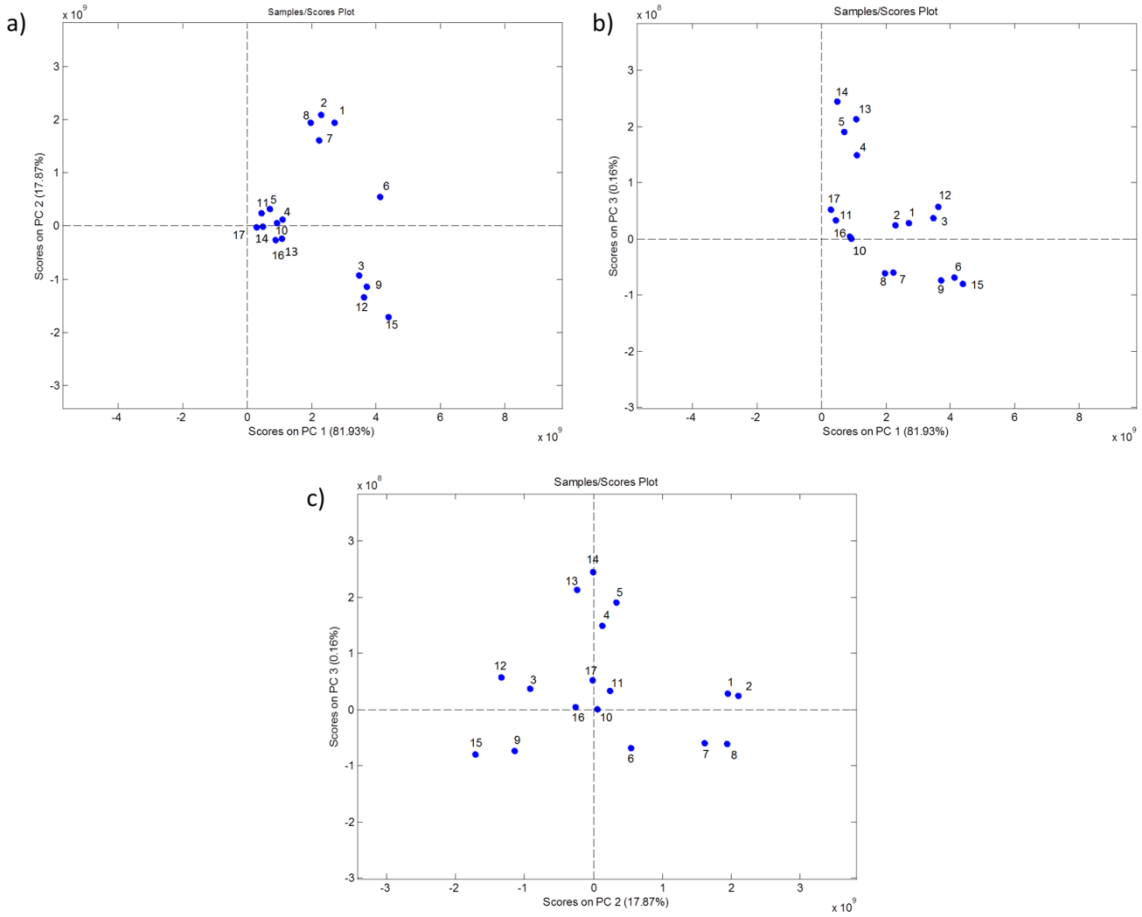


Figure III.24 : Scores plots of a) PC1 and PC2, b) PC1 and PC3 and c) PC2 and PC3.

Obscurity of the resolution is also clearly visible in Figure III.25, especially in the plots of the loadings. The six peaks in the loading plot of PC1 at variables 3240, 3331, 3422, 3515, 3612 and 3710 Gauss correspond well to Mn^{2+} sextet but important admixture of Cu^{2+} band is evident. On the other hand, the main band in the loading plot of PC2 at 3171G represents the EPR signal of the copper ion but the inverted peaks of Mn^{2+} (sextet) are present as well. These results show that the signals of Mn^{2+} and Cu^{2+} are considerably entangled. Finally, the loading of PC3 is composed of eight peaks corresponding to the signature of VO^{2+} (compare with Figure III.8) although the relative intensities of the peaks are not well expressed and the negative values of the signal have not chemical meanings. Therefore, the loadings of PC1, PC2 and PC3 represent similar spectral signature to Mn^{2+}, Cu^{2+} and VO^{2+}, respectively.

The scores on the Principal Components resemble the tendency of the concentration profile of each component (Figure III.3). For the same reason, the scores of PC2 and PC3 cannot be considered the very good representations of the actual concentration profiles which have to be positive.

The above analysis indicates that PCA suffers some important drawbacks in case of strongly overlapped 3-component EPR data set. Nevertheless, this technique should not be discredited as a preliminary analytical method. More specifically the loadings, despite their mixing character, clearly suggest that three different components with distinctive spectral signatures are present in the mixture. Therefore, PCA can be very useful first approach to estimate and visualize the series of the concentration and the spectral profiles before further advanced methods such as MCR-ALS.

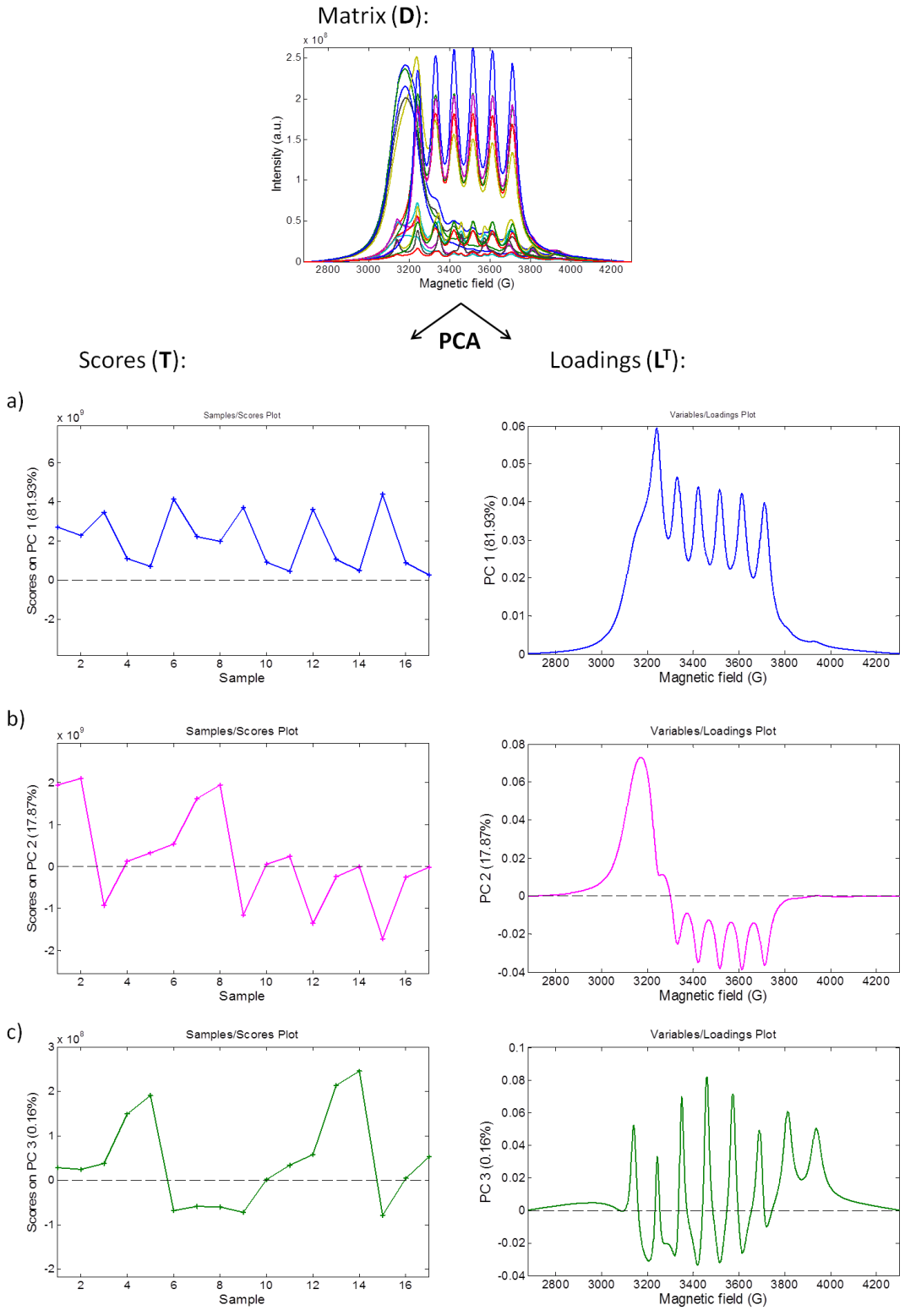


Figure III.25 : Bilinear decomposition of the matrix **D** to give scores **T** and their corresponding loadings plots $\mathbf{L}^{\mathbf{T}}$ a) on PC1, b) on PC2 and c) on PC3.

III.6.2 MCR-ALS

As with the binary mixtures data, a 3-component system was analyzed and resolved by the MCR-ALS procedure. The data matrix **D** was, thus, composed of 17 rows and 2048 columns. However as mentioned above, the integrated spectra were taken to build matrix **D** in this case.

III.6.2.1 Rank estimation

In order to determine the number of pure components, SVD was applied on the data set **D**. Figure III.26 indicates three significant singular values of the data matrix **D** i.e. the values which are clearly higher than others. Thus, the results demonstrate that the total number of species present in the studied system can be correctly evaluated by SVD method with no prior knowledge, from the sole information included in **D**.

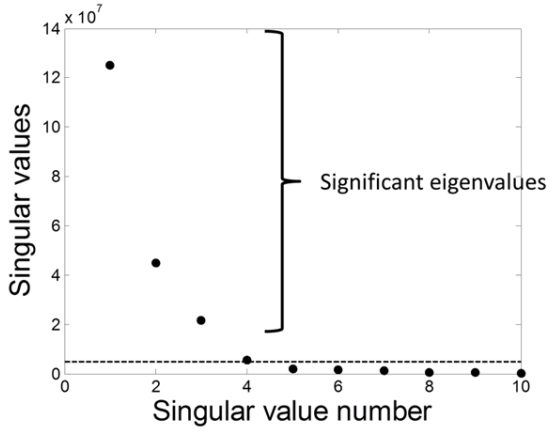


Figure III.26 : Eigenvalues obtained by Singular Value Decomposition of spectral data matrix **D**.

III.6.2.2 Initial estimation

As previously described for the binary mixtures, the initial estimates in the case of a 3-component system were obtained using the same method SIMPLISMA applied on the data

matrix **D**. The columns of the most dissimilar concentration profile were obtained and the concentration profiles were used as initial estimates (i.e. $\mathbf{C}_{ini}$) for further optimization process (Figure III.27).

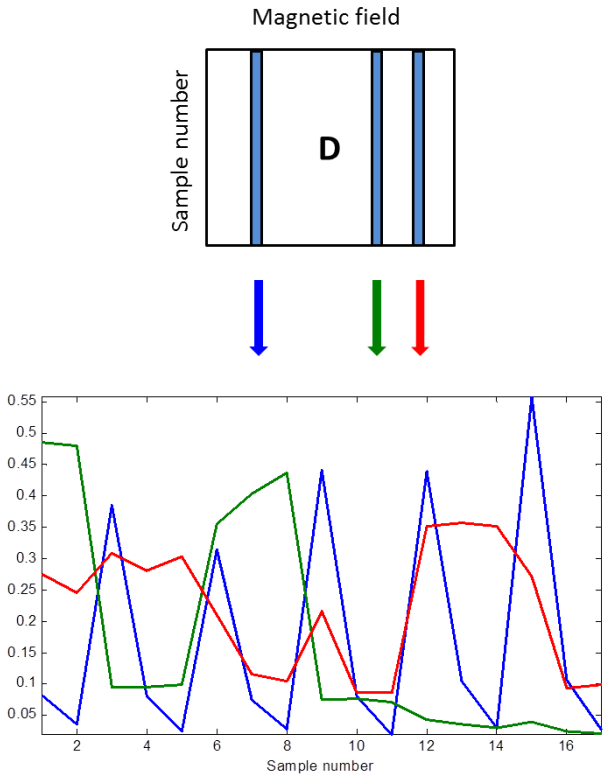


Figure III.27 : SIMPLISMA analysis on EPR data set. Representation of the concentration profile of the purest magnetic fields $\mathbf{C}_{ini}$.

III.6.2.3 Constrained ALS

Knowing the number of components and the initial concentration profiles, ALS optimization of data matrix **D** was run. As mentioned previously, ALS was applied on the integrated data, under the non-negativity constraint in both spectral and concentration directions as well as normalization in the spectral direction. Figure III.28 shows the final results obtained by MCR-ALS with the simultaneous extraction of the pure concentration profiles **C** and their corresponding spectra $\mathbf{S}^{\mathbf{T}}$ from the experimental matrix **D** only. The first concentration profile (blue) in the matrix **C** corresponds to the first spectral profile in the matrix $\mathbf{S}^{\mathbf{T}}$and so on.

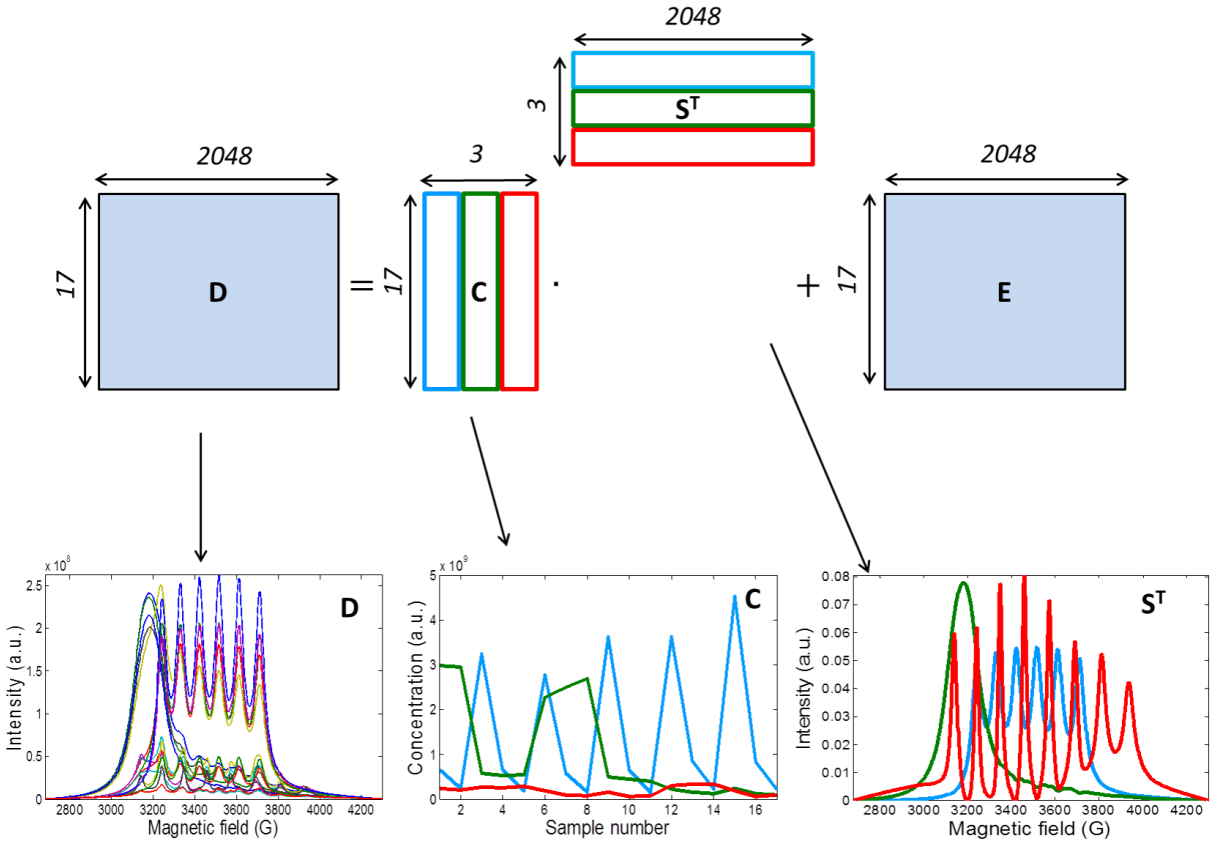


Figure III.28 : Application of MCR-ALS on the integrated data set **D** to obtain the spectral contribution $\mathbf{S^T}$ and the concentration profiles **C** of three paramagnetic ions, Cu^{2+}, VO^{2+} and Mn^{2+}. The first column (blue) in matrix **C** is the concentration profile of manganese ions and corresponds to the first row in the spectral profile $\mathbf{S^T}$ (blue) that is the spectral signature of this species. The second column in the concentration profile **C** (green) represents the concentration profile of copper ions and its corresponding spectral signature in matrix $\mathbf{S^T}$ is the second row (green). The third column in the concentration profile **C** (red) represents the concentration profile of vanadium ions and its corresponding spectral signature in matrix $\mathbf{S^T}$ is the third row (red).

III.6.2.4 Reliability of MCR-ALS solution

The reliability of the obtained solution was evaluated in the same way as for the binary mixtures. The lack of fit and the explained variance give the values of 2% and 99.96%, respectively. Moreover, no components or significant structure have been detected in the residual matrix **E** (Figure III.29).

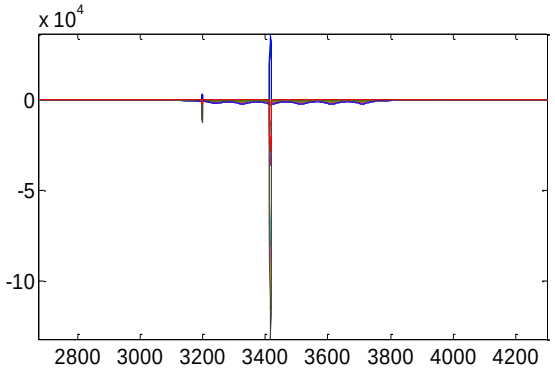


Figure III.29 : Residuals obtained at the end of MCR procedure applied on the ternary mixtures.

Analogously to the two-component system, MCR bands were also applied to the three component data to assess ambiguity of the solution. The minimum and maximum boundaries were calculated and the difference between the maximum and the minimum boundaries is equal to zero for the three components. Then, the results previously shown are unique for the three paramagnetic ions in their concentration and spectral directions.

The final argument in favor of the applicability of MCR-ALS approach for discussed dataset is the comparison of the resolved profiles with the actual known data. Figure III.30 presents the pure spectral contributions extracted by MCR-ALS method (solid red line) overlapping with the reference spectra (dotted black line) related to the three paramagnetic ions (manganese, copper and vanadium ions). Even if the dataset is significantly more complex than the 2-component one and thus, more challenging for the analytical methods, all the pure spectra are satisfactorily retrieved.

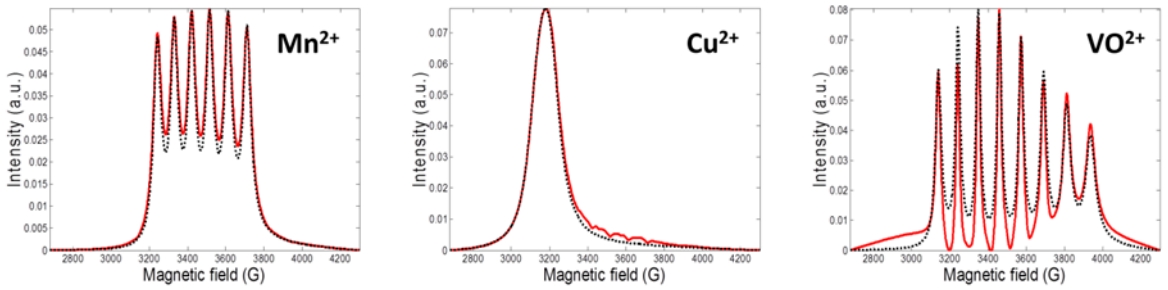


Figure III.30 : Pure spectra extracted by MCR-ALS (solid red line) and the corresponding experimentally measured spectra of pure manganese, copper and vanadium ions (dotted black line).

The pure concentration profiles extracted by MCR-ALS method (solid red line) with the experimentally measured data (dotted black line) are presented in Figure III.31. In this case some differences between the calculated and real concentration profiles are more pronounced but their global shapes are consistent with the reference values. The best agreement was achieved for concentration profile of Cu^{2+} which is the most concentrated ion in most of the mixtures.

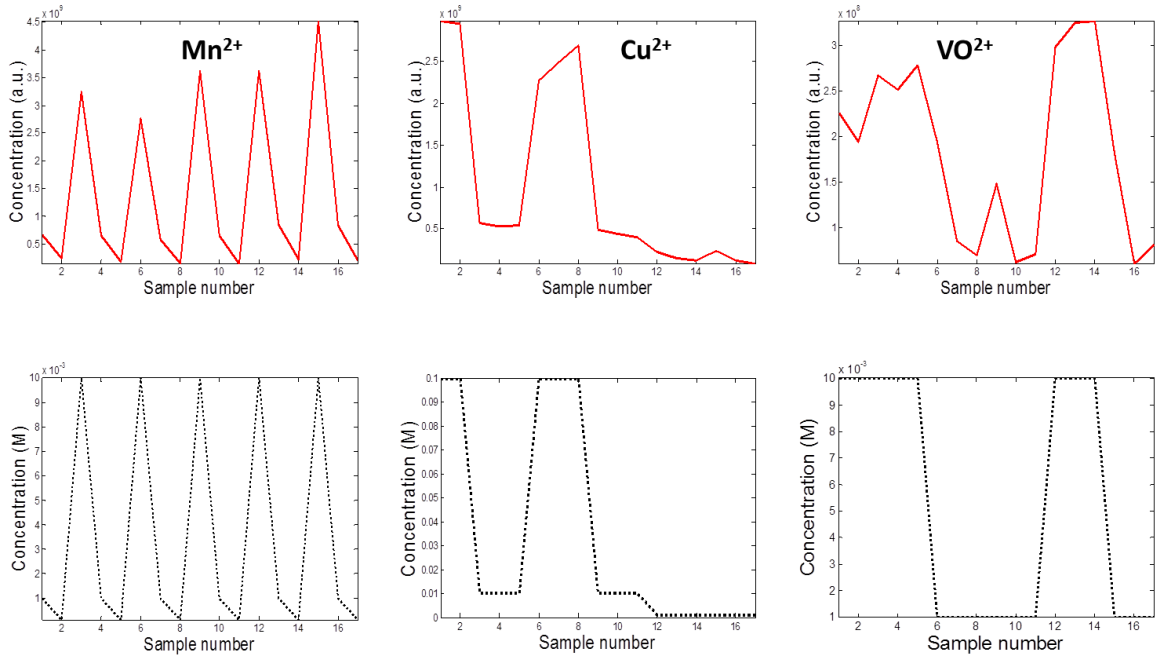


Figure III.31 : Pure concentration profiles extracted by MCR-ALS (solid red line) with the corresponding reference concentrations of manganese, copper and vanadium ions (dotted black line).

It becomes clear that using the intensity at a particular magnetic field as a way to estimate the concentration for a given spectrum can be risky because of this lack of spectral specificity. From this point of view, the multivariate MCR-ALS approach is a much better alternative to analyze EPR datasets. MCR method is different than other bilinear methods, e.g. PCA, in the way the matrix decomposition is performed. The former applies softer and natural constraints but with a more physical meaning (non-negativity). As a consequence, the solutions obtained by this method are easily interpretable and closer to the true sources of data variation.

III.7 Conclusion

For the first time, MCR-ALS has been applied on EPR spectra coming from paramagnetic ions mixtures with severely overlapped multipeak signals and has provided very satisfactory chemical results. Based on the good agreement between experimental and theoretical data, it can be concluded that MCR-ALS is suitable technique to such kind of complex data. The presented results show that MCR-ALS can be an important step forward in the data treatment of EPR spectroscopy because it enables the identification of the different species present in the mixtures i.e. the extraction of the pure concentration profile and their corresponding spectral contributions even without any prior knowledge about the investigated system. A future promising pathway in the resolution of EPR spectra can be concluded from the results above. However, troubleshooting in this study emphasized the key role of the appropriate constraints in MCR-ALS procedure as well as the requirement of a correct data pretreatment. In the case of the two-component system, the spectra were kept in their experimentally acquired first derivative form which was enough for successful resolution. Only one constraint (non-negativity on the concentration profile) was sufficient to be implemented in the procedure which lead to a low ambiguity. In the case of a three-component system, where one of the spectra was much broader than others, the overlap increased and it was necessary to integrate experimental spectra to use additional constraint (non-negativity on the spectral profile). Hence, the feasibility of the EPR study by MCR-ALS method was proved to resolve successfully problems associated with the presence of severely overlapped spectra in the unknown systems, and opened a perspective for further challenging investigations in the EPR domain such as spectral / spatial imaging or kinetics studies. The presented results represent a fundamental benchmark for further studies on sophisticated EPR data sets. In particular, they were the inspiration for the analysis of a data cube presented in the next chapter.

Chapter IV: Application of Advanced Chemometric Tool MCR-ALS to EPR Hyperspectral Imaging Data Sets

IV.1 Introduction

Hyperspectral imaging is a kind of instrumental spectroscopic measurement that contains both spatial and spectral information about the sample [164]. It represents a group of techniques which are based on the acquisition of spectral data at each pixel (or voxel) in the scanned image to form a three (or four) -dimensional array of data called cube with spatial dimensions (*x, y, or z*) and one spectral direction (*n*). Compared with the traditional spectroscopic methods, hyperspectral techniques allow the identification of chemical components in the studied sample as well as to map their spatial distribution. Therefore, this technique is advantageous to analyze heterogeneous samples with a single measurement giving a high amount of data. Simultaneous acquisition of spatial and spectral data makes hyperspectral imaging very fast and non-destructive method. For this reason nowadays, such techniques are an active area of research where a drastic growth can be noticed in the last few years due to their application in a wide range of fields such as biomedical diagnostic, material science, analytical chemistry, environmental analysis, pharmaceutical and food industry [173-180]. Hyperspectral images can be acquired by several spectroscopies such as mid and near infrared, Raman, fluorescence and electron paramagnetic resonance (EPRI).

Although hyperspectral images are very informative type of data and have the ability to solve a specific analytical problem, but these images are complex and have large size. Thus, advanced analytical tools have to be developed to extract knowledge on chemical information. During the last years, different multivariate analytical methods were proposed and developed to resolve spectroscopic images [188, 191, 214-223]. Some promising results have been obtained especially for Raman [101, 102, 216, 224-226] and IR [188, 227, 228] imaging. Interestingly, multivariate chemometric techniques have never been applied to EPRI data, whereas it is a potent tool for investigating aspects ranging from medicine and biological studies to applications in material science including polymers, ceramics, semiconductors, inorganic materials as well as their defects and impurities [46-48, 229]. This fact as well as successful resolution of a multicomponent spectroscopic paramagnetic system by MCR-ALS method (see Chapter III) became motivation to apply this methodology on more advanced data set obtained from EPR imaging of photo damaged calcium fluoride (CaF_2) sample.

CaF_2 is a representative alkaline earth fluoride that has been the subject to many experimental and theoretical studies for years. It attracted much attention due to its good

transparency from UV to IR, being used as optical components as well as dielectric medium for semiconductors [230-233]. One of the most common application of CaF_2 plates is the generation of the white-light continuum pulses spanning the ultraviolet (UV) to the near infrared (IR) that are induced, by the nonlinear interaction of ultra-intense laser pulse with the condensed transparent media [234]. The conversion of monochromatic pulse to a white broadband one allows the implementation and the development of different time-resolved spectroscopic techniques [235]. However, under high-energy photon flux and laser irradiation, all metal halides have the tendency to form defects due to the damage and degradation of material properties after short exposure time [236-238]. Such micro-damages strongly influence the optical parameters of the material and therefore, the precise characterization of the photo induced defects in the CaF_2 plates is indispensable. Since these defects very often exhibit paramagnetic properties, EPRI is an ideal experimental technique for such study.

IV.2 Objective

The aim of this chapter is to demonstrate the applicability of multivariate chemometric approach in the analysis of data coming from EPR imaging experiments. Since MCR-ALS method has been successfully applied on 2D dataset containing entangled EPR spectra (see chapter III) the same approach, with some extensions, is proposed here to resolve more complex 3D data sets from EPR imaging. A photo damaged CaF_2 plate containing some paramagnetic defects and impurities was chosen as an experimental model system for this purpose.

The three dimensional cube acquired by EPRI is formed of a spectrum per sample surface unit (pixel). Each acquired spectrum comes potentially from a mixture. Therefore, the objective of this study is to identify the number of various unknown components and to explore spatially their location as well as to extract and characterize the corresponding spectra of each pure constituent from the spectral-spatial EPR data set by MCR-ALS algorithm (see Figure IV.1).

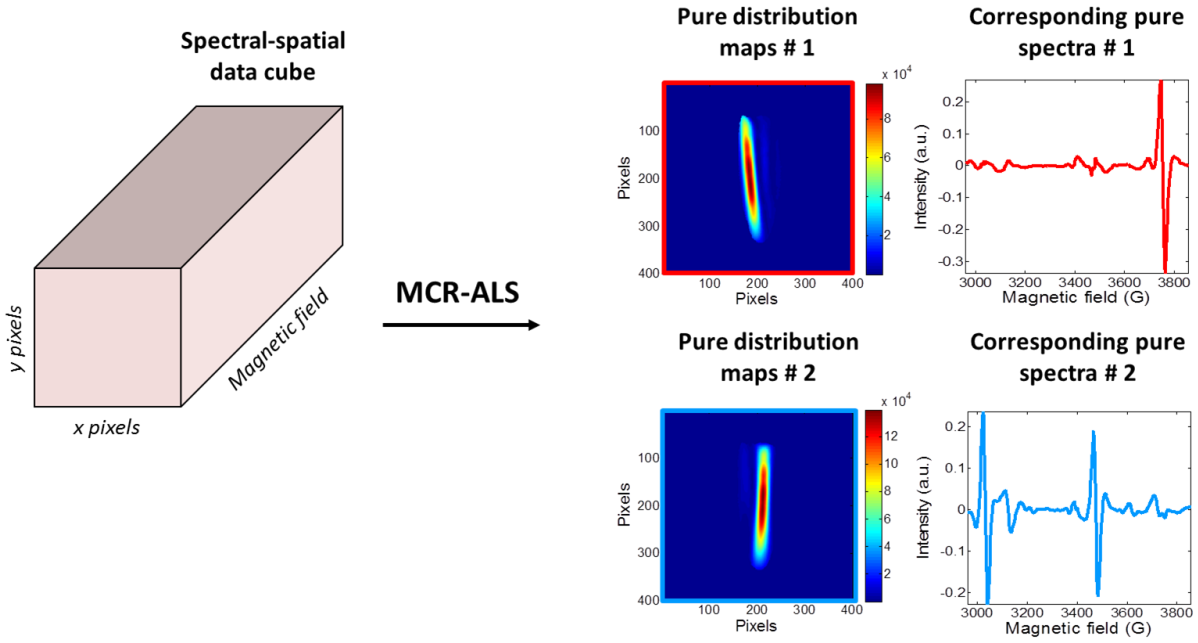


Figure IV.1 : The concept of resolving hyperspectral EPR imaging data cube presenting an example of a data cube with two pure contributions.

To apply this method (as well as preliminary PCA examination), 3D data were first unfolded to the corresponding 2D form and refolded after subsequent analysis. Moreover, the additional constraints such as local rank and selectivity, imposed on some analyzed pixels during MCR-ALS procedure, are presented and discussed. The results presented in this chapter represent the first application of MCR-ALS method on spectral-spatial EPR data set and show the benefits of applying PCA and MCR-ALS on complex 3D EPR data set. These results have been partially published in [239].

IV.3 Sample preparation

A CaF_2 plate was obtained from a femtosecond transient absorption setup where it was routinely used to convert the monochromatic 800 nm pulses into a light with a broad spectral bandwidth. The plate was 1-mm-thick, 7 mm wide and 20 mm high. The femtosecond experiments were carried out by using a 1kHz Ti:Sapphire laser system based upon a Coherent (MIRA 900D) oscillator and a BM Industries (ALPHA 1000) regenerative amplifier providing 1mJ, 100 fs pulses at a wavelength of 800 nm. The small part (a few microjoules) of the fundamental energy was used to generate a white continuum in the studied CaF_2 plate.

After several experiments when the CaF_2 plate was irradiated by the pulsed laser some bump-like features and cavities were observed on the exposed material. These bumps are related to compressive stresses due to a pressure build-up induced by fast laser heating and their subsequent relaxation. The formation of these nanocavities is attributed to the explosive expansion generated by shock waves due to laser-induced plasma after the nonlinear absorption of the laser energy by the material. The CaF_2 plate was, then, examined by EPRI spectroscopy to explore intrinsic damages.

IV.4 Image acquisition

For data acquisition, the CaF_2 plate was analyzed with a continuous wave-electron paramagnetic resonance spectrometer (CW-EPR). EPR images were acquired at room temperature of 20°C using an ELEXSYS E-580 (Bruker) EPR spectrometer operating at the X-Band. Spectra were recorded at modulation field frequency of 100 kHz and a microwave frequency of 9.80 GHz, with an amplitude modulation set at 0.1 mT and a microwave power of 12 mW corresponding to non-saturation conditions. The spatial window (field of view) was 20 mm and the weak pitch from Bruker was used as standard reference and contained a known concentration of spin/mass (1.29×10^{13} spins/g). The spin concentration is given by the double integration of the first derivative of the EPR signal. Receiver gain was set at 60 dB, time constant of 10.24 ms, conversion time of 20.84 ms. Images were reconstructed from a complete set of projections collected as a function of the magnetic field gradient. EPR spectra were recorded in the interval between 2930 and 3880 Gauss. A data cube (see Figure IV.2) with a pixel size of 50 μm was acquired giving the dimensions $400 \times 400 \times 400$ ($x \times y \times n$, where x and y are the number of pixels in the two directions of the plate surface and n the magnetic field).

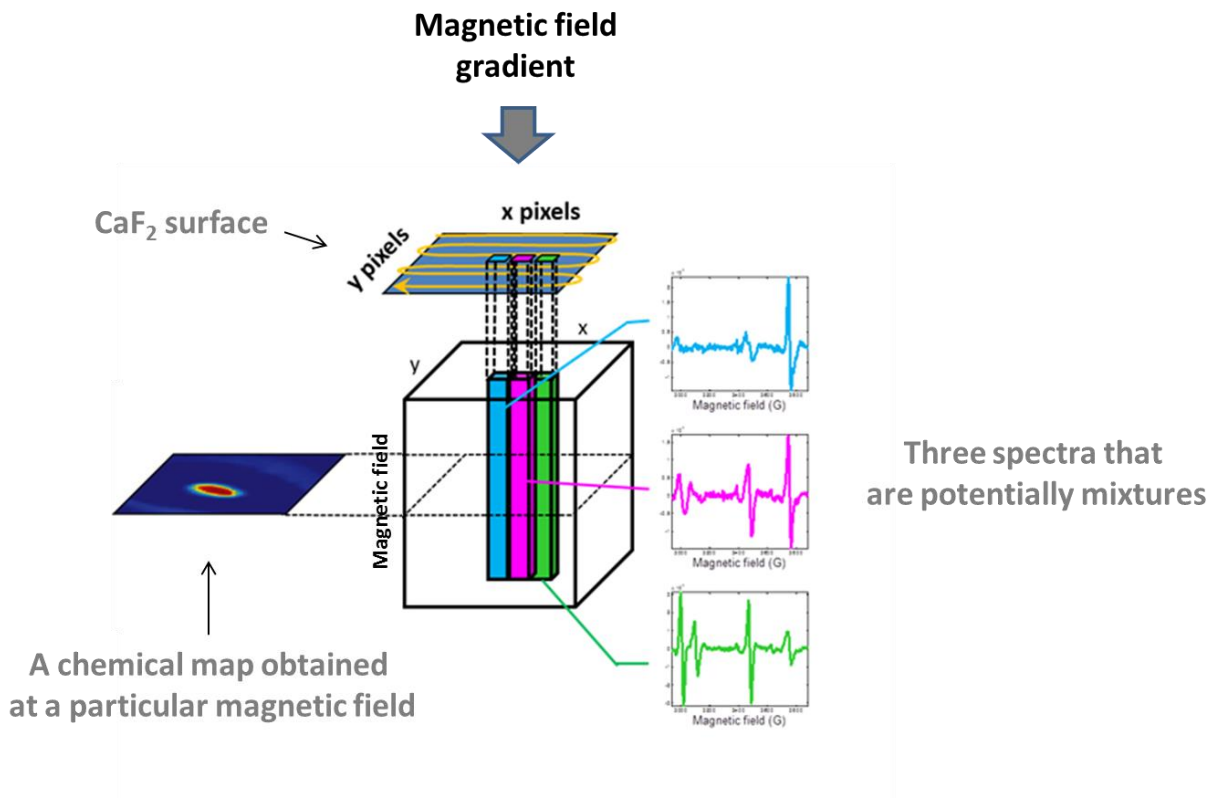


Figure IV.2 : Three-dimensional cube obtained by EPRI spectroscopy.

IV.5 Structure of data sets

Given the 3D structure of the considered EPR data cube, each pixel is a micro zone of analysis for which a characteristic spectrum is obtained. From another point of view, the cube possesses a collection of images forming many layers, where each layer represents an image recorded at a certain magnetic field. The acquired EPR spectral-spatial data set of the CaF_2 plate is in the form of a cube whose dimensions are 400 pixels × 400 pixels × 400 magnetic fields. Figure IV.3a shows an overlay of the 160,000 EPR spectra corresponding to all pixels. General inspection of this set of spectra indicates that different paramagnetic centers are present but it is very difficult at this stage to establish the number of pure contributions in the system.

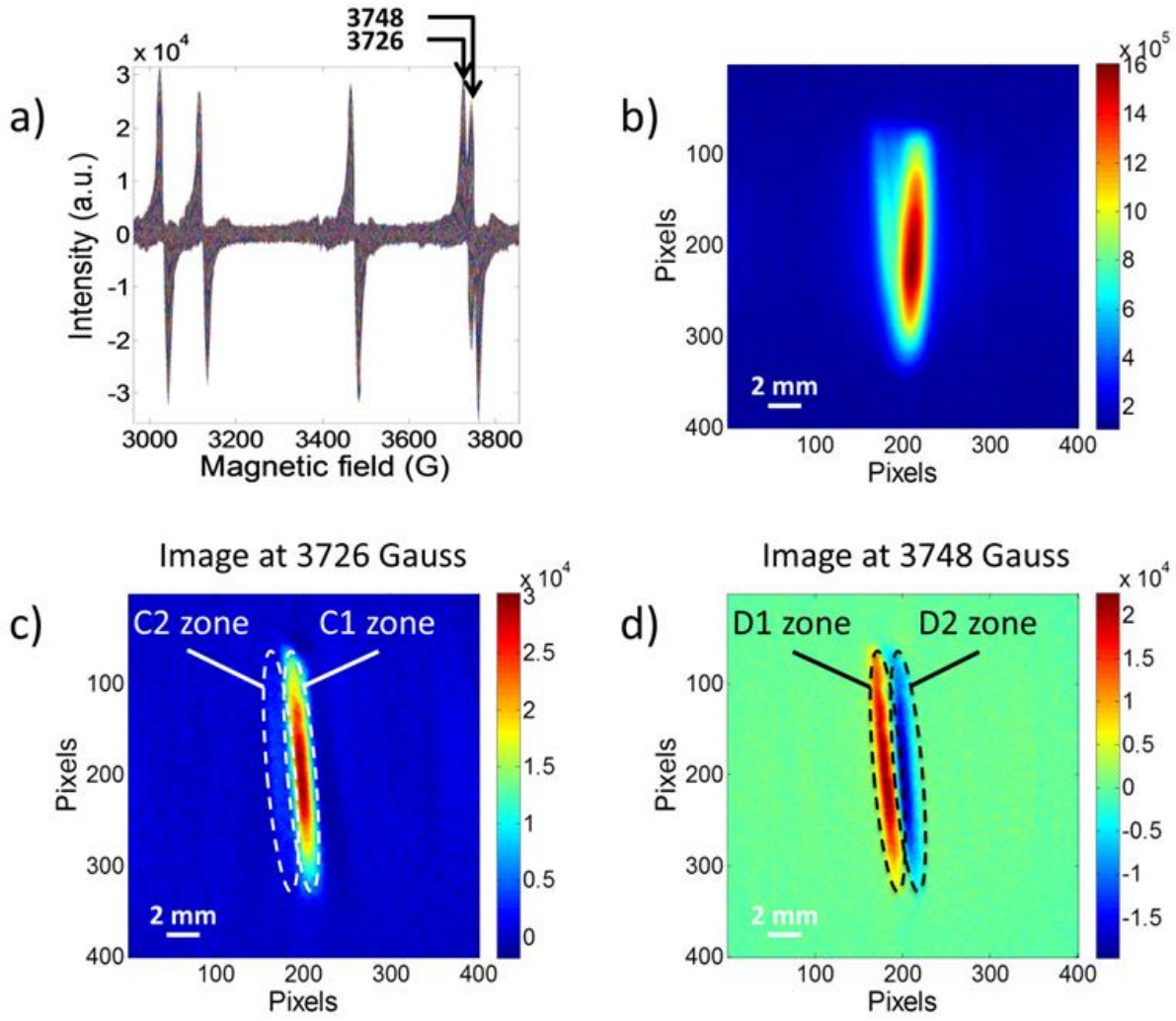


Figure IV.3 : Conventional EPR data set exploration a) overlay of raw spectra b) global integration 2D image c) single integration 2D image at 3726 Gauss d) single integration 2D image at 3748 Gauss.

In hyperspectral imaging, the global intensity map is often calculated in order to have a first insight about the signal distribution over the sample surface. In this case, EPR signal integration is done over the whole spectral domain for each pixel. Figure IV.3b indicates that the highest EPR contributions (red of the colorbar) are located in the center of the CaF_2 plate corresponding to location of the laser beam strike. However, one can notice that the signal is not homogenously distributed and different zones are present. Considering the conventional univariate way of such data analysis in EPR imaging, Figure IV.3c and Figure IV.3d show images generated as integration of the single selected spectral lines at the magnetic field of 3726 Gauss and 3748 Gauss, respectively.

At first sight, Figure IV.3c seems to be composed of intense contribution in C1 zone and a lower one in C2 zone. On the other hand, Figure IV.3d presents particular features with positive and negative contributions in D1 and D2 zones, respectively. This last observation can indicate that spectral interferences are present at 3748 Gauss. Considering the proximity of the two signals at 3726 Gauss and 3748 Gauss and, therefore, the possible lack of selective EPR channels, doubts could be cast upon considering the plot in Figure IV.3c as a pure compound image. This simple two magnetic fields example demonstrates the difficulty to generate with a univariate approach real chemical maps from highly overlapping EPR imaging data set without prior knowledge about the system.

In consequence, the chemometric methods (e.g. PCA and MCR-ALS) are needed to resolve the system. However, the original 3D structure of a data set is not convenient to such analysis techniques. Therefore, 3D cube has to be first adapted for the subsequent examination.

For this purpose, the cube can be unfolded to create a two-dimensional spectral data matrix **D** ($m\ spectra \times n\ magnetic\ fields$) where each spectrum is acquired at a certain ($x \times y$) pixel set in order to make use of second order procedures i.e. PCA and MCR-ALS. As demonstrated in Figure IV.4, the set of spectra are taken consecutively from the hyperspectral image pixel by pixel and are arranged sequentially. The *x* first spectra corresponding to the *x* first pixel will be the *x* first lines in the matrix **D**.

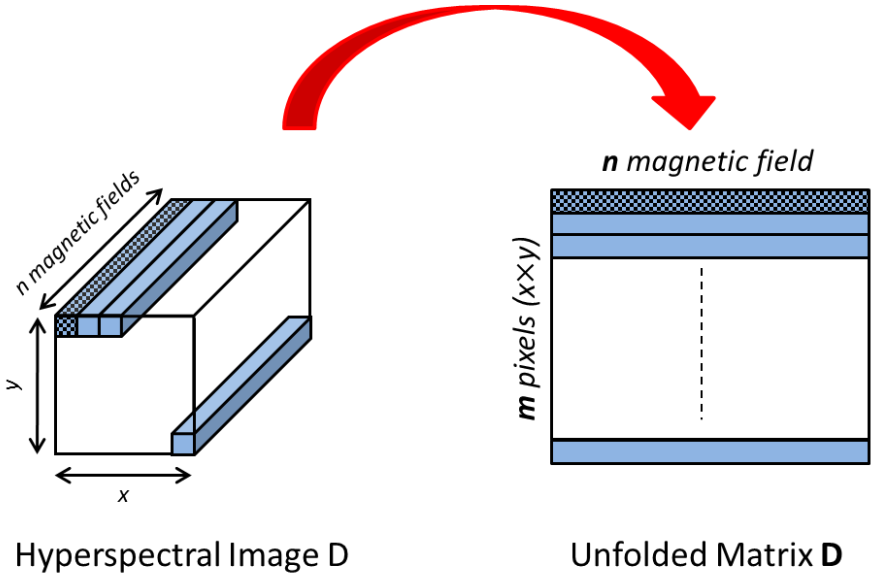


Figure IV.4 : Unfolding the 3D cube.

The unfolded matrix contains *m* rows and *n* columns. By conserving the pixel location information of each spectrum, the obtained results after the resolution process can be mapped back to a spatial location creating an image map with higher dimensional spatial ordering. The refolding procedure demonstrates one of the primary advantages of the spatial information in the hyperspectral images over other the classical spectroscopic techniques.

IV.6 Results

The EPRI data cube representing the surface of studied CaF_2 plate was transformed to 2D unfolded data matrix ($160{,}000\ spectrum \times 400\ magnetic\ fields$), and the measurement variation in the image data set is assumed to follow a bilinear model that is the underlying chemical model of many spectroscopic measurement. It means that mixed signal measured at every pixel is the concentrated weighted sum of the pure signals from the chemical constituents present in the system. Preliminary spectral and spatial explorations were carried by PCA to get a general view of the number of pure components, their spatial location as well as spectral signature. Finally, MCR-ALS algorithm was applied for the final resolution.

IV.6.1 PCA

PCA was first applied on the unfolded matrix **D** ($160{,}000 \times 400$). Four Principal Components were retained with the values 53.72%, 17.53%, 8,92% and 3,62%, respectively, from the total variance. The fifth Principal Component represents 1.85% but exhibits a slight level of noise in its loading. Therefore, the presence of four pure species in data set is assumed with some doubts to be five. Nevertheless, to remove the doubts SVD algorithm was also employed and the plot of the obtained eigenvalues is shown in Figure IV.5. It presents the plot of the first ten eigenvalues obtained from the experimental data matrix **D**. The SVD analysis indicates the presence of four distinct contributions on the hyperspectral CaF_2 image. Indeed, four significant eigenvalues are clearly distinguished from the rest due to the sudden

change in the slopes. In fact, all eigenvalues below the fourth one express variance associated with the experimental noise.

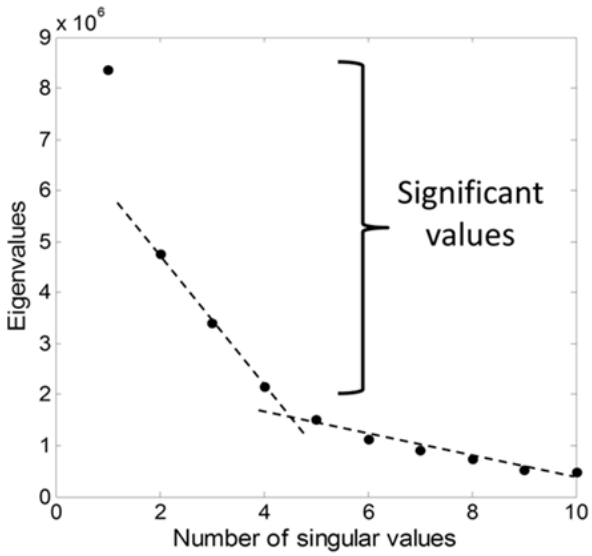


Figure IV.5 : Eigenvalues obtained by Singular Value Decomposition of the experimental data matrix **D**.

Thus finally, the number of pure components present in the image is assumed to be four. After the bilinear decomposition of the matrix **D** into scores **T** and loadings $\mathbf{L}^{\mathbf{T}}$, the columns in the scores matrix can be folded back to form score maps, whereas the rows in the loadings matrix present the PCs. Figure IV.6 shows the four different scores (in refolded forms) and their corresponding loadings on PC1, PC2, PC3 and PC4, respectively.

Since Principal Components are calculated to keep null correlation among them (see Chapter II), the obtained scores and loadings cannot be often associated with the real chemical compounds. This is particularly visible in the case of PC2 and PC3. The loadings (allegedly representing the spectra) share some spectral lines. Moreover, the corresponding scores maps reveal some important similarities. Especially negative signal zone (blue colored) exhibits the very similar shape and location. Thus, it can be concluded that two species are still strongly mixed and shared between PC2 and PC3 components. Hence, at least two species cannot be satisfactory resolved by PCA likely due to a large overlap among these components in the spectral direction. Therefore, more advanced methods have to be used to disentangle the studied data set.

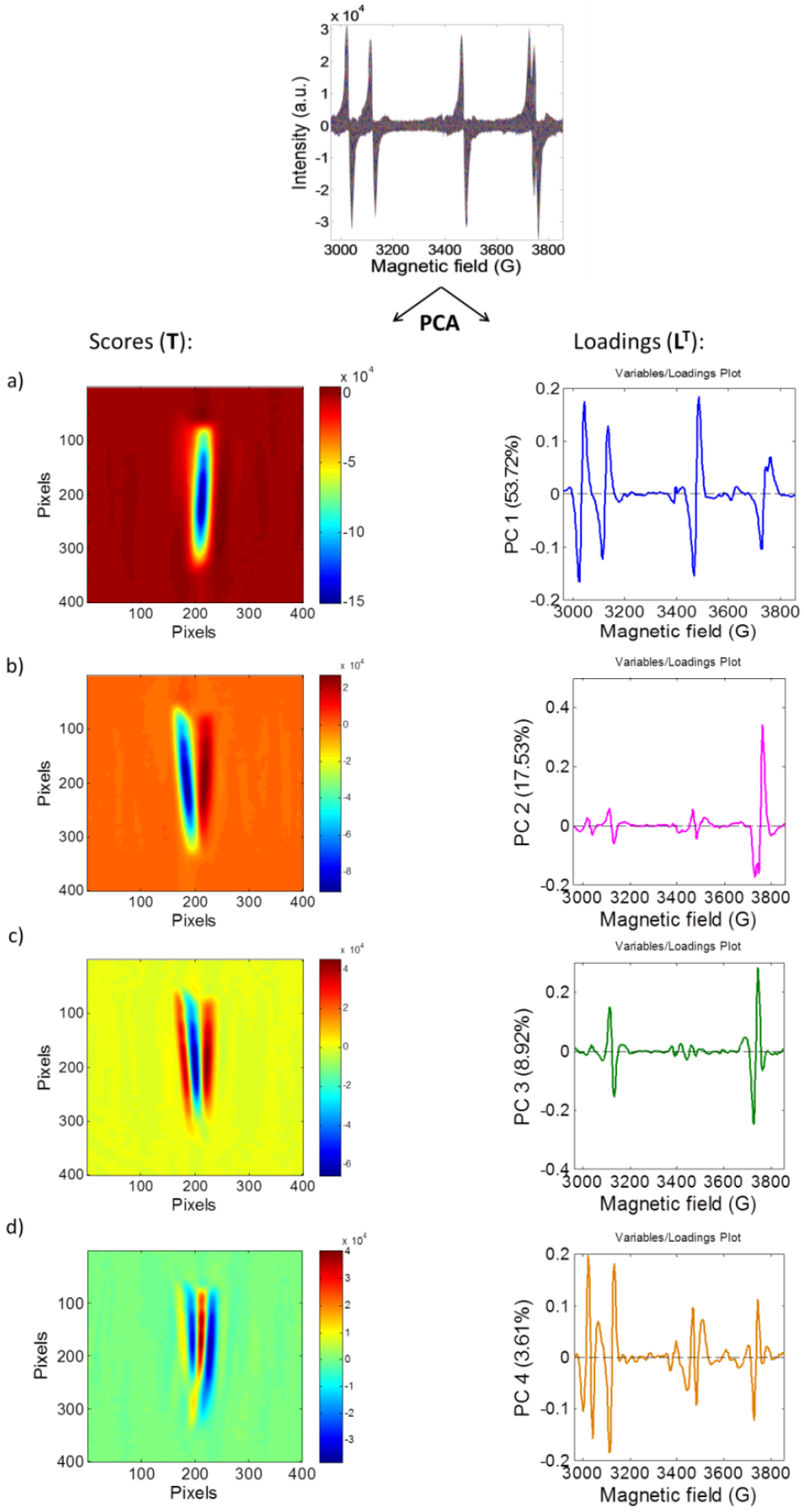


Figure IV.6 : Bilinear decomposition of the matrix **D** to give scores **T** and their corresponding loadings plots $\mathbf{L}^{\mathbf{T}}$ a) on PC1, b) on PC2, c) on PC3 and d) on PC4.

IV.6.2 Local rank analysis

As presented in the previous paragraph, PCA analysis on the global data set does not provide satisfactory resolution of the pure components. Alternative or supplementary approach in such case can be realized by the analysis of smaller parts of a global data set i.e. local analysis. In the simplest but not useful form such examination could be performed on the individual pixels. For instance, Figure IV.7 shows the spectra corresponding to four selected pixels. It is worthy to note that even in the high intensity (red color) region (pixel 1) the estimation of the number of spectral contribution is not obvious. It is even more difficult or impossible in regions with lower global intensity (pixels 2 and 3) especially where the signal is about noise level (pixel 4).

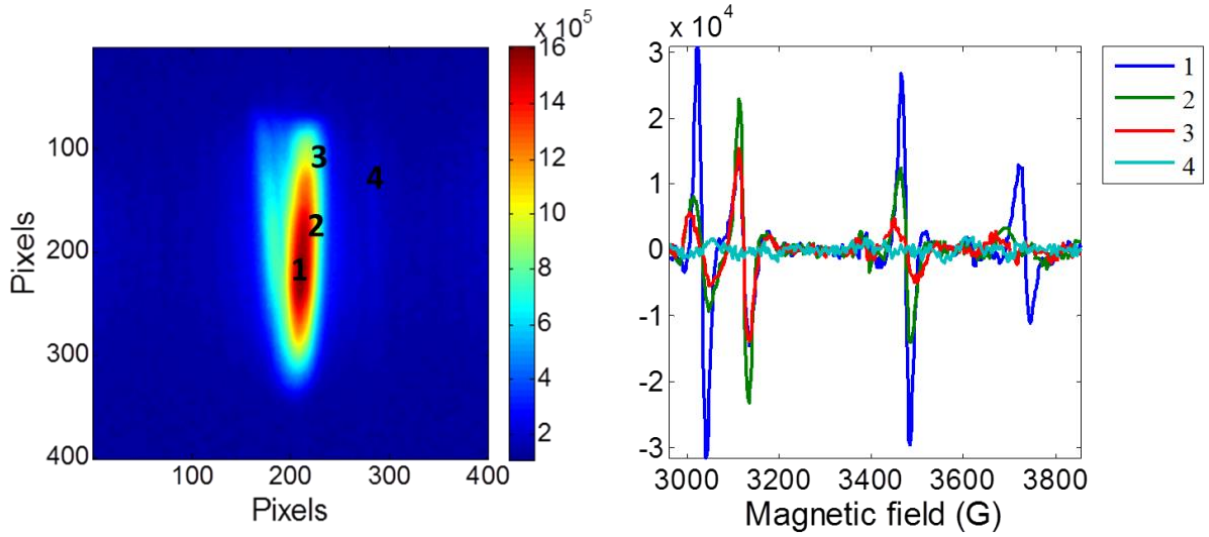


Figure IV.7 : Exploration of the image.

On the other hand, PCA on the whole image provides information on the global number of components present in the dataset (see previous part) but this information comes from the total image analysis and does not give specific information on each pixel, whereas not all components are present in all the pixels. The good solution can be PCA analyses performed on the small areas of the image (called windows) selected in a way that the whole image is covered.

The method which is based on such idea is Fixed Size Image Window-Evolving Factor Analysis (FSIW-EFA) which is in turn the modification of the local rank algorithm called Fixed Size Moving Windows-EFA (FSMW-EFA) [146]. It performs local PCA (more

precisely it is SVD in practice) in the whole image by moving small fixed windows around each individual pixel area in the two spatial dimensions of the image until the entire image is scanned. The PCA analyses performed on each of the pixel areas provide information about the local complexity of the image. It is important to know the number of components (in this case is 4 from PCA on the global data set). The number of pixels in the window should be as small as possible to preserve the spatial resolution of the image but should be equal or exceeds the total number of compounds in the image to allow for local rank values representing all possible situations of compound overlap in the image. Thus, in the studied case, the window size is fixed (2×2) (Figure IV.8a). Similarly to global data set analysis, the selected windows i.e. sub-cubes are first unfolded to 2D form and then PCA (SVD) analyses are performed on the unfolded sub-matrices containing the spectra in the related windows.

Singular values obtained in these local analyses can be displayed as plots where the spatial structure of the image is preserved (2-dimensional image). Each singular value plot is displayed as 400 pixels versus 400 pixels and third flattened color coded dimension representing magnitude of a given singular value at the appropriate pixel location (Figure IV.8b). As it can be seen, warm colors (large values) indicate zones of the image where the related singular value presents significant contributions to the signal related to the presence of chemical constituent whereas cold colors (small values) describe experimental noise. The high intensity zones in the singular value plot number one indicate the presence of one component, however, in the singular value plot number two, high intensity regions indicate the overlap of two different components and so on. It can be noticed from Figure IV.8b that the first two singular values exhibit relatively well localized distribution around the center of surface. On the contrary, the third and fourth singular value plots show nearly homogenous and low intensity scattering along the entire image surface. Therefore, it seems to indicate that the most complex zones in the image (oval longitudinal shape around the center) can show dominating two overlapping components and very minor presence of a third and fourth overlap. The plot of singular value number two can help us deduce information about the system since it represents maximum overlap in our image.

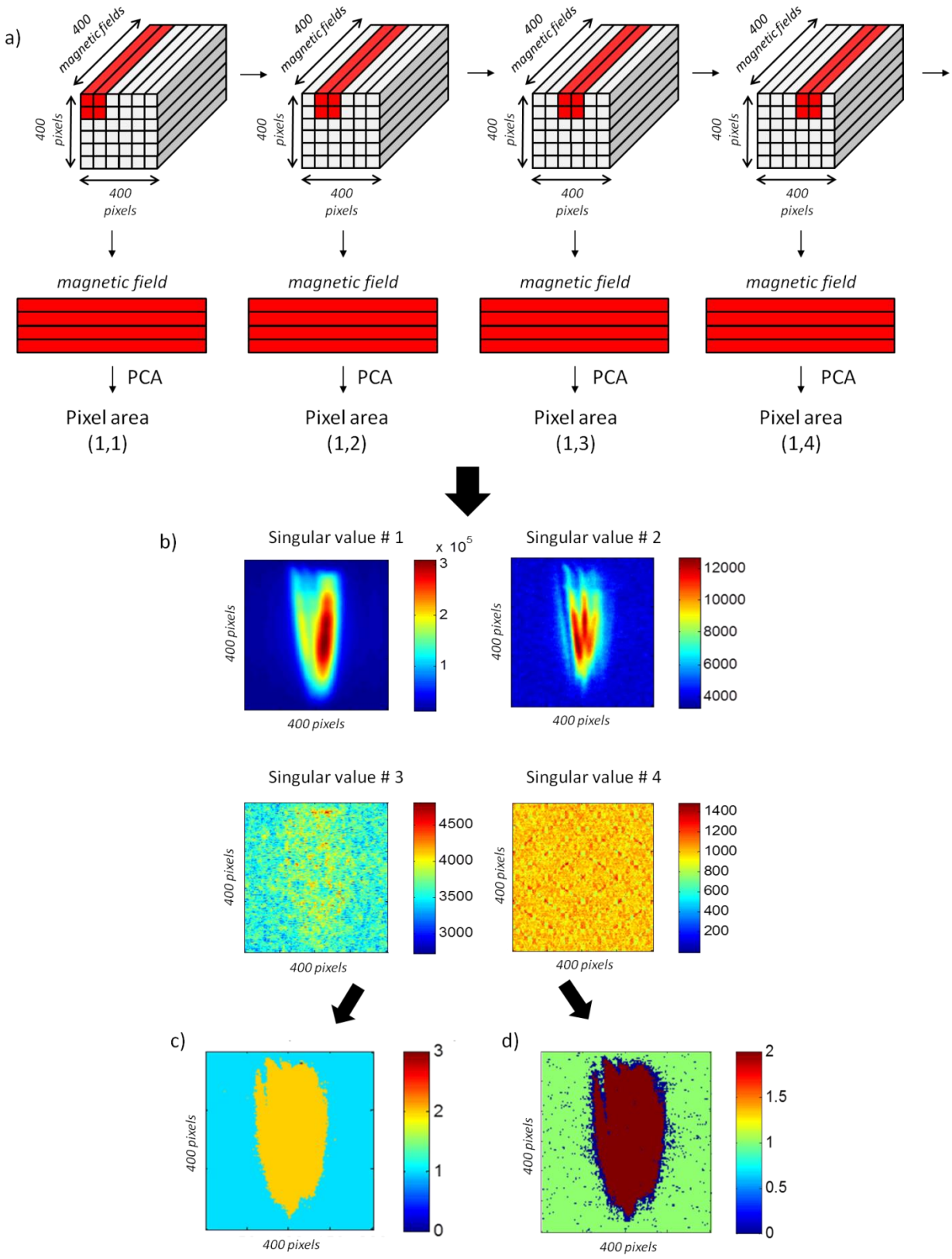


Figure IV.8 : FSIW-EFA a) Construction of pixel windows. b) Singular value plots of local PCA analyses. c) Complete local rank map. d) Partial local rank map.

Knowing that the number of pure components is four (obtained previously by PCA) and the presence of three red oval shape zones in the singular value plot number 2 (two different components overlap in these zones), we can infer the position of the four components illustrated in Figure IV.9.

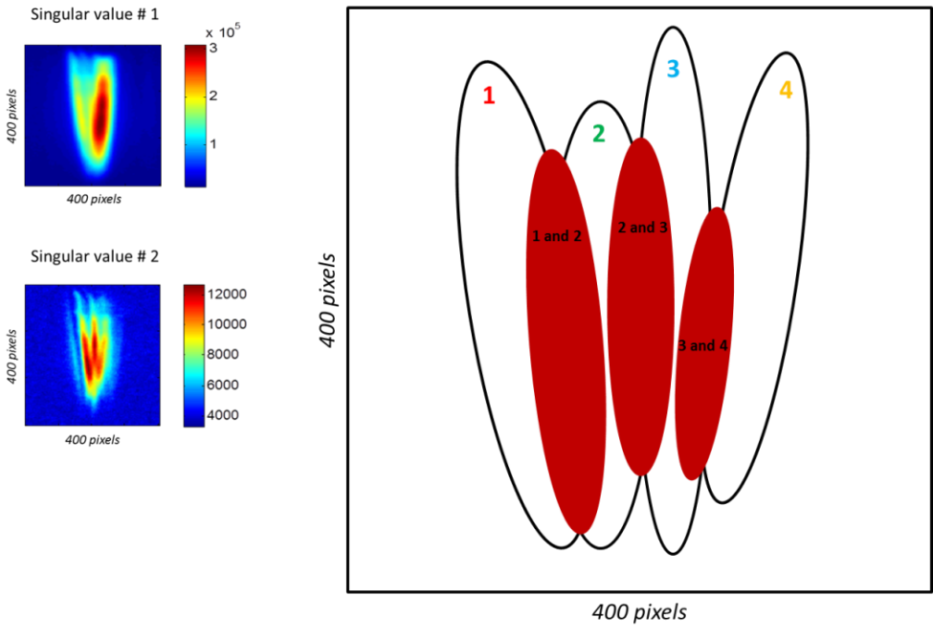


Figure IV.9: Illustration of the position of the four components in the sample deduced from the results obtained by FSIW-EFA.

Such knowledge can be translated to so called local rank maps possibly used as additional constraints in MCR-ALS analysis (see paragraph IV.6.3.2). Local rank maps are obtained by assigning the rank value (i.e. the number of significant singular values) to every pixel. Thus, it is necessary to decide which singular values are above noise threshold in a given pixel. It can be assessed from the singular value plots by displaying the evolution of the singular values versus the unfolded pixels in an image. Such plot has as many lines as the singular values calculated in each PCA window analysis. A threshold value or band is set to mark a limit between the significant singular values from the noise-related singular values. Depending on the threshold selection, complete local map (Figure IV.8c) as well as partial local map (Figure IV.8d) can be distinguished.

Complete local rank map displays how many singular values are above the threshold value for each pixel in the image i.e. they show the number of components that overlap in each pixel. This map gives an indication on which zones of the image have a lower or higher composition complexity. As in the singular values plot, the local rank map recovers the 2D

structure of the original surface image. The map is displayed as 400 pixels (x- and y- axes, respectively) and a color bar that shows the number of singular values exceeding the threshold value in each pixel. In our CaF_2 sample, zones with higher local rank values (yellow) appear in the middle of the sample where the main components are present and overlap (Figure IV.8c).

Partial local rank map is constructed using a threshold band to separate significant from non-significant singular values. The local rank determined using the upper and lower boundaries of the boundaries of the threshold band can change for some pixels but remain invariant for others. These latter pixels have consistent rank information and are the only ones displayed in the partial local rank map of the image. The map representation is analogue to the complete local rank map i.e. 400 pixels (x- and y- axes, respectively) and a color bar that shows the number of singular values exceeding the threshold value in each pixel. In our sample, dark red color represents the high local rank values, whereas the blue pixels are these where the rank is not unequivocally determined.

IV.6.3 Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS)

The ultimate goal of the analysis of hyperspectral images is to provide reliable distribution maps and to characterize each individual pure spectral (signature) component present in the image. For this purpose, MCR-ALS is the proposed approach.

As previously mentioned, the spectral data cube is first unfolded into the matrix **D** (160000 × 400) where each line represents a spectrum, which is potentially a mixture of various constituents. The first step in MCR-ALS is the application of SVD on the matrix **D** to determine the number of pure components in the data set. As presented in Figure IV.4, the total number of four pure compounds was found.

IV.6.3.1 Initial estimation

After determining the number of pure components, the next step in MCR-ALS process was the use of SIMPLISMA algorithm to generate the initial estimates (see Chapter II). In the

unfolded data set, initial estimates may be found either as the purest magnetic fields (i.e. the purest columns) or the purest pixels (i.e. purest rows). The rows related to the purest pixels give an initial approximation of the pure spectra of the constituent in an image, whereas the column related to the purest spectral channel allow the building of the approximated initial distribution maps. In order to find the purest pixels, the purity is calculated per each pixel spectrum as according to eq. II.11 in chapter II, and the purest spectra $\mathbf{S}^{\mathbf{T}}_{\mathbf{ini}}$ are obtained. Similarly, when magnetic field direction (spectral channel) is chosen, the purest distribution maps $\mathbf{C}_{\mathbf{ini}}$ are obtained. The purest spectra and distribution maps obtained by SIMPLISMA cannot be simply associated to the pure components since the selective pixels and the spectral channels in the measurements are not ensured. However, the results obtained by SIMPLISMA give initial guesses to start the ALS iteration.

In this work, selecting the purest pixels giving an initial guess $\mathbf{S}^{\mathbf{T}}_{\mathbf{ini}}$ matrix (purest spectra) was chosen. Knowing that SIMPLISMA is designed to work with positive data, initial estimate of EPR spectral profiles were obtained by applying the method on the matrix containing the absolute values of the derivative spectra in **D**. In this way, the rows of the most dissimilar spectra could be obtained and then original EPR spectra related to these rows were used as the initial estimate (see Figure IV.10).

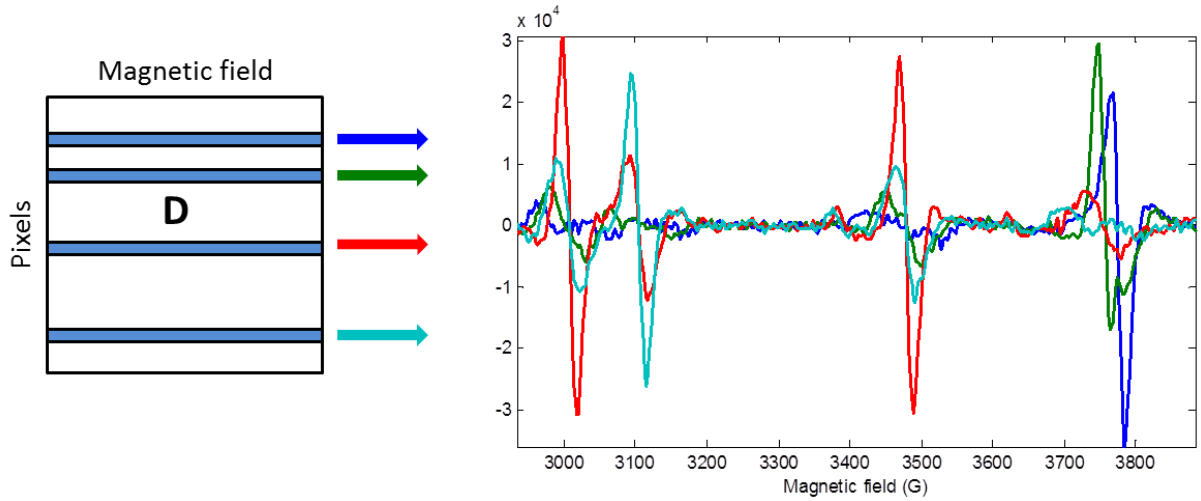


Figure IV.10 : SIMPLISMA analysis on EPR data set. Representation of the spectra of the purest pixels $\mathbf{S}^{\mathbf{T}}_{\mathbf{ini}}$.

IV.6.3.2 Alternating Least Squares optimization with different constraints

As detailed in Chapter II, different constraints on the spectral and concentration profiles can be imposed during ALS optimization to decrease the ambiguity of the solution. However, only few constraints can be applied to the images because of the lack of specific pattern in the concentration direction of this kind of data set. In addition to the universal constraint such as non-negativity, the local rank constraint can be applied for image analysis. This constraint includes well-condition of selectivity and set the absence of one or more components in certain pixels in the concentration profiles **C**.

In the first attempt of ALS, non-negativity on the concentration profile is the only constraint applied for the considered EPR data set in addition to the normalization (spectral direction). Indeed, the contribution of any constituent to the signal (concentration) is always positive, whereas the raw EPR spectra in their first derivative form can contain negative and positive values. The alternating least squares procedure involves optimization of the equation $\boldsymbol{D = CS^T}$ by the following operation $\boldsymbol{C = DS(S^TS)^{-1}}$ and $\boldsymbol{S^T = (C^TC)^{-1}C^TD}$, respectively (Figure IV.11). The iteration process ends when the reproduction of the original image from the product of the resolved concentration profiles and spectra has enough quality and there is no significant variation among the results of the consecutive cycles.

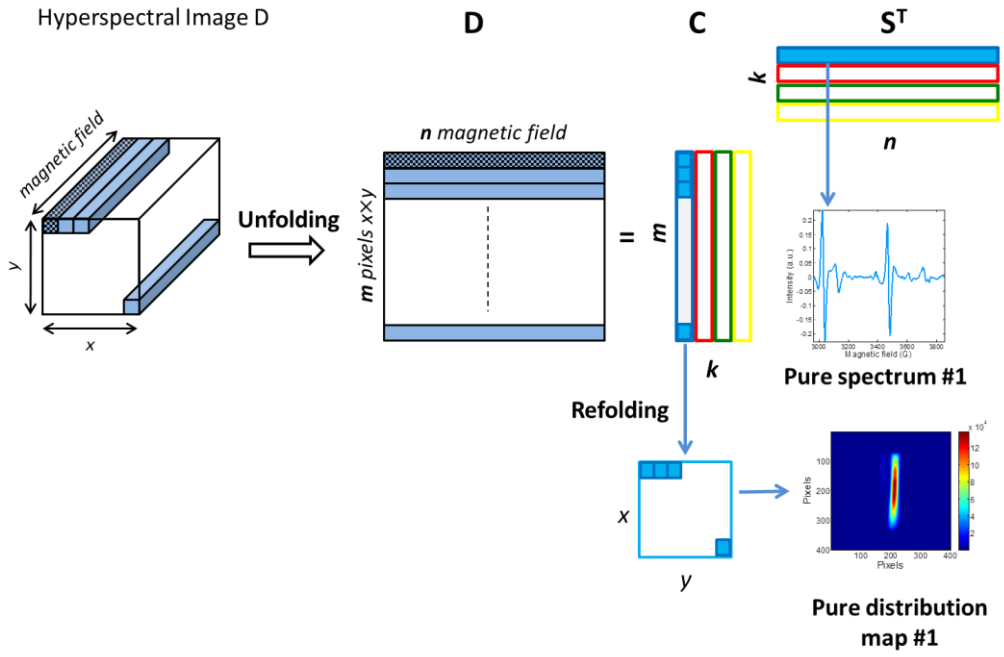


Figure IV.11 : Multivariate curve resolution applied on EPR images.

Figure IV.12 shows the final results obtained after the application of MCR-ALS on the sole raw EPR spectra i.e. the pure spectral profiles $\mathbf{S}^T$ and the corresponding concentration profiles **C** that are folded back to give the distribution map of each pure contributions in the CaF_2 plate. Four different pure spectral signatures and the corresponding distribution maps are represented. Given the signal to noise ratio in the raw EPR data, a lack of fit of 25 % and explained variance of 94 % is a good figure of merit for the proposed extraction. The absence of signals or structures in the error matrix **E** is another good sign of the quality of the results. Moreover, MCR bands were calculated for all the concentration and spectral profiles. The difference between the minimum and maximum boundaries is equal to zero hence the solution does not exhibit any rotational ambiguity.

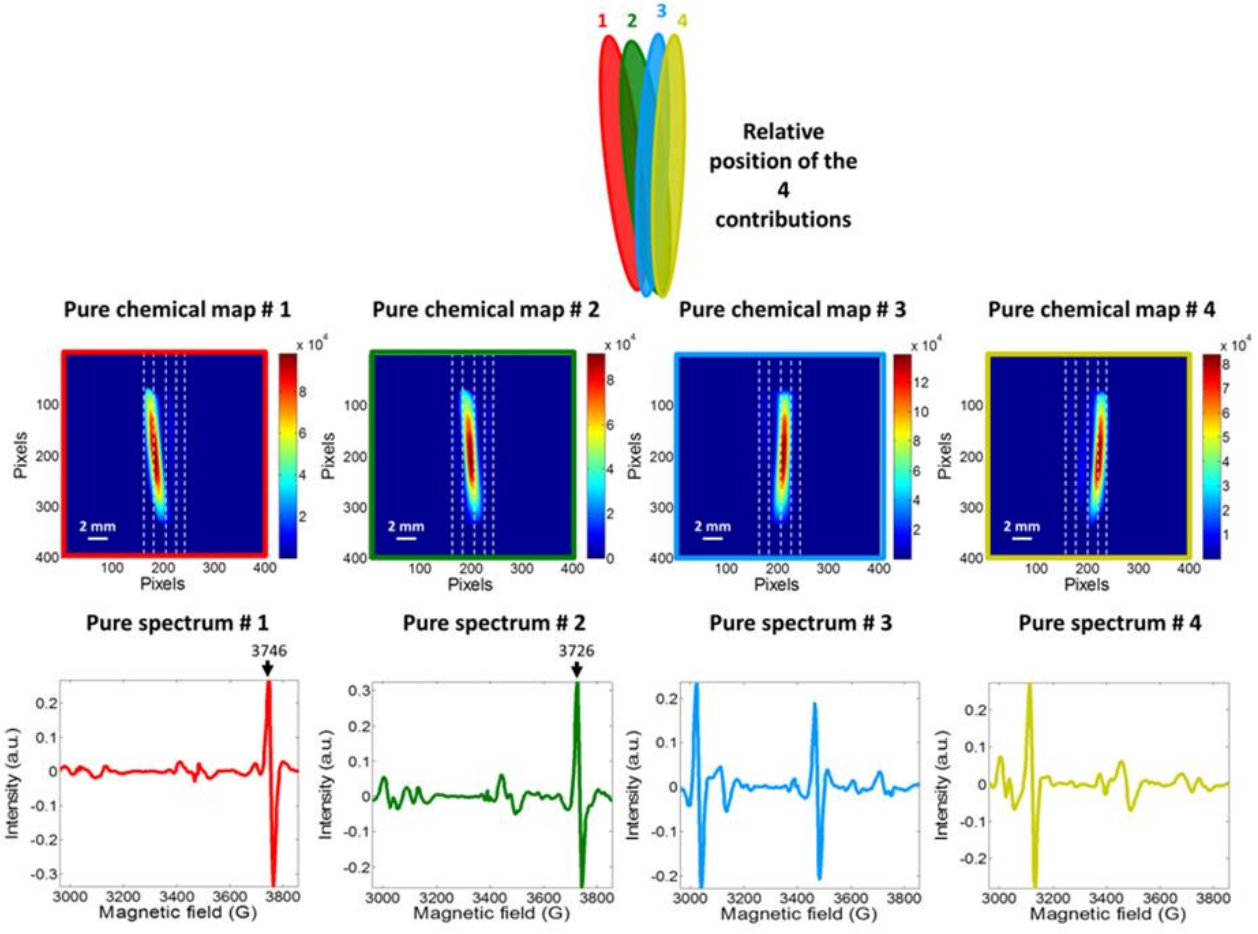


Figure IV.12 : MCR-ALS extraction: the four distribution maps **C** (top) and their corresponding pure spectra $\mathbf{S}^T$ (bottom).

In general, when the solution has no rotational ambiguity, there is no need to apply additional constraints. Nevertheless, aiming to ameliorate our results an additional constraint

i.e local rank in addition to the non-negativity on the concentration profile was applied during the second attempt of ALS optimization.

The idea of the local rank constraint is based on the estimation of the number of missing components in every pixel [191]. For any pixel, the number of missing components is equal to the difference between the image total rank (obtained by SVD on the global data set) and the local rank pixel. Local rank information is obtained from FSIW-EFA as described in section IV.6.2. This information is typically presented as a partial local rank map (see Figure IV.13a) where some pixels with not well defined rank are excluded. Such pixels are not a subject of the constraint. To identify which component(s) is absent in a given pixel, the reference spectral information is still needed. It can be extracted from the approximations of the pure spectra using purest variable selection (SIMPLISMA) or the results of another MCR-ALS analysis e.g. where only non-negativity has been applied (see Figure IV.13b).

The identification of the absent component(s) starts by calculating the correlation coefficient between the raw pixel spectrum and each of the reference spectra of the different constituents. The component(s) with the lowest correlation(s) is assumed to be the missing one(s) [191]. The maps of the individual missing components can be thus drawn (Figure IV.13c). In order to use this information as a constraint in ALS procedure, it can be encoded as a matrix called Csel sized same as **C** (see Figure IV.13d). The latter indicates which components are missing (value zero in red) in every pixel. After building such a matrix, the local rank and non-negativity constraints (on concentration profile) as well as normalization (on spectral profile) were used in the ALS algorithm.

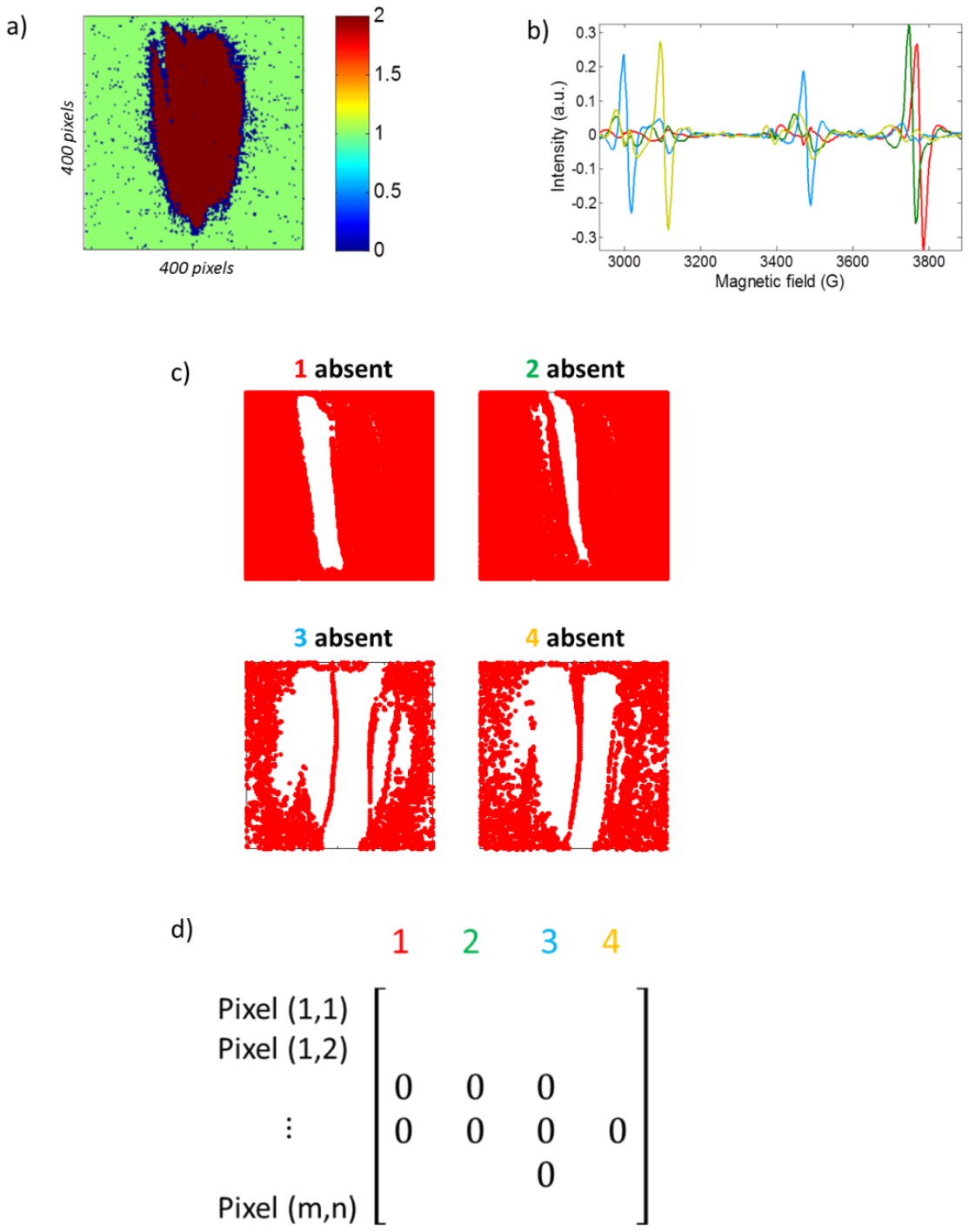


Figure IV.13 : Procedure followed to incorporate local rank constraints in resolution a) Partial local rank map, b) Reference spectral information (spectral profile obtained from MCR-ALS), c) Pixels masks for absence of components in the image and d) transfer of information from c) to a matrix (csel matrix).

Similarly to the first ALS optimization the pure concentration and spectral profiles of the four different components were obtained (Figure IV.14). Surprisingly, it can be noticed that the distribution maps of the four components seem to be extracted with a bit worse quality than in the first attempt. It is especially visible on the surroundings of the first and

second component which are slightly deformed. On the other hand, comparing the spectra obtained with and without the application of local rank constraint the results are similar and no important improvements is observed. Finally, slightly lower explained variance 92% and a higher lack of fit 28% are noted what indicates that the algorithm is not able to converge possibly due to irrelevant constraints in some pixels.

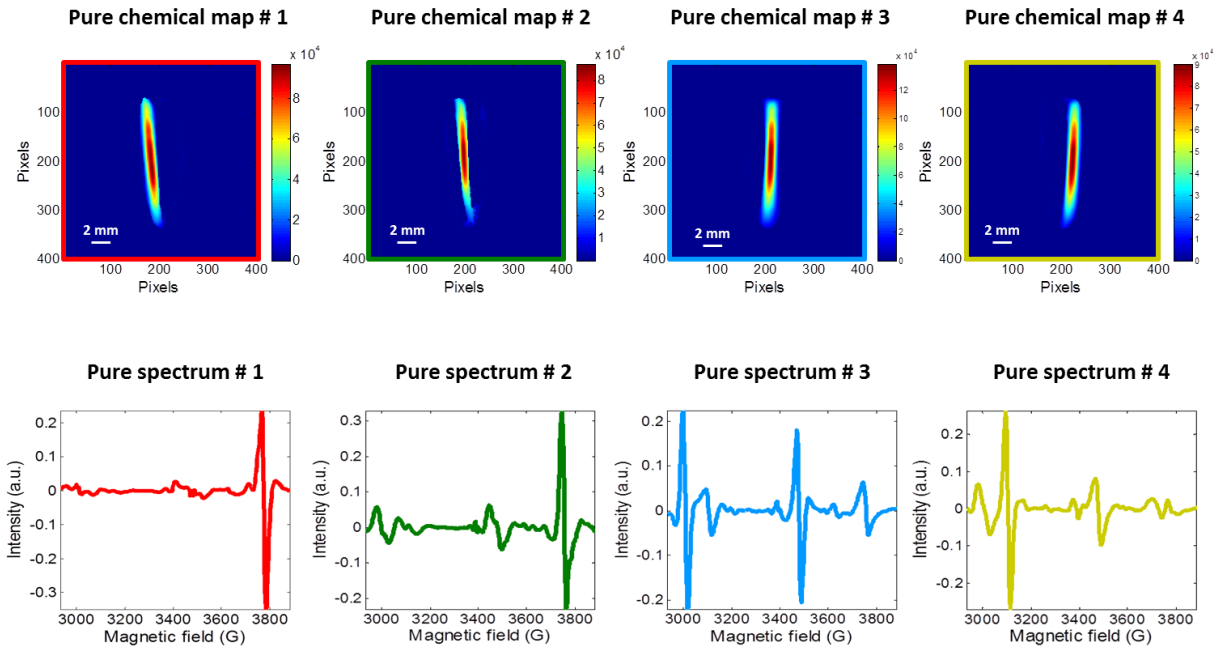


Figure IV.14 : Distribution maps (top) and spectral signatures (bottom) of four constituents obtained by MCR-ALS with the implementation of local rank and non-negativity constraints (concentration direction) as well as normalization (spectral direction) on the matrix **D**.

It can be, thus, concluded that the results obtained without local rank constraint are better for the studied data set. Although the usefulness and advantages of the local rank constraint were thoroughly tested and confirmed in the literature [102, 146, 191], it turned out not suitable to be used in the studied case since some pixels are harshly deformed. Therefore, the results obtained without the implementation of the local rank will be the subject of further interpretation (Figure IV.12).

IV.6.3.3 Identification of paramagnetic centers

Precise description and spectral assignment of the defects created in the ionic crystals is very challenging and typically demands the combining of EPR with different experiments

at various temperatures what is beyond the scope of the present work [240]. Nevertheless, considering the extracted pure spectra, a tentative attribution of the defects generated during the laser beam exposure can be proposed. The fluorite structure possesses the smaller cations Ca^{2+} that form a face centered cubic lattice with the larger anions F^- situated at the corners of the eight cubes of the elementary cell [240, 241] (see Figure IV.15a). Therefore, each anion is surrounded by four cations in a tetrahedral manner. The F^- sub lattice is considered a simple cubic structure. Consequently, only the center of every second cube is occupied by a cation and the remaining empty cubes (interstitial) are important for defect formation and the accommodation of impurities like rare earth ions. From this description, we can note that non-damaged CaF_2 surface is transparent to EPR spectroscopy due to the lack of any paramagnetic center.

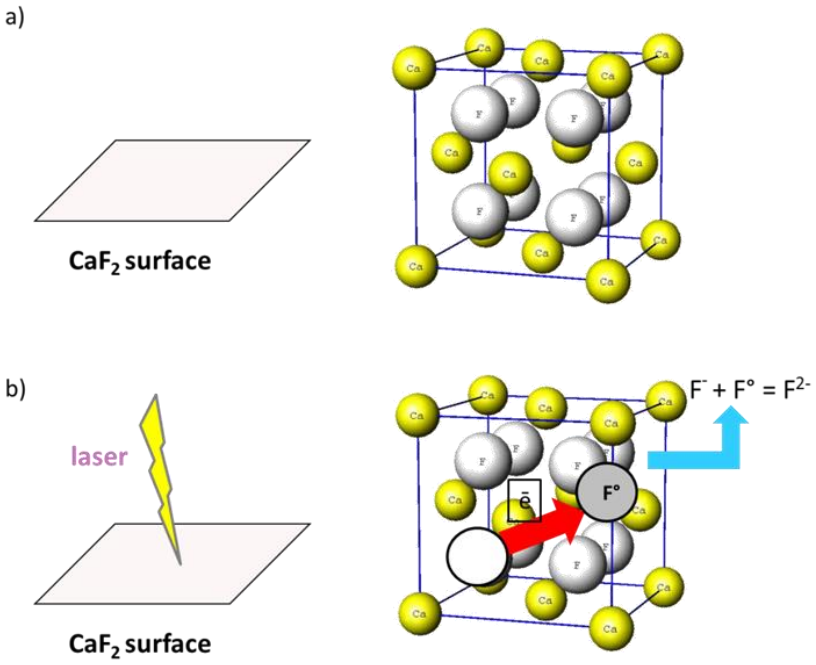


Figure IV.15 : Crystallographic structure of CaF_2 a) initial state b) after laser striking the surface.

However, when the intense laser beam strikes the sample, the energy of the photons (*hv*) induces the liberation of electrons from the F^- anions. The various paramagnetic defects can be created due to the formation of electron-hole pairs, their mobility and trapping by means of lattice deformation [242-244]. One of the most important defects are so called *F-centers* (V_F^x) (see equation (IV.1)), which are electrons trapped in a fluoride ion vacancies [242, 243, 245].

$$\mathbf{CaF_2 + h\nu \rightarrow F^{\circ} + V_F^x} \quad \text{(IV.1)}$$

Comparison with EPR results reported previously for X-irradiated CaF_2 [240, 246] as well as the fact that such defects should be the most abundant paramagnetic centers in the photo damaged plate, the first two components extracted by MCR-ALS (red and green lines) in Figure IV.12 are assigned to *F-centers*. The presence of two such centers exhibiting similar spectra with slightly shifted main peaks (3746 G vs 3726 G) can be attributed to the fact that electrons are trapped in the vacancies with two different microenvironments.

The assignment of the third and fourth extracted components is more problematic. Nevertheless, at least two other types of paramagnetic species should be considered. The first one is the fluorine neutral atom that liberated the electrons. In fact, fluorine atom possesses nine electrons therefore has five electrons on their outermost energy level. When the fluoride anion (F^-) loses one electron, it has the same electronic configuration as the fluorine neutral atom (F°) (see Figure IV.15). Therefore, it becomes a paramagnetic atom that can be detected by EPR spectroscopy.

Another group of defects in CaF_2 are self-trapped holes which have the structure of molecular ion F_2^-. Such coupling can take place between two adjacent F^- ions on a normal lattice site (V_k center) or when one of them is an interstitial fluorine atom (H center) [246-249] (see Figure IV.12 and Figure IV.15).

$$\mathbf{F^{\circ} + F^{-} \rightarrow F_2^{-}} \quad \text{(IV.2)}$$

The third (blue line) and fourth (yellow line) spectra extracted by MCR-ALS can be thus, tentatively assigned to what was mentioned above fluorine defects although the precise attribution is very difficult without additional experimental data. In general, the defects formed on the alkaline earth fluoride CaF_2 are associated with the changes in the valence state of the anion sub-lattice as well as intrinsic damages. Figure IV.16 shows the different types of paramagnetic ions emphasizing on the fact of the presence of one free electron in their outermost energy levels.

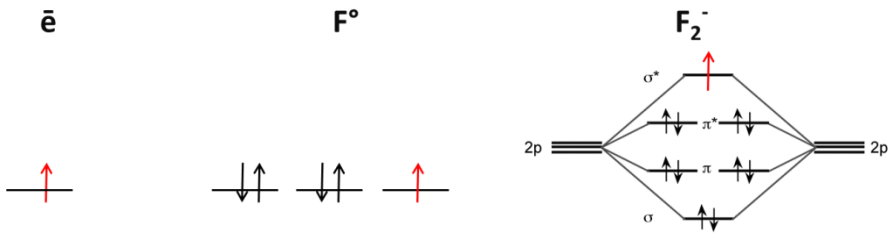


Figure IV.16 : Three types of paramagnetic ionization.

It should be also stressed that commercially available CaF_2 plates contains various impurities giving rise to EPR signal. In consequence, the smaller signals visible in every extracted spectrum in addition to the main resonance lines can be assigned to the impurities such as Y, Mn, Gd and other rare earths surrounded by fluorine ions [246]. Unfortunately, the spatial resolution of such impurities in photo damaged plate is very difficult due to their small concentration in comparison to predominant paramagnetic species i.e. defects.

The resolution of the pure concentrations maps in Figure IV.12 demonstrates the advantage of MCR-ALS analysis over the univariate data exploration exemplified. One can observe that the localization of the paramagnetic species is significantly more complex than the one observed on the global integration image (Figure IV.3b). Indeed comparing the second MCR extracted chemical map with the one in Figure IV.3c, it can be concluded that the light blue C2 zone was an overestimation of the concentration due to a spectral interference between the two *F-centers* spectra. This issue is even more pronounced when comparing between the chemical map extracted at 3748 Gauss (Figure IV.3d) and the first MCR extracted map. D1 zone is the most intense contribution and almost no contribution should have been present in D2 zone. These examples demonstrate the importance of applying appropriate multivariate curve resolution methods to extract the correct pure spectra for species identification and to generate unbiased distribution maps.

IV.7 Conclusions

The results presented in this chapter demonstrate the applicability of multivariate approach in analysis of hyperspectral imaging data. The novelty of this work is that, for the first time, the set of EPR images acquired at different magnetic fields were analyzed by MCR-ALS method to retrieve the pure distribution maps and the pure spectra of each constituent on

the CaF_2 surface that was irradiated by a laser. The most important achievement is the extraction of the spatial distribution maps presenting precise location of different light induced paramagnetic centers on the studied surface. Moreover, the identification of each created paramagnetic species was achieved from the resolved pure spectra. It should be emphasized that all the extractions were obtained without prior knowledge about the chemical system, which reflects the power of MCR-ALS methodology. The obtained results are convincing that the proposed methodology will provide new trends for data analysis in EPR imaging, particularly when the complex and unknown samples are explored.

Chapter V: Application of MCR-ALS to Time Evolving EPR data

V.1 Introduction

The study of the kinetics of chemical reactions is very often based on a univariate approach. Typically, a series of spectra are measured as a function of time and the evolution of single peaks is independently analyzed (see Figure V.1) such that the amount of a particular compound is followed by the increase or decrease in that band.

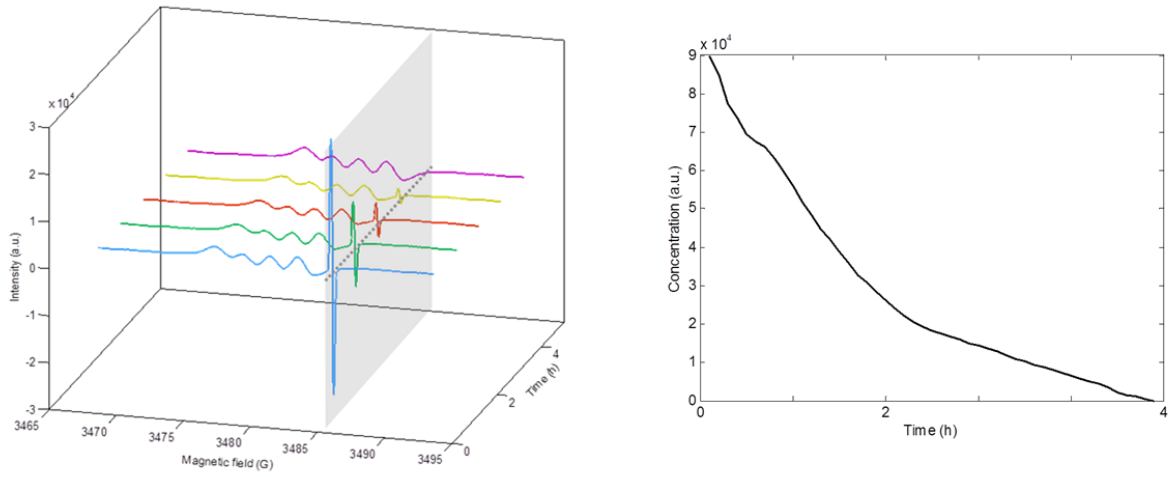


Figure V.1 : Univariate analysis of single peak at 3485 G in series of time-resolved EPR spectra (left) and its corresponding kinetic profile (right).

This approach can lead to biased interpretation when the signals from different species are overlapped i.e. when the spectral variable is not selective. Thus, multivariate methods are indispensable in the analysis of complicated systems and can overcome most of the problems reported by univariate approaches. In most chemometric methods, knowledge about reactants and the order of the reaction is required. A given model is, then, assumed and can be implemented in chemometric analysis (hard modeling based methods) [250]. However, the acquaintance of the kinetic model is sometimes very uncertain or not known, at all. Thus, soft modeling method such as MCR-ALS can provide a solution to such problems. MCR-ALS was successfully applied to different time evolving data [107] as well as other kinds of sequence ordered data sets in chromatographic runs [251-253], series of titrations [129, 254, 255] and voltammetry cycles [256]. It is, hence, a useful tool in kinetics study.

EPR is an irreplaceable tool in the study of kinetics for many kinds of paramagnetic centers. It is especially a valued technique in the examination of reactions of radicals. In addition, it was demonstrated in Chapter III that MCR-ALS is an effective approach in the

analysis of entangled EPR signal from paramagnetic mixtures. Therefore, the application of MCR-ALS on EPR time evolving data seems to be a promising tool in the study of the reactions of radicals [257]. However, in contrast to a mixture of non-interacting paramagnetic species, as studied in Chapter III, a reactive mixture can result in highly correlated concentration profiles among the different species. In such a case, chemometric analysis can be more challenging and more restrictive constraints are needed to obtain correct MCR-ALS resolution and quantitation. In the analysis of sequence ordered data set, very often, an additional constraint which imposes monotonic character of profile, i.e. unimodality (see paragraph II.3.2.6), is applicable. Nevertheless, in complicated systems even this additional constraint is not enough to obtain satisfactory resolution. To decrease more rotational ambiguity without imposing a specific model on the profiles (hard modeling) it is possible to use multiset data analysis (see paragraph II.3.2.8). Such a possibility is one of the biggest advantages of MCR-ALS method. To benefit this property, data sets coming from independent experiments on the same system are required. Simultaneous analysis of different data sets which share common spectral or/and concentration profiles can drastically decrease the rotational ambiguity and help to find more consistent solutions [258].

The objective of this chapter is twofold; to confirm the applicability of MCR-ALS to EPR data set from reactive mixture of paramagnetic species and to demonstrate the advantage of multiset analysis in decreasing the rotational ambiguity in this type of data. As a model system for such study, one of the derivatives of 3-hydroxycoumarin (see below) was chosen because it forms different radicals in alkaline conditions. It is also intensively studied due to its antioxidant properties [259]. The presented results are the first MCR-ALS application on EPR kinetics data.

V.2 3-hydroxycoumarins as antioxidants

Antioxidants are widely spread chemicals that interact with the free radicals and prevents them from causing damage [260]. The former is also known as free radical scavenger. Free radicals are highly reactive metabolites and are naturally produced by the body as a result of normal metabolism and energy production. In addition, these radicals are biological response to environmental toxins such as radiations, smoke, chemicals and even

pharmaceutical drugs. These incomplete molecules aggressively attack other molecules in the body to refill the missing electrons. This oxidation severely affects the cells and leads to different diseases, antioxidants defend the body against such damages [261]. The body produces some of the antioxidants that are called endogenous antioxidants. On the other hand, the body relies also on external (exogenous) sources, primarily the diet, called dietary antioxidants in order to obtain the rest of the required antioxidants [262]. For instance, rich sources of dietary antioxidants are fruits, vegetables and grains [262]. Antioxidants play an important contributory role in the treatment of many degenerative or chronic diseases such as atherosclerosis, brain dysfunction, immune system decline and cancer, since, considerable experimental evidence links the production of reactive oxygen species to the initiation and or progression of those pathologies [263-266]. There exists a wealth of antioxidants, for instance, ascorbic acid (vitamin C), α-tocopherol (vitamin E), glutathione (GSH), carotenoids, lycopene and flavonoids are some examples [267]. 3-hydroxycoumarin is an antioxidant that belongs to the coumarin-related compounds. The former is rare by nature and is formed from the hydroxylation of the coumarins at the 3-position by liver microsomes [268]. In 1990, Aihara *el al.* described their first chemical preparation from coumarins [269]. The 3-hydroxylation turned out to significantly improve antioxidant activity and, for instance, increased the 5-lipoxygenase and α-glycosidase inhibitory activities of coumarins. One of the derivatives of this molecule which exhibits very promising antioxidant properties is 3,6,7-trihydroxycoumarin presented in Figure V.2. In this work, EPR spectra of this molecule measured in alkaline conditions as a function of time are analysed by MCR-ALS.

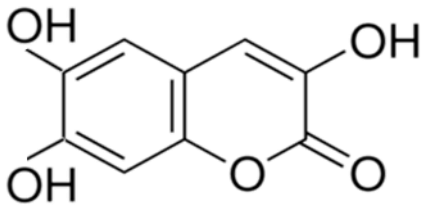


Figure V.2 : 3,6,7-trihydroxycoumarin.

V.3 Sample preparation

3,6,7-trihydroxycoumarin was synthesized as described by Cotelle *et al.* [257]. A weight of 1 mg of the compound was mixed with 500 µL solution of sodium hydroxide (1M).

In alkaline conditions, the compound has propensity to be oxidized and form several radicals. The samples (potentially mixtures of radicals) were monitored by EPR as a function of time to obtain a set of 80 time-resolved spectra (first set). Analogously, the second set was prepared by mixing the same compound with a solution of NaOH (2M). In this case, 80 EPR spectra were also recorded as a function of time.

V.4 Spectral acquisition

Samples were analyzed with a continuous wave-electron paramagnetic resonance spectrometer (CW-EPR). Once conditions have been optimized for measurement of the signal of interest, a spectrum required only a few seconds to be recorded. EPR spectra were recorded at a constant room temperature of 20°C, using a Bruker ELEXSYS 500 spectrometer operating at the X-Band. All spectra were recorded at a modulation field frequency of 100 kHz and a microwave frequency of 9.80 GHz, with an amplitude modulation of 0.002 mT and a microwave power of 0.1 mW corresponding to non-saturation conditions. The weak pitch from Bruker was used as standard reference and contained a known concentration of spin/mass (1.29×10^{13} spins/g). The following spectrometer parameters were used during the spectral acquisition: receiver gain 68 dB, time constant 40.96 ms, conversion time 81.92 ms. The spin concentration is given by the double integration of the first derivative of the EPR signal. Spectra were measured over a 1024 point range (from 3473 G to 3486 G). The delay between each spectrum is 900 seconds which covered 20 hours as a time range for the first set and 20 hours for the second set.

V.5 Data treatment

The preliminary analysis attempts were performed by using the spectra in their raw first derivative form. Similar to Chapter III (section III.6), the results obtained by the chemometric methods were not consistent due to strong overlap. Therefore, integrated spectra (Figure V.3) were used for the analysis what made it possible to insert additional constraints (non-

negativity on spectral profile) and to decrease the rotational ambiguity in MCR-ALS resolution.

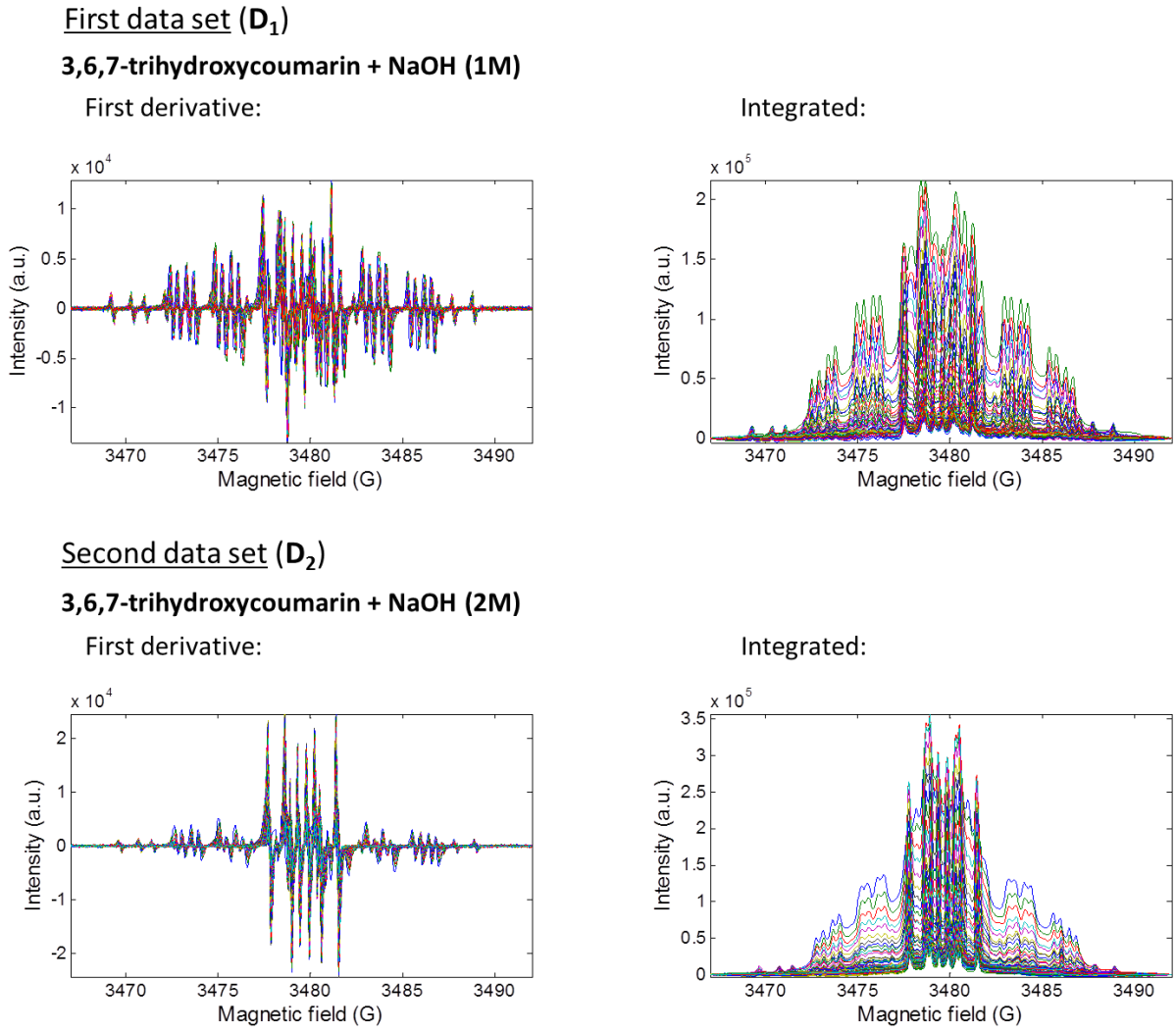


Figure V.3 : 80 EPR time-resolved spectra obtained for 3,6,7-trihydroxycoumarin in 1M NaOH in their first derivative form (upper left) and integrated form (upper right). 80 EPR time-resolved spectra obtained for 3,6,7-trihydroxycoumarin in 2M NaOH in their first derivative form (lower left) and integrated form (lower right).

The first data set is represented in a matrix $\mathbf{D_1}$ (80×1024), it consists of 80 rows that corresponds to different spectra each acquired at a certain time t, and 1024 different columns representing 1024 different magnetic fields. The second data set is represented in a matrix $\mathbf{D_2}$ (80×1024) and possesses identical structure as first data set.

In the first approach, PCA was applied to the first individual data set $\mathbf{D_1}$ as well as to the augmented matrix $\mathbf{D_1}$ and $\mathbf{D_2}$ to get preliminary insight about the data structure and to estimate the number of pure components. Next, MCR-ALS was applied to the same data.

V.5.1 PCA for individual matrix

The first goal to achieve was the estimation of the number of pure components in the data set i.e. number of radicals possibly formed during the reaction of 3,6,7-trihydroxycoumarin with NaOH (1M). For this purpose, PCA was first applied to the data set $\mathbf{D_1}$. Figure V.4 (on the left) shows the plot of the first 20 eigenvalues. Two different components are certainly present in this data set, the logarithm plot (on the right) suggests the presence of three significant singular values due to the sudden change in the slopes.

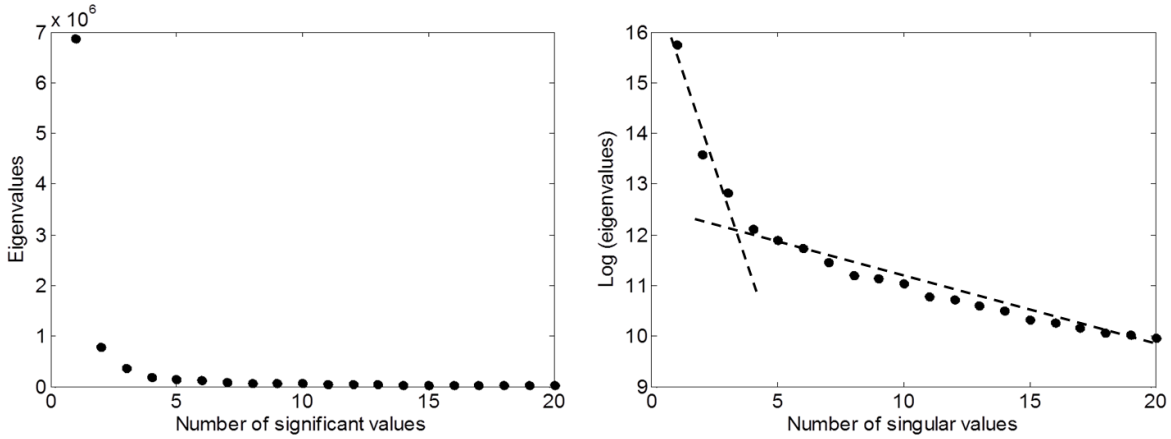
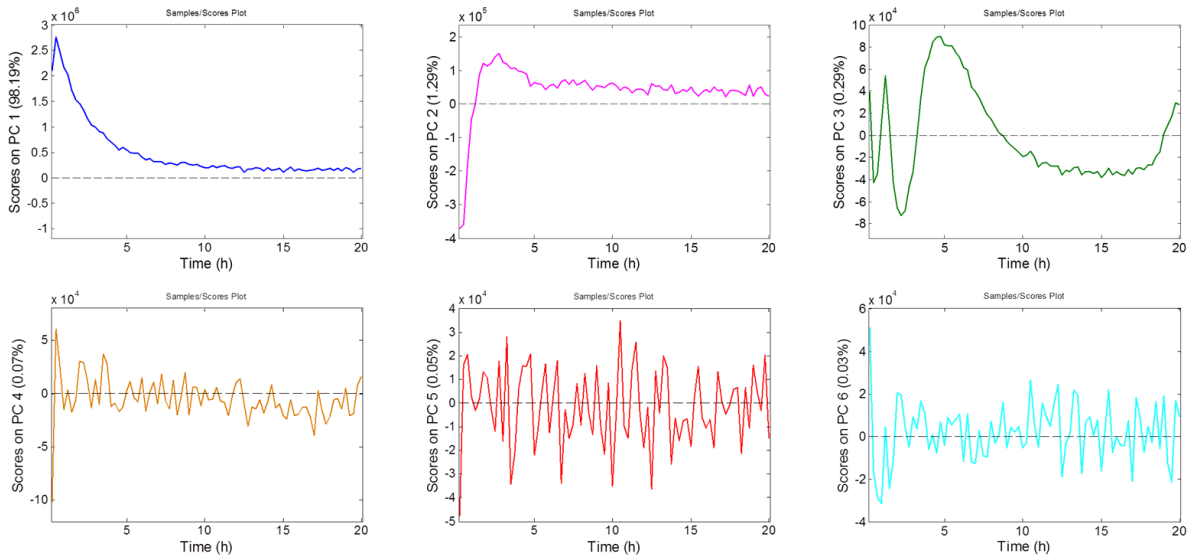


Figure V.4 : Eigenvalues (left) and their logarithms (right) obtained by Singular Value Decomposition of the spectral data matrix $\mathbf{D_1}$.

For further investigations about the number of components, the score $\mathbf{T_1}$ and the corresponding loading plots $\mathbf{L^T_1}$ on the first six Principal Components are presented in Figure V.5. The first three score plots corresponding to PC1, PC2 and PC3 show a clear insight of time profile, whereas the scores on PC4, PC5 and PC6 are not really structured. In addition, the first three Principal Components represent spectral information whereas the others are less consistent with EPR spectral shapes. Moreover, the first three Principal Components represent the values 98.19%, 1.29% and 0.29%, respectively from the total variance. The fourth Principal Component represents 0.07%, the fifth represents 0.05% and the sixth 0.03% which are negligible values compared to the values of the first three components. Therefore, the first three Principal Components represent the majority of the chemical variance, whereas the others explain noise variance.

Scores ($\mathbf{T_1}$):



Loadings ($\mathbf{L^T_1}$):

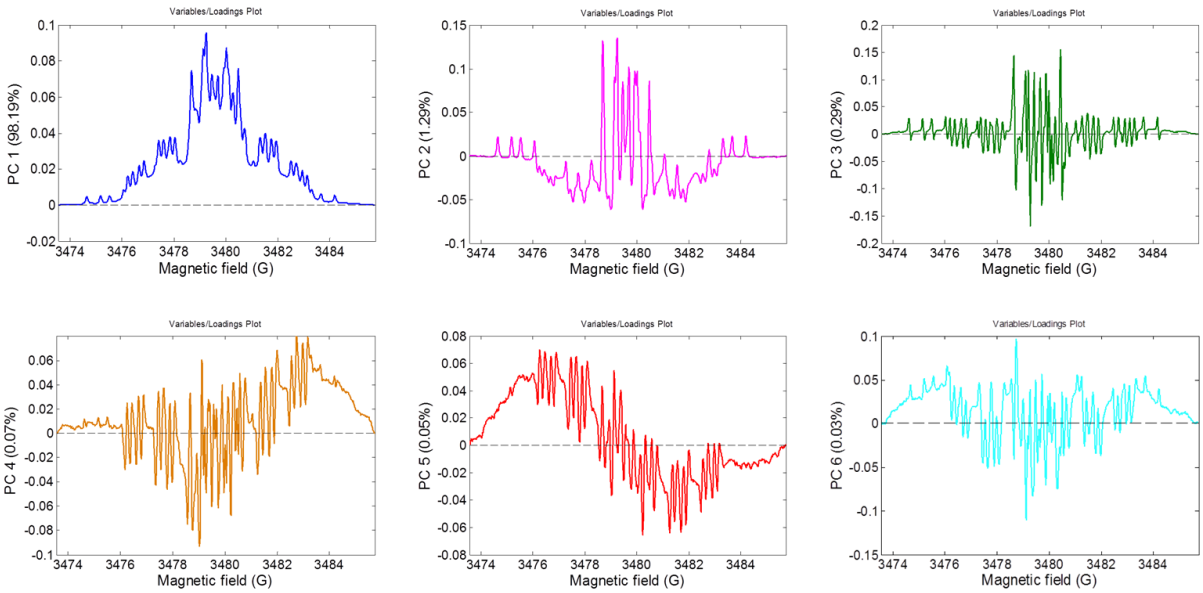


Figure V.5 : Scores $\mathbf{T_1}$ (top) and their corresponding loadings $\mathbf{L_1}^{\mathbf{T}}$ (bottom) plots for the first six Principal Components obtained by PCA on $\mathbf{D_1}$.

In conclusion, the rank of the matrix $\mathbf{D_1}$ seems to be three and therefore, three chemically meaningful contributions are deduced. In the next step, the MCR-ALS method is required to extract the pure time and spectral profiles.

V.5.2 MCR-ALS for individual matrix

The first step in MCR-ALS is the determination of the number of pure component and according to the previous section, three pure contributions are assumed in the experimental mixture. The second step is the generation of initial estimates, in this case Evolving Factor Analysis (EFA) was applied as suitable method for the study of time evolving processes (see section II.3.2.3). Figure V.6 presents the initial estimation $\mathbf{C_{ini}}$ obtained by EFA.

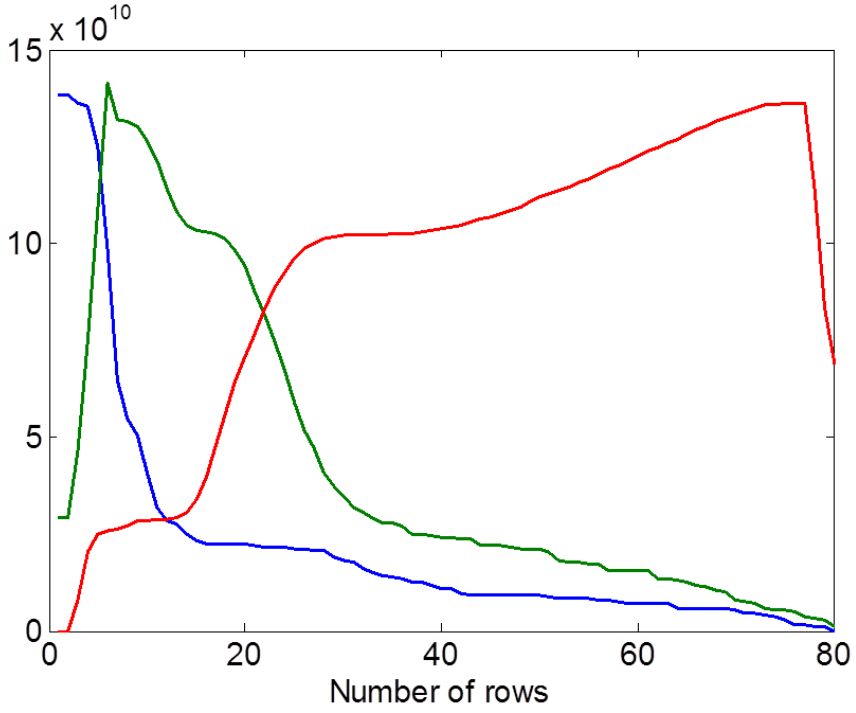


Figure V.6: Initial estimate $\mathbf{C_{ini}}$ obtained as a result of EFA on $\mathbf{D_1}$.

Next, alternating least-squares is started by using the integrated matrix $\mathbf{D_1}$ and the initial estimate $\mathbf{C_{ini}}$ with the implementation of the non-negativity constraint on both spectral and concentration profiles as well as spectral normalization. Individual resolution of the experiment $\mathbf{D_1}$ by MCR-ALS is presented in Figure V.7. The concentration and spectral profile of three contributions being possibly pure radicals or their mixtures (see below) are extracted from the sole matrix $\mathbf{D_1}$.

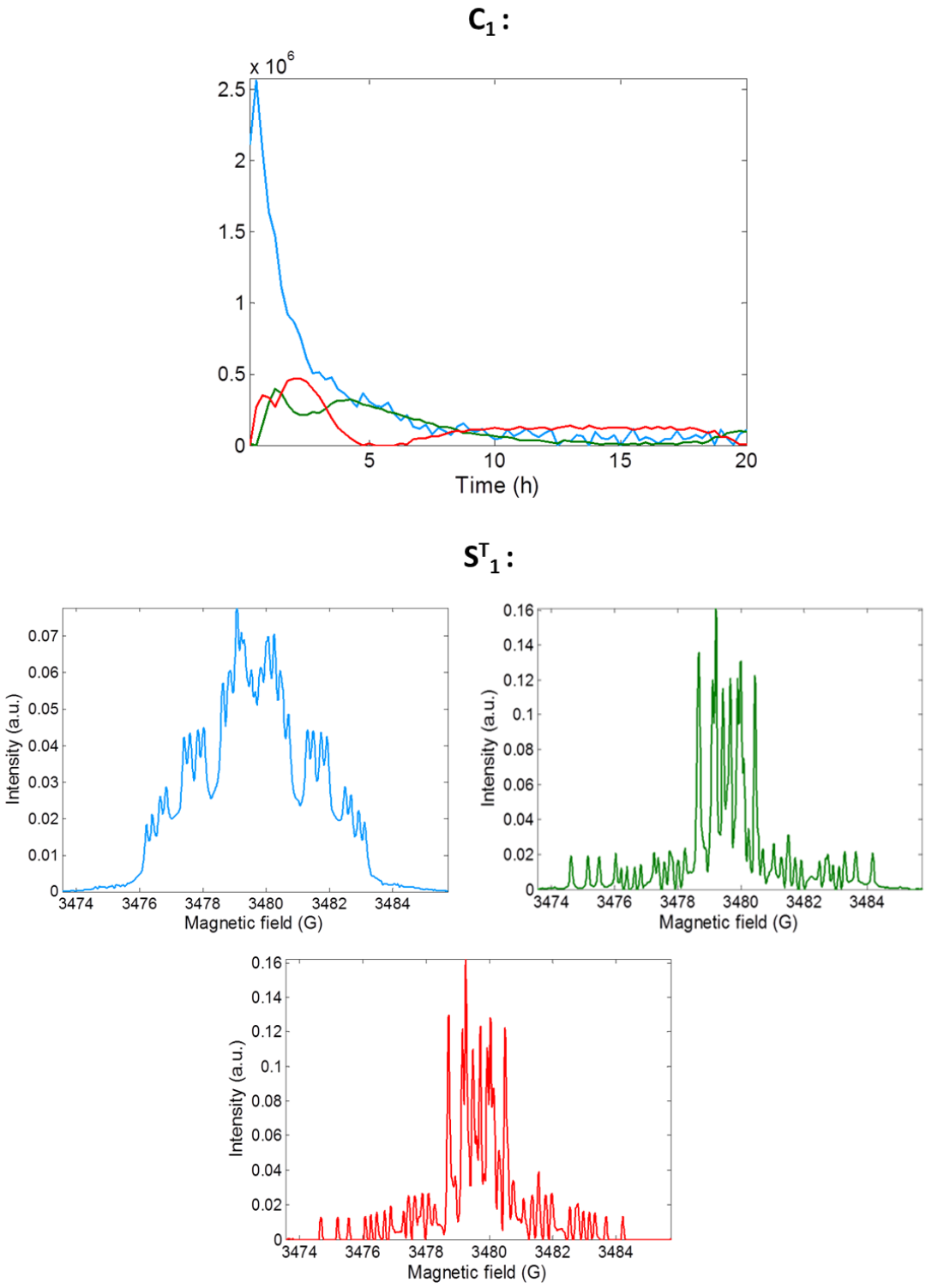


Figure V.7 : Concentration profiles $\mathbf{C_1}$ (top) and corresponding spectra $\mathbf{S^T_1}$ (bottom) of three components resolved by MCR-ALS on data set $\mathbf{D_1}$.

Even a rough inspection of spectra obtained by MCR-ALS shows that the second (green) and the third (red) spectra are very similar. This is clearly visible when superposing

the second and the third spectra in Figure V.8 where both spectra are nearly identical but slightly shifted.

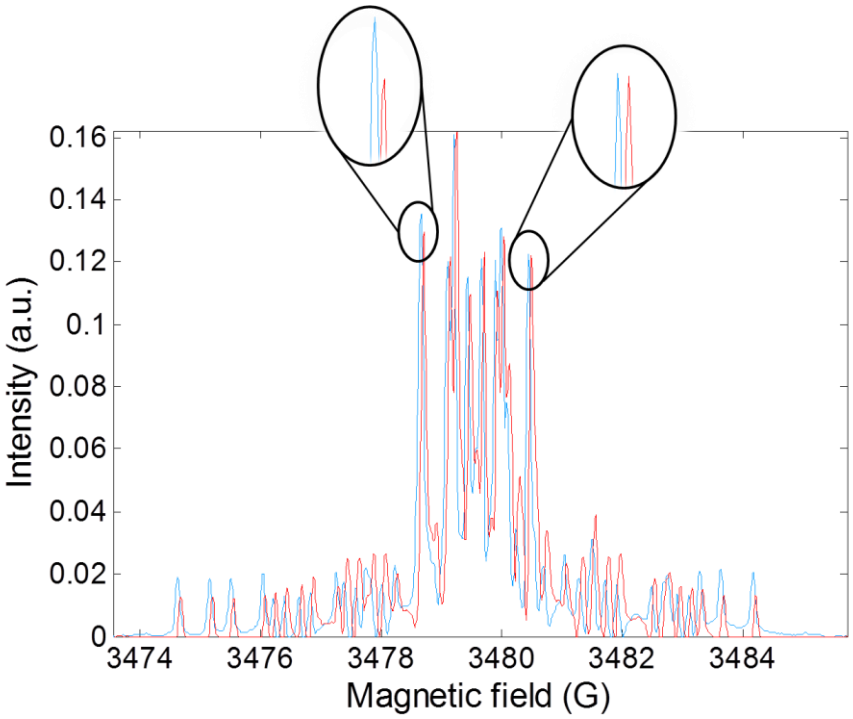


Figure V.8: Superposition of the second and the third pure spectra obtained by MCR-ALS.

Reliability of the obtained solution was evaluated by the lack of fit and the explained variance giving reasonable values of 4.8% and 99.78%, respectively. We can consider that MCR-ALS needs two spectral contributions to take into account a spectral shift of one compound along time. The fact of the presence of such shifts is highly undesired (as shown previously in chapter III in section III.3) in MCR-ALS analysis since the data set loses its bilinear character. Therefore, in order to achieve reliable results, the spectral alignment of shifted spectra has to be performed.

V.5.3 Alignment of the spectral shifts

During a long EPR experiment, several factors can fluctuate in time (temperature, magnetic field intensity etc.). Consequently the spectra, representing the same species but measured at different moments, can exhibit some spectral shifts even if the global shape of the spectrum is conserved. Thus, a spectral line in one spectrum can be detected at magnetic field shifted of few Gauss in relation to the corresponding line in another spectrum. If the shift of

all the lines in a given spectrum is the same, a simple procedure of manual alignment of the spectra is a sufficient solution. The procedure starts first by fixing a reference spectrum among the series of spectra, and then the position (in Gauss) of each spectrum is compared with the analog peak in the reference spectrum. As a result, the given spectrum is shifted by several gauss in relation to the reference one making the spectra aligned. Careful inspection reveals that the observed shifts are homogenous in entire spectral range. In this way, we consider that a simple manual shift correction can be enough. All the spectra in each data set were corrected in the same way in Matlab. In addition the data set $\mathbf{D_2}$ was spectrally aligned with data set $\mathbf{D_1}$ for multiset analysis purpose (see below).

V.5.4 Exploration of the corrected data

After the alignments of the spectra in the matrix $\mathbf{D_1}$ the chemometric analysis was repeated on the corrected data set $\mathbf{D_1'}$.

V.5.4.1 PCA on corrected individual data set

In spite of data alignment, SVD analysis presented in Figure V.9 seems to indicate the same number of three significant singular values separated from the noise level in the data matrix $\mathbf{D_1'}$. However, a fourth contribution is not excluded.

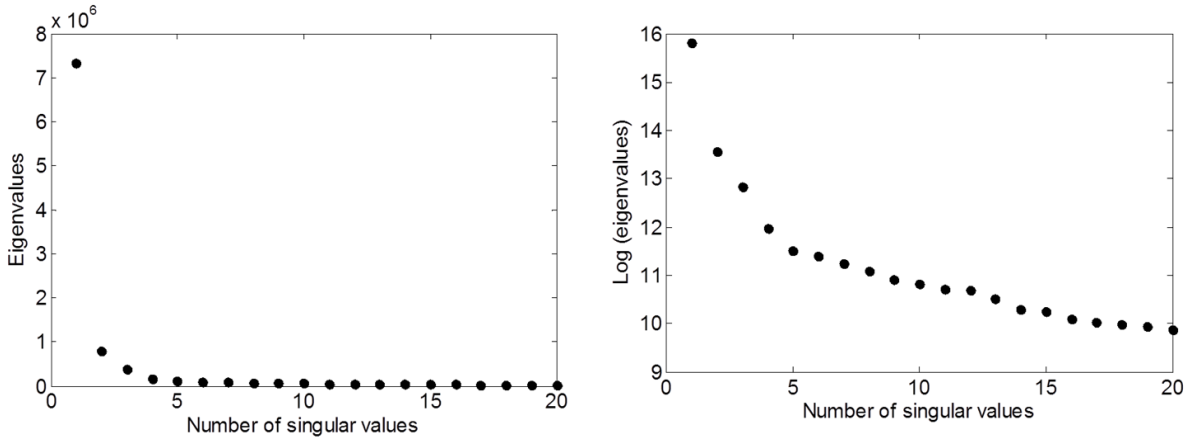


Figure V.9 : Eigenvalues (left) and their logarithms (right) obtained by Singular Value Decomposition of the spectral data matrix $\mathbf{D_1'}$.

The score and corresponding loading plots on the first six Principal Components obtained by PCA are shown in Figure V.10 for further inspection about the number of significant components. The first three scores on PC1, PC2 and PC3 exhibit more or less regular insight of time profile pattern whereas the scores on PC4, PC5 and PC6 are not structured. The first three Principal Components represent the values 98.50%, 1.10% and 0.26%, respectively from the total variance. The fourth Principal Component represents 0.04%, the fifth represents 0.02% and the sixth 0.01% which are negligible values compared to the values of the first three.

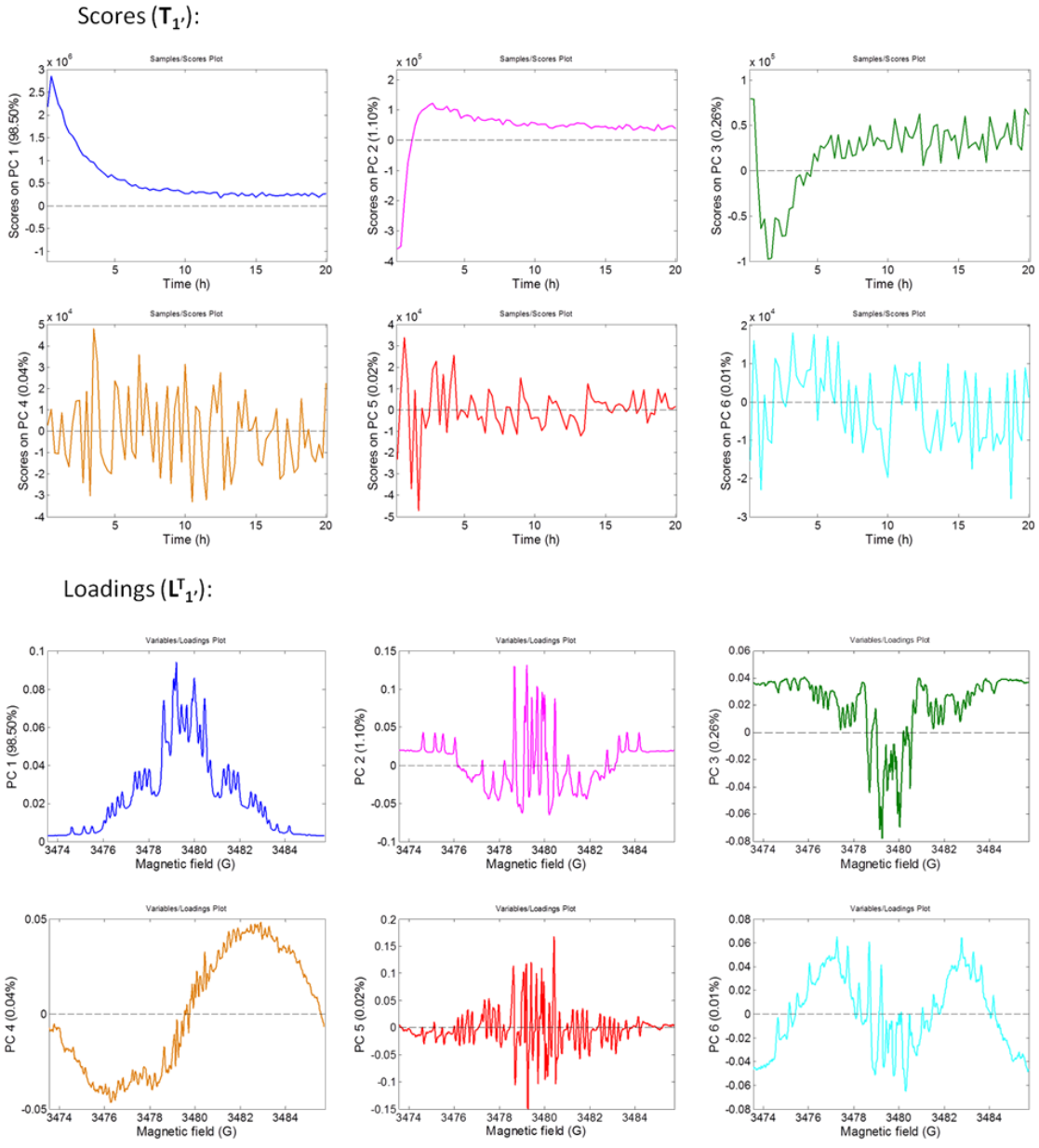


Figure V.10 : Scores $\mathbf{T_1}$, (top) and their corresponding loadings $\mathbf{L^T_1}$, (bottom) plots for the first six Principal Components obtained by PCA on $\mathbf{D_1}$'.

Thus, the results indicate that three seems to be the total number of species present in the studied data. Nevertheless, the loadings and scores of PC2 and PC3 cannot represent meaningful spectra and kinetic curves due to their negative values and orthogonality property. Hence, MCR-ALS analysis is needed to impose adequate constraints.

V.5.4.2 MCR-ALS on corrected individual data set

Knowing that the rank of the matrix is three, initial estimation of the concentration profile was obtained by EFA (Figure V.11) like in paragraph V.5.2. It is worth noting that the correction of data set resulted in different initial time profiles (compare with Figure V.6). In this case, one can observe more consistent profiles for the red and green profiles.

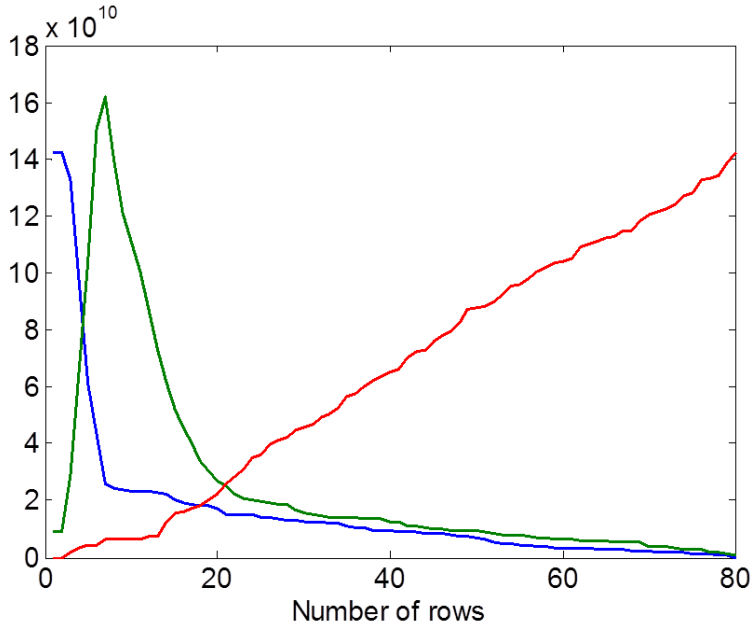


Figure V.11: Initial estimate $\mathbf{C_{ini}}$, obtained as a result of EFA on $\mathbf{D_1}$'.

Alternating least squares optimization was then started with the implementation of the non-negativity constraint on spectral and concentration profiles as well as spectral normalization. The results obtained by MCR-ALS are shown in Figure V.12. In case of spectra of second (green) and third (red) components, the similarity is still observed but the spectra are certainly not identical as for an uncorrected data set (compare with Figure V.7). However, an improvement in the resolution is observed by uncertainty factors. Given the signal-to-noise ratio in the raw EPR data, a lack of fit of 3.62 % and explained variance of

99.87 % is a good figure of merit for the proposed extraction. We can conclude from the above results that the procedure of alignment did not completely resolve the ambiguity problem since the components (red and green) are still similar.

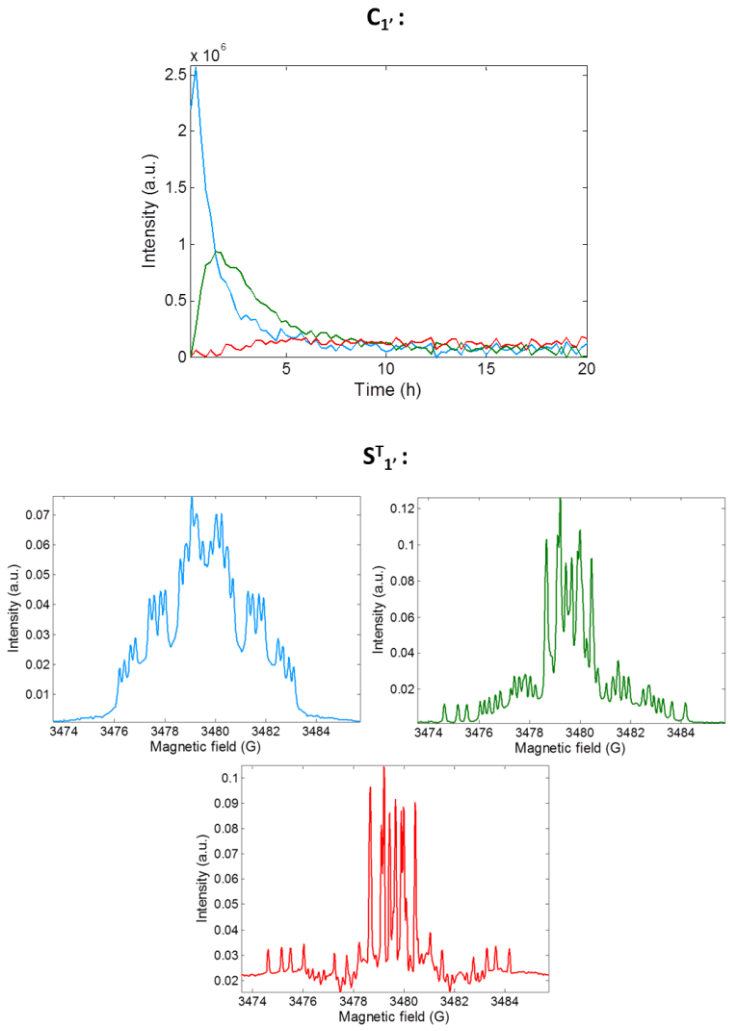


Figure V.12 : Concentration profiles $\mathbf{C_1}$, (top) and corresponding spectra $\mathbf{S^T_1}$, (bottom) of three components resolved by MCR-ALS on data set $\mathbf{D_1}$,.

As mentioned in the introduction, the procedure applied above can be easily extended to the simultaneous analysis of multiple data sets to improve resolution. This is possible as long as there is at least one data mode (direction) in common. In this case, two different data sets $\mathbf{D_1}$, and $\mathbf{D_2}$, were analyzed by the same EPR spectrometer in identical spectral range. Therefore, an extension of the bilinear model, i.e. multi-set MCR analysis, can be applied in order to decrease the ambiguity that is observed in the spectral profiles from the previous experiment.

V.5.5 MCR-ALS: Multiset data analysis

The two individual data sets $\mathbf{D_1}$, and $\mathbf{D_2}$, obtained using the experimental procedure described previously share the same spectral range and are supposed to exhibit different kinetics time profiles due to different alkaline conditions (see paragraphs V.3 and V.4). Multivariate curve resolution method has the possibility to be applied to the simultaneous analysis of more than one data matrix. In this case, matrices are arranged in an augmented matrix. Since the matrices have the same spectral dimension therefore, a column augmented matrix can be constructed in the magnetic field direction (Figure V.13). It can also be called magnetic field wise-augmented matrix, keeping the magnetic fields values in common.

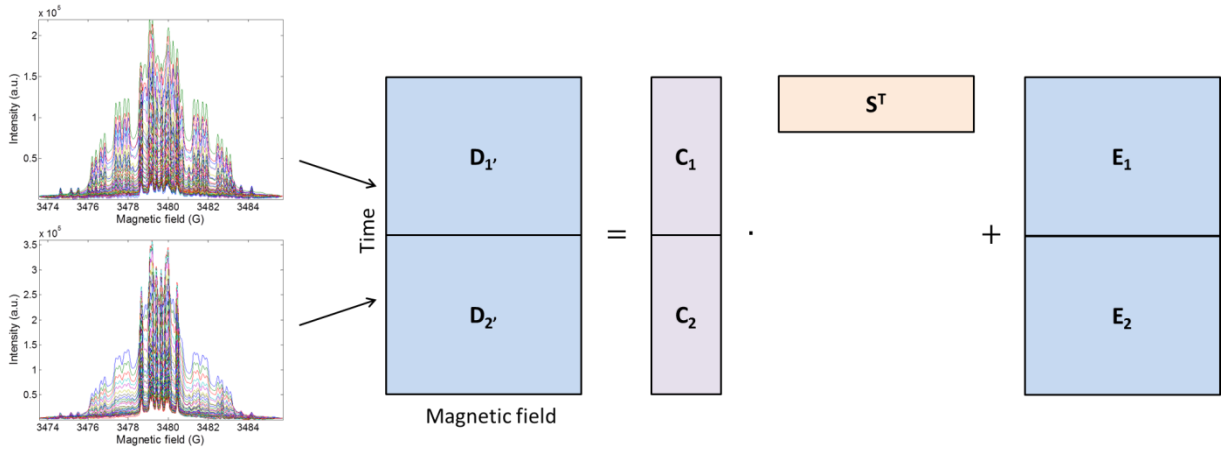


Figure V.13 : Column augmented matrix.

This magnetic field wise-augmented matrix denoted by $\mathbf{D_{aug}}$, formed by $\mathbf{D_1}$, and $\mathbf{D_2}$, has a number of rows equal to the total number of acquired spectra in two kinetic experiments (160 spectra) and a number of columns equal to the number of magnetic fields present in each single matrix (1024 magnetic fields). The time-wise augmentation (also called row-wise augmented matrix) is only possible if the same time evolution in both matrices is expected what is not the case in the studied data sets. It is also often used when different spectroscopic techniques are used to probe the same sample with a perfect synchronization of time.

For the augmented matrix $\mathbf{D_{aug}}$, the linear model can be written as

$$\boldsymbol{D_{aug}} = \boldsymbol{C_{aug}} \cdot \boldsymbol{S^T} + \boldsymbol{E_{aug}} \qquad \text{(V.1)}$$

Multiset MCR-ALS consists of several steps detailed below. The augmented concentration matrix $[\mathbf{C_1};\mathbf{C_2}]$ denoted as $\mathbf{C_{aug}}$ describes the concentration changes of the resolved

components in each of the related $\mathbf{D_{aug}}$ data matrix, respectively. In contrast, the spectral matrix $\mathbf{S^T}$ is single and presents the pure spectra of the different components present in all the considered $\mathbf{D_{aug}}$. This bilinear model assumes that the components in the $\mathbf{D_{aug}}$ share the same pure spectra, whereas they have different concentration profiles. At least some of the components, if not all, in the different data matrices ($\mathbf{D_1}$, and $\mathbf{D_2}$,) are the same and therefore, they have the same pure spectra, whereas their concentration profile can be different in scale and shape.

V.5.5.1 Rank estimation of the augmented matrix

In analogy to single matrix analysis, the first step performed on augmented matrix $\mathbf{D_{aug}}$ is rank estimation. The aim of this step is again to determine the number of independent components present in the experimental mixture. Thus, PCA was applied on the augmented matrix and the plot of the first 20 eigenvalues is presented in Figure V.14.

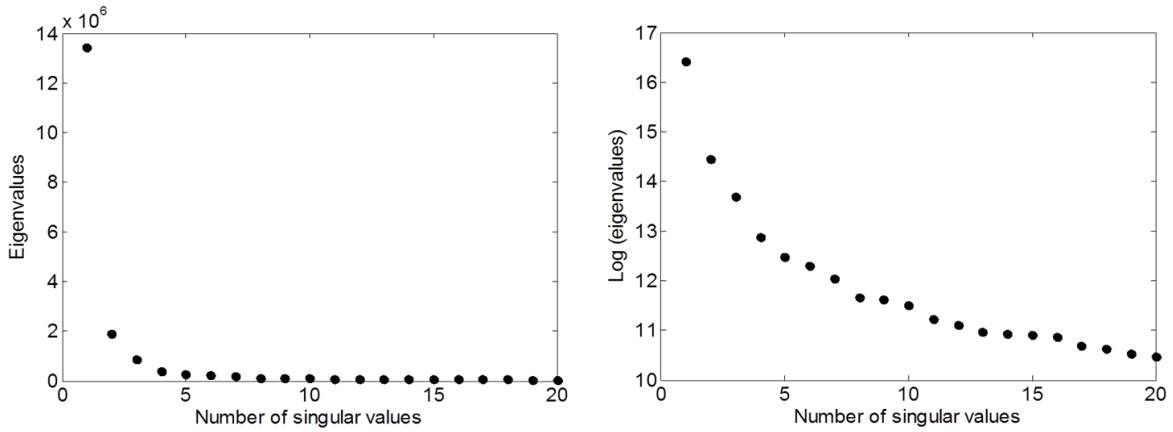


Figure V.14: Eigenvalues (left) and their logarithms (right) obtained by Singular Value Decomposition of the spectral data matrix $\mathbf{D_{aug}}$.

It seems to indicate the presence of three significant singular values but the presence of a fourth component cannot be excluded. It should be stressed here that the presence of an additional component in the augmented matrix in comparison to single matrix can sometimes be due to spectral misalignment between two data sets. However, in this case the matrices $\mathbf{D_1}$, and $\mathbf{D_2}$, were mutually aligned and corrected.

To more reliably judge the number of components, the score and corresponding loading plots on the first six Principal Components, obtained by PCA, were drawn in Figure

V.15. Interestingly, in contrast to single matrix analysis not only the first three scores on PC1, PC2, PC3 but also score on PC4 show a structured time profile (PC5 and PC6 are noisy). This is thus an argument for assuming the presence of four components in the mixture.

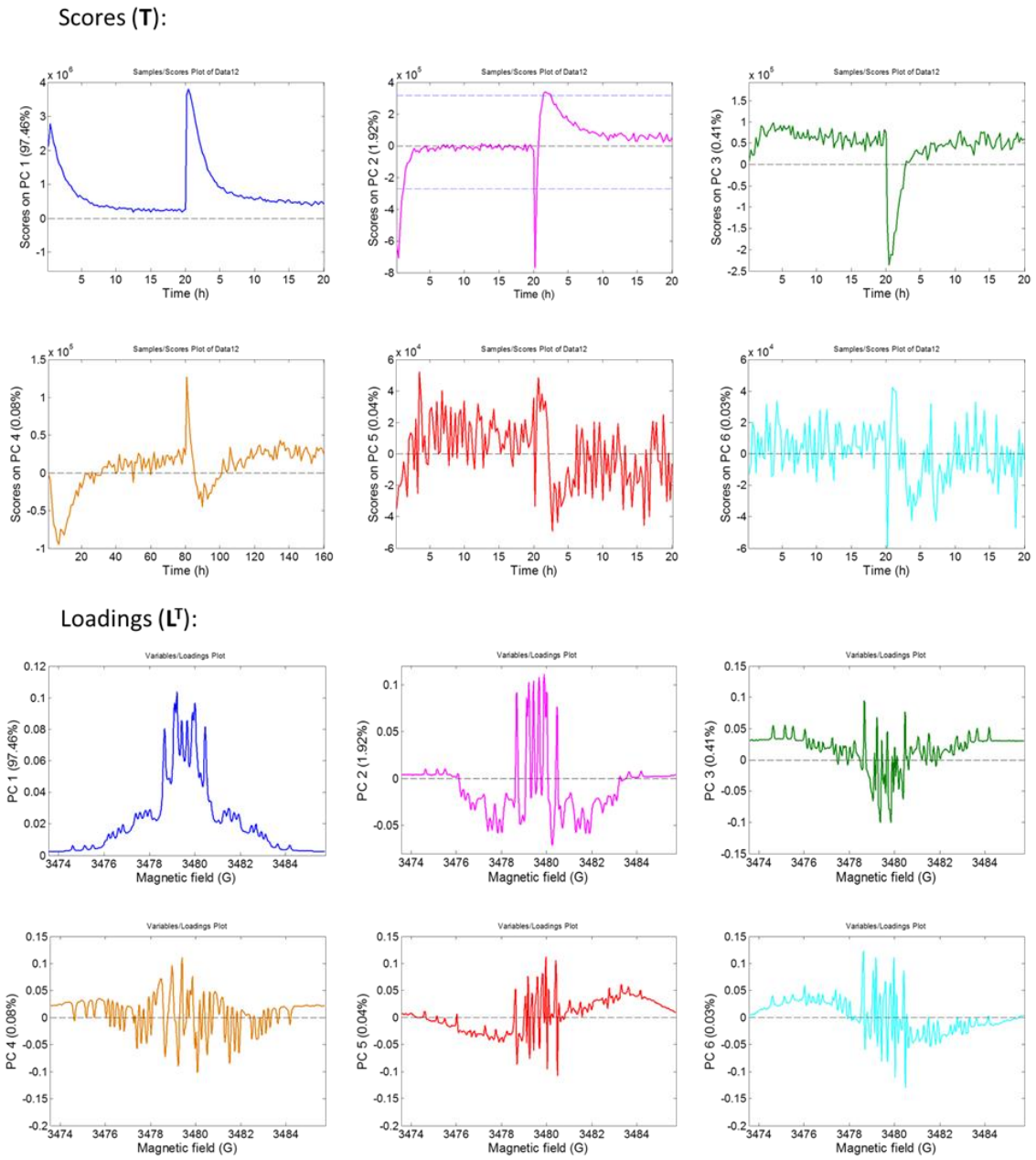


Figure V.15 : Scores **T** (top) and their corresponding loadings $\mathbf{L}^{\mathbf{T}}$ (bottom) plots for the first six Principal Components obtained by PCA on $\mathbf{D}_{\mathbf{aug}}$.

V.5.5.2 Initial estimation

The second step in the MCR-ALS process was the initial estimation of the concentration or spectral profiles. In the case of sequence ordered data, EFA is commonly used for initial estimation (see above). However, in this context, EFA is not suitable to be used in the augmented matrix $\mathbf{D_{aug}}$ since the two sequence ordered matrices $\mathbf{D_1}$, and $\mathbf{D_2}$, have no sequence between each other. Therefore, SIMPLISMA is a more convenient tool to be applied in this case in order to find initial estimates. Initial estimates may be found either from the purest rows or purest columns in the augmented and integrated data $\mathbf{D_{aug}}$. Such rows, then, give an approximation of the pure spectra and the columns are related to the purest concentration profiles. In the studied case, four purest spectra $\mathbf{S^T_{ini}}$ were obtained by SIMPLISMA (Figure V.16) and were used as initial estimates for further ALS optimization.

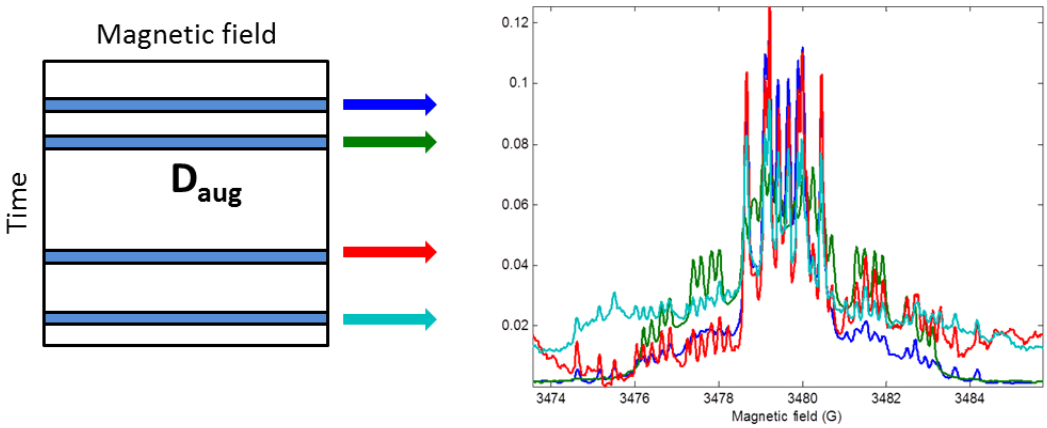


Figure V.16 : SIMPLISMA analysis on the augmented matrix. Representation of the spectral profile of the purest time $\mathbf{S^T_{ini}}$.

V.5.5.3 ALS optimisation

After the estimation of the matrix rank $k = 4$ as well as the initial spectra, alternating least-squares optimization process according to the equation (V.1) was started by using the extended and integrated data $\mathbf{D_{aug}}$ and the initial guess of spectra $\mathbf{S^T_{ini}}$ matrix. Non-negativity in the augmented concentration profile $\mathbf{C_{aug}}$ as well as spectral profile $\mathbf{S^T}$, in addition to spectral normalization, were the constraints implemented during the optimization step. The

iterative process of alternating least squares was stopped when the convergence criteria were achieved. The results obtained from overall MCR-ALS analysis of $\mathbf{D_{aug}}$ are presented in Figure V.17. The final results represent the pure spectra matrix $\mathbf{S^T}$ and the related pure augmented concentration matrix $\mathbf{C_{aug}}$ of the studied kinetic data set. Figure V.17 shows in blue the first component, its time profile in $\mathbf{C_1}$ for the first experiment $\mathbf{D_1}$, and its time profile in $\mathbf{C_2}$ in the second experiment $\mathbf{D_2}$, as well as its spectral signature. The second component is represented in green, the third one in red and the fourth one in yellow.

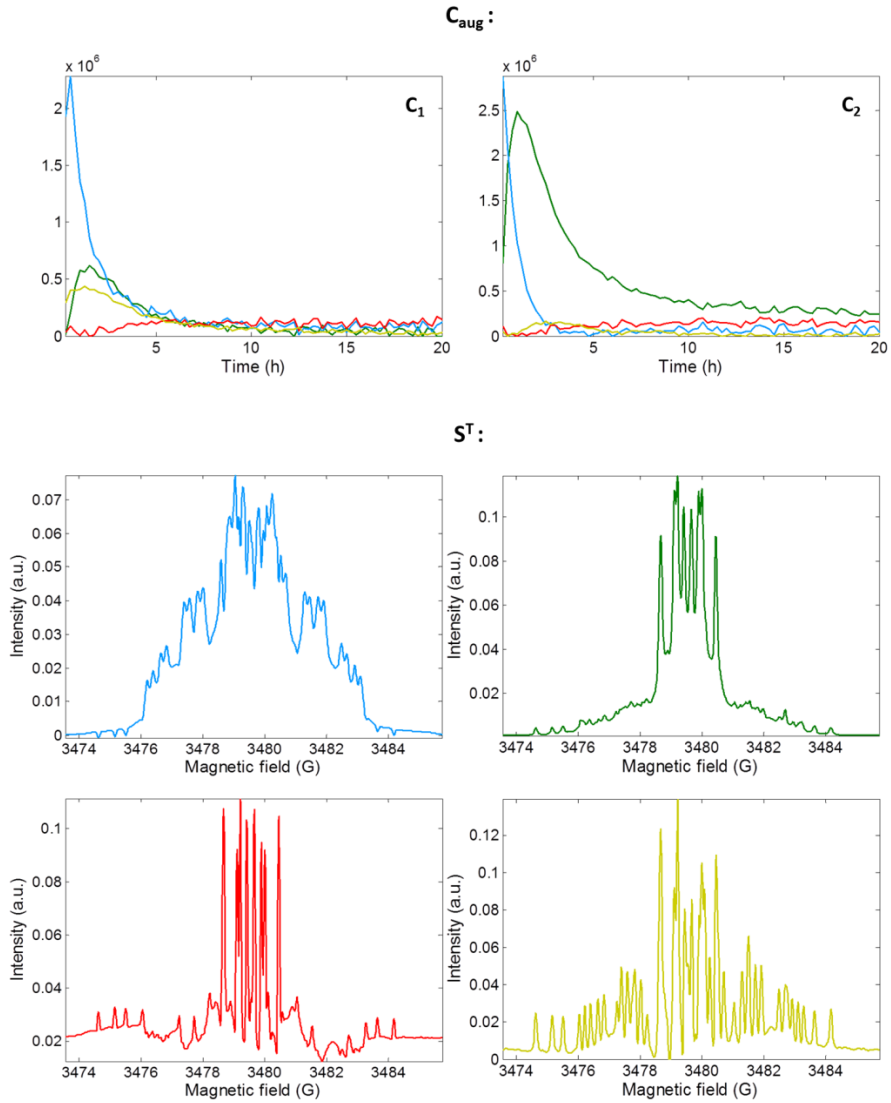


Figure V.17 : Concentration profiles $\mathbf{C_{aug}}$ (top) and corresponding spectra $\mathbf{S^T}$ (bottom) of four components resolved by MCR-ALS on augmented data set $\mathbf{D_{aug}}$.

The uncertainty factors, i.e. a lack of fit of 3.45 % and explained variance of 99.88 %, are slightly better than in the individual matrix analysis. From the concentration profiles, it can be understood that the component (green) is better extracted from $\mathbf{D_2}$, whereas the red and

yellow ones are better extracted from $\mathbf{D_1}'$. This is the reason why multiset analysis is a valuable tool.

V.6 Interpretation of the MCR-ALS results and identification of the radicals

As mentioned in paragraph V.2 and previously described in the literature [257] 3-hydroxycoumarin can be oxidized under alkaline conditions and form various radicals. They can be further detected by EPR spectroscopy. According to Cotelle *et al.* six radicals can be formed in the reaction of 3,6,7-trihydroxycoumarin (**1**) in the presence of NaOH. The structure and reaction paths for these radicals are presented in Figure V.18.

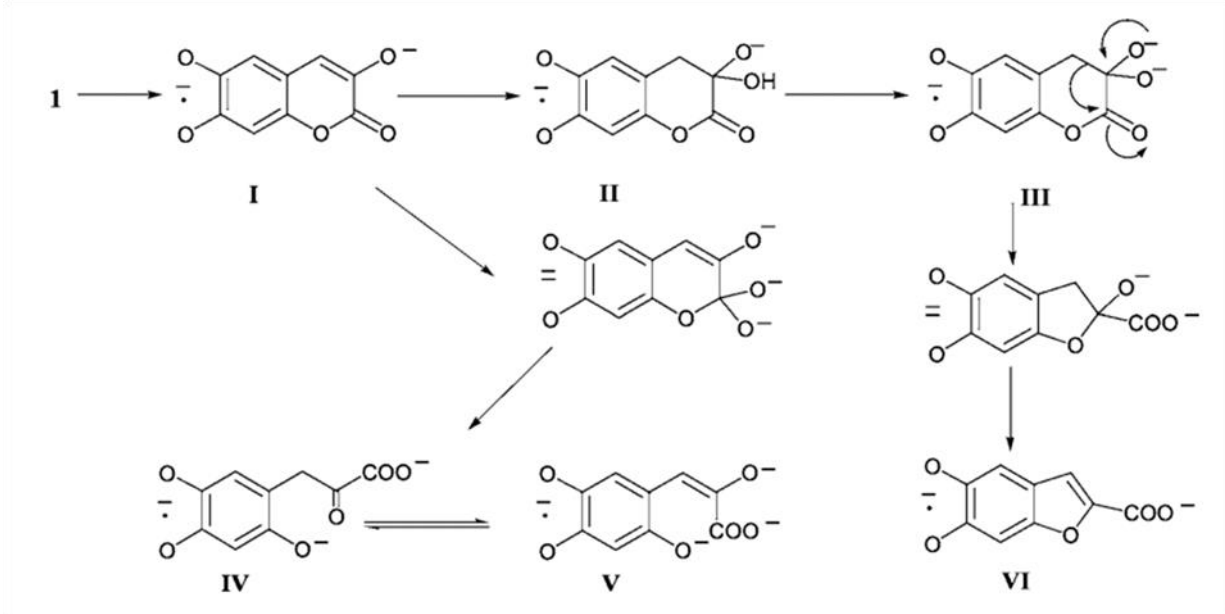


Figure V.18: Scheme representing the radicals formed in the reaction of 3,6,7-trihydroxycoumarin in alkaline solution [257].

Six radicals may be formed by transformation from one to another. The theoretical calculations for the postulated radicals revealed that they should exhibit characteristic patterns in EPR spectra. In addition to the previous literature, a new set of simulations of the radical spectra was performed using WinSim software considering the hyperfine constants provided by Cotelle *et al.* [257] (Figure V.19).

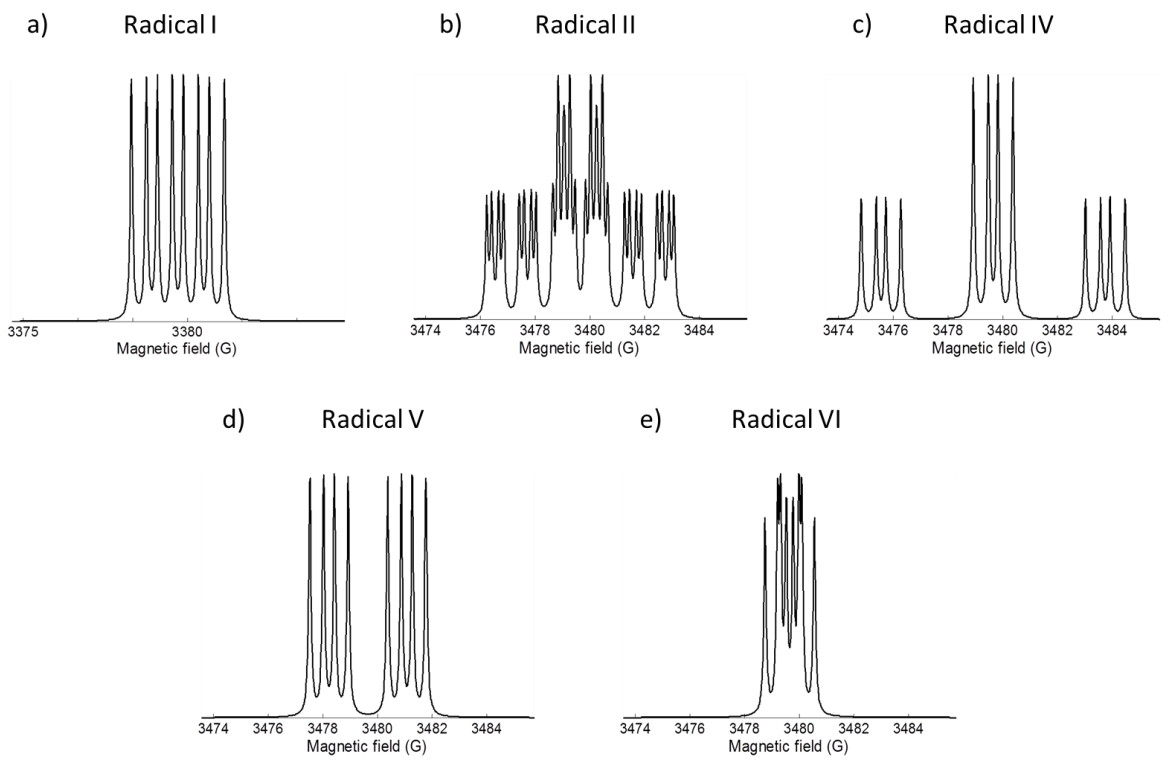


Figure V.19: Simulated EPR spectra of the pure radicals discussed in this chapter (see Figure V.18) in their integrated form.

Radical I is represented by an 8-line EPR spectrum (Figure V.19a) and can only be observed in weakly alkaline solution (example 0.1M of NaOH). It possesses three doublets that are attributed to the coupling between the unpaired electron and the three protons (H4, H5 and H8). Therefore, radical I may not be observed in the presence of more concentrated base due to very rapid transformation into the further forms. After, two possible pathways are postulated. The first one assumes that radical II is formed, however, it rapidly decays and is replaced by radical III that is more stable. The EPR spectra of radicals II and III are very similar so they are difficult to distinguish. Actually, in highly basic medium (1 or 2M of NaOH) the spectra of these two radicals are nearly identical. They possess characteristic EPR pattern with 32-line spectrum due to the interaction between the unpaired electron and five protons, however, only four quartets and two pentets can be shown due to the overlap of the bands (see Figure V.19b). Radical III can be further transformed into radical VI which exhibits an 8-line spectrum presented in Figure V.19e. The parallel reaction path includes the transition from Radical I to radicals IV and V which are in equilibrium. They have different EPR signals presented in Figure V.19c and Figure V.19d. The former possesses triplet of quartet due to the coupling between the unpaired electron and four protons, among which two protons are equivalent. The latter has an 8-line EPR spectrum due to the interaction between

the unpaired electron and three protons. In the context of this established reactions scheme, the presence of signal from radicals III, IV, V and VI can be observed in the studied dataset.

Single matrix D_1,

Comparing the spectral profile extracted by MCR-ALS (Figure V.12) and the theoretical spectra obtained from the article (Figure V.19) the assignment of the radicals is possible. According to expectations, the spectrum of radical I (Figure V.19a) does not appear at all in the data set due to its very rapid transformation (faster than detection time) in strongly alkaline solution. Since radical II and III possess nearly identical spectrum they should appear as one non distinguishable component. Indeed, the first extracted EPR spectrum (blue one in Figure V.12) reveals characteristic four quartets and two pentets patterns and closely corresponds to the simulated spectrum of radical II and/or III presented in Figure V.19b. The second spectrum (in green) corresponds perfectly to the spectrum of radical VI in its middle part. However, the presence of the small peaks on its side indicates that this component is not completely pure and perfectly extracted. The third extracted spectrum (red one in Figure V.12) also seems to be a mixture of radicals. The central part of the spectrum resembles radical VI but the characteristic quartets typical for radical IV are clearly visible on the wings of the spectrum. It indicates that these radicals co-exist which corresponds to their parallel formation.

In spite of the fact that the second and third extracted spectra are not completely pure, the extracted concentration profiles can provide valuable information about kinetics of the reactions. Looking at the time profile $\mathbf{C_1}$, (blue line Figure V.12) that corresponds to the pure spectrum of radical II (or III), it can be noticed that it is formed in the first hour (increase) and afterwards disappears. After 5 hours, the remaining concentration of this radical decreases further, very slowly. The time profile of the second radical (in green in Figure V.12), which is predominantly assigned to Radical VI, is clearly related to the first component. Namely, a decay of radical II (or III) corresponds to the growth of Radical VI. It confirms the postulated reaction pathway presented in Figure V.18. Afterwards, a decrease of this component is noticed with the rate constant slightly lower than in case of Radical II. The third concentration profile (in red in Figure V.12) starts with a zero concentration, it slowly increases within 10 hours and afterwards stays nearly constant. Such kinetics behavior can partially confirm postulated reaction path (Figure V.18) for radical IV which is formed parallel to radical II. It is clearly visible that the second process is slower. It means that the formation of Radical IV

is less competitive than Radical II what explains the small concentration of this component. It should be, however, emphasized here that the third component extracted from MCR-ALS is not pure and it seems to be a mixture of radical IV and VI. Thus, kinetics curve of radical IV is not pure and it is affected by kinetics of radical VI.

Augmented Matrix

The results for the augmented matrix can be interpreted in a very similar way to that of the single matrix. The first spectrum corresponds very well to initially formed pure radical II. The second spectrum (in green) in Figure V.17 can be undoubtedly attributed to the EPR spectrum of radical VI. It is worth noting that it is a purer spectrum than the previously extracted one in single matrix analysis. The admixture of other radical signals on the wings is much less pronounced. It confirms the advantage of multiset analysis. The third spectrum (in red) is not pure and it is a mixture of radical VI (central part) and radical IV recognized by characteristic four peaks at the left and the right of the central part. The fourth spectrum (in yellow) cannot be easily assigned and it is probably a mixture of several radicals.

The concentration profiles of radical II (blue) and radical VI (green) in $\mathbf{D_1}'$ and $\mathbf{D_2}'$ are similar to these ones observed in the single matrix analysis. In both cases, a decay of radical II corresponds to the growth of radical VI which confirms a postulated reaction path (Figure V.18). However in $\mathbf{D_2}'$, a decay of radical II is clearly faster than in $\mathbf{D_1}'$. This is logical because data set $\mathbf{D_2}'$ represents an experiment with more alkaline conditions (2M NaOH instead of 1M NaOH). It is worth adding that the kinetics profile for radical VI is more reliable than in the single matrix analysis due to a purer spectral character of the extracted component. The concentration profile of the third component, which consists of strong admixture of radical IV and VI, slowly increases and reaches a quasi constant value. Such initial slow growth and low intensity of this profile confirm that the formation of radical IV is probably less competitive process than the formation of radical II. In the case of $\mathbf{D_1}'$, the concentration profile of the fourth component (in yellow) is very similar to the second component (green line) and in $\mathbf{D_2}'$, it is nearly identical to the third contribution (red line). Thus, it seems that the presence of the fourth component, which was tentatively deduced after PCA analysis of augmented matrix, can be questionable and the fourth component can be the result of small a misalignment between matrices $\mathbf{D_1}'$ and $\mathbf{D_2}'$.

Nevertheless, the results obtained by the augmented matrix analysis seem to be the most convincing. The spectral signatures of two pure radicals were fully extracted (radicals II and VI) without important admixture of other species. Moreover, the concentration profiles are informative and confirm the hypothetical reaction paths. However, rotational ambiguity still exists and the obtained results are not yet perfect.

V.7 Conclusion

The results presented in this chapter demonstrate the applicability of MCR-ALS in the analysis of multicomponent time evolving EPR data. The presented results are the first MCR-ALS application on this type of EPR data sets. The analysis confirmed that a complicated system of interacting paramagnetic species can be resolved by multivariate analysis giving valuable information about the reaction pathways. In addition, it was demonstrated that the application of simultaneous analysis of multiple data matrices can provide advantages over the resolution of individual matrices. In particular, it was shown that EPR spectra resolved from multiset analysis are purer than their single matrix counterparts. It was also demonstrated that EFA is a very efficient method for the initial estimation of ordered EPR data sets. Furthermore, results obtained by MCR-ALS (multiset analysis) seem appropriate but not yet optimal due to the presence of rotational ambiguity that still exists. In order to fully resolve the pure components from this challenging system, richer experimental data, such as more matrices, are required. An experiment with varied acquisition speed is needed for further exploration of the reaction, especially in its initial part, since most of the reactions occur in this part.

Conclusions and Perspectives

In the presented thesis, chemometric techniques have been used to analyze various experimental data sets obtained by EPR spectroscopy. In particular, for the first time MCR-ALS has been applied on EPR spectra to verify the usefulness of this approach for such type of data. The complex spectra coming from paramagnetic ions mixtures with severely overlapped multipeak signals were analyzed to extract the pure concentration profiles and the spectral signatures of each component free of the contribution of the others. The results obtained by MCR-ALS are very satisfactory due to the agreement between the theoretical and the experimental data. It confirms that MCR-ALS can be used to analyze strongly entangled paramagnetic mixtures even without any prior knowledge about the system. However, it was noticed that a particular care has to be taken with regard to the accuracy and the reproducibility of the experimental data. Spectral shifts between the corresponding peaks in the spectra can cause a loss of the bilinearity in the data set leading to an overestimation of the number of the components. It was also demonstrated that the application of suitable constraints play a key role in the amelioration of the results and the decrease of the rotational ambiguity. In the case of three-component system, where one of the spectrum was much broader than the others causing very strong overlap, the analysis of the original data set (first derivative) did not provide satisfactory results but the integration of the experimental spectra allowed using an additional constraint (non-negativity on the spectral profile) what resulted in a successful resolution.

In addition to 2D data sets, the first multivariate approach in the EPR hyperspectral imaging data demonstrates the feasibility of MCR-ALS for the 3D data sets. The pure distribution maps and the pure spectra of each constituent on the CaF_2 surface that was irradiated by a laser were retrieved without prior knowledge about the chemical system. The spatial distribution maps present a precise location of different light induced paramagnetic centers on the studied surface. On the other hand, from the resolved spectra, the identification of each created paramagnetic species can be achieved. The obtained results open a perspective of new trends for the data analysis in EPR imaging.

In the last part of this work, the application of MCR-ALS on the time evolving EPR spectra illustrates the usefulness of this chemometric method on the sequence ordered data

sets especially in the kinetics studies. The reaction of the antioxidant, 3-hydroxycoumarin, with the base NaOH (1M) and (2M), were studied with the individual data sets analyses as well as the multiset simultaneous analysis of the data from the independent experiments. In the case of the multiple data analysis, the spectral profiles of the radicals formed during the reaction and the corresponding time-evolving concentration profiles (kinetic curves) were retrieved with lower ambiguity than the results obtained by the individual data set resolution.

In general conclusion, it was evidenced that MCR-ALS is a very flexible technique able to adapt to different kind of scenarios and types of data (two as well as three dimensional data). A future promising pathway in the resolution of EPR data is now opened and a perspective for further challenging investigations in this domain are expected. The work presented in this thesis is mainly focused on the qualitative aspects of the potential of the resolution methods. Nevertheless, chemometrics is a field vaster than what is presented. For instance, classification and multivariate regression methods are suitable tools for the quantitative characterization of the complex mediums. It is expected that the application of these methods in EPR spectroscopy will open a new analytical dimension for EPR specialists. One of the very promising subjects is the application of chemometrics on the Micro-ESR data. In fact, Micro-ESR spectrometers are modern devices with reduced size, cost and complexity compared to the classical EPR spectrometer. They are suitable to measure short-lived radicals and spin-trapping in biomedical field to measure oxidative stress. Industrial process applications of the on-line Micro-ESR sensor for the real time monitoring of free radicals, lubricant conditions and crude oil analysis can be done as well. In addition, the micro spectrometer is privileged for *in situ* measurements of the radiation dosimetry and the irradiated food, plastics and carbon coal. Thus, providing fast and effective analytical methods for the exploration of such data sets and the identification of different constituents should open a perspective for modern, cheap and efficient diagnostic tools.

References

[1] A. Abragam, B. Bleaney, *electron paramagnetic resonance of transition ions*, Oxford, **1970**.
[2] S. Wold, Spline Functions, a New Tool in Data-Analysis, Kemisk Tidskrift 84 (1972) 34-&.
[3] K. Pearson, On lines and planes of closest fit to systems of points in space, Philosophical Magazine 2 (1901) 559-572.
[4] R. Tauler, Multivariate curve resolution applied to second order data, Chemometrics and Intelligent Laboratory Systems 30 (1995) 133-146.
[5] J.A. Weil, J.R. Bolton, *Electron Paramagnetic Resonance: Elementary Theory and Practical Apllications, 2nd Edition*, John Wiley and Co., New York, **2007**.
[6] A. Lund, M. Shiotani, S. Shimada, *Principles and Applications of ESR Spectroscopy*, In, Springer, New York, **2011**.
[7] W. Gerlach, O. Stern, The experimental proof of magnetic moment of silver atoms, Zeitschrift Fur Physik 8 (1921) 110-111.
[8] G.E. Uhlenbeck, S. Goudsmit, Spinning electrons and the structure of spectra, Nature 117 (1926) 264-265.
[9] I.I. Rabi, J.R. Zacharias, S. Millman, P. Kusch, A new method of measuring nuclear magnetic moment, Physical Review 53 (1938) 318-318.
[10] G. Breit, I.I. Rabi, Measurement of nuclear spin, Physical Review 38 (1931) 2082-2083.
[11] E. Zavoisky, Paramagnetic Relaxation of Liquid Solutions for Perpendicular Fields, Journal of Physics USSR 9 (1945) 211-216.
[12] E. Zavoisky, Spin-Magnetic Resonance in Paramagnetics, Journal of Physics USSR 9 (1945) 245.
[13] F. Bloch, W.W. Hansen, M. Packard, nuclear induction, Physical Review 69 (1946) 127-127.
[14] E.M. Purcell, H.C. Torrey, R.V. Pound, resonance absorption by nuclear magnetic moments in a solid, Physical Review 69 (1946) 37-38.
[15] R.L. Cummerow, D. Halliday, paramagnetic losses in 2 manganous salts, Physical Review 70 (1946) 433-433.
[16] D.M.S. Bagguley, J.H.E. Griffiths, paramagnetic resonance and magnetic energy levels in chrome alum, Nature 160 (1947) 532-533.
[17] R.J. Blume, electron spin relaxation times in sodium-ammonia solutions, Physical Review 109 (1958) 1867-1873.
[18] K. Holczer, D. Schmalbein, ESP 380: a high power, general purpose pulsed EPR spectrometer, BRUKER Report 1 (1987) 22.
[19] P. Höfer, G.G. Maresch, D. Schmalbein, H. K., BRUKER Report 142 (1996) 15.
[20] T. Yoshimura, H. Yokoyama, S. Fujii, F. Takayama, K. Oikawa, H. Kamada, In vivo EPR detection and imaging of endogenous nitric oxide in lipopolysaccharide-treated mice, Nat Biotech 14 (1996) 992-994.
[21] R. Pievo, B. Angerstein, A.J. Fielding, C. Koch, I. Feussner, M. Bennati, A Rapid Freeze-Quench Setup for Multi-Frequency EPR Spectroscopy of Enzymatic Reactions, Chemphyschem 14 (2013) 4094-4101.
[22] D. Gourier, A. Bessiere, S.K. Sharma, L. Binet, B. Viana, N. Basavaraju, K.R. Priolkar, Origin of the visible light induced persistent luminescence of Cr3+ -doped zinc gallate, Journal of Physics and Chemistry of Solids 75 (2014) 826-837.

[23] K. Abbas, M. Hardy, F. Poulhes, H. Karoui, P. Tordo, O. Ouari, F. Peyrot, Detection of superoxide production in stimulated and unstimulated living cells using new cyclic nitrone spin traps, Free Radical Biology and Medicine 71 (2014) 281-290.
[24] H.M. Swartz, B.B. Williams, B.I. Zaki, A.C. Hartford, L.A. Jarvis, E.Y. Chen, R.J. Comi, M.S. Ernstoff, H.G. Hou, N. Khan, S.G. Swarts, A.B. Flood, P. Kuppusamy, Clinical EPR: Unique Opportunities and Some Challenges, Academic Radiology 21 (2014) 197-206.
[25] Y. Hernandez, J.G. Carriazo, O. Almanza, Characterization by XRD and electron paramagnetic resonance (EPR) of waster materials from "Cerro Matoso" Mine (Colombia), Materials Characterization 57 (2006) 44-49.
[26] W. Nolting, A. Ramakanth, *Quantum Theory of Magnetism*, Springer, Heidelberg, Germany, **2009**.
[27] M. Gomberg, An instance of trivalent carbon triphenylmethyl, Journal of the American Chemical Society 22 (1900) 757-771.
[28] J.C. Walton, *Analysis of Radicals by EPR in Encyclopedia of Radicals in Chemistry, Biology and Materials.*, New York, **2012**.
[29] W. Kaim, B. Schwederski, H. Fischer (Ed.), *Magnetic properties of free radicals*, Berlin, Germany, **2006**.
[30] R.W. Fessenden, R.H. Schuler, EPR Observation of Steady-State Ethyl Radical Concentration in Radiolysis of Liquid Ethane, Journal of Chemical Physics 33 (1960) 935-936.
[31] S.P. McGlynn, T. Azumi, M. Kinoshita, *Molecular spectroscopy of the triplet state*, Prentice-Hall, Englewood Cliffs, N.J., **1969**.
[32] T.D. Smith, J.R. Pilbrow, Determination of Structural-Properties of Dimeric Transition-Metal Ion Complexes from EPR-Spectra, Coordination Chemistry Reviews 13 (1974) 173-278.
[33] K. Dyrek, M. Che, EPR as a tool to investigate the transition metal chemistry on oxide surfaces, Chemical Reviews 97 (1997) 305-331.
[34] C. Calle, A. Sreekanth, M.V. Fedin, J. Forrer, I. Garcia-Rubio, I.A. Gromov, D. Hinderberger, B. Kasumaj, P. Leger, B. Mancosu, G. Mitrikas, M.G. Santangelo, S. Stoll, A. Schweiger, R. Tschaggelar, J. Harmer, Pulse EPR methods for studying chemical and biological samples containing transition metals, Helvetica Chimica Acta 89 (2006) 2495-2521.
[35] S.A. Kazanskii, A.I. Ryskin, A.E. Nikiforov, A.Y. Zaharov, M.Y. Ougrumov, G.S. Shakurov, EPR spectra and crystal field of hexamer rare-earth clusters in fluorites, Physical Review B 72 (2005).
[36] A. Abragam, M.H.L. Pryce, theory of the nuclear hyperfine structure of paramagnetic resonance spectra in crystals, Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences 205 (1951) 135-153.
[37] B. Odom, D. Hanneke, B. D'Urso, G. Gabrielse, New measurement of the electron magnetic moment using a one-electron quantum cyclotron, Physical Review Letters 97 (2006).
[38] W.R. Hagen, *Biomolecular EPR spectroscopy*, CRC Press, Boca Raton, **2009**.
[39] R. Weber, J. Jiang, D.P. Barr, in, Bruker Instruments, Inc., Billerica, USA, 1998.
[40] G.R. Eaton, S.S. Eaton, D.P. Barr, R.T. Weber, *Quantitative EPR*, Springer, New York, **2010**.
[41] G.R. Eaton, S.S. Eaton, K. Ohno, C. press (Ed.), *EPR imaging and in vivo EPR*, Boca Raton, Florida, **1991**.
[42] P.C. Lauterbur, image formation by induced local interactions - examples employing nuclear magnetic-resonance, Nature 242 (1973) 190-191.

[43] M.J.R. Hoch, A.R. Day, imaging of paramagnetic centers in diamond, Solid State Communications 30 (1979) 211-213.
[44] S.S. Eaton, G.R. Eaton, EPR imaging, Journal of Magnetic Resonance (1969) 59 (1984) 474-477.
[45] K. Ohno, ESR Imaging and Its Applications, Applied Spectroscopy Reviews 22 (1986) 1-56.
[46] S. Schlick, K. Kruczala, M.V. Motyakin, J.L. Gerlock, Spectral profiling of radicals in polymer degradation based on electron spin resonance imaging (ESRI), Polymer Degradation and Stability 73 (2001) 471-475.
[47] R. Krzyminiewski, T. Kubiak, B. Dobosz, G. Schroeder, J. Kurczewska, EPR spectroscopy and imaging of TEMPO-labeled magnetite nanoparticles, Current Applied Physics 14 (2014) 798-804.
[48] G.A. Watt, M.E. Newton, J.M. Baker, EPR and optical imaging of the growth-sector dependence of radiation-damage defect production in synthetic diamond, Diamond and Related Materials 10 (2001) 1681-1683.
[49] G. Redler, E.D. Barth, K.S. Bauer, J.P.Y. Kao, G.M. Rosen, H.J. Halpern, In Vivo Electron Paramagnetic Resonance Imaging of Differential Tumor Targeting Using cis-3,4-Di(Acetoxymethoxycarbonyl)-2,2,5,5Tetramethyl-1-Pyrrolidinyloxyl, Magnetic Resonance in Medicine 71 (2014) 1650-1656.
[50] P. Leveque, Q. Godechal, A. Bol, F. Trompier, B. Gallez, X-band EPR imaging as a tool for gradient dose reconstruction in irradiated bones, Medical Physics 36 (2009) 4223-4229.
[51] X.L. Wang, M. Emoto, A. Sugimoto, Y. Miyake, K. Itto, M. Amasaka, S. Xu, H. Hirata, H. Fujii, H. Arimoto, Synthesis of N-15-labeled 4-oxo-2,2,6,6-tetraethylpiperidine nitroxide for EPR brain imaging, Tetrahedron Letters 55 (2014) 2146-2149.
[52] P. Kuppusamy, M. Chzhan, K. Vij, M. Shteynbuk, D.J. Lefer, E. Giannella, J.L. Zweier, Three-dimensional spectral-spatial EPR imaging of free radicals in the heart: a technique for imaging tissue metabolism and oxygenation, Proceedings of the National Academy of Sciences of the United States of America 91 (1994) 3388-3392.
[53] P. Kuppusamy, J.L. Zweier, Cardiac applications of EPR imaging, NMR in Biomedicine 17 (2004) 226-239.
[54] T. Czechowski, W. Chlewicki, M. Baranowski, K. Jurga, P. Szczepanik, P. Szulc, P. Kedzia, M. Szostak, P. Malinowski, S. Wosinski, W. Prukala, J. Jurga, Two-dimensional spectral–spatial EPR imaging with the rapid scan and modulated magnetic field gradient, Journal of Magnetic Resonance 243 (2014) 1-7.
[55] M.M. Maltempo, Differentiation of spectral and spatial components in EPR imaging using 2-D image reconstruction algorithms, Journal of Magnetic Resonance 69 (1986) 156-161.
[56] G.R. Eaton, S.S. Eaton, Introduction to EPR imaging using magnetic-field gradients, Concepts in Magnetic Resonance 7 (1995) 49-67.
[57] M.M. Maltempo, S.S. Eaton, G.R. Eaton, Spectral-spatial two-dimensional EPR imaging, Journal of Magnetic Resonance (1969) 72 (1987) 449-455.
[58] B.B. Williams, X. Pan, H.J. Halpern, EPR imaging: The relationship between CW spectra acquired from an extended sample subjected to fixed stepped gradients and the Radon transform of the resonance density, Journal of Magnetic Resonance 174 (2005) 88-96.
[59] P. Kuppusamy, J. Zweier, Hyperfine artifacts in electron paramagnetic resonance imaging, Res. Chem. Intermed. 22 (1996) 593-604.
[60] P. Kuppusamy, J.L. Zweier, A forward-subtraction procedure for removing hyperfine artifacts in electron paramagnetic resonance imaging, Magnetic Resonance in Medicine 35 (1996) 316-322.

[61] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, *Chemometrics: A textbook*, Elsevier, Amsterdam, **1988**.
[62] P. Oliveri, M. Forina, *Data Analysis and Chemometrics*, In: Chemical Analysis of Food Techniques and Application, Y. Pico (Ed.), Elsevier, Oxford, U.K., **2012**, pp. 25-57.
[63] Student, The probable error of a mean, Biometrika 6 (1908) 1-25.
[64] B.K. Lavine, A.C.S.D.o.C.i. Chemistry, A.C.S. Meeting, *Chemometrics and chemoinformatics*, American Chemical Society, **2005**.
[65] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.d. Jong, P.J. Lewi, J. Smeyers-Verbeke, *Hanbook of Chemometrics and Qualimetrics: Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, Amsterdam, **1997**.
[66] P. Paatero, U. Tapper, Positive Matrix Factorization - A Nonnegative Factor Model with Optimal Utilization of Error-Estimates of Data Values, Environmetrics 5 (1994) 111-126.
[67] P. Comon, Independent Component Analysis, a New Concept, Signal Processing 36 (1994) 287-314.
[68] G.H. Dunteman, *Principal Components Analysis*, SAGE, Newbury Park, California, **1989**.
[69] S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58 (2001) 109-130.
[70] I.T. Jolliffe, A Note on the Use of Principal Components in Regression, Applied Statistics-Journal of the Royal Statistical Society Series C 31 (1982) 300-303.
[71] W. Svante, S. Michael, *SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy*, In: Chemometrics: Theory and Application, American Chemical Society, **1977**, pp. 243-282.
[72] R.A. Fisher, The use of multiple measurements in taxonomic problems, Annals of Eugenics 7 (1936) 179-188.
[73] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics*, Elsevier, Amsterdam, **1998**.
[74] T. Kohonen, *Self-Organizing Maps*, Third Edition Springer, New York, **2001**.
[75] J. Zupan, Introduction to artificial neural network (ANN) methods: what they are and how to use them, Acta Chimica Slovenica 41 (1994) 327-327.
[76] R.A. Fisher, W.A. MacKenzie, Studies in crop variation II The manurial response of different potato varieties, Journal of Agricultural Science 13 (1923) 311-320.
[77] H. Wold, *nonlinear estimation by iterative least squares procedures*, Research Papers in Statistics, Wiley, New York, **1966**.
[78] H. Hotelling, Analysis of a complex of statistical variables into principal components, Journal of Educational Psychology 24 (1933) 417-441 and 498-520.
[79] G.D. Franco, A. Marradi, *Factor Analysis and Principal Factor Analysis*, FrancoAngeli, Milan, Italy, **2013**.
[80] E.R. Malinowski, D.G. Howery, *Factor Analysis in Chemistry*, Wiley, New York, **1980**.
[81] L.S. Ramos, K.R. Beebe, W.P. Carey, E. Sanchez, B.C. Erickson, B.E. Wilson, L.E. Wangen, B.R. Kowalski, CHEMOMETRICS, Analytical Chemistry 58 (1986) R294-R315.
[82] K.G. Jöreskog, J.E. Klovan, R.A. Reyment, *Geological Factor Analysis*, Elsevier, Amsterdam, **1976**.
[83] J. Davis, *Statistics and Data Analysis in Geology*, Wiley, New York, **1973 and 1986**.
[84] G.H. Golub, C.F. VanLoan, *Matrix Computations*, The Johns Hopkins University Press, Oxford, **1983**.
[85] J. Mandel, use of the singular value decomposition in regression-analysis, American Statistician 36 (1982) 15-24.
[86] M. Loéve, *Fonctions aleatoires de seconde ordre*, In: Processus Stochastiques et Mouvement Brownien, P. Levy (Ed.), Hermann, Paris, **1948**.

[87] I.T. Joliffe, *Principal Component Analysis*, Springer, Berlin, **1986**.
[88] S. Wold, *soft-modeling by latent variables: the non-linear iterative partial least squares (nipals) algorithm.*, In: Perspectives in Probability and Statistics., J. Gani (Ed.), Academic Press, London, **1975**.
[89] G.H. Golub, C. Reinsch, Singular Value Decomposition and Least Squares Solutions, Numerische Mathematik 14 (1970) 403-420.
[90] A. Hyvarinen, E. Oja, Independent component analysis: algorithms and applications, Neural Networks 13 (2000) 411-430.
[91] Patel, II, J. Trevisan, G. Evans, V. Llabjani, P.L. Martin-Hirsch, H.F. Stringfellow, F.L. Martin, High contrast images of uterine tissue derived using Raman microspectroscopy with the empty modelling approach of multivariate curve resolution-alternating least squares, Analyst 136 (2011) 4950-4959.
[92] J.-H. Jiang, Y. Ozaki, Self-Modeling Curve Resolution (SMCR): Principles, Techniques, and Applications, Applied Spectroscopy Reviews 37 (2002) 321-345.
[93] A. De Juan, E. Casassas, R. Tauler, *Soft modeling of analytical data*, In: Encyclopedia of analytical chemistry: instrumentation and applications, **2000**, pp. 9800-9837.
[94] P.J. Gemperline, A Priori Estimates of the Elution Profiles of the Pure Components in Overlapped Liquid-Chromatography Peaks Using Target Factor-Analysis, Journal of Chemical Information and Computer Sciences 24 (1984) 206-212.
[95] B.G.M. Vandeginste, W. Derks, G. Kateman, Multicomponent Self-Modeling Curve Resolution in High-Performance Liquid-Chromatography by Iterative Target Transformation Analysis, Analytica Chimica Acta 173 (1985) 253-264.
[96] C. Mason, M. Maeder, A. Whitson, Resolving factor analysis, Analytical chemistry 73 (2001) 1587-1594.
[97] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, a graphical user-friendly interface for mcr-als: a new tool for multivariate curve resolution in matlab, Chemometrics and Intelligent Laboratory Systems 76 (2005) 101-110.
[98] A. de Juan, R. Tauler, Chemometrics applied to unravel multicomponent processes and mixtures - Revisiting latest trends in multivariate resolution, Analytica Chimica Acta 500 (2003) 195-210.
[99] R. Tauler, B. Kowalski, S. Fleming, Multivariate Curve Resolution Applied to Spectral Data from Multiple Runs of an Industrial-Process, Analytical Chemistry 65 (1993) 2040-2047.
[100] M. Ando, H.O. Hamaguchi, Molecular component distribution imaging of living cells by multivariate curve resolution analysis of space-resolved Raman spectra, Journal of Biomedical Optics 19 (2014).
[101] S. Piqueras, L. Duponchel, R. Tauler, A. de Juan, resolution and segmentation of hyperspectral biomedical images by multivariate curve resolution-alternating least squares, Analytica Chimica Acta 705 (2011) 182-192.
[102] S. Piqueras, L. Duponchel, R. Tauler, A. de Juan, Monitoring polymorphic transformations by using in situ Raman hyperspectral imaging and image multiset analysis, Analytica Chimica Acta 819 (2014) 15-25.
[103] S. Piqueras, J. Burger, R. Tauler, A. de Juan, Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution, Chemometrics and Intelligent Laboratory Systems 117 (2012) 169-182.
[104] M. Garrido, M.S. Larrechi, F.X. Rius, R. Tauler, Calculation of band boundaries of feasible solutions obtained by Multivariate Curve Resolution Alternating Least Squares of multiple runs of a reaction monitored by NIR spectroscopy, Chemometrics and Intelligent Laboratory Systems 76 (2005) 111-120.

[105] G.L. Alexandrino, R.J. Poppi, NIR imaging spectroscopy for quantification of constituents in polymers thin films loaded with paracetamol, Analytica Chimica Acta 765 (2013) 37-44.
[106] S. Piqueras, L. Duponchel, M. Offroy, F. Jamme, R. Tauler, A. de Juan, Chemometric Strategies To Unmix Information and Increase the Spatial Description of Hyperspectral Images: A Single-Cell Case Study, Analytical Chemistry 85 (2013) 6303-6311.
[107] M. Garrido, F.X. Rius, M.S. Larrechi, Multivariate curve resolution-alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes, Analytical and Bioanalytical Chemistry 390 (2008) 2059-2066.
[108] E. Widjaja, G.H. Lim, P.S. Chow, S. Tan, Multivariate data analysis as a tool to investigate the reaction kinetics of intramolecular cyclization of enalapril maleate studied by isothermal and non-isothermal FT-IR microscopy, European Journal of Pharmaceutical Sciences 32 (2007) 349-356.
[109] S.A. Bortolato, K.E. Thomas, K. McDonough, R.W. Gurney, D.M. Martino, Evaluation of photo-induced crosslinking of thymine polymers using FT-IR spectroscopy and chemometric analysis, Polymer 53 (2012) 5285-5294.
[110] Y. Wang, R.R. Zhu, Y.N. Ni, S. Kokot, Competitive interactions of anti-carcinogens with serum albumin: A spectroscopic study of bendamustine and dexamethasone with the aid of chemometrics, Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy 123 (2014) 241-248.
[111] Y. Wang, Y.N. Ni, Combination of UV-vis spectroscopy and chemometrics to understand protein-nanomaterial conjugate: A case study on human serum albumin and gold nanoparticles, Talanta 119 (2014) 320-330.
[112] A. Jayaraman, S. Mas, R. Tauler, A. de Juan, Study of the photodegradation of 2-bromophenol under UV and sunlight by spectroscopic, chromatographic and chemometric techniques, Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 910 (2012) 138-148.
[113] C. Ruckebusch, L. Duponchel, J.P. Huvenne, A. Caudron, L. Boilet, J.P. Cornard, J.C. Merlin, A. de Juan, Chemometric strategies for the study of the complexation of Al(111) ions with model molecule of humic substances from UV-vis data sets, Analytica Chimica Acta 544 (2005) 337-344.
[114] G. Vivo-Truyols, M. Ziari, P. Magusin, P.J. Schoenmakers, Effect of initial estimates and constraints selection in multivariate curve resolution-Alternating least squares. Application to low-resolution NMR data, Analytica Chimica Acta 641 (2009) 37-45.
[115] T.K. Karakach, R. Knight, E.M. Lenz, M.R. Viant, J.A. Walter, Analysis of time course H-1 NMR metabolomics data by multivariate curve resolution, Magnetic Resonance in Chemistry 47 (2009) S105-S117.
[116] T. Salomonsen, H.M. Jensen, F.H. Larsen, S. Steuernagel, S.B. Engelsen, Direct quantification of M/G ratio from C-13 CP-MAS NMR spectra of alginate powders by multivariate curve resolution, Carbohydrate Research 344 (2009) 2014-2022.
[117] J. Jaumot, V. Marchan, R. Gargallo, A. Grandas, R. Tauler, Multivariate curve resolution applied to the analysis and resolution of two-dimensional H-1,N-15 NMR reaction spectra, Analytical Chemistry 76 (2004) 7094-7101.
[118] V. Boeris, J.A. Arancibia, A.C. Olivieri, Determination of five pesticides in juice, fruit and vegetable samples by means of liquid chromatography combined with multivariate curve resolution, Analytica Chimica Acta 814 (2014) 23-30.
[119] K. Chen, F. Lynen, L. Hitzel, M. Hanna-Brown, R. Szucs, P. Sandra, A New Strategy for Fast Chiral Screening by Combining HPLC-DAD with a Multivariate Curve Resolution-Alternating Least Squares Algorithm, Chromatographia 76 (2013) 1055-1066.

[120] M.J. Culzoni, A.M. de Llanos, M.M. De Zan, A. Espinosa-Mansilla, F. Canada-Canada, A.M. de la Pena, H.C. Goicoechea, Enhanced MCR-ALS modeling of HPLC with fast scan fluorimetric detection second-order data for quantitation of metabolic disorder marker pteridines in urine, Talanta 85 (2011) 2368-2374.
[121] J. Omar, M. Olivares, J.M. Amigo, N. Etxebarria, Resolution of co-eluting compounds of Cannabis Sativa in comprehensive two-dimensional gas chromatography/mass spectrometry detection with Multivariate Curve Resolution-Alternating Least Squares, Talanta 121 (2014) 273-280.
[122] H. Parastar, R. Tauler, Multivariate Curve Resolution of Hyphenated and Multidimensional Chromatographic Measurements: A New Insight to Address Current Chromatographic Challenges, Analytical Chemistry 86 (2014) 286-297.
[123] Z.D. Zeng, J. Li, H.M. Hugel, G.W. Xu, P.J. Marriott, Interpretation of comprehensive two-dimensional gas chromatography data using advanced chemometrics, Trac-Trends in Analytical Chemistry 53 (2014) 150-166.
[124] G.M.S. Alves, J. Magalhaes, R. Tauler, H. Soares, Simultaneous Anodic Stripping Voltammetric Determination of Pb and Cd, Using a Vibrating Gold Microwire Electrode, Assisted by Chemometric Techniques, Electroanalysis 25 (2013) 1895-1906.
[125] P.D. Wentzell, T.K. Karakach, S. Roy, M.J. Martinez, C.P. Allen, M. Werner-Washburne, Multivariate curve resolution of time course microarray data, BMC Bioinformatics 7 (2006) 1-19.
[126] S. Navea, R. Tauler, A. de Juan, Monitoring and modeling of protein processes using mass spectrometry, circular dichroism, and multivariate curve resolution methods, Analytical Chemistry 78 (2006) 4768-4778.
[127] J.H. Jiang, Y.Z. Liang, Y. Ozaki, Principles and methodologies in self-modeling curve resolution, Chemometrics and Intelligent Laboratory Systems 71 (2004) 1-12.
[128] M. Amrhein, B. Srinivasan, D. Bonvin, M.M. Schumacher, On the rank deficiency and rank augmentation of the spectral measurement matrix, Chemometrics and Intelligent Laboratory Systems 33 (1996) 17-33.
[129] A. IzquierdoRidorsa, J. Saurina, S. HernandezCassou, R. Tauler, Second-order multivariate curve resolution applied to rank-deficient data obtained from acid-base spectrophotometric titrations of mixtures of nucleic bases, Chemometrics and Intelligent Laboratory Systems 38 (1997) 183-196.
[130] J. Saurina, S. Hernandez-Cassou, R. Tauler, A. Izquierdo-Ridorsa, Multivariate resolution of rank-deficient spectrophotometric data from first-order kinetic decomposition reactions, Journal of Chemometrics 12 (1998) 183-203.
[131] X.-Y. Song, S.-Y. Lee, Bayesian estimation and model selection of multivariate linear model with polytomous variables, Multivariate Behavioral Research 37 (2002) 453-477.
[132] H. Akaike, A new look at the statistical model identification, Automatic Control, IEEE Transactions on 19 (1974) 716-723.
[133] H. Akaike, Fitting autoregressive models for prediction, Annals of the institute of Statistical Mathematics 21 (1969) 243-247.
[134] C. Biernacki, G. Celeux, G. Govaert, Assessing a mixture model for clustering with the integrated completed likelihood, Pattern Analysis and Machine Intelligence, IEEE Transactions on 22 (2000) 719-725.
[135] S. Cubaynes, C. Lavergne, E. Marboutin, O. Gimenez, Assessing individual heterogeneity using model selection criteria: how many mixture components in capture–recapture models?, Methods in Ecology and Evolution 3 (2012) 564-573.
[136] G. McLachlan, D. Peel, *Finite mixture models*, John Wiley & Sons, **2004**.
[137] J. Yu-Yan, H. Yi-Jian, Y. Xiu-Cheng, in: Fuzzy Systems and Knowledge Discovery, 2009. FSKD'09. Sixth International Conference on, IEEE, 2009, pp. 567-572.

[138] W. Windig, J. Guilment, interactive self-modeling mixture analysis, Analytical Chemistry 63 (1991) 1425-1432.
[139] W. Windig, D.A. Stephenson, self-modeling mixture analysis of 2nd-derivative near-infrared spectral data using the simplisma approach, Analytical Chemistry 64 (1992) 2735-2742.
[140] H.R. Keller, D.L. Massart, Evolving Factor-Analysis, Chemometrics and Intelligent Laboratory Systems 12 (1992) 209-224.
[141] M. Maeder, A.D. Zuberbuehler, the resolution of overlapping chromatographic peaks by evolving factor-analysis, Analytica Chimica Acta 181 (1986) 287-291.
[142] H. Gampp, M. Maeder, C.J. Meyer, A.D. Zuberbuhler, calculation of equilibrium-constants from multiwavelength spectroscopic data .3. model-free analysis of spectrophotometric and electron-spin-resonance titrations, Talanta 32 (1985) 1133-1139.
[143] M. Maeder, evolving factor-analysis for the resolution of overlapping chromatographic peaks, Analytical Chemistry 59 (1987) 527-530.
[144] H. Gampp, M. Maeder, C.J. Meyer, A.D. Zuberbuhler, calculation of equilibrium-constants from multiwavelength spectroscopic data .4. model-free least-squares refinement by use of evolving factor-analysis, Talanta 33 (1986) 943-951.
[145] H.R. Keller, D.L. Massart, Peak purity control in liquid chromatography with photodiode-array detection by a fixed size moving window evolving factor analysis, Analytica Chimica Acta 246 (1991) 379-390.
[146] A. de Juan, M. Maeder, T. Hancewicz, R. Tauler, local rank analysis for exploratory spectroscopic image analysis. fixed size image window-evolving factor analysis, Chemometrics and Intelligent Laboratory Systems 77 (2005) 64-74.
[147] W.H. Lawton, Sylvestr.Ea, self modeling curve resolution, Technometrics 13 (1971) 617-&.
[148] M. Vosough, C. Mason, R. Tauler, M. Jalali-Heravi, M. Maeder, on rotational ambiguity in model-free analyses of multivariate data, Journal of Chemometrics 20 (2006) 302-310.
[149] J. Jaumot, R. Tauler, MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution, Chemometrics and Intelligent Laboratory Systems 103 (2010) 96-107.
[150] R. Tauler, Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution, Journal of Chemometrics 15 (2001) 627-646.
[151] R. Tauler, M. Maeder, *two-way data analysis: multivariate curve resolution-error in cruve resolution*, In: Comprehensive Chemometrics, S. Brown, R. Tauler, R. Walczak (Eds.), Elsevier, Oxford, **2009**, pp. 345-363.
[152] R. Tauler, A. Smilde, B. Kowalski, Selectivity, Local Rank, 3-Way Data-Analysis and Ambiguity in Multivariate Curve Resolution, Journal of Chemometrics 9 (1995) 31-58.
[153] A. deJuan, Y. VanderHeyden, R. Tauler, D.L. Massart, Assessment of new constraints applied to the alternating least squares method, Analytica Chimica Acta 346 (1997) 307-318.
[154] A. de Juan, R. Tauler, Multivariate curve resolution (MCR) from 2000: Progress in concepts and applications, Critical Reviews in Analytical Chemistry 36 (2006) 163-176.
[155] R. Tauler, M. Maeder, A. De Juan, *Multiset data analysis: extended multivariate curve resolution*, In: Comprehensive Chemometrics S.D. Brown, R. Tauler, B. Walczak (Eds.), Elsevier, Oxford, **2009**, pp. 473-505.
[156] S. Navea, A. De Juan, R. Tauler, three-way data analysis applied to multispectroscopic monitoring of protein folding, Analytica Chimica Acta 446 (2001) 187-197.

[157] C. Ruckebusch, A. De Juan, L. Duponchel, J.P. Huvenne, matrix augmentation for breaking rank-deficiency: a case study, Chemometrics and Intelligent Laboratory Systems 80 (2006) 209-214.
[158] A. Savitzky, M.J.E. Golay, smoothing and differentiation of data by simplified least squares procedures, Analytical Chemistry 36 (1964) 1627-1639.
[159] B. Walczak (Ed.), *Wavelets in Chemistry*, Elsevier, Amsterdam, **2000**.
[160] F. Ehrentreich, L. Summchen, Spike removal and denoising of Raman spectra by wavelet transform methods, Analytical Chemistry 73 (2001) 4364-4373.
[161] P.H.C. Eilers, Parametric time warping, Analytical Chemistry 76 (2004) 404-411.
[162] NASA, National Aeronautics and Space Administration Remote Sensing Tutorial, http://rst.gsfc.nasa.gov, 2006.
[163] D.M. Schotton, *Electronic Light Microscopy: Techniques in Modern Biomedical Microscopy*, Wiley-Liss, New York, **1993**.
[164] H.F. Grahn, P. Geladi, *techniques and applications of hyperspectral image analysis*, Chichester, **2007**.
[165] D.-W. Sun, *Hyperspectral Imaging for Food Quality Analysis and Control*, Elsevier, Oxford, U.K., **2010**.
[166] Y.R. Chen, K.L. Chao, M.S. Kim, Machine vision technology for agricultural applications, Computers and Electronics in Agriculture 36 (2002) 173-191.
[167] P.M. Mehl, Y.R. Chen, M.S. Kim, D.E. Chan, Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations, Journal of Food Engineering 61 (2004) 67-81.
[168] S.M. Schweizer, J.M.F. Moura, Efficient detection in hyperspectral imagery, Ieee Transactions on Image Processing 10 (2001) 584-597.
[169] D.P. Kise, D. Magana, M.J. Reddish, R.B. Dyer, Submillisecond mixing in a continuous-flow, microfluidic mixer utilizing mid-infrared hyperspectral imaging detection, Lab on a Chip 14 (2014) 584-591.
[170] S.S. Chen, F.F. Zhang, J.F. Ning, X. Liu, Z.W. Zhang, S.Q. Yang, Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging, Food Chemistry 172 (2015) 788-793.
[171] C.P. Marshall, A.O. Marshall, Raman Hyperspectral Imaging of Microfossils: Potential Pitfalls, Astrobiology 13 (2013) 920-931.
[172] C.D. Everard, M.S. Kim, H. Lee, A comparison of hyperspectral reflectance and fluorescence imaging techniques for detection of contaminants on spinach leaves, Journal of Food Engineering 143 (2014) 139-145.
[173] L.A. Paluchowski, M. Milanic, A. Bjorgan, B. Grandaunet, A. Dhainaut, M. Hoff, L.L. Randeberg, Identification of inflammation sites in arthritic joints using hyperspectral imaging, Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues Xii 8947 (2014).
[174] J.P. Ryan, C.O. Davis, N.B. Tufillaro, R.M. Kudela, B.C. Gao, Application of the Hyperspectral Imager for the Coastal Ocean to Phytoplankton Ecology Studies in Monterey Bay, CA, USA, Remote Sensing 6 (2014) 1007-1025.
[175] N. Zabcic, B. Rivard, C. Ong, A. Mueller, Using airborne hyperspectral data to characterize the surface pH and mineralogy of pyrite mine tailings, International Journal of Applied Earth Observation and Geoinformation 32 (2014) 152-162.
[176] B.D. Cook, L.A. Corp, R.F. Nelson, E.M. Middleton, D.C. Morton, J.T. McCorkel, J.G. Masek, K.J. Ranson, V. Ly, P.M. Montesano, NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager, Remote Sensing 5 (2013) 4045-4066.
[177] J.M. Amigo, Practical issues of hyperspectral imaging analysis of solid dosage forms, Analytical and Bioanalytical Chemistry 398 (2010) 93-109.

[178] C. Gendrin, Y. Roggo, C. Collet, Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review, Journal of Pharmaceutical and Biomedical Analysis 48 (2008) 533-553.
[179] A. Iqbal, D.W. Sun, P. Allen, An overview on principle, techniques and application of hyperspectral imaging with special reference to ham quality evaluation and control, Food Control 46 (2014) 242-254.
[180] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, G. Downey, J.M. Frias, Hyperspectral imaging - an emerging process analytical tool for food quality and safety control, Trends in Food Science & Technology 18 (2007) 590-598.
[181] P. Geladi, S. Wold, K. Esbensen, image-analysis and chemical information in images, Analytica Chimica Acta 191 (1986) 473-480.
[182] K. Esbensen, P. Geladi, strategy of multivariate image-analysis (mia), Chemometrics and Intelligent Laboratory Systems 7 (1989) 67-86.
[183] P. Geladi, H. Isaksson, L. Lindqvist, S. Wold, K. Esbensen, principal component analysis of multivariate images, Chemometrics and Intelligent Laboratory Systems 5 (1989) 209-220.
[184] P. Geladi, K. Esbensen, regression on multivariate images - principal component regression for modeling, prediction and visual diagnostic-tools, Journal of Chemometrics 5 (1991) 97-111.
[185] K.H. Esbensen, P.L. Geladi, H.F. Grahn, strategies for multivariate image regression, Chemometrics and Intelligent Laboratory Systems 14 (1992) 357-374.
[186] T.T. Lied, P. Geladi, K.H. Esbensen, multivariate image regression (mir): implementation of image plsr-first forays, Journal of Chemometrics 14 (2000) 585-598.
[187] C. Shao-Shan, I.C. Chein, I.W. Ginsberg, in: Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, 2000, pp. 3136-3138 vol.3137.
[188] L. Duponchel, W. Elmi-Rayaleh, C. Ruckebusch, J.P. Huvenne, Multivariate curve resolution methods in imaging spectroscopy: Influence of extraction methods and instrumental perturbations, Journal of Chemical Information and Computer Sciences 43 (2003) 2057-2067.
[189] P. Geladi, H.F. Grahn, *multivariate image analysis*, John Wiley & Sons Ltd., Canada, **1997**.
[190] J.P. Hornak, *Encyclopedia of Imaging Science and Technology*, Wiley, New York, **2002**.
[191] A. de Juan, M. Maeder, T. Hancewicz, R. Tauler, use of local rank-based spatial information for resolution of spectroscopic images, Journal of Chemometrics 22 (2008) 291-298.
[192] L.C. Brunel, High Frequency Tunable Continuous Wave ESR Spectroscopy, Applied Magnetic Resonance 3 (1992) 83-97.
[193] S. Stoll, A. Schweiger, EasySpin, A Comprehensive Software Package for Spectral Simulation and Analysis in EPR, Journal of Magnetic Resonance 178 (2006) 42-55.
[194] D.R. Duling, Simulation of Multiple Isotropic Spin-Trap EPR-Spectra, Journal of Magnetic Resonance Series B 104 (1994) 105-110.
[195] L. Dumas, B. Albela, L. Bonneviot, D. Portinha, E. Fleury, Electron spin resonance quantitative monitoring of five different radicals in gamma-irradiated polyvinylidene fluoride, Radiation Physics and Chemistry 86 (2013) 102-109.
[196] O. Steinbock, B. Neumann, B. Cage, J. Saltiel, S.C. Muller, N.S. Dalal, a demonstration of principal component analysis for epr spectroscopy: identifying pure component spectra from complex spectra, Analytical Chemistry 69 (1997) 3708-3713.

[197] K.R. Beebe, B.R. Kowalski, An Introduction to Multivariate Calibration and Analysis, Analytical Chemistry 59 (1987) 1007A-1017A.
[198] J.P. Barnes, W.A. Bernhard, the distribution of electron trapping in DNA - one-electron-reduced oligodeoxynucleotides of adenine and thymine, Radiation Research 143 (1995) 85-92.
[199] E.H. Novotny, H. Knicker, L. Martin-Neto, R.B.V. Azeredo, M.H.B. Hayes, Effect of residual vanadyl ions on the spectroscopic analysis of humic acids: a multivariate approach, European Journal of Soil Science 59 (2008) 439-444.
[200] N. Nagai, Y. Yamaguchi, R. Saito, S. Hayashi, M. Kudo, IR and EPR study of the Na ion-implanted SiO2/Si system, Applied Spectroscopy 55 (2001) 1207-1213.
[201] P. Moens, P. Devolder, R. Hoogewijs, F. Callens, R. Verbeeck, Maximum-Likelihood Common-Factor Analysis as a Powerful Tool in Decomposing Multicomponent EPR Powder Spectra, Journal of Magnetic Resonance Series A 101 (1993) 1-15.
[202] G. Vanhaelewyn, J. Sadlo, F. Callens, W. Mondelaers, D. De Frenne, P. Matthys, A decomposition study of the EPR spectrum of irradiated sucrose, Applied Radiation and Isotopes 52 (2000) 1221-1227.
[203] A.G. Savic, M. Mojovic, Free Radicals Identification from the Complex EPR Signals by Applying Higher Order Statistics, Analytical Chemistry 84 (2012) 3398-3402.
[204] J.Y. Ren, C.Q. Chang, P.C.W. Fung, J.G. Shen, F.H.Y. Chan, Free radical EPR spectroscopy analysis using blind source separation, Journal of Magnetic Resonance 166 (2004) 82-91.
[205] C.Q. Chang, J.Y. Ren, P.C.W. Fung, Y.S. Hung, J.G. Shen, F.H.Y. Chan, Novel sparse component analysis approach to free radical EPR spectra decomposition, Journal of Magnetic Resonance 175 (2005) 242-255.
[206] R. Tauler, E. Casassas, A. Izquierdoridorsa, Self-Modeling Curve Resolution in Studies of Spectrometric Titrations of Multi-Equilibria Systems by Factor-Analysis, Analytica Chimica Acta 248 (1991) 447-458.
[207] P. Tauler, E. Casassas, Application of Principal Component Analysis to the Study of Multiple Equilibria Systems - Study of Copper(II) Salicylate Monoethanolamine, Diethanolamine and Triethanolamine Systems, Analytica Chimica Acta 223 (1989) 257-268.
[208] M. Abou Fadel, A. de Juan, N. Touati, H. Vezin, L. Duponchel, New chemometric approach MCR-ALS to unmix EPR spectroscopic data from complex mixtures, Journal of Magnetic Resonance 248 (2014) 27-35.
[209] M. Mazur, A dozen useful tips on how to minimise the influence of sources of error in quantitative electron paramagnetic resonance (EPR) spectroscopy - A review, Analytica Chimica Acta 561 (2006) 1-15.
[210] K. Berg, M. Ericsson, M. Lindgren, H. Gustafsson, A High Precision Method for Quantitative Measurements of Reactive Oxygen Species in Frozen Biopsies, PLoS ONE 9 (2014) e90964.
[211] D.B. McPhail, D.J. Linehan, B.A. Goodman, An Electron-Paramagnetic Resonance (Electron-Paramagnetic-Res) Study of the Uptake of Vanadyl by Wheat Plants, New Phytologist 91 (1982) 615-620.
[212] G.H. Reed, M. Cohn, Electron Paramagnetic Resonance Spectra of Manganese(II)-Protein Complexes . Manganese(II)-Concanavalin-A, Journal of Biological Chemistry 245 (1970) 662-664.
[213] V. Haber, Study of the amorphous product of dehydration of copper (II) sulfate pentahydrate, Chemical Papers 31 (1977) 190-196.
[214] M.M. A. de Juan, T. Hancewicz, L. Duponchel, R. Tauler, in: R. Salzer, H.W. Siesler (Eds.), Wiley-VCH (Ed.), *chemometric tools for image analysis in infrared and raman spectroscopic imaging*, **2009**.

[215] A. de Juan, R. Tauler, R. Dyson, C. Marcolli, M. Rault, M. Maeder, spectroscopic imaging and chemometrics: a powerful combination for global and local sample analysis, Trac-Trends in Analytical Chemistry 23 (2004) 70-79.
[216] J.H. Wang, P.K. Hopke, T.M. Hancewicz, S.L.L. Zhang, Application of modified alternating least squares regression to spectroscopic image analysis, Analytica Chimica Acta 476 (2003) 93-109.
[217] K. Artyushkova, J.E. Fulghum, Identification of chemical components in XPS spectra and images using multivariate statistical analysis methods, Journal of Electron Spectroscopy and Related Phenomena 121 (2001) 33-55.
[218] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, J.F. Huntington, ICE: a statistical approach to identifying endmembers in hyperspectral images, IEEE Transactions on Geoscience and Remote Sensing 42 (2004) 2085-2095.
[219] X. Zhang, F. Liu, Y. He, X. Li, Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds, Sensors 12 (2012) 17234-17246.
[220] X. Zhang, R. Tauler, Application of multivariate curve resolution alternating least squares (MCR-ALS) to remote sensing hyperspectral imaging, Analytica chimica acta 762 (2013) 25-38.
[221] C. Ferrari, G. Foca, A. Ulrici, Handling large datasets of hyperspectral images: Reducing data size without loss of useful information, Analytica Chimica Acta 802 (2013) 29-39.
[222] M.A. Bendoumi, M.Y. He, S.H. Mei, Hyperspectral Image Resolution Enhancement Using High-Resolution Multispectral Image Based on Spectral Unmixing, Ieee Transactions on Geoscience and Remote Sensing 52 (2014) 6574-6583.
[223] M. Gaskell, K.I.E. McLuckie, P.B. Farmer, Comparison of the repair of DNA damage induced by the benzene metabolites hydroquinone and p-benzoquinone: a role for hydroquinone in benzene genotoxicity, Carcinogenesis 26 (2005) 673-680.
[224] B. Vajna, A. Farkas, H. Pataki, Z. Zsigmond, T. Igricz, G. Marosi, Testing the performance of pure spectrum resolution from Raman hyperspectral images of differently manufactured pharmaceutical tablets, Analytica Chimica Acta 712 (2012) 45-55.
[225] S. Sobanska, G. Falgayrac, J. Laureyns, C. Brémard, Chemistry at level of individual aerosol particle using multivariate curve resolution of confocal Raman image, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 64 (2006) 1102-1109.
[226] Y. Batonneau, J. Laureyns, J.-C. Merlin, C. Brémard, Self-modeling mixture analysis of Raman microspectrometric investigations of dust emitted by lead and zinc smelters, Analytica chimica acta 446 (2001) 23-37.
[227] J.M. Amigo, J. Cruz, M. Bautista, S. Maspoch, J. Coello, M. Blanco, Study of pharmaceutical samples by NIR chemical-image and multivariate analysis, Trac-Trends in Analytical Chemistry 27 (2008) 696-713.
[228] J. Cruz, M. Bautista, J.M. Amigo, M. Blanco, Nir-chemical imaging study of acetylsalicylic acid in commercial tablets, Talanta 80 (2009) 473-478.
[229] F. Trompier, A. Romanyukha, R. Reyes, H. Vezin, F. Queinnec, D. Gourier, State of the art in nail dosimetry: free radicals identification and reaction mechanisms, Radiation and Environmental Biophysics 53 (2014) 291-303.
[230] H. Nagayoshi, S. Nishimura, K. Terashima, Ieee, *Optical properties of Eu doped CaF2 crystal for solar cell application*, Conference Record of the 2006 IEEE 4th World Conference on Photovoltaic Energy Conversion, Vols 1 and 2, Ieee, New York, **2006**, pp. 150-153.
[231] L.W. Sang, M.Y. Liao, Y. Koide, M. Sumiya, High-performance metal-semiconductor-metal InGaN photodetectors using CaF2 as the insulator, Applied Physics Letters 98 (2011).

[232] Z.W. Quan, D.M. Yang, P.P. Yang, X.M. Zhang, H.Z. Lian, X.M. Liu, J. Lin, Uniform Colloidal Alkaline Earth Metal Fluoride Nanocrystals: Nonhydrolytic Synthesis and Luminescence Properties, Inorganic Chemistry 47 (2008) 9509-9517.
[233] A.J. Stevenson, H. Serier-Brault, P. Gredin, M. Mortier, Fluoride materials for optical applications: Single crystals, ceramics, glasses, and glass-ceramics, Journal of Fluorine Chemistry 132 (2011) 1165-1173.
[234] R.R. Alfano, *The Supercontinuum Laser Source*, Springer, New York, **1989**.
[235] L. Wang, J.X. Yan, X.H. Chen, Y.X. Weng, S.H. Zhou, *Experimental study on the chirp of the white-light continuum generation in CaF2 and sapphire*, In: Nonlinear Optical Phenomena and Applications, Q. Gong, Y.P. Cui, R.A. Lessard (Eds.), Spie-Int Soc Optical Engineering, Bellingham, **2005**, pp. 108-111.
[236] C.H. Anderson, E.S. Sabisky, EPR Studies of Photochromic CaF2, Physical Review B 3 (1971) 527-536.
[237] A.V. Egranov, T.Y. Sizova, Configurational instability at the excited impurity ions in alkaline earth fluorites, Journal of Physics and Chemistry of Solids 74 (2013) 530-534.
[238] H. Barfuss, G. Bohnlein, W. Kreische, H. Niedrig, H. Tauwaldt, G. Thiele, Lattice-Defects in CaF2 And SrF2 After Proton Irradiation, Zeitschrift Fur Physik B-Condensed Matter 50 (1983) 9-16.
[239] M. Abou Fadel, X. Zhang, A. de Juan, R. Tauler, H. Vezin, L. Duponchel, Extraction of Pure Spectral Signatures and Corresponding Chemical Maps from EPR Imaging Data Sets: Identifying Defects on a CaF2 Surface Due to a Laser Beam Exposure, Analytical Chemistry 87 (2015) 3929-3935.
[240] W. Hayes, *Crystals with the Fluorite Structure: Electronic, Vibrational, and Defect properties*, Oxford University Press, **1974**.
[241] A. Kelly, K.M. Knowles, *Crystal Structures. In Crystallography and Crystal Defects*, John Wiley and sons, **2012**.
[242] R.C. Alig, S. Bloom, Electron-Hole-Pair Creation Energies in Semiconductors, Physical Review Letters 35 (1975) 1522-1525.
[243] R.W. Ure, Ionic Conductivity of Calcium Fluoride Crystals, Journal of Chemical Physics 26 (1957) 1363-1373.
[244] K. Schwartz, *Excitons and radiation damage in alkali halides*, In: Atomic Physics Methods in Modern Research, Springer, **1997**, pp. 351-366.
[245] B.C. Cavenett, W. Hayes, I.C. Hunter, A.M. Stoneham, Magneto Optical Properties of F Centres in Alkaline Earth Fluorides, Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences 309 (1969) 53-68.
[246] R. Nakata, K. Khono, M. Sumita, E. Higuchi, New ESR Center In X-Irradiated CaF2 Crystals, Journal of the Physical Society of Japan 40 (1976) 1328-1332.
[247] J. Sierro, Paramagnetic Resonance of Vf Center in CaF2, Physical Review 138 (1965) A648-A650.
[248] J. Arends, Color Centers in Additively Colored CaF2 And BaF2, Physica Status Solidi 7 (1964) 805-815.
[249] Y. Kazumata, ESR of Neutron-Irradiated BaF2 Crystals, physica status solidi (b) 34 (1969) 377-384.
[250] A. de Juan, M. Maeder, M. Martinez, R. Tauler, Combining hard- and soft-modelling to solve kinetic problems, Chemometrics and Intelligent Laboratory Systems 54 (2000) 123-141.
[251] S. Mas, A. Carbo, S. Lacorte, A. de Juan, R. Tauler, Comprehensive description of the photodegradation of bromophenols using chromatographic monitoring and chemometric tools, Talanta 83 (2011) 1134-1146.

[252] R. Gargallo, R. Tauler, F. CuestaSanchez, D.L. Massart, Validation of alternating least-squares multivariate curve resolution for chromatographic resolution and quantitation, Trac-Trends in Analytical Chemistry 15 (1996) 279-286.
[253] K. De Braekeleer, A. de Juan, D.L. Massart, Purity assessment and resolution of tetracycline hydrochloride samples analysed using high-performance liquid chromatography with diode array detection, Journal of Chromatography A 832 (1999) 67-86.
[254] J. Saurina, S. Hernandez-Cassou, R. Tauler, A. Izquierdo-Ridorsa, Procedure for the quantitative determination of mixtures of nucleic acid components based on multivariate spectrophotometric acid-base titrations, Analytical Chemistry 71 (1999) 126-134.
[255] J. Diewok, A. De Juan, R. Tauler, B. Lendl, Quantitation of mixtures of diprotic organic acids by FT-IR flow titrations and multivariate curve resolution, Applied Spectroscopy 56 (2002) 40-50.
[256] A. Safavi, B. Hemmateenejad, F. Honarasa, Chemometrics assisted resolving of net faradaic current contribution from total current in potential step and staircase cyclic voltammetry, Analytica Chimica Acta 766 (2013) 34-46.
[257] P. Cotelle, H. Vezin, EPR of free radicals formed from 3-hydroxyesculetin and related derivatives, Res. Chem. Intermed. 29 (2003) 365-377.
[258] A. De Juan, A. Izquierdo-Ridorsa, R. Tauler, G. Fonrodona, E. Casassas, A soft-modeling approach to interpret thermodynamic and conformational transitions of polynucleotides, Biophysical journal 73 (1997) 2937.
[259] F. Bailly, C. Maurin, E. Teissier, H. Vezin, P. Cotelle, Antioxidant properties of 3-hydroxycoumarin derivatives, Bioorganic & Medicinal Chemistry 12 (2004) 5611-5618.
[260] B. Halliwell, Free radicals, antioxidants, and human disease: curiosity, cause, or consequence?, The Lancet 344 (1994) 721-724.
[261] B.N. Ames, M.K. Shigenaga, T.M. Hagen, Oxidants, antioxidants, and the degenerative diseases of aging, Proceedings of the National Academy of Sciences 90 (1993) 7915-7922.
[262] J. Bouayed, T. Bohn, Exogenous antioxidants—double-edged swords in cellular redox state: health beneficial effects at physiologic doses versus deleterious effects at high doses, Oxidative Medicine and Cellular Longevity 3 (2010) 228-237.
[263] M. Fontecave, J.L. Pierre, Oxygen Activation and Toxicity - Principles for an Efficient Antioxidant Therapy, Bulletin De La Societe Chimique De France (1991) 505-520.
[264] B. Halliwell, Drug Antioxidant Effects - A Basis for Drug Selection, Drugs 42 (1991) 569-605.
[265] C. Behl, Amyloid beta-protein toxicity and oxidative stress in Alzheimer's disease, Cell and Tissue Research 290 (1997) 471-480.
[266] C.A. Riceevans, A.T. Diplock, Current Status of Antioxidant Therapy, Free Radical Biology and Medicine 15 (1993) 77-96.
[267] M. Valko, D. Leibfritz, J. Moncol, M.T.D. Cronin, M. Mazur, J. Telser, Free radicals and antioxidants in normal physiological functions and human disease, The International Journal of Biochemistry & Cell Biology 39 (2007) 44-84.
[268] C. Zdero, F. Bohlmann, H.M. Niemeyer, A Heliangolide, 3-Hydroxyumbelliferone Derivatives and Diterpenes from Bahia-Ambrosioides, Phytochemistry 29 (1990) 205-208.
[269] K. Aihara, T. Higuchi, M. Hirobe, 3-Hydroxycoumarins - 1st Direct Preparation From Coumarins Using a Cu-2+ Ascorbic-Acid O2 System, And Their Potent Bioactivities, Biochemical and Biophysical Research Communications 168 (1990) 169-175.

Title: Chemometric Contributions to Electron Paramagnetic Resonance Spectroscopy: New Perspectives for Processing of Spectral Data with Spatial Dimension (Imaging) and / or Time Evolution.

Abstract

Electron Paramagnetic Resonance (EPR) Spectroscopy has undoubtedly become the first-choice technique for the characterization of complex materials containing some unpaired electrons (transition metal ions, rare earth ions, defects, organic radicals ...). Similarly to nuclear magnetic resonance spectroscopy, EPR generates multidimensional (2D, 3D…) spectral and recently also spatial (imaging) data as well as spectral/spatial ones. It is thus, surprising that despite the large amount of spectral data to be explored and complexity of the EPR signals, there hardly exist at the international level of exploitation the multivariate data processing methods that are widely available in chemometrics. The objective of this thesis is thus, to develop new tools for the treatment of these EPR spectral data, to establish new analytical methodologies and to evaluate their performance. The two main areas that will be studied are spectroscopic imaging and time-resolved spectroscopy. In this work, we will show that the implementation of the methods known as "multivariate curve resolutions" can extract, simultaneously, and without a priori all chemical maps and their corresponding spectra of pure compounds present in the studied sample. This methodology will also be exploited to extract the EPR spectra of intermediate species during a kinetic monitoring.

Keywords: multivariate analysis of spectral data, chemometrics, imaging, electron paramagnetic resonance, Multivariate Curve Resolution (MCR-ALS), Principal Component Analysis (PCA).

Titre : Apports de la chimiométrie à la spectroscopie de Résonance Paramagnétique Electronique : nouvelles perspectives de traitement de données spectrales à dimensions spatiales (imagerie) et / ou temporelles.

Résumé

La spectroscopie de Résonance Paramagnétique Electronique (RPE) est incontestablement devenue une technique de choix pour la caractérisation des matériaux complexes quand ceux-ci possèdent des électrons non appariés (ions de transitions, terres rares, défauts, radicaux organiques…). A l'instar de la résonnance magnétique nucléaire, la RPE génère des données spectrales multidimensionnelles (2D, 3D…) et depuis peu une imagerie spatiale mais aussi spectrale/spatiale. Il est ainsi étonnant de voir que malgré la grande quantité de données spectrales à explorer et la complexité des signaux RPE, il n'existe quasiment pas au niveau international d'exploitation des méthodes de traitements de données multivariées, méthodes largement proposées en chimiométrie. L'objectif de cette thèse est ainsi de développer de nouveaux outils pour le traitement de ces données spectrales RPE, d'établir de nouvelles méthodologies d'analyse et d'en évaluer leurs performances. Les deux axes principalement étudiés seront l'imagerie spectroscopique et la spectroscopie résolue en temps. Nous démontrerons dans ce travail que la mise en œuvre de méthodes dites de «résolutions multivariées de courbes» permet d'extraire simultanément et sans a priori toutes les cartographies chimiques et les spectres associés des composés purs présents dans l'échantillon étudié. Cette méthodologie sera aussi exploitée afin d'extraire les spectres RPE d'espèces intermédiaires lors d'un suivi cinétique.

Mots clés : analyse multivariée des données spectrales, chimiométrie, imagerie, résonance paramagnétique électronique, résolution multivariée de courbes (MCR-ALS), analyse en composantes principales (ACP).