UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1672

Signal Processing Tools for Electron Microscopy

JAKOB SPIEGELBERG

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2018

ISSN 1651-6214
ISBN 978-91-513-0345-1
urn:nbn:se:uu:diva-348264

Dissertation presented at Uppsala University to be publicly examined in Å2001, Ångströmlaboratoriet, Lägerhyddsvägen 1, Uppsala, Tuesday, 12 June 2018 at 09:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Johan Verbeeck (EMAT, University of Antwerp).

Abstract
Spiegelberg, J. 2018. Signal Processing Tools for Electron Microscopy. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1672. 60 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0345-1.

The detection of weak signals in noisy data is a problem which occurs across various disciplines. Here, the signal of interest is the spectral signature of the electron magnetic chiral dichroism (EMCD) effect. In principle, EMCD allows for the measurement of local magnetic structures in the electron microscope, its spatial resolution, versatility and low hardware requirements giving it an eminent position among competing measurement techniques. However, experimental shortcomings as well as intrinsically low signal to noise ratio render its measurement challenging to the present day.

This thesis explores how posterior data processing may aid the analysis of various data from the electron microscope. Following a brief introduction to different signals arising in the microscope and a yet briefer survey of the state of the art of EMCD measurements, noise removal strategies are presented. Afterwards, gears are shifted to discuss the separation of mixed signals into their physically meaningful source components based on their assumed mathematical characteristics, so called blind source separation (BSS).

A data processing workflow for detecting weak signals in noisy spectra is derived from these considerations, ultimately culminating in several demonstrations of the extraction of EMCD signals. While the focus of the thesis does lie on data processing strategies for EMCD detection, the approaches presented here are similarly applicable in other situations. Related topics such as the general analysis of hyperspectral images using BSS methods or the fast analysis of large data sets are also discussed.

Jakob Spiegelberg, Department of Physics and Astronomy, Materials Theory, Box 516, Uppsala University, SE-751 20 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-0345-1
urn:nbn:se:uu:diva-348264 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-348264)

There is a T-rex in this thesis!

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **Detecting magnetic ordering with atomic size electron probes**
J. C. Idrobo, J. Rusz, J. Spiegelberg, M. A. McGuire, C. T. Symons, R. R. Vatsavai, C. Cantoni and A. R. Lupini
Adv. Struct. Chem. Imag. **2**, 5 (2016).

II **Magnetic measurements with atomic-plane resolution**
J. Rusz, S. Muto, J. Spiegelberg, R. Adam, K. Tatsumi, D. E. Bürgler, P. M. Oppeneer and C. M. Schneider
Nature Communications **7**, 12672 (2016).

III **Can we use PCA to detect small signals in noisy data?**
J. Spiegelberg and J. Rusz
Ultramicroscopy **172**, 40-46 (2017).

IV **Analysis of Electron Energy Loss Spectroscopy Data using Geometric Extraction Methods**
J. Spiegelberg, J. Rusz, T. Thersleff and K. Pelckmans
Ultramicroscopy **174**, 14-26 (2017).

V **Tensor Decompositions for the Analysis of Atomic Resolution Electron Energy Loss Spectra**
J. Spiegelberg, J. Rusz and K. Pelckmans
Ultramicroscopy **175**, 36-45 (2017).

VI **Localization of magnetic circular dichroic spectra in transmission electron microscopy experiments with atomic plane resolution**
J. Rusz, J. Spiegelberg, S. Muto, T. Thersleff, K. Tatsumi, M. Ohtsuka, K. Leifer and P. M. Oppeneer
Phys. Rev. B **95**, 174412 (2017).

VII **Towards sub-nanometer real-space observation of spin and orbital magnetism at the Fe / MgO interface**
T. Thersleff, S. Muto, M. Werwiński, J. Spiegelberg, Y. Kvashnin, B. Hjörvarsson, O. Eriksson, J. Rusz and K. Leifer,
Scientific Reports **7**, 44802 (2017).

VIII **The usage of data compression for the background estimation of electron energy loss spectra**
J. Spiegelberg, J. Rusz, K. Leifer and T. Thersleff
Ultramicroscopy **181**, 117-122 (2017).

IX **Fully non-local inelastic scattering computations for spetroscopical TEM methods**
J. Rusz, A. Lubk, J. Spiegelberg and D. Tyutyunnikov
Phys. Rev. B **96**, 245121 (2017).

X **Unmixing Hyperspectral Data by using Signal Subspace Sampling**
J. Spiegelberg, S. Muto, M. Ohtsuka, K. Pelckmans and J. Rusz
Ultramicroscopy **182**, 205-211 (2017).

XI **Local low rank denoising for enhanced atomic resolution imaging**
J. Spiegelberg, J. C. Idrobo, A. Herklotz, T. Z. Ward, W. Zhou and J. Rusz
Ultramicroscopy **187**, 34-42 (2018).

XII **On the usage of ICA for the analysis of EELS and EDX data**
J. Spiegelberg
Ultramicroscopy, submitted.

XIII **Low dose STEM imaging via inpainting of regularly undersampled images**
J. Spiegelberg, J. Hachtel, J. C. Idrobo and J. Rusz
Scientific Reports, submitted.

XIV **Real space mapping of magnetism at atomic resolution using APR-EMCD**
J. Spiegelberg, S. Muto, M. Ohtsuka, and J. Rusz
Microscopy, submitted.

XV **Blind identification of magnetic signals in electron magnetic chiral dichroism using independent component analysis**
J. Spiegelberg, D. Song, R. E. Dunin-Borkowski, J. Zhu, and J. Rusz
Ultramicroscopy, submitted.

XVI **Data fusion approaches for electron microscopy**
J. Spiegelberg, Y. Yamamoto, S. Muto, M. Ohtsuka, and J. Rusz
Ultramicroscopy, submitted.

Papers that were (co-)authored by me, but are not included in this thesis:

XVII **Validity of the rotating wave approximation in nonadiabatic holonomic quantum computation**
J. Spiegelberg and E. Sjökvist
Phys. Rev. A **88**, 054301 (2013).

XVIII **A multislice theory of electron scattering in crystals including backscattering and inelastic effects**
J. Spiegelberg and J. Rusz
Ultramicroscopy **159**, 11-18 (2015).

Contents

1. Introduction

Reader! Let me lay out the center of my interest over the last few years to you. Alas, not by dialogue, but by these very lines. Much like when skipping stones, we shall touch many topics briefly, and without much depth.

Since you took this little booklet from the shelf, it is clear that you have at least a rudimentary idea of the functioning and applications of an electron microscope. I need not praise its capabilities nor potential in front of you! From our position, it is, however, still necessary to take a step backwards in order to progress forwards: All too frequently, measurements in the microscope lead to data with low signal to noise ratios, such as in low dose imaging or high resolution spectrum imaging. Also having an eye on ever increasing amounts of data, the need to develop robust routines to detect the signal of interest, automated routines to tackle the analysis of tomorrow's large data sets or generally a data processing combining both, robust to shortcomings of the data, yet scalable to fully exploit all of the data correlations at hand, is apparent. The third step of a microscopy experiment after sample preparation and measurement, the analysis of the obtained data, needs to be refined.

This is the ground on which this thesis stands. While most of the results and considerations presented here are straightforwardly generalizable to other applications, the signal of interest shall be the spectral signature of the electron magnetic chiral dichroism (EMCD) [1, 2] effect. EMCD is a scanning transmission electron microscope (STEM) technique, a daughter of electron energy loss spectroscopy (EELS). The EMCD signal gives information on local magnetic structures. Combined with the capabilities of modern STEM instruments, EMCD offers thus an opportunity to study nano-sized magnetic devices while simultaneously offering the information wealth of an EELS measurement.

Considering the analysis of EEL spectra, one can identify two challenges. First, the signal, which is often excruciatingly weak, has to be recognized beneath the noise it is hidden by. Second, signals may appear mixed, it may be necessary to invert the mixing before they can be interpreted. The former boils down to the challenge of noise removal, denoising of the signal. Since the noise in EEL spectra ideally follows a Poisson distribution, this can be achieved by averaging of identical spectra, by application of local filters, or also more sophisticated approaches, discussed below, which avoid deterioration of the structures by blurring under milder mathematical assumption. The latter is the much discussed problem of blind source separation (BSS). By assuming suitable, ideally generally applicable mathematical constraints, one

tries to solve the otherwise ill-conditioned task of matrix factorization. E.g., statistical independence or non-negativity combined with sparsity are often assumed one way or the other to solve the BSS problem.

A third challenge is the size and format of data coming from electron microscopes. Modern instruments allow for acquisition of 10^5 to 10^6 spectra per measurement session, data can be anything from one dimensional single spectra to two dimensional images, three dimensional hyperspectral images, or also higher dimensional formats when considering spectrum tomography [3, 4, 5], measurements with pixelated cameras [6, 7] or segmented annular dark field (ADF) detectors [8], or generally measurements with added parameters (e.g., time in a time series). Fortunately, while the increasing data size puts limitations on the computational efficiency of applied algorithms, BSS methods typically benefit from large data sets so that the noise is suppressed further.

On this backdrop, consider the declared goal of measuring EMCD: The signal has a notoriously low variance, some few percent of the non-magnetic signal's strength [9]. Correspondingly, one expects a low signal to noise ratio (SNR). Especially when taken to the extreme of atomic resolution, the predicted SNR drops to the limit of detectability. The signal occurs necessarily mixed, only in ideal situations it can be obtained by subtracting normalized spectra from each other, making some signal separation necessary. However, the EMCD spectrum defies the mathematical assumptions of several more common BSS methods. It is neither independent from the non-magnetic signal, nor is it non-negative. Being highly sample dependent it can not be parametrized straightforwardly either...

You see, reader, EMCD extraction implies challenges, it is an ideal playground to test and develop novel approaches. After getting a most brief overview over different signals that can be measured in the microscope in section 2, a selection of data processing tools useful for the task of EMCD extraction and general EEL spectrum analysis are presented in section 3. At last, these two threads are joined, new methods furthering the development of EELS analysis are introduced, ultimately culminating in the demonstration of different EMCD measurements. In the coda of the thesis, a few related data processing strategies are discussed.

2. Signals in the electron microscope

The versatility of electron microscopy is directly coupled to the broad range of signals that can be measured on a suitably equipped instrument. Besides the basic imaging modes in transmission and scanning transmission electron microscopy (TEM and STEM, respectively), an ADF detector can be used to obtain an additional image for electrons scattered to larger angles. When using a segmented ADF detector [8], one immediately obtains an additional mode for the data, the two-dimensional image becomes a three-dimensional tensor, the third mode describing the detector segments. Spectrometers allow for the detection of fluorescent X-rays, cathodoluminescence (CL), or also the energy loss of the electron after scattering in the sample. Focusing on STEM, a fast CCD camera can be used to acquire a diffraction pattern at every image pixel, again leading to a large increase of information and data complexity (and size) [6, 7]. Additional complexity is gained by introducing apertures selecting only electrons scattered to certain angles, by tailoring the (phase) profile of the electron beam changing its scattering behavior, or multiple other parameters which can be considered at liberty.

It must be clear that this thesis can not give an exhaustive overview over suitable data processing techniques for all of these signals. Instead, the focus lies on the analysis of spectral data sets. The three spectroscopic techniques in the microscope, EELS, CL and energy dispersive X-ray spectroscopy (EDX), and proposed data processing tools for either of them will be briefly discussed in the sections below, followed by a short introduction of the thesis' main objective, the EMCD signal, and miscellaneous signals of interest, selected by the author in a biased manner.

Before introducing the different spectroscopic techniques it is worthwhile considering the overall data format. Not considering additional measurement parameters (time, additional spatial mode for spectroscopic tomography, ...), a STEM measurement where a spectrum is measured at each pixel will yield a three-dimensional data set, $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E}$, where n_x and n_y are the spatial coordinates and n_E denotes the number of energy channels. This data structure occurs in multiple applications, e.g., in the remote sensing community, it is referred to a *hyperspectral image* (HI, e.g. [10] and references therein), in the image processing community it would more likely be referred to as *multi-channel image*, in the microscopy community the terms *data cube* or *spectrum image* are more commonly used.

From seeing how the same data format occurs in different disciplines, one can recognize that some methods developed for one are probably applicable also in the other. However, here the focus will be kept on STEM experiments.

2.1 The spectral trinity

2.1.1 Cathodoluminescence

CL is considerably less prominent than the other two spectral techniques, but nevertheless finds its application range for signals in the low energy range. The signal is caused by radiative energy relaxation of the sample which has previously been excited by the electron radiation. Collection of the radiation emitted from the sample then yields the CL emission spectrum. Typical wavelengths lie around the visible spectrum, which corresponds to outer shell excitations in the emitting atom. Given this energy range, CL is affected by the chemical environment and can be used to, e.g., study defects in (In,Ga)N/GaN nanowires [11], or to study plasmonic structures of nano-particles [12, 13]. The spatial resolution of CL experiments is more limited due a larger interaction volume. Since the energy range of CL radiation lies around few eV, also scattered electrons are able to excite it. Hence, the interaction volume increases well beyond the order of magnitude of the converged STEM probe to several nanometers (e.g. [14]).

Given the emission spectrum character, the CL spectrum shares some characteristics with the EDX spectra: Both consist of a series of peaks, in the CL spectrum broad and featureless, often overlapped which may complicate the analysis. Given this obstacle, BSS methods seem a good candidate to facilitate the analysis of CL data, but are not yet spread in the community. An example can be found in [15], where principal component analysis (PCA) is used for classification of spectra from different types of river sediments. The inherent lower resolution of CL spectra, however, implicates often enough a relatively high SNR so that basic analysis methods such as integration over individual peaks (e.g. [16]) suffices to characterize the sample. It remains subject of future explorations to demonstrate the usage of BSS for CL spectra. E.g., by enabling the interpretation of heavily overlapped signals or boosting the signal to noise ratio to detect also faint signals.

2.1.2 Energy dispersive X-ray spectroscopy

Similar to CL, fluorescent radiation is collected in EDX spectroscopy. However, as suggested by the name, radiation at much higher energies, resulting from core-shell excitations, is collected. Consequently, much of the information in the CL spectrum is lost. Unless a wavelength dispersive detector is used, EDX spectra only contain elemental information without fine-structures or chemical shifts distinguishing stoichiometrically equivalent, but chemically different compounds [17]. Acquisition of X-ray spectra at high energy resolutions using wavelength dispersive X-ray spectroscopy (WDX or WDS) overcomes this limitation, but is typically intractable in a STEM experiment as single pixel dwell times do not permit acquisition of the full WDX spectrum. When operated in STEM mode, WDX measurements typically only acquire

intensities at selected, pre-defined energies determined by the diffraction angle of the dispersion grating (i.e., the crystal used as prism) (see, e.g., [18] for an application of WDX). The increased energy of the radiation also implies a smaller interaction volume. If sufficiently thin specimens are used, atomic resolution maps can be achieved (see, e.g., [19]). Notably, the EDX signal appears less delocalized than its EELS counterpart [20], which may improve the contrast of elemental maps at atomic resolution.

Different data processing techniques have been proposed for EDX spectra. For example, besides noise filters [21], PCA has been applied both for a denoising of the data as well as subsequent analysis of rotated factor loadings [22]. Straightforwardly interpretable components are obtained. Given that the EDX spectrum consists of a broad Bremsstrahlung background and sharp peaks, non-negative matrix factorization (NMF) appears to be a suitable candidate for the analysis of EDX spectra. Besides the consideration in paper X, where source components of EDX data are identified using NMF on data generated by drawing random samples from the signal subspace while enforcing non-negativity, which amounts to an assumption of stand-alone peaks in the source spectra, NMF's applicability to EDX data has been demonstrated, e.g., in [23], where meaningful source spectra of different phases in a $LaSrMnO_3$-Y doped ZrO_2 are extracted using a sparsity constrained NMF algorithm.

2.1.3 Electron energy loss spectroscopy

At last, EELS has to be discussed. More in depth overviews can be found in [17, 24]. In contrast to CL or EDX, the energy loss of transmitted electrons rather than emitted photons is measured in EELS. This energy loss is highly characteristic of the sample, and results in a spectrum with largely different features. Consider the stereotypical EEL spectrum in Fig.2.1: The spectrum's most prominent feature is the zero-loss peak consistent of electrons which have passed through the sample without loosing any energy. Since it has by far the highest intensity, it dominates the image contrast in STEM. If a monochromator is employed, the features on its tails can be studied. E.g., small energy gains by thermal excitation can be studied on the negative energy loss (i.e., energy gain) side [25]. On the energy loss side, phonons and bandgaps are the first features to appear in so-called valence EELS (VEELS, [26]). Next, the plasmonic peaks dominate the spectrum. Especially at larger thicknesses, they may dominate the spectrum. Multiple scattering events may become pronounced also for electrons which have undergone a high-energy loss scattering event. In this case, a deconvolution of the spectrum with the low-loss spectrum may correct for the multiple scattering contributions [24]. Beyond the plasmonic spectrum the high-loss region spans, which is the region where most analyses are carried out. Core-loss edge signals do not only allow for an elemental characterization of the sample, but also carry fine-structures which

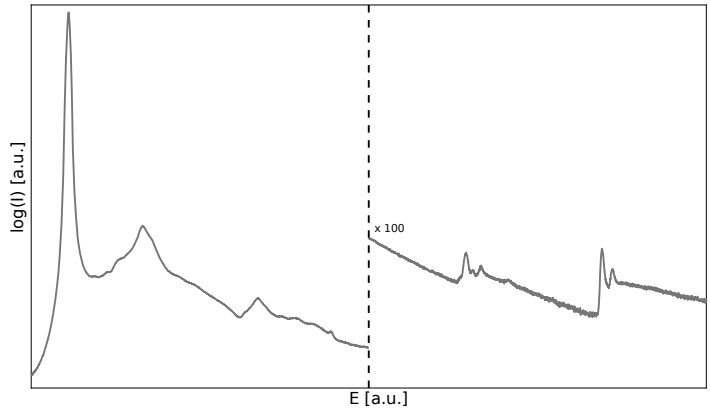

Figure 2.1. Schematic of a stereotypical EELS spectrum displaying zero-loss and Plasmon peaks on the low energy side and element characteristic core loss spectra at high energies.

are characteristic of the chemical environment, the magnetic structure, etc. Especially when considering further enriching the spectrum information by placing a collection aperture in the microscope's backfocal plane so that momentum dependent scattering contributions can be observed, the full wealth of information of an EEL spectrum becomes apparent.

While this information wealth offers great opportunities to study various material properties at high resolutions, the mathematical characteristics of the core-loss spectrum make EELS analysis more challenging compared to EDX or CL. The spectrum contains a large background, typically modeled as a power law decay. The edge signals on top of the background can in first approximation be modeled as saw-tooth function, also contributing with a long tail to the background signal in energetically close lying edges. On the edge onset, fine-structures are typically observed. Due to the large variety of chemical environments (and other fine-structure contributions), the shape of the fine-structures eludes a poignant parametric description. Apart from these features of interest, EEL spectra may contain multiple artifacts. Spurious X-ray impacts may lead to sharp spikes in the measured spectrum. More severely, since it is less easily corrected for a posteriori, the detection system can lead to different additive components ranging from a correlated noise like contribution to intensity fluctuations either in time or across the energy channels. Assuming a stable microscope, these artifact components can be corrected for by subtracting a dark and gain reference from the measured spectra. At last, it has to be mentioned that EEL spectra often suffer from a low SNR. Considering the logarithmic intensity scale in Fig.2.1, it becomes apparent that the core-loss region has much fewer counts than other, low energy-loss features

of the spectrum. When measuring with short dwell times, e.g., at high resolutions where sample damage puts a limiting factor on the maximal dwell time, this implies very noisy spectra, which poses a challenge to analysis afterward.

Given the multitude of signals and challenges of the SNR, various data processing approaches have been proposed. Narrowing the discussion to analysis by BSS techniques, some examples may be found in the attached papers, other, selected examples are listed below: NMF, optionally sparsity constrained, has been demonstrated to be applicable to EEL spectra [27, 23]. The goal of identifying meaningful source spectra (and abundances) is always considered first for NMF on data including large backgrounds. Independent component analysis (ICA) was also proposed for EEL spectrum analysis [28, 29, 30], with limitations for the application to spectra including fine-structures, see paper XII. A workhorse for spectrum denoising has been PCA [31, 32, 33, 34]. Limitations of PCA for spectrum denoising have been discussed in [35, 36, 37] or also in paper III. A more detailed discussion of EELS analysis follows in the remainder of this thesis.

2.2 The holy grail

2.2.1 Brief theoretical background of EMCD

The story of EMCD starts with the desire to characterize magnetic structures at high spatial resolutions. Neutron scattering experiments can be used to study magnetic structures accurately, but the spatial resolution is rather limited. Unless high demands are put on the sample preparation, i.e., to place a distinct isotope layer at the exact position in the sample which should be studied [38], neutron scattering is a bulk measurement. As first improvement on the spatial resolution, X-ray magnetic circular dichroism (XMCD) enables resolutions of approx. 10 nanometers if synchrotron radiation is used [39, 40, 41]. A quantitative analysis of the measured signals is then enabled by use of the so called sum-rules [42, 43], which relate the spectral shape of the XMCD signal to spin and orbital magnetic moment.

Modern electron microscopes, on the other side, routinely achieve spatial resolutions in the order of few tens of picometers. Considering EELS measurement, magnetic signals can be assessed using EMCD. EMCD dates back to 2003 [1] and has been demonstrated experimentally shortly after in 2006 [2]. In close analogy to XMCD, sum-rules were formulated also for EMCD [44, 45] enabling a quantitative analysis. Since then, progress has been made to push the achievable resolution, SNR or generality of the measurement by refined experiments [9, 46, 47, 48, 49, 50, 51, 52], to further the theoretical understanding and optimize experimental parameters [53, 54, 55, 56, 57, 58, 59], as well as to study the quantitativeness of the method potentially improved by suitable data processing [60, 61, 62, 63, 64, 65]. New measurement schemes for an optimized SNR or experimental feasibility have been proposed as well.

E.g., the use of aberrated electron probes for EMCD measurement was discussed and demonstrated [66, 67, 68]. Electron vortex beams [69, 70, 71] were proposed for EMCD measurements on crystals with arbitrary symmetry [72, 73, 74]. Despite recent advances achieving atomically sized vortex beams [75, 76, 77], a conclusive measurement of EMCD using vortex beams has not been reported so far. Other advances include the proposal of alternative selection apertures in atomic plane resolution EMCD (APR-EMCD) [78, 79], the use of alternative apertures in the microscope for further generalization of the technique [80], or the use of spatially resolved EELS (SREELS) for EMCD [81].

The remainder of this section is dedicated to a quick introduction to the underlying scattering theory of EMCD followed by a survey of different EMCD detection geometries. A more detailed introduction to dynamical diffraction can be found in [82, 83], an introduction to EMCD in [53].

The double differential scattering cross-section $\frac{\partial^2\sigma}{\partial E\partial\Omega}$ contains the EMCD signal. It can be expressed as a sum of so called mixed dynamic form factors $S(\mathbf{q},\mathbf{q}',E)$ (MDFF) [2]. In two beam orientation, dynamical diffraction effects are particularly simple and the scattering cross section can be written as:

$$\frac{\partial^2\sigma}{\partial E\partial\Omega} = \frac{4\gamma^2}{a_0^2}\frac{k_f}{k_i}\left(S(\mathbf{q},\mathbf{q},E)+S(\mathbf{q}',\mathbf{q}',E)+2\mathfrak{Im}\left[S(\mathbf{q},\mathbf{q}',E)\right]\right) \tag{2.1}$$

where $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ is the relativistic correction factor, a_0 is the Bohr radius and the MDFF

$$S(\mathbf{q},\mathbf{q}',E) = \sum_{i,f}\left\langle i\left|\frac{e^{i\mathbf{q}\cdot r}}{q^2}\right|f\right\rangle\left\langle f\left|\frac{e^{-i\mathbf{q}'\cdot r}}{q'^2}\right|i\right\rangle\delta\left(E_f-E_i-E\right) \tag{2.2}$$

where $|i\rangle$, $|f\rangle$, E_i and E_f denote the state and energy of initial and final wavefunction, respectively. $\mathbf{q}$ denotes the momentum transfer in the scattering event, E the energy loss. After recasting the equations, one notes that the magnetic signal is only contained in the imaginary part of the MDFF [53]. Assuming that non-dipole transitions are negligible:

$$S(\mathbf{q},\mathbf{q}',E) \approx \frac{1}{q^2q'^2}\left[N(E)\mathbf{q}\cdot\mathbf{q}'+i\mathbf{M}(E)\cdot(\mathbf{q}\times\mathbf{q}')\right] \tag{2.3}$$

where $\mathbf{M}(E)$ and $N(E)$ denote magnetic and non-magnetic contribution, respectively, and are purely real. The classical EMCD geometry then follows via assuming magnetization along z and a Bragg vector $\mathbf{G} = (G,0,0)$ with $\mathbf{q}' = \mathbf{q}-\mathbf{G}$ [66]

$$\mathfrak{Im}\left[S(\mathbf{q},\mathbf{q}',E)\right] = \frac{GM_z(E)q_y}{|\mathbf{q}|^2|\mathbf{q}-\mathbf{G}|^2} \tag{2.4}$$

When moving the detector from the wavevector $\mathbf{k}^f = (k_x^f,k_y^f,k_z^f)$ to $\mathbf{k}'^f = (k_x^f,-k_y^f,k_z^f)$ the sign of q_y is inverted what corresponds to an inversion of the

perceived magnetic signal's sign. The EMCD spectrum, the magnetic signal, is thus obtained by acquiring the spectra for both these wavevectors and then taking their difference.

2.2.2 Survey of experiment settings

Finding measurement settings suitable to get a sizable EMCD signal boils down to maximizing contributions of the imaginary parts of the MDFF to the spectrum. Either one selects electrons scattered into suitable angles by means of a collection aperture, or one manipulates the phase of the incoming electrons. Having realized this, all proposed measurement schemes become variations on a theme:

Direct difference EMCD: DD-EMCD, sometimes referred to as classical EMCD, stems from the original EMCD proposal [1, 2]. The sample is oriented in 2 (or 3) beam case orientation, as indicated by the derivations above, subsequent placement of a collection aperture on two opposing positions on the Thales circle will yield spectra with opposing EMCD signs. *Direct difference* of these two spectra will then obtain the EMCD signal. When extended to spectrum imaging, a DD-EMCD measurement leads to a four dimensional data set $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E \times n_a}$, n_a being the aperture placement index. Although demonstrated experimentally repeatedly, this scheme bears the disadvantage that signal fractions drop when reaching atomic resolution and the same sample area needs to be scanned several times to obtain the EMCD signal, which may lead to problems with sample damage or microscope stability issues due to the prolonged measurement time.

SREELS-EMCD: Being quite recent [81], SREELS-EMCD is closely related to DD-EMCD in the sense that it also utilizes a collection aperture for selection of suitably scattered electrons. Now, however, a second selection slit is placed in the image plane (or the detector entrance slit acts as such an image plane aperture) and the spectrometer lenses are used to disperse the electrons by energy. One obtains a three dimensional data set, $\mathscr{X} \in \mathbb{R}^{n_y \times n_E \times n_a}$, information along the second spatial mode is lost. A benefit of this scheme is the applicability also to atomic resolution imaging, the SNR is greatly improved by the implicit spatial averaging and possibility of prolonged measurement times. However, one is still required to measure the same sample area several times. Furthermore, dedicated equipment with non-standard aberration correctors (cc-correctors) is needed in order to avoid some artifacts occurring in the energy dispersive plane.

EFTEM-EMCD: If a signal can be measured by placing a collection aperture in STEM mode, it is similarly contained in the energy filtered diffraction pattern (EFTEM). Hence, acquisition of an EFTEM series across the magnetic species edge obtains a three dimensional data set, $\mathscr{X} \in \mathbb{R}^{k_x \times k_y \times n_E}$, where k_x and k_y are the k-space grid, which contains the EMCD spectra at the local-

ization determined by the sample structure and orientation. While the EMCD signal in EFTEM-EMCD can be obtained from a single EFTEM series, the restriction to a single site looses the spatial resolution. Naturally, convergent beams can be used to measure EFTEM series at neighboring positions, which recovers the spatial information. As discussed in paper XV, a 4D-STEM measurement employing an energy filter can be used to obtain an EMCD weight map (but not spectrum), across a region of interest.

Vortex filter EMCD: A slightly different Ansatz was presented in [80]. Instead of using the sample as beam splitter, a vortex mask is included in the diffraction plane so that contributions which have undergone magnetic scattering (and have thus obtained an orbital moment) are scattered to different angles than non-magnetic contributions. Using this approach, more general samples can be studies, no specific sample orientation or even crystallinity is assumed. However, experimental limitations, such as a finitely large beam size, render the EMCD signal faint, the vortex filter scheme has not been demonstrated experimentally yet.

Vortex EMCD: Already using a vortex mask, one can similarly create a vortex beam in front of the specimen so that an EMCD signal results [72, 73, 74]. Curse and blessing of the vortex beam approach is the necessity to go to large convergence angles. On the one hand, simulations predict sizable EMCD contributions also at large convergence angles allowing for high resolution measurements. However, production of such highly converged, small, yet simultaneously high quality vortex beams is challenging to be realized experimentally. Furthermore, the inclusion of the vortex mask leads to a substantial loss of beam intensity reducing the SNR (this can potentially be avoided by usage of magnetic needles for the creation of electron vortices). At last, also the vortex EMCD scheme necessitates repeated measurements of the same sample area.

Aberrated probe EMCD: Instead of using vortex beams, the lens system of an aberration corrected microscope can similarly be used to introduce customized aberrations to the electron probe [68]. If the aberrations are suitably tuned to the symmetry of the magnetic lattice, an EMCD signal can arise. Since tuning of aberrations is more feasible than the production of a high quality vortex mask, this EMCD approach has already been demonstrated experimentally. The achievable beam intensities are higher than for the vortex approach, but the expected EMCD signal fractions are lower. As a result EMCD detection remains challenging also here. As before, unless the system studied naturally provides magnetic signals of opposing signs, such as in anti-ferromagnets, repeated scans are necessary.

APR-EMCD: The last method mentioned here is APR-EMCD [67], which achieves sizable EMCD contributions by positioning a rectangular aperture above the systematic row reflections in the backfocal plane of the objective lens for a sample oriented in 3 beam case orientation. Due to the special sample orientation, atomic columns are no longer observed, instead they are

averaged (by projection) to atomic planes. As selected by the aperture, EMCD signals with opposing signs occur on the sides of these lattice planes overcoming the limitation of repeated scans shared by all other methods above, i.e., $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E}$. Furthermore, the rectangular aperture is larger than typically used circular collection apertures, which boosts the SNR. A drawback of this approach is the loss of sample orientation flexibility. The features of interest need to be identifiable also when rotated to three beam case (3BC) orientation.

2.3 Apocryphal signals

So far, only spectral signals have been considered, although it has already been mentioned above how diverse signals in the electron microscope can be. An outlook beyond the canon of this thesis must be mentioned here: Being microscopy's main objective many image manipulation and/or analysis methods have been suggested. Many of the latter have been introduced outside of the microscopy community and can impossibly all be incorporated into this thesis. Since it coincides with the main topic of this thesis, some examples of analysis by BSS methods have to be pointed out, e.g., [84, 85]. The two more recent instrumental developments in STEM are the segmented ADF detector and the fast, pixelated CCD camera.

The former leads to a data set $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_a}$, n_a being the number of detector segments. Currently, no dedicated signal processing strategies have been suggested for these data tensors. The signal is also non-negative, although intensity shifts may be more easily described also with negative amplitudes. In the lack of a data set to try them on, some ideas may be mentioned as open questions: Can ICA identify shift components (which are approximately orthogonal to the mean intensity and thus independent) and be used for an automated differential phase contrast analysis? Can BSS methods generally be used to overcome the necessity to study (artificial) bright and dark field separately by obtaining a *best contrast* component? Given their wide application range it is almost sure that geometric blind source separation can be used to obtain an identification of different phases in the sample.

At last, consider 4D-STEM measurements. Here, a four-dimensional data set $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times k_x \times k_y}$ is measured, which, much like a segmented ADF detector allows for the formation of virtual dark or bright field image and DPC measurements, suggesting that some of the ideas from above are also applicable here. Given that the 4D-STEM technique is still quite recent, not much dedicated signal processing literature has been published. A sole example can be found in [86], where neural networks are used for the analysis of position averaged convergent beam electron diffraction patterns. In paper XV it was identified how ICA can be used to identify EMCD signals in 4D-STEM data. This suggests that BSS methods might also benefit the analysis of these type of data.

3. A selection of data processing tools for a microscopist

After the preceding brief overview of signals we can expect to process, it is now time to consider what we can actually do with them. Naturally, it is again neither feasible nor necessary to give an exhaustive overview here. For a reader interested in a general introduction to basic analysis techniques common to the community textbooks are readily at hand (e.g., [24, 17]). Instead, the focus will lie on factorization approaches for the signal matrices (and tensors), (and) on BSS techniques.

I.e., given a measured signal matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$, which can, e.g., be a matricized EELS tensor where each row contains a spectrum and each column a vectorized energy slice of the tensor, we seek a factorization into k source signals $\mathbf{A} \in \mathbb{R}^{k \times m}$, here the source spectra, and associated scores $\mathbf{M} \in \mathbb{R}^{n \times k}$, here the abundances of the source spectra:

$$\mathbf{X} = \mathbf{MA} \tag{3.1}$$

Note how this matrix factorization model implies a linear mixing model. For tensors and coupled factorization a wider variety of factorization models can be set up. The linear mixing model assumption implies absence of multiple scattering effects for spectroscopic data, which holds for thinner specimen and can be generally achieved by means of a deconvolution in practice. For other applications in imaging, the linear mixing model has not been discussed or tested yet, a point to consider in the future.

In the following, noise removal approaches for spectra and images are discussed followed by a discussion of ICA and NMF, the two most commonly used BSS methods in electron microscopy. Afterwards, geometric BSS methods, much more common in the remote sensing community, are surveyed. Tensor factorization and coupled factorizations are included as outlook going beyond classical matrix factorization approaches.

3.1 Noise removal approaches

As hinted at above, raw data from microscopes can be excruciatingly noisy. Especially when considering high resolution experiments or generally measurements on beam sensitive specimen, where beam damage becomes a limiting factor, the maximal achievable SNR in a single spectrum or image is dictated by the sample studied. The raw data may no longer be directly interpretable, a denoising needs to be performed to appreciate the signal features.

3.1.1 Principal Component Analysis

The workhorse of spectrum denoising is PCA, and thereby likely the most widely applied BSS method in electron microscopy. The method dates back to 1901 [87], but took its time to spread into the electron microscopy community [88], where it lay unnoticed for another decade, until various research groups picked up on the method, first for denoising purpose only in the early 2000s [89, 90, 91, 31]. Since then, PCA has established itself predominantly for spectrum denoising [33, 22], but also for analysis of spectrum images, either by rotation of the factor loadings prior to analysis [32], or by performing PCA only on sufficiently small clusters of the original data set so that the principal components become interpretable [34]. Limitations of PCA for EELS analysis were discussed in [35, 36, 92, 37].

Given the age of the method and spread in various fields, comprehensive introductions can be readily found. Here another description of key ideas of PCA is offered: Thinking in geometric terms, we want to find a cartesian (i.e., orthonormal) coordinate system that describes the data best, i.e., in as few components as possible. This can be achieved by seeking a vector that maximizes the overlap with the data matrix. I.e., given the overlap of the ith component $\mathbf{a}_i$

$$\mathbf{m}_i = \mathbf{a}_i^T \mathbf{X} \tag{3.2}$$

we maximize

$$\mathbb{E}\left[\mathbf{m}_i^2\right] = \mathbb{E}\left[\mathbf{a}_i^T \mathbf{X}\mathbf{X}^T \mathbf{a}_i\right] \tag{3.3}$$

$$= \mathbf{a}_i^T \mathbf{C}\mathbf{a}_i \tag{3.4}$$

and additionally enforce $\mathbf{a}_i^T \mathbf{a}_i = 1$. A greedy algorithm can be formulated, where the data matrix is reduced by the previously computed component

$$\mathbf{X} \leftarrow \mathbf{X} - \mathbf{m}_i \mathbf{a}_i \tag{3.5}$$

ensuring that successively computed components describe variance orthogonal to the already computed space.

Note that one can already identify that $\mathbf{C} = \mathbf{X}\mathbf{X}^T$ is indeed the covariance matrix of $\mathbf{X}$ and all $\mathbf{a}_i$ can be found via its eigenvalue decomposition. Alternatively, a singular value decomposition (SVD) contains the same information, albeit delivers simultaneously eigenvectors (or rather singular vectors) of row and column space, $\mathbf{U}$ and $\mathbf{V}$, respectively:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{3.6}$$

Different algorithms have been proposed to compute a PCA. PCA via variance maximization was discussed early by Hotelling [93], today's algorithm for a SVD dates back to the 70s [94], the Karhunen-Loéve transformation connects PCA to second order statistics [95],[96]. More interesting to the microscopist are probably extensions of PCA to deal with large or corrupted data:

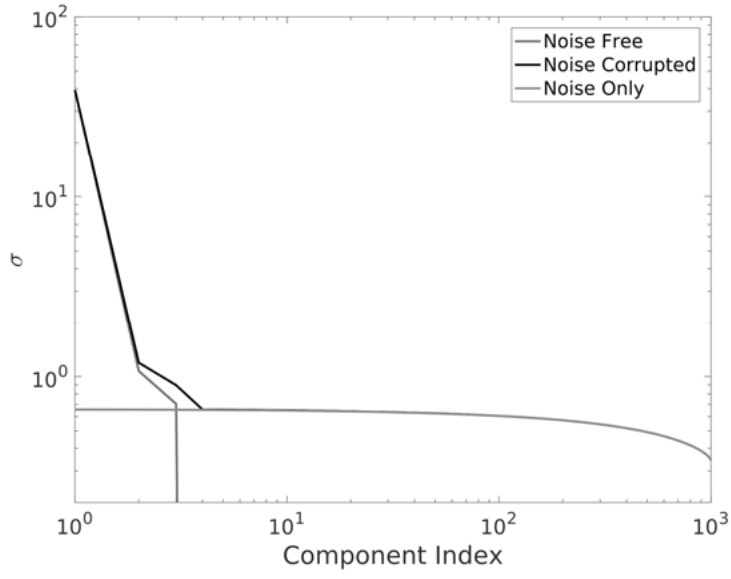

Figure 3.1. Example for a scree plot displaying a kink after the third component revealing the intrinsic dimension of the data. The scree plots of the noise free data and the noise spectrum are indicated as well.

Robust principal component analysis (RPCA, e.g., [97] and references therein) has been formulated to deal with data containing outliers, either via identifying a sparse corruption or reducing the estimation of the principal components to a least outlying subspace. Random PCA (RandPCA, see [98] and references therein) has been formulated to compute PCA also on large matrices in limited time and with limited memory requirements. Other extensions, e.g., to enforce sparsity on the components (e.g., [99]) or to perform PCA online (e.g., [100]) can similarly be found, but are probably of minor interest to microscopists who typically deal with dense matrices in an offline analysis process.

A last important feature of PCA to be discussed is its rank revealing nature. When computing the eigenvalue (or singular value) decomposition of the covariance matrix, one not only obtains the eigenvectors, but also the eigenvalues which indicate the components variance. By convention, the eigenvalues are sorted in a descending manner, a plot of the entire range of eigenvalues results in a plot like Fig.3.1, the so called scree plot. For a noise free data set all eigenvalues but the first n significant ones will be zero. The rank of the data, n, can be read directly from the scree plot. In case noise is included, a (in first approximation) flat noise spectrum is added. Now always of full rank, a truncation of the data is nevertheless meaningful. Assuming that the signal components' variance lies above the noise variance (or rather fulfills certain detection criteria [37]), only the first n components will contain the signal variance. Omission of the remaining components can be used for com-

pression of the data, or most commonly for denoising. The rank of the data can be identified as the point beyond which the spectrum becomes flat. In case no clear kink is observed, care has to be taken. Components may be lost since their variance is too weak, artifacts may be contained in the data, or the linear mixing model may be violated.

3.1.2 Image denoising

Microscopy deals not only with spectra, but also with images, let us briefly consider image denoising. There is a whole ocean of image filters out there, but microscopists tend to be conservative when it comes to their usage. Often, nothing more sophisticated than a Gaussian filter or a Fourier filter is used. The latter often not by selecting only the n largest Fourier coefficients in the reconstruction, but by applying circular masks to selected areas in the Fourier transformed image, or even more straightforwardly by only considering the n lowest frequencies in the image, which is motivated by the instrumental analogy of placing a collection aperture in the backfocal plane of the objective lens. The same philosophy of thresholding of large coefficients is followed when using wavelet transforms for image denoising (e.g., [101] and references therein), but offers a better general applicability also for non-periodic signals.

For electron microscopy, some image filters have been proposed previously [102, 103], but recent efforts have been centered around non-local filters both in the image processing community [104, 105, 106, 107] and the microscopy community [108, 109, 110, 85] or also in paper XI. The key idea here to achieve the denoising by truncation of the expansion coefficients in a transformation the signal is (approximately) sparse in, like in Fourier filtering, but to use the inherent self-similarity of images to gain a denoising via averaging similar patches, computing a low rank approximation of similar patches or to compute a common basis for all patches.

To understand the concept of self-similarity, consider Fig.3.2: An image which has sharp edges and is of high rank can still be approximated by a low number of local environments. After segmenting the image as indicated, similar patches can be automatically recognized via some clustering step. Afterward averaging similar patches for noise reduction, of computing a low rank approximation of a patch class via PCA obtains the denoised patches. Straightforward inversion of the segmentation obtains then a denoised image. Note that typically patches are drawn around every pixel and not only in a rigid grid as indicated in Fig.3.2.

3.2 Matrix factorizations I - ICA and NMF

Sometimes more than a denoising is needed. I.e., Eq.3.1 needs to be solved in a manner that allows not only to separate signal and noise, but to identify

individual components as well. This, however, brings problems, the matrix factorization as formulated in Eq.3.1 is an ill-conditioned problem. Any unitary rotation matrix $\mathbf{U} \in \mathbb{R}^{k \times k}$ can be induced changing the components found without affecting the variance captured by the model

$$\mathbf{X} = \mathbf{MA} = \mathbf{MU}^T\mathbf{UA} = \tilde{\mathbf{M}}\tilde{\mathbf{A}} \tag{3.7}$$

Interpretability of the components is thus easily lost. PCA achieved a unique factorization by constraining the components to be orthogonal and ordered by variance, which is not a physically meaningful constraint in practice. If one of the factors is known, e.g., in the form of reference spectra for EELS unmixing, Eq.3.1 reduces to regression and can be solved with standard methods. Unsurprisingly, this is a popular approach, see, e.g., [111] for a recent example of this referencing approach.

But what if no reference spectra are at hand? Other constraints have to be included to solve the BSS problem in a unique, robust manner, or at least to achieve good convergence into physically meaningful solutions.

3.2.1 Independent Component Analysis

ICA was proposed for EELS analysis in [28], but afterward greatly popularized in the community by the HyperSpy program package [112]. A variety of algorithms has been proposed [113, 114, 115, 116, 117, 118, 119] (or [120, 121, 122] and references therein), but the underlying assumption on the data remains unchanged: As the name suggests, ICA assumes statistically independent variables.

Although this assumption is somewhat opaque for the application to unknown source components, i.e., statistical independence is only translated into

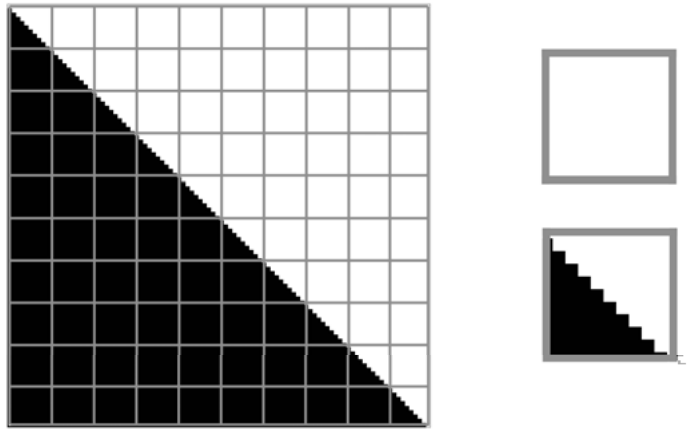

Figure 3.2. Schematic representation of an image with high self-similarity. The entire image consists only of two different blocks as indicated.

constraints on spectral or spatial features with difficulty, some generalities for the application to hyperspectral images can be derived. In [123] it was recognized that spatial correlations as they occur in hyperspectral images greatly impact the applicability of ICA to these data. In [28] and later in [29] it was noticed that the performance of ICA on EEL spectra is greatly enhanced by working with the derivative of the signal instead. As discussed in paper XII, this enhancement is limited to cases without overlapping fine-structures (with selected exceptions).

While the application of ICA is thus not entirely unambiguous, satisfactory unmixing performances have been reported (e.g. [29],[30]). Especially in cases where independence (or orthogonality) of the components can be assumed, ICA is an elegant method to estimate the data's source components.

3.2.2 Non-negative Matrix Factorization

Considering that spectra represent counts, they naturally are non-negative. Hence, it appears intuitive to constrain the BSS problem such that both source components and their scores are non-negative, i.e., to perform NMF. The resulting optimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{M},\mathbf{A}} \quad & ||\mathbf{X}-\mathbf{MA}||_2^2 \\ \text{s.t.} \quad & \mathbf{M} \geq 0 \\ & \mathbf{A} \geq 0 \end{aligned}$$

Unfortunately, the uniqueness conditions of NMF are closely connected to the sparsity of the signal [124, 125, 126]. Since observed hyperspectral images typically are not sparse, uniqueness conditions are more often than not not met.

Trying to deal with this non-uniqueness, one can do an exhaustive search over local minima, which is not elegant, but feasible given that hyperspectral images are often sufficiently small to allow for this procedure. As pointed out in [124], inclusion of additional constraints can alleviate this situation. Additionally enforcing orthogonality or sparsity of the components, or to apply minimum volume constraints improves the uniqueness conditions of NMF. Alternatively, sparsity enhancing pre-processings can be applied (see also paper X).

The limitations imposed by uniqueness of NMF is also reflected in the publications on NMF: Initially focused to achieve an unconstrained NMF [127, 128], constraints were added later. For hyperspectral images, orthogonality constraints can be found in [129], a minimum volume constraint was first proposed in [130] but has sparked development of a multitude of algorithms since. Surprisingly, the microscopy community has still be largely silent on the development of customized NMF algorithms for the analysis of

their data. Besides the work in this thesis, a sparsity constrained algorithm can be found in [23]. By enforcing sparsity on the abundance maps, meaningful source spectra could be extracted.

Stepping back to the considerations in [124], one can recognize that sparsity on one factor matrix behaves similarly to enforcing minimum volume on the other. Can one achieve an unmixing by solely enforcing the minimum volume? Here we step into the field of geometric BSS methods.

3.3 Matrix factorizations II - using geometry

Consider a phase diagram for a ternary mixture. All physically meaningful mixtures (with molar fractions ranging from 0 to 100%) fall within a triangle, a simplex region, formed by the pure substances. Now consider a data set of rank 3. Just as for the ternary mixture, there are three pure source components. As long as we can assume that the scores can be interpreted as *molar fractions*, i.e., sum to one, the pure substances will form the corners of a triangle in which all other, mixed data points fall. Further relaxing the sum to one constraint, the data now fall into a cone spanned by the vectors defined by the pure source spectra. If one generalizes to include more source components, the data will fall into a high-dimensional simplex (or cone).

Can one use these geometric considerations to identify the source components? What if the source components are included in the data set? More importantly, what if not? This is the topic of geometric BSS methods.

3.3.1 Finding pure pixels

First, let us consider the case where the pure source spectra, in the remote sensing community called pure pixels, are contained in the data. In a slight twist of events, the easiest way to identify the source spectra is not by assuming a simplex structure, but by means of classification of different phases in the data. Especially if not only a single pixel can be considered pure in the spectrum image, but an extended region (think of a multilayer system with different chemical composition), automatic identification of these phases (e.g., by a clustering algorithm of your choice) followed by averaging of the corresponding data points gives the source components. Afterward, these components can be fitted to the data to complete the BSS. Maybe this is not the most elegant strategy to unmixing, but it is effective (see paper IV) and useful for identifying different phases in the sample. Since many clustering algorithms utilize distance measures for identifying the different signal classes, this approach fits well into the geometric BSS section.

Now coming back to simplex structure BSS methods, identification of pure pixels was firstly done by inflating (or shrinking) a simplex in the data. Once the maximum volume within the data has been reached, the vertices of the

simplex are the pure source spectra. E.g., the *pure pixel index* (PPI) [131], AVMAX and SVMAX [132], or the N-FINDR algorithm [133] use this approach. A more effective approach was proposed in [134], vertex component analysis (VCA): By first identifying the signal subspace by PCA and the projecting the data onto orthogonal, random vectors and selecting the data points with maximal amplitude along each of these projections, the vertices of the simplex can be extracted efficiently. A robust version of VCA was proposed in paper IV, but also all methods used for pure pixel estimation in data where they are missing (see below) can be used for data containing the source components.

Note how all of the aforementioned approaches only yield estimates of the source components, but not their scores. These have to be estimated in a second step via any regression approach. For the context of hyperspectral images dedicated regression techniques constraining the abundances to be non-negative and to fulfill the sum-to-one constraint have been developed (e.g., [135]).

While geometric BSS methods were mostly developed in the remote sensing community they do not make explicit assumption on the data, rather on their degree of mixing. As such they are straightforwardly applicable also in other disciplines. For an application of VCA to EEL spectra see, e.g., [136].

3.3.2 Estimating pure pixels

Finding the pure pixels in your data suffices for the analysis of some data sets, but may be lacking in other situations given that in BSS we would like to unmix data also in situations where a straightforward analysis by manually browsing through the data and identifying suitable reference spectra is no longer sufficient. I.e., we would like to be able to analyze data sets without pure pixels. Nevertheless, pure pixels methods are valuable for quick analyses of the data (see paper VIII) or automated analyses of large data sets.

If no pure pixels are contained in the data, this amounts to chopping off the corners of the data simplex (or cone). This implies that pure pixel methods will fail, the pure pixels are no longer contained in the data set. But what about the notion to estimate a minimum volume simplex enclosing the data? Even if the corners of the simplex are gone, if the faces a sufficiently populated, the minimum volume simplex is still uniquely defined!

Based on these very considerations, uniqueness conditions for volume minimization approaches can be formulated [137]. The mathematically versed reader is referred back to [137], for others let me paraphrase the conditions: Consider the inner circle of a triangle. If the data not only fill out this circle, but exceed its boundaries within the triangle, the triangle itself becomes the minimum volume enclosing simplex. This considerations can be formulated for arbitrary dimensions and also in terms of cones rather than simplexes, further

generalizing minimum volume constrained methods. Essentially, geometric BSS methods achieve uniqueness not by assuming certain mathematical characteristics of the source components, but by assuming a certain minimal purity (degree of mixing) in the data. The more components are in the data, the more mixed they can occur. As long as the facets of the data simplex (cone) are populated, the vertices (extreme rays) can be identified.

A variety of algorithms working with this minimum volume approach has been proposed. Algorithms working purely with convex geometry are, e.g., minimum volume enclosing simplex (MVES, [138]), minimum volume simplex analysis (MVSA, [139]) or also the VolMin algorithm of [137] and other algorithms in the remote sensing community. However, minimum volume constraints have also been incorporated in different settings. E.g., Bayesian linear unmixing (BLU, [140]) has been proposed to estimate minimum volume simplexes with additionally enforced non-negativity in a Bayesian setting, which is very similarly constrained as the multitude of minimum volume constrained NMF algorithms which have been proposed, e.g., MVC-NMF [130] or (R)CoNMF [141].

Recent works take steps beyond the pure minimum volume constrained. E.g., extreme direction analysis (EDA, [142]) wagers a minimum volume term against a sparsity measure of the source components, which achieves good performances for the unmixing of natural images, suggesting the methods usefulness especially for microscope data at low magnifications. In [143], inspiration was taken from minimum volume constraints to propose an algorithms to uniquely identify non-negative source components with an arbitrary mixing matrix, which generalizes geometric BSS methods further to be also applicable to data with non-negativity destroying pre-processings, which may open further possibilities to boost unmixing performances.

In the electron microscopy community, geometric BSS methods are not yet very wide-spread. In [136], VCA and BLU were demonstrated on an EELS data set, paper IV further includes MVSA. Some applications of VCA and BLU can be found in [84, 144]. This is somewhat surprising given that they have explicitly been designed for hyperspectral images and do not have explicit constraints on the spectral shape of the source components, which makes them applicable to any not too strongly mixed data set, regardless whether it is CL, EDX or EELS.

3.4 *Matrix* factorizations III - beyond matrices

Consider Eq.3.1: We have already at this first step limited ourselves to the analysis of three- (or multi-) dimensional data sets in a matricized form where we obtain spectral source components and associated weights. But what is gained by this matricization? Could it vice versa be beneficial to omit the matricization and instead factorize the tensor directly? This is the topic of this

section, factorization approaches beyond matrices, i.e., tensor factorizations and coupled factorizations.

Before starting to introduce different factorization models, it is useful to introduce the n-mode product:

$$\mathscr{A} \times_n \mathbf{B} = \sum_{i_n} a_{i_1 i_2 \ldots i_{n-1} i_n i_{n+1} \ldots i_N} b_{j_n i_n} = \mathscr{C}, \tag{3.8}$$

The multiplication of the tensor $\mathscr{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_{n-1} \times I_n \times I_{n+1} \times \cdots \times I_N}$ with the matrix $\mathbf{B} \in \mathbb{R}^{J_n \times I_n}$ yields the tensor $\mathscr{C} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N}$. While somewhat abstract at first, this notation may become more clear when rewriting the classical SVD, Eq.3.6 using the n-mode product. The SVD then reads:

$$\mathbf{X} = \mathbf{S} \times_1 \mathbf{U} \times_2 \mathbf{V} \tag{3.9}$$

I.e., the components stored in the columns of $\mathbf{U}$ are multiplied with the first mode (the rows) of $\mathbf{S}$, the columns of $\mathbf{V}$ are multiplied with the second mode (the columns) of $\mathbf{S}$. It is clear how this notation allows for generalizations also to higher dimensions.

3.4.1 Tensor factorizations

Comprehensive reviews of tensor decomposition models can be found [145, 146, 147]. Despite their relative age, they have not yet been used much for the analysis of hyperspectral images. For the context of microscopy data, paper V is an exception, for remote sensing images, non-local tensor models have been proposed for denoising and classification (e.g., [148, 149]), a dedicated non-negativity constrained decomposition model was proposed in [150]. Here, only three tensor decompositions are considered, i.e., the Tucker decomposition (TKD), the canonical polyadic decomposition (CPD) and the block term decomposition (BTD). Note, however, that many more models have been proposed, such as tensor trains (TT) [151, 152], ...

Tucker Decomposition

Let us first consider the TKD. The TKD model of a tensor $\mathscr{X} \in \mathbb{R}^{n_1 \times n_2 \times \ldots}$ reads

$$\mathscr{X} = \mathscr{S} \times_1 \mathbf{A}_1 \times_2 \mathbf{A}_2 \times \ldots \tag{3.10}$$

where $\mathscr{S} \in \mathbb{R}^{k_1 \times k_2 \times \cdots}$ is the so called core tensor and $\mathbf{A}_i \in \mathbb{R}^{n_i \times k_i}$ are the components matrices of the ith mode.

For the case of three-dimensional hyperspectral images, a TKD would not only yield spectral source components, but also source components for the spatial features of the tensor. Abundance maps $\mathscr{M}$ of the matrix factorization can, however, be retained (assuming the third mode is the energy mode):

$$\mathscr{X} = \mathscr{S} \times_1 \mathbf{A}_1 \times_2 \mathbf{A}_2 \times_3 \mathbf{A}_3 \tag{3.11}$$

$$= \mathscr{M} \times_3 \mathbf{A}_3 \tag{3.12}$$

Importantly, where before all components of the spatial components were considered in the factorization model, one now has the opportunity to truncate the model further, which may help to suppress the influence of noise or lead to better convergence of the extraction also of spectral features.

The TKD dates back to 1966 [153], where it was computed without imposing any constraints onto factors or core tensor. This lack of constraints unfortunately leads to poor uniqueness properties. I.e., it has to be assumed that the TKD is not unique, multiple computations may be necessary to find a global optimum, which is typically only found up to rotation [154], which complicates the interpretation of the obtained factors. Furthermore, the TKD model has unfavorably scaling memory demands for very high-dimensional data. For the typically 3 to 4 dimensional tensors of microscopists, however, this should not be a limiting factor.

Constrained TKDs may naturally have better uniqueness properties. Non-negativity and sparsity constrained algorithms have been formulated (e.g., [155, 156]), imposing orthogonality constraints on both factors and core tensor leads to the multilinear singular value decomposition (MLSVD) or also higher order singular value decomposition (HOSVD) [157], which is unique in absence of degenerate eigenvalues along any mode. As compared to the classical matrix SVD in Eq.3.9, additional factors are simply added, the core matrix is replaced by a core tensor (which is no longer (super)diagonal, but orthogonal).

This uniqueness of MLSVD makes it a valuable denoising tool, but direct analysis of factors is prohibited due to the orthogonality constraint. As compared to PCA denoising an additional subtlety occurs: Where before the number of components was truncated to achieve a denoising, one now stands in front of the dilemma that one either can truncate the number of components along all modes, or the maximum number of core components to the n largest amplitudes (or both). The latter truncation reminds of denoising via thresholding in sparsity promoting transformations, the former amounts to a limitation of the *multilinear rank*, the rank of the matricized tensor for an unfolding along the respective mode, as in contrast to the *tensor rank*, the number of rank-1 tensors that capture all information of the original tensor.

Canonical polyadic Decomposition

The tensor rank closely connects to the CPD, first introduced in [158] although the name CPD first occurred in [159]. A rank-1 tensor is a tensor which is described by a single component along each of its modes. A CPD is then a factorization into rank-1 tensors:

$$\mathscr{X} = \sum_i \mathbf{a}_i \times_2 \mathbf{b}_i \times_3 \mathbf{c}_i \times \dots \tag{3.13}$$

$$= \mathscr{I} \times_1 \mathbf{A}_1 \times_2 \mathbf{A}_2 \times_3 \mathbf{A}_3 \times \dots \tag{3.14}$$

where the latter equation with $\mathscr{I}$ as tensor which is 1 along the superdiagonal and 0 otherwise highlights the connection between CPD and TKD.

Important differences between CPD and TKD are that the estimation of the number of components becomes much harder for CPD, as estimation of the tensor rank can only be done via tracking the reconstruction error with increasingly large ranks, which may be costly for large tensors. On the other hand, CPD promises more relaxed uniqueness conditions by only imposing assumptions about minimal ranks of the factor matrices and not on other mathematical properties of the factors itself. The earliest conditions were formulated in 1977 [160], but increasingly relaxed conditions have been found since, more recent ones include [161],[162].

Similar as for the TKD, a CPD model can be reformulated to resemble Eq.3.1:

$$\mathscr{X} = \sum_i \mathbf{a}_i \times_2 \mathbf{b}_i \times_3 \mathbf{c}_i \tag{3.15}$$

$$= \sum_i \mathbf{M}_i \times_3 \mathbf{c}_i \tag{3.16}$$

However, the implicit constraint that abundance maps of each source spectrum are of rank 1 is likely not met by many samples in the electron microscope. However, if the source data have additional modes so that the spatial modes can be unfolded into a single mode, CPD can become a useful tool, as discussed in section 4.2.

Block Term Decomposition

The last factorization considered here is the BTD [163, 164, 165]. In some sense, BTD weds the notions of multilinear rank and tensor rank and thereby TKD and CPD by offering a decomposition into a series of Tucker models:

$$\mathscr{X} = \sum_i \mathscr{T}_i = \sum_i \mathscr{S}^{(i)} \times_1 \mathbf{A}_1^{(i)} \times_2 \mathbf{A}_2^{(i)} \times \ldots. \tag{3.17}$$

Naturally, this can be recast to obtain abundance maps and source spectra for application to (three-dimensional) hyperspectral images

$$\mathscr{X} = \sum_i \mathscr{S}^{(i)} \times_1 \mathbf{A}_1^{(i)} \times_2 \mathbf{A}_2^{(i)} \times_3 \mathbf{A}_3^{(i)} \tag{3.18}$$

$$= \sum_i \mathscr{M}^{(i)} \times_3 \mathbf{A}_3^{(i)} \tag{3.19}$$

Much as the TKD, a general BTD offers no unique solution, becomes computationally demanding to estimate and requires large memory allocations for larger tensors, which makes the method less attractive. However, in [166] it was recognized that less strict uniqueness criteria apply for a decomposition (of three-dimensional tensors) into $(L \times L \times 1)$-terms, i.e., tensors of rank L along two modes while having rank 1 along the third. This LL1-decomposition can be considered a relaxed form of the CPD allowing also for more complex patterns under the assumption that the joined factor matrices are of full rank.

This feature may be useful for the analysis of some hyperspectral images with features of low rank, but not necessarily of rank 1 so that a CPD is not sufficiently flexible to capture the signal information.

3.4.2 Coupled factorizations

The next generalization moves from single matrix or tensor decomposition to coupled factorization. Consider, e.g., a joined EELS/EDX measurement, at every pixel of the STEM image both spectra are measured. One thus obtains two tensors whose spectral modes are quite different, but since both spectra originate from the same pixel, one can expect that the spatial structures of both tensors are highly correlated. If one would then analyze the tensors in matricized form, it would make sense to assume that some (or all) of the spatial factors are shared between the two factorizations:

$$\mathbf{X}_{EDX} = \mathbf{S}\mathbf{A}_{EDX} \qquad \text{and} \qquad \mathbf{X}_{EELS} = \mathbf{S}\mathbf{A}_{EELS}$$

One can then further imagine a coupling of different matrix or tensor factorization with different factor matrices (partially) shared, with different constraints on either factor. Can factors be estimated more accurately? Does one get better uniqueness results?

There are different approaches in which to deal with such coupled factorizations. E.g., a time series measurement of 2D data could be cast into a tensor decomposition so that the factors of individual 2D measurements are shared across all times. Dealing with independence constraints, different joint BSS approaches were formulated, see [167, 168] and references therein. Here, the discussion is constrained to discuss the PARAFAC2 decomposition [169, 170] as well as the coupled matrix and tensor factorization (CMTF) [171] and structured data fusion (SDF) [172, 173].

PARAFAC2

PARAFAC2, the name deriving as extension to PARAFAC (parallel factor analysis, a different name for a CPD), decomposes a number of slabs $\mathbf{X}_k \in \mathbb{R}^{n_k \times m}$ jointly into

$$\mathbf{X}_k = \mathbf{E}_k \mathbf{D}_k \mathbf{S}^T \tag{3.20}$$

where $\mathbf{D}_k$ are diagonal matrices. In the context of microscopy, joined spatial factor matrices can be derived for hyperspectral images (or any other measurement in STEM where several detectors yield signals at the same pixel).

Sparing the reader details of the computation of the model (can be found in [170]), an important feature of PARAFAC2 is its uniqueness criteria. I.e., a sufficient conditions for a unique result is that the correlation matrix of the detector channel component, $\mathbf{E}_k^T \mathbf{E}_k$, is constant along all k and the rank of the model R fulfills the following condition (K being the total number of slabs) $K \geq \frac{1}{24} R(R+1)(R+2)(R+3)$ [174]. Simulations in [170] even indicated

uniqueness for larger R if $K \geq 4$, falling much below this sufficient condition, also admitting the hope for a tighter bound in future works.

The assumption that $\mathbf{E}_k^T \mathbf{E}_k$ stays constant for all k is somewhat restrictive for the application to coupled STEM data. E.g., the spectral components of a coupled EELS/EDX measurement likely defy this assumption due to their largely different mathematical characteristics, although their spatial components may be exactly shared. On the other hand, the lack of restrictions on the exact shape of $\mathbf{E}_k$ admits the possibility that PARAFAC2 is applicable across data from many different sensor types.

For some applications, the $\mathbf{E}_k^T \mathbf{E}_k$ constancy even appears plausible: E.g., for coupling of EEL spectra measured at different energy dispersions, or for monitoring of a fine-structure component across a time series, or cases where $\mathbf{E}_k^T \mathbf{E}_k$ becomes block-diagonal so that a component of interest (or a subspace of interest) can be extracted reliably.

Coupled Matrix and Tensor Factorization - Structured Data Fusion

CMTF and SDF were proposed separately and indeed differ in their algorithms and intent. However, a joining feature is that in both cases the optimization criterion is formulated as sum over l_2 (or Frobenius) norms of the residuals between data matrices and their models. The models are then formulated thus that several factor matrices are shared between the individual terms of the optimization objective, as the user defines. Furthermore, individual constraints can be applied to individual factors, further generalizing the model.

While CMTF and SDF are thus fully general tools meeting the requirements of a microscopist, they bear two disadvantages: First, uniqueness conditions vary for any specified model. Likely not generally unique anymore, inclusion of customized constraints largely benefits the convergence of the model into meaningful solutions. Second, this need to minutely tailor the model to the data at hand requires a dedicated user. It is not possible to use CMTF or SDF as blackbox algorithm to be applied to a (general) class of problems. This is not a conceptual problem of the approach, but affects its attractiveness to a user less used to identifying suitable model structure and constraints for a given application.

4. Looking for the holy grail

After the preceding review of signals in the microscope and different matrix and tensor factorization approaches, it is now time to put them to practice and find the EMCD signal. I hope, reader, I will be able to convince you that the EMCD is much like the holy grail elusive, but, unlike it, real.

In section 4.1 some considerations on how to find weak signals in noisy hyperspectral images are presented. These findings are then applied to propose measurement and extraction strategies for EMCD signals in section 4.2, which is followed by a demonstration of the proposed schemes in section 4.3. After this climax, section 4.4 is dedicated to selected (un)related approaches for different data processing problems in electron microscopy as a bridge into the thesis' coda.

4.1 Finding signals in noisy hyperspectral images

A common problem is that whatever signal one is interested in has a low variance (or energy for non-zero means) and is thus obscured by the noise. For the case of hyperspectral images in microscopy, simple strategies to improve upon this situation are to change the measurement conditions if possible, to measure the same sample area repeatedly to average the thus obtained data sets into a single high-quality set, or to denoise the data. The former is clearly a safe choice, but we will here assume that the SNR in individual spectra has already been optimized omitting this path. Averaging of identical data (up to the noise realization) places demands on the microscope hardware to be able to deal with larger amounts of data and works with implicit assumptions that nothing changes over the course of acquisition, which (with some post-processing, i.e., spatial alignment of the data) is a good assumption for static structures, but is not suitable for any measurement where time is a variable. Similarly, if accidental or unexpected structure changes occur, these will deteriorate the resulting analysis without the operator noticing it. Lastly, a denoising retains more control over the data, but is limited to a maximal denoising (under some assumptions) depending on the original SNR in the raw data.

Here, the latter path is explored, following the content of papers III, V, VIII, XI.

4.1.1 Detecting signals with PCA

Already in section 3.1, PCA was introduced as a denoising tool for spectral data. However, to fully understand its potentials and limitations, it is necessary

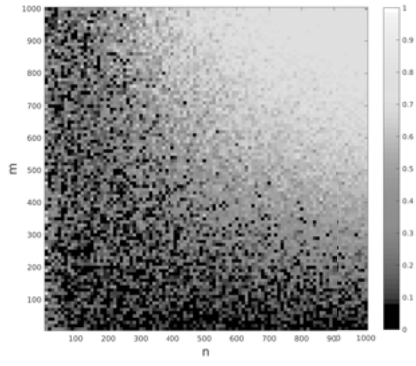

Figure 4.1. Dot-product between true and recovered components of a uniform random variable retrieved from data with fixed noise variance using PCA at different matrix sizes n and m.

to consider the boundaries beyond which components can not be retrieved by PCA. To understand the asymptotics of PCA, it is useful to consider the bias of eigenvalues in PCA. Already in [175] it was identified that PCA of a data matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$ gives rise to the following bias b_λ in each eigenvalue λ:

$$b_\lambda = (n+m)\sigma^2 \tag{4.1}$$

where σ^2 is the noise variance. This equation looks troublesome at first glance: The larger the data matrix, the larger the bias of the PCA model! However, when considering reconstructions, one needs to consider the bias normalized by the size of the data matrix in order to assess the error at any point of the matrix. One obtains:

$$\tilde{b}_\lambda = \left(\frac{1}{n} + \frac{1}{m}\right)\sigma^2 \tag{4.2}$$

which reveals a much more pleasing asymptotic. For infinitely large matrices, PCA recovers the true data exactly, unaffected by the noise level. For finite sizes, especially in situations where $n \gg m$, the error is determined by noise variance and size of the other mode alone. The description of [175] gives correct results only asymptotically, a more accurate description also around the information loss threshold is given in [176]. A cutoff criterion is found beyond which components are lost:

$$\frac{m}{n} \geq \left(\frac{\sigma^2}{\lambda^*}\right)^2 \tag{4.3}$$

where λ^* denotes eigenvalues normalized by m. Since λ^* still scales with n, one thus obtains a cutoff where n and m enter symmetrically.

This behavior is also observed when retrieving a uniformly distributed random variable at constant noise level using PCA for data with different n and m

(see Fig.4.1). A cutoff can be observed beyond which the component is lost. This cutoff is symmetric with respect to n and m where these variables are of approximately same size. For tall matrices the asymptotics derived from Fabers algorithm is reached, since there the cutoff is determined by only the smaller mode size.

But what does this imply for the application of PCA to hyperspectral images? The number of energy channels is usually fixed by the detection system, the noise variance (or rather the SNR) can be considered to have been optimized and is also fixed. The number of pixels, however, can be varied at will. Hence, PCA dictates a detection limit based on the optimized SNR and number of detector channels. Given that signals typically spread over several 100 channels, this implies that the minimum signal variance can be a factor 100 to 1000 below the noise variance, which suffices for most applications. Hence, if the microscopist only remembers to measure enough spectra, PCA is an easy to use tool to estimate source spectra and their scores.

A drawback is the remaining noise in the components. Many pixels imply almost noise free spectra, but a finite (small) number of energy channels leads to considerable noise in the scores. PCA nicely detects spectral shapes of the components, but not their localization. Instead, one needs to average the denoised spectra spatially, filter the score maps, or exploit further correlations in the data set.

4.1.2 Mapping signals with MLSVD and LLR

An obvious strategy to improve on the denoising of PCA is to denoise the score maps further using any state-of-the-art image denoising algorithm, which is, coincidentally, one of the most recently published strategies for denoising of hyperspectral images [177]. There is no use in discussing performance of different image denoising algorithms here, instead, the questions on how low-rank approximations can be extended to spatial correlations will be discussed. This has the advantage that it generalizes also to multi-modal data straightforwardly.

Moving from purely spectral denoising to spectral-spatial denoising amounts to moving from matrix denoising to tensor denoising. As discussed in section 3.4, MLSVD is a good candidate for tensors whose spatial structure allows for a low rank approximation. Due to the inherent uniqueness properties a near optimal denoising can be achieved reliably also at low SNRs, truncation of multilinear ranks or active core elements followed by reconstruction of the full tensor is a direct route to denoise the data further beyond the limits of PCA. Examples can be found in papers V and VIII.

Despite this progress, MLSVD is still limited to tensors which allow for a (global) low rank approximation. If modes do not allow for this compression, like more complex spatial structures in hyperspectral images do,

MLSVD denoising approaches the denoising of PCA and little is gained. A strategy to alleviate this dilemma was found in paper XI: While a noisy tensor $\mathscr{X} \in \mathbb{R}^{I_1 \times \cdots \times I_i \times \cdots \times I_N}$ might have a high multilinear rank along the ith mode prohibiting MLSVD to give a good denoising, local segments of this mode may allow for a low rank approximation. I.e., when drawing subtensors $\mathscr{X}_{sub} \in \mathbb{R}^{I_1 \times \cdots \times n_i \times \cdots \times I_N}$ along every pixel in the ith mode and stacking these in an additional mode to $\mathscr{Y} \in \mathbb{R}^{I_1 \times \cdots \times n_i \times \cdots \times I_N \times M_i}$, one obtains a tensor which is small or compressible along all modes. MLSVD can now be expected to give good denoisings independent of the structure along the ith mode. Some further subtleties for an optimal denoising of $\mathscr{Y}$ are discussed in paper XI.

Using MLSVD and LLR one can thus extend the low-rank denoising of PCA also to the full tensorial hyperspectral image or even further multi-modal data (time series,...). A data processing workflow to both detect source spectra and their localization is thus to apply PCA to sufficiently many measured spectra (also from multiple data sets) and then denoise the remaining modes further using tensor factorizations. If the spatial structures are of low rank, MLSVD can achieve excellent denoising performances. If the spatial structures of the hyperspectral image do not allow for a global low-rank approximation, one utilizes the LLR segmentation to segment both spatial modes and applies LLR on the stacked tensor, again achieving a low-rank approximation, now of the local spatial-spectral features.

4.2 Learning how to find EMCD

Having learned how to identify signals with small variance, we have now tools to detect also faint signals - as needed for EMCD detection. However, one additional step has to be taken: Instead of merely denoising the data faithfully, the EMCD spectrum has to be extracted. Here, three extraction strategies are listed. A fourth one fits into the introduction of this section:

In a classical EMCD measurement, but also any other EMCD measurement where the same sample area is measured multiple times resulting in data sets with opposing EMCD signs, the spectra can be background subtracted and normalized in the post-edge. Subsequent subtraction of the data sets with opposing EMCD signs yields a data tensor only containing the EMCD signal. This direct difference (DD) procedure is straightforward, but sensitive also to other fine-structures contributing to differences between the data sets. E.g., if some sample damage occurs or contamination is building up on the sample between the scans, this DD approach may fail.

Hence, consider how DD can be automatized and robustified: I.e., the content of papers XIV, XV.

4.2.1 Geometric blind source separation for EMCD

Given that the DD procedure works, geometric BSS methods are a natural candidate to automatize this analysis. Instead of subtracting the data with opposing EMCD contributions form each other, they can be joined in a single data set. Methods constrained to estimate pure pixels can then identify spectra carrying the positive and negative EMCD contributions, taking the difference between these spectra yields the EMCD signal, the difference of their abundance maps the EMCD structure.

A subtle detail needs to be considered: If pure pixel estimating methods are used which base their estimation on a single spectrum, e.g., (R)VCA, the noise in the data set will lead to an overestimation of the EMCD amplitude scaling with the noise variance. As the noise gives rise to random shifts of the data inside the simplex, it effectively increases the simplex volume and thus the amplitude of the data features. This spurious amplitude increase is less noticeable when using clustering techniques to identify different phases in the sample. Here, source spectra are estimated as averages over extended regions, thus attenuating the noise.

The advantage of using geometric BSS methods over a DD approach is not only the automatization aspect, but also that the approach generalizes straightforwardly to single data sets already containing different EMCD contributions like in APR-EMCD or measurement of magnetic structures across domain walls or in different grains. Furthermore, if some regions of the data set are degraded by some unwanted components, the BSS methods will pick up uncontaminated components from other parts, thus also alleviating this problem.

Usage of pure pixel estimating algorithms does typically give no improvement for the task of identifying different phases in the data, as these (by definition) occur pure. However, an interesting approach is to apply minimum volume constrained algorithms to mean subtracted data sets. I.e., if some other component occurs between the data sets, this components may not appear highly mixed, but only as small perturbation (while dominating in other parts of the data set). Using, e.g., MVSA to estimate pure pixels in this data set may extrapolate the EMCD spectrum beyond this unwanted component.

A drawback of this approach is that the underlying assumptions of MVSA, in particular non-negativity of the abundances, is violated for a data set containing EMCD signals with opposing signs. Estimating EMCD signals of opposing signs, however, leads to rank-deficiency also rendering the algorithm useless. Nonetheless, for some cases satisfactory unmixing was achieved. Future geometric BSS approaches pushing beyond the limits of MVSA may further improve this approach.

4.2.2 EMCD and ICA

The EMCD signal is typically correlated to the non-magnetic signal, independence is thus violated and ICA seems like a bad idea. However, it has to be remembered that while microscopists tend to forget that BSS methods can not only be used to estimate source spectra, we can similarly attempt to unmix the EMCD signal by assuming statistical independence of its scores from the scores of other signals.

I.e., if the data is represented as a four-dimensional tensor $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E \times n_a}$, n_a may be too small to allow for unmixing of all spectral components due to rank deficiency (if n_a is smaller than the multilinear rank along the energy mode), but matricization into $\mathbf{X} \in \mathbb{R}^{n_x \cdot n_y \cdot n_a \times n_E}$ may allow for extraction of the EMCD component since it is (ideally) the only component changing with n_a. Its sign flip between different n_a ideally makes it orthogonal (and thereby independent) of the other signals. If other components varying with n_a occur (asymmetry, contamination, sample damage, ...), independence might still hold (allowing for EMCD extraction unaffected by artifacts!), or be destroyed necessitating other approaches.

For data sets including different EMCD signs in real space, ICA should similarly be applicable. E.g., the shift of non-magnetic and EMCD fluctuations in APR-EMCD should ideally make the signals independent from each other. However, I observe that imperfections of the data (noise, drift, ...) often do not permit this extraction to be successful.

At last, EFTEM EMCD should be considered. As pointed out in paper XV, the k-space maps of EMCD are orthogonal to their non-magnetic associates. ICA is thus a method well suited to extract the EMCD contributions regardless of sample orientation and magnetization direction.

4.2.3 Extracting EMCD straight from the tensor

The last strategy included here is to get the EMCD via a tensor decomposition method. Given that the spatial structures are likely of high rank, it makes sense to unfold these two modes into a single spatial mode, i.e., $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E \times n_a}$ is transformed to $\mathscr{X} \in \mathbb{R}^{n_x \cdot n_y \times n_E \times n_a}$. Can a CPD be used to extract pairs of spectral source components with associated abundances (vectorized) and weights at different aperture positions?

A possible obstacle is that n_a is typically small (i.e., 2 or 4). Hence uniqueness criteria of CPD are better fulfilled if we keep the ranks of the other modes also small. This can be achieved by a suitable pre-processing of the data. E.g., data can be background subtracted, normalized and mean subtracted (each reducing the rank of the energy mode by typically 1). CPD would likely be even better suited to analyze data sets where at every pixel of a $k_y \times E$ maps is acquired by projecting the diffraction plane onto the spectrometer entrance slit, giving a data set $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E \times k_y}$ (then transformed to $\mathscr{X} \in \mathbb{R}^{n_x \cdot n_y \times n_E \times k_y}$).

Since $k_y \gg a$ and the EMCD signal distribution along k_y allows for an easy separation of the non-magnetic signals (for some sample orientations and k-space slit placements), CPD will likely achieve a correct unmixing more easily.

This scheme was already successfully tested on simulated data in my licentiate thesis [178], it has not yet been published and tested also on experimental data.

4.3 Finding EMCD

Oh reader, you have reached the strangely anti-climatic climax of this thesis, the demonstration of the proposed processing strategies on experimental data. However, if you have read attentively until here, you may skip this section altogether, for it can be summarized in two words: It works.

4.3.1 EMCD with abberated probes

This demonstration refers back to the contents of paper I. Aberrated electron probes were used to induce EMCD contributions in the antiferromagnet LaMnAsO, experimental details can be found in paper I.

In short, this project demonstrates the usage of PCA for detection of EMCD. I.e., pre-processed (spike removal, alignment of energy axis,...) and then subjected to PCA. By performing PCA over multiple data sets, the number of spectra was increased to approx. 300000 allowing also faint components to be detected. Selecting components according to the now clear scree plot resulted in essentially noise free spectra. The remaining noise in the scores, however, prohibited a direct interpretation of the spatial maps regardless which extraction strategy was chosen. Thus, we resorted to averaging spectra according to the known magnetic structure, thus obtaining two averaged spectra with opposing EMCD signals, DD yielded the magnetic spectrum.

4.3.2 APR-EMCD

Extraction of the EMCD signal from APR-EMCD signals was addressed in detail in paper XIV. Geometric BSS methods as well as tensor decomposition can be used to extract the EMCD signals here.

Since the EMCD signal changes sign within the same data set, no repeated measurements of the same sample area are needed to obtain the signal, pure pixel methods can identify the phases with opposing signals, both clustering and VCA perform well. A successful data processing workflow is thus the following: Pre-processing (spike removal, alignment, background removal, optionally normalization of the post-edge), PCA jointly on several data sets to obtain the spectral source components, denoising of the score maps by LLR (or MLSVD on aligned data), geometric BSS (or tensor decompositions).

As demonstrated in paper XIV, geometric BSS methods show an excellent performance here, with a LLR denoising the weights of the EMCD signal in single pixels can be investigated. For an ideal data set without distortions, CPD is not applicable since one mode of the tensor is only of rank 1 creating problems with the uniqueness of CPD. PARAFAC2 may be an attractive candidate for the analysis of these sets as well, otherwise a Tucker model was demonstrated to extract qualitatively correct components. Corresponding analyses are contained in paper XIV.

4.3.3 EFTEM-EMCD

Since the k-space distribution of the EMCD signal is orthogonal to (and thus independent of) the non-magnetic signal, ICA can be used to distinguish the k-space components of EMCD for various sample orientations.

This has been demonstrated in paper XV. The data were pre-processed similarly to STEM-EMCD, i.e., spikes were removed, a background model was estimated and subtracted. Afterward, ICA was applied to the centered data's k-space basis (estimated by PCA). The EMCD components can be identified. However, a quantitative analysis is so far prohibited. Due to artifact contributions from the camera and microscope instabilities, the observed rank of the data exceeded the expected number. These artifact components are not necessarily independent of the EMCD signal, thus decreasing the unmixing performance. Future measurements on more stable microscopes may improve this approach. In particular in combination with recent 4D STEM measurements, this may open an attractive road to study non-trivial spin arrangements in the microscope.

4.3.4 DD-EMCD

The DD approach has been used one way or the other in almost all measurements of EMCD so far. An example can be found in paper VII. The data were subjected to the standard pre-processing (spike removal, energy axis alignment, background removal, post-edge normalization), the data sets with opposing EMCD signs were subtracted from each other, the resulting tensor only contains EMCD signal and possibly other contaminations from changing contributions between the scans.

4.4 What has been learned from the quest for EMCD

At last, miscellaneous (un)related approaches deserving mentioning are listed in this section. How do the considerations above apply to other problems? More importantly, how can they be extended? The former question is easily

answered: The detection of any fine-structure of interest follows essentially the same steps, denoising and extraction methods have anyway been developed not for EMCD measurements, but a much wider class of problems. The latter question is the topic of this section.

4.4.1 A faster way of finding pure pixels

Already in paper IV, the computational efficiency of different geometric algorithms was compared. Clustering was identified to be a time and cost efficient step for the extraction of pure pixels in the data set. However, the speed determining step still lies in the computation of the SVD. Can this step be sped up further?

Re-reading section 3.1, one quickly finds randomized PCA as opportunity for further reducing computation time (and RAM needs). Substituting the SVD step by a random SVD (see [98] for some algorithms) reduces the computational complexity of the SVD from $\mathcal{O}(mn^2)$ to $\mathcal{O}(mn\log(k))$ (k being the number of estimated components), thus promising a considerable speed up.

Already considering randomized algorithms, a randomized interpolative decomposition (rID), e.g., $\mathbf{X} \in \mathbb{R}^{n\times m}$ with rank k is decomposed by identifying k columns of $\mathbf{X}$ to satisfy

$$\mathbf{X} = \mathbf{X}_{\cdot,j}\mathbf{M} \tag{4.4}$$

where $\mathbf{M} \in \mathbb{R}^{k\times m}$. While IDs have made their impact mostly in the compression of large matrices, the parallels to pure pixel estimation are apparent. To deal also with large matrices efficiently, randomized algorithms have been proposed [179, 180].

Lastly, taking inspiration from such techniques for matrix compression, greedy algorithms for the submatrix selection can perform at even lower costs and higher speeds. E.g., the algorithm published in [181] promises a reduction of computation costs to $\mathcal{O}(mk^2)$!

To do a quick proof of principle, let us consider a simplistic data set: 5 normally distributed random variables with 1000 observation channels which occur in 5 *phases* in the sample as pure pixels, i.e., the total of 50000 simulated spectra can be distributed into 5 times 10000 pure spectra. Afterwards applying VCA, clustering using a rSVD (with a gaussian matrix as initial compression step), an ID (no randomized algorithm used) and the algorithm of [181] retrieves the correct components in all cases. The Amari indices are below 10^{-10} in all cases, further underlining the good unmixing (Amari indices are vanishing for exact unmixing and increase with a decreasing unmixing performance). However, computation times differ: Normalized to the VCA computation time, the times are 13.2 for ID, 1.0 for VCA, 0.09 for clustering following rSVD, 0.05 for the greedy column selection.

The speed of the ID is not impressive. Since no randomized algorithm is used, the speed hinges on the intrinsic QR factorization. Better results can

be expected from the algorithms in [179, 180]. For clustering with rSVD and greedy column selection, however, significant speed ups are observed. If the matrix size is increased further, the differences to VCA become even more apparent, a speed up of two orders of magnitude can be achieved.

Not having been used in the context of microscopy yet, these algorithms give hope that also larger data sets on tomorrow's microscopes will be analyzable using the methods presented in this thesis, or may be useful for compression during acquisition to enable measurement of larger data sets.

4.4.2 Analysis of compressed signals

Pure pixel estimating methods are quickly computed and yield source spectra close to the spectra in the original data set. Also considering the above parallels to matrix compression, the question of how else one can use the compression for the analysis of spectra arises.

In paper VIII, an efficient background subtraction was proposed: Basis vectors $\mathbf{A}$ are estimated using VCA, a background model $\mathbf{A}_B$ is then estimated on $\mathbf{A}$. The background subtracted data set $\mathbf{X}_{BGS}$ is then obtained the following way:

$$\begin{aligned} \mathbf{X} &= \mathbf{MA} \\ &= \mathbf{M}(\mathbf{A}_S + \mathbf{A}_B) \\ \mathbf{X}_{BGS} &= \mathbf{X} - \mathbf{MA}_B \end{aligned}$$

Straightforward and easily implemented, this scheme can be extended to many more traditionally conducted analyses. E.g., an elemental analysis can be carried out on the basis vectors, estimated elemental contents can then be extrapolated to the entire data set by expanding with the abundances. A correction for multiple scattering effects can be achieved by deconvolving the basis vectors with the low-loss spectrum, estimating the difference spectrum between original and corrected spectrum and subtracting the difference spectrum from the data set. Besides speed and advantage of performing analyses on the basis vectors rather than individual spectra is the inherent noise reduction. Especially when correcting noisy data, the risk of introducing noise related artifacts into the data may outweigh the advantage of the correction. This is side-stepped when performing the operation on the compressed data.

4.4.3 Exploring the signal subspace

An important lesson from section 4.1 is the usefulness of PCA to reliably find vanishingly little noise biased source spectra from hyperspectral images by simply increasing the number of spectra sufficiently. An excellent basis for the spectral signal space is obtained. Geometric BSS methods typically

operate on PCA compressed data, in this signal (sub)space. Can one use other properties of this space for a better, more general unmixing?

In paper X steps towards that direction were taken. If the source signals can be assumed to be non-negative and orthogonal in some observation channel (e.g., for spectra with a stand-alone peak), the extreme rays of the cone spanned by the non-negative spectra in the signal subspace are the sought after source components. An algorithm to their extraction is thus PCA, random sampling of scores, only those scores that lead to non-negative spectra are retained. As a result, one finds the cone of non-negative spectra in the signal subspace. Application of NMF or pure pixel methods obtains then the source spectra.

Can one use this signal subspace sampling (SSS) idea more generally? Before BSS methods were introduced to the microscopy community, the BSS problem was often referred to as multivariate curve resolution (MCR), to be solved via alternating least squares. Addition of physically motivated constraints leads to better convergence, but local minima have to be feared. This strategy of finding source spectra by included meaningful constraints can be extended to a little elegant, but straightforward SSS algorithm. Essentially any constraint can be enforced as they occur only as conditional statements when deciding whether an artificially generated score vector should be retained or not. The likelihood in which meaningful spectra are sampled determines the usefulness of this approach, but given that a normal distribution seemed an efficient prior in paper X one can hope that this will not become a limiting factor. E.g., non-negativity and post-edge constancy could prove useful constraints for EELS fine-structure analysis, much like in MCR.

4.4.4 Unmixing hyperspectral data by exploiting spatial periodicity

Another strategy to unmix spectrum images may be to exploit the spatial periodicity of images at atomic resolution. In paper V tensor decomposition methods were already proposed for the analysis of tensors which are globally of low rank. However, using the non-local approach of, e.g., LLR, also high-rank images can be described in a low rank setting. Considering atomic resolution images only, the LLR assumption can be sharpened even further: Since crystals admit for a description with a unit cell, a block sampling approach should lead to similarly good results and lead to even lower multilinear ranks.

I.e., consider a hyperspectral images of a crystal $\mathscr{X} \in \mathbb{R}^{n_x \times n_y \times n_E}$ with a unit cell of $u_x \times u_y$ pixels. Either by application of template matching algorithms, or by performing block segmentation on the image (i.e., form a regular grid of non-overlapping blocks), individual unit cells can be identified and stacked in the tensor $\mathscr{Y} \in \mathbb{R}^{u_x \times u_y \times n_E \times N}$, where N is the number of blocks. If the unit

cell is well aligned with the image borders, the two spatial modes of the tensor likely allow for a rank 1 approximation, for most crystals without rank deficiency of the factor matrices. Also assuming that the energy mode is not rank deficient (and some of the components change across the hyperspectral image so that the stacking mode is not too rank deficient), the uniqueness criteria of a CPD are likely met, the source spectra can be extracted without any further assumption on their spectral shape! In case the unit cell is not aligned with the image borders (e.g., the crystal appears rotated in the image), one can either move to LL1-decompositions (but needs to reconsider whether the factor matrices will be of full rank), or one can reshape $\mathscr{Y}$ to $\tilde{\mathscr{Y}} \in \mathbb{R}^{u_x \cdot u_y \times n_E \times N}$. In the latter case, the spatial rank of the unit cell structure can increase arbitrarily, but one needs to reconsider rank deficiency along the stacking mode. Unless the lattice changes across the image (leading to non-uniform changes of the factor matrices along the stacking mode), uniqueness conditions may not be met anymore.

Let us briefly demonstrate this method on two simulated, abstract data set. Source components and source structures are displayed in Fig.4.2. Afterward, the image was segmented into blocks of 10×10 (which in this case matches exactly with the *unit cell* of the structure). Due to the chosen structure, the spatial structures were exactly of rank 1 CPD on $\mathscr{Y}$ thus successfully extracts the source components. If the spatial rank would be higher, it is worth considering applying CPD to $\tilde{\mathscr{Y}}$. Here, however, one observes that only the second set with changes of the lattice within the image leads to the correct unmixing. For the other case, the stacking mode is too rank deficient (only rank 2, the uniform lattice accounts for one components, the gradient for the other) to allow for the lattice components to be separated. The gradient component is still found correctly.

While the source components do not resemble real microscopy data, they were also not constrained in any way. The proposed approach should thus be transferable to different signals in the microscope. Using pattern matching instead of rigid segmentation grids likely will increase the flexibility further, if care is taken to track the multilinear ranks of the tensor when applying this mosaicking approach, a fairly general analysis of atomic resolution spectrum images is possible. The possibility to separate lattice signal space from other components changing on larger scales over several unit cells may also be useful to correct for artifacts (gradual energy drift, beam damage, ...).

4.4.5 Inpainting instead of denoising

In some cases, the sensitivity of the specimen to damage through illumination with the electron beam becomes the limiting factor for the achievable image quality. It has been proposed [182, 183] to reduce the dose the sample is exposed to during image acquisition by intentionally omitting some pixels and

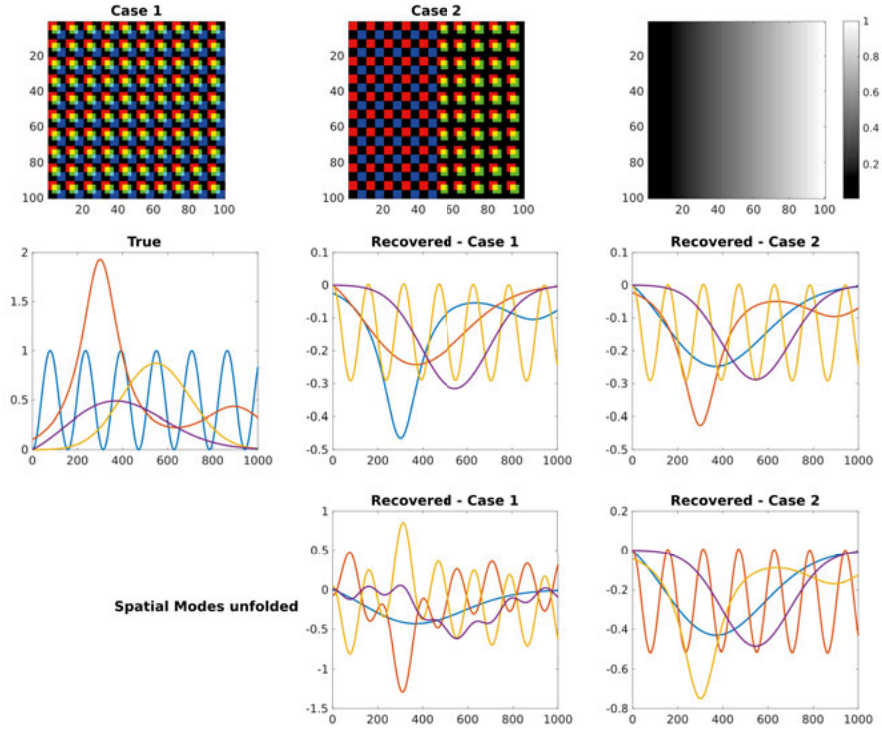

Figure 4.2. Top row: Source abundances for the two sets studied here, the first three components are displayed as RGB image (intensities ranging from 0 to 1), the last gradient component is shared between the sets. *Center row:* Source components and recovered components for both sets when applying CPD to $\mathscr{Y}$. *Bottom row:* Recovered source components when applying CPD to $\tilde{\mathscr{Y}}$.

then inpainting the missing intensities in an offline process. For this inpainting task, basis pursuit approaches dominate today.

However, any denoising algorithm can easily be transformed into an inpaiting tool when reformulating the algorithm to estimate its components only on the measured pixels, e.g., by assigning the missing observations zero weight when computing the covariance matrix of PCA. Hence, a hidden contribution of section 4.1 was the development of inpainting tools. Considering images, LLR is readily rewritten as inpainting tool. As small difference to image denoising, in our experience, one often obtains better results when computing a high-rank CPD rather than a low rank Tucker model on the LLR stacking tensor (with missing pixels). Especially in cases where the spatial ranks of the stacking tensor can not be compressed considerably, CPD outperforms the TKD, which tends to over-represent the measured pixels on cost of an overall worse reconstruction.

As discussed at length in paper XIII, an inpainting algorithm for regularly (sub)sampled images can also be formulated using LLR. Since LLR allows for a low rank approximation of images, one can estimate a local feature basis on a set of training images and then apply this basis to the undersampled image. Since the basis is small, it avoids coherence artifacts otherwise occurring in basis pursuit inpainting naturally.

5. Conclusions

The majority of this thesis' content was focused on the extraction of the EMCD signal using BSS methods. After giving an overview over different signals occurring in the microscope, different classes of BSS methods, both matrix and tensor factorizations, were briefly sketched. These threads were then combined to a data processing workflow to detect weak signals in noisy hyperspectral images. Applying this knowledge, data sets containing EMCD signals can be reliably denoised. Based on known signal distributions, different BSS methods were suggested for different types of EMCD measurements. Afterward a selection of the proposed methods was demonstrated experimentally.

Besides this focus, one must not forget that the methods are applicable to a much wider range of signals. This is reflected in the variety of fields the literature of section 3 stems from, but also in the rhapsodic section 4.4 summarizing some other ideas of usage of data processing methods in microscopy (and beyond).

For a reader thinking beyond the applications discussed here, focusing on underlying assumption on mathematical characteristics of the data, parallels to many other applications may open. As the size and types of data sets are ever increasing there will be ample space for the different factorization techniques presented here to shine.

6. Sammanfattning på svenska

Elektronmikroskopi använder sig av högt accelererade elektroner i mindre än nanometer tjocka strålar för att utföra spridningsexperiment på olika material. Från dessa kan man lära sig om materialets struktur, defekter, sammansättning, elektroniska struktur och magnetiska egenskaper. Studerar man dessutom en effekt som kallas för "elektromagnetisk kiral dikroism" (EMCD), kan man mäta en enskild atoms spinn och magnetiska moment i ett visst material. I ideala fall kan EMCD-effekten tillåta att man kombinerar magnetisk analys av material med alla fördelar som elektronmikroskop kan tillåta, t. ex. hög upplösning och låga instrumentella krav, jämfört med andra mätmetoder. Dock så är det svårt att mäta signaler i data med dålig upplösning – ett problem som förekommer i många olika vetenskapliga discipliner, antagligen bl.a. också i forskningen om T-rexer.

I detta doktorsarbete analyseras magnetiska strukturer med hjälp av elektronmikroskop, med fokus på EMCD-effekten. Specifikt handlar doktorsarbetet om att undersöka hur moderna metoder inom dataanalys kan underlätta upptäckten av EMCD-signalen, liksom andra signaler från elektronmikroskop. Till en början diskuteras olika signaler som uppstår i mikroskopen liksom deras matematiska egenskaper, följt av en kort sammanfattning av EMCD-teorin och de experimentella mätmetoder som används. För att reducera brus i mätningarna har jag använt Principal Component Analysis (PCA), följt av tensorfaktorisering och ytterligare brusreducering. Förutom olika brusreduceringsmetoder, diskuteras även metoder för en direkt identifikation av relevanta signaler. En översikt av olika metoder ges, inklusive automatiska metoder som jag utvecklat.

I arbetets sista del, används de metoder vi föreslagit för att detektera EMCD-signaler i specifika experiment. Dessa demonstrerar hur EMCD kan användas i högupplösta mätningar av magnetiska strukturer för att karakterisera material med atom-atom-upplösning. Semiautomatiska detektionsmetoder testas på samma sätt. Med verktyg från modern dataanalys kan därmed EMCD:ns potential för att mäta magnetism i olika material utvecklas till fullo.

7. Acknowledgements

At last, it is time to express my gratitude to the numerous people which helped to shape my PhD to what it has become. Firstly, my two main supervisors, Ján Rusz and Kristiaan Pelckmans have to be mentioned for great advice, patience for many failed attempts and discussions on anything ranging from scientific issues to the joys of contemporary music. None of this thesis work would be half as relevant without the contributions from all the experimental collaborators, in random order my thanks goes to T. Thersleff, J. C. Idrobo, J. Hachtel, D. Song, D. Pohl, S. Schneider, M. Ohtsuka. A special thanks goes to S. Muto for scientific discussions, hospitality and once more Hibiki - the little bottle has since become vessel for various other whiskeys for tasting portions.

The members of materials theory wonderfully created an atmosphere able to accommodate a great deal of insanity in form of lunch discussions, coffee breaks and the wonderful chocolate-rich office! The local hedonist deserve great praise for upholding the ideals of pleasure and ethical egocentrism. Marco, Iulia, Beatriz, Raquel, ... a long list has to be written here. Special praise to the Ulleråker gang for covering all those topics which were not fit for the fine saloons of the university.

Blank pages and typographical errors I have to dedicate again to E. Makkonen, I may have plagiarized the former. At last, thank you, Chenjuan, for being there and for being you!

References

[1] C. Hébert, P. Schattschneider, *Ultramicroscopy* **96**, 463 (2003).

[2] P. Schattschneider, S. Rubino, C. Hébert, J. Rusz, J. Kunes, P. Novak, E. Carlino, M. Fabrizioli, G. Panaccione, G. Rossi, *Nature* **441**, 486 (2006).

[3] O. Nicoletti, F. de la Pena, R. K. Leary, D. J. Holland, C. Ducati, P. A. Midgley, *Nature* **502**, 80 (2013).

[4] K. Jarausch, P. Thomas, D. N. Leonard, R. Twesten, C. R. Booth, *Ultramicroscopy* **109**, 326 (2009).

[5] B. Goris, S. Turner, S. Bals, G. V. Tendeloo, *ACS Nano* **8**, 10878 (2014).

[6] C. Ophus, P. Ercius, M. Sarahan, C. Czarnik, J. Ciston, *Microscopy and Microanalysis* **20**, 62 (2014).

[7] H. Yang, L. Jones, H. Ryll, M. Simson, H. Soltau, Y. Kondo, R. Sagawa, H. Banba, I. MacLaren, P. D. Nellist, *Journal of Physics: Conference Series* **644**, 012032 (2015).

[8] N. Shibata, Y. Kohno, S. D. Findlay, H. Sawada, Y. Kondo, Y. Ikuhara, *Journal of Electron Microscopy* **59**, 473 (2010).

[9] P. Schattschneider, C. Hébert, S. Rubino, M. Stöger-Pollach, J. Rusz, P. Novak, *Ultramicroscopy* **108**, 433 (2008).

[10] J. M. Bioucas-Dias, A. Plaza, in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pp. 1135–1138 (2011).

[11] J. Lahnemann, C. Hauswald, M. Wolz, U. Jahn, M. Hanke, L. Geelhaar, O. Brandt, *Journal of Physics D: Applied Physics* **47**, 394010 (2014).

[12] A. Losquin, L. F. Zagonel, V. Myroshnychenko, B. Rodríguez-González, M. Tencé, L. Scarabelli, J. Förstner, L. M. Liz-Marzán, F. J. García de Abajo, O. Stéphan, M. Kociak, *Nano Letters* **15**, 1229 (2015), pMID: 25603194.

[13] C. Colliex, M. Kociak, O. Stéphan, *Ultramicroscopy* **162**, A1 (2016).

[14] L. F. Zagonel, S. Mazzucco, M. Tence, K. March, R. Bernard, B. Laslier, G. Jacopin, M. Tchernycheva, L. Rigutti, F. H. Julien, R. Songmuang, M. Kociak, *Nano Letters* **11**, 568 (2011).

[15] A. Quaranta, G. Valotto, A. D. L. Pezzolo, G. Mazzocchin, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **121**, 1 (2014).

[16] M. Willander, O. Nur, Q. X. Zhao, L. L. Yang, M. Lorenz, B. Q. Cao, J. Z. Perez, C. Czekalla, G. Zimmermann, M. Grundmann, A. Bakin, A. Behrends, M. Al-Suleiman, A. El-Shaer, A. C. Mofor, B. Postels, A. Waag, N. Boukos, A. Travlos, H. S. Kwack, J. Guinard, D. L. S. Dang, *Nanotechnology* **20**, 332001 (2009).

[17] L. Reimer, H. Kohl, *Transmission Electron Microscopy - Physics of Image Formation*, Springer (2008).

[18] J. L. Knipping, L. D. Bilenker, A. C. Simon, M. Reich, F. Barra, A. P. Deditius, M. Walle, C. A. Heinrich, F. Holtz, R. Munizaga, *Geochimica et Cosmochimica Acta* **171**, 15 (2015).

[19] A. J. D'Alfonso, B. Freitag, D. Klenov, L. J. Allen, *Phys. Rev. B* **81**, 100101 (2010).
[20] M. P. Oxley, E. C. Cosgriff, L. J. Allen, *Phys. Rev. Lett.* **94**, 203906 (2005).
[21] P. Potapov, P. Longo, E. Okunishi, *Micron* **96** (2017).
[22] G. Lucas, P. Burdet, M. Cantoni, C. Hébert, *Micron* **52**, 49 (2013).
[23] M. Shiga, K. Tatsumi, S. Muto, K. Tsudac, Y. Yamamoto, T. Morie, T. Tanji, *Ultramicroscopy* (2016).
[24] R. F. Egerton, *Electron Energy Loss Spectroscopy in the Electron Microscope*, Springer (2011).
[25] F. J. G. de Abajo, M. Kociak, *New Journal of Physics* **10**, 073035 (2008).
[26] L. Gu, V. Srot, W. Sigle, C. Koch, P. van Aken, F. Scholz, S. B. Thapa, C. Kirchner, M. Jetter, M. Rühle, *Phys. Rev. B* **75**, 195214 (2007).
[27] S. Muto, T. Yoshida, K. Tatsumi, *Materials Transactions* **50**, 964 (2009).
[28] N. Bonnet, D. Nuzillard, *Ultramicroscopy* **102**, 327 (2005).
[29] F. de la Pena, *Advanced methods for Electron Energy Loss Spectroscopy core-loss analysis*, Ph.D. thesis, L'Universite Paris Sud 11 (2010).
[30] F. de la Pena, M.-H. Berger, J.-F. Hochepied, F. Dynys, O. Stephan, M. Walls, *Ultramicroscopy* **111**, 169 (2011).
[31] M. Bosman, M. Watanabe, D. Alexander, V. Keast, *Ultramicroscopy* **106**, 1024 (2006).
[32] C. Witte, N. Zaluzec, L. Allen, *Ultramicroscopy* **110**, 1390 (2010).
[33] K. Dudeck, M. Couillard, S. Lazar, C. Dwyer, G. Botton, *Micron* **43**, 57 (2012), applications of EELS in Materials and Physics Research.
[34] P. Potapov, *Ultramicroscopy* **160**, 197 (2016).
[35] P. Cueva, R. Hovden, J. A. Mundy, H. L. Xin, D. A. Muller, *Microscopy and Microanalysis* **18**, 667 (2012).
[36] S. Lichtert, J. Verbeeck, *Ultramicroscopy* **125**, 35 (2013).
[37] P. Potapov, *Ultramicroscopy* **182**, 191 (2017).
[38] G. Sharma, A. Gupta, P. Rajput, R. Rüffer, *Journal of Physics D: Applied Physics* **47**, 345305 (2014).
[39] A. Gloskovskii, J. Barth, B. Balke, G. H. Fecher, C. Felser, F. Kronast, R. Ovsyannikov, H. Dürr, W. Eberhard, G. Schönhense, *Journal of Physics D: Applied Physics* **40**, 1570 (2007).
[40] M. J. Robertson, C. J. Agostino, A. T. N'Diaye, G. Chen, M.-Y. Im, P. Fischer, *Journal of Applied Physics* **117** (2015).
[41] A. DiLullo, N. Shirato, M. Cummings, H. Kersell, H. Chang, D. Rosenmann, D. Miller, J. W. Freeland, S.-W. Hla, V. Rose, *Journal of Synchrotron Radiation* **23**, 574 (2016).
[42] B. T. Thole, P. Carra, F. Sette, G. van der Laan, *Phys. Rev. Lett.* **68**, 1943 (1992).
[43] P. Carra, B. T. Thole, M. Altarelli, X. Wang, *Phys. Rev. Lett.* **70**, 694 (1993).
[44] J. Rusz, O. Eriksson, P. Novák, P. M. Oppeneer, *Phys. Rev. B* **76**, 060408 (2007).
[45] L. Calmels, F. Houdellier, B. Warot-Fonrose, C. Gatel, M. J. Hÿtch, V. Serin, E. Snoeck, P. Schattschneider, *Phys. Rev. B* **76**, 060409 (2007).
[46] B. Warot-Fonrose, F. Houdellier, M. Hytch, L. Calmels, V. Serin, E. Snoeck, *Ultramicroscopy* **108**, 393 (2008).

[47] P. Schattschneider, M. Stöger-Pollach, S. Rubino, M. Sperl, C. Hurm, J. Zweck, J. Rusz, *Phys. Rev. B* **78**, 104413 (2008).
[48] J. Salafranca, J. Gazquez, N. Pérez, A. Labarta, S. T. Pantelides, S. J. Pennycook, X. Batlle, M. Varela, *Nano Letters* **12**, 2499 (2012).
[49] T. Thersleff, J. Rusz, S. Rubino, B. Hjörvarsson, Y. Ito, N. J. Zaluzec, K. Leifer, *Scientific Reports* **5**, 13012 (2015).
[50] S. Schneider, D. Pohl, S. Löffler, J. Rusz, D. Kasinathan, P. Schattschneider, L. Schultz, B. Rellinghaus, *Ultramicroscopy* **171**, 186 (2016).
[51] D. Song, J. Rusz, J. Cai, J. Zhu, *Ultramicroscopy* **169**, 44 (2016).
[52] D. Song, A. Tavabi, Z. an Li, A. Kovacs, J. Rusz, W. Huang, G. Richter, R. Dunin-Borkovski, J. Zhu, *Nature Communications* **8** (2017).
[53] S. Rubino, *Magnetic Circular Dichroism in the Transmission Electron Microscope*, Ph.D. thesis, Vienna University of Technology (2007).
[54] J. Rusz, S. Rubino, P. Schattschneider, *Phys. Rev. B* **75**, 214425 (2007).
[55] C. Hébert, P. Schattschneider, S. Rubino, P. Novak, J. Rusz, M. Stöger-Pollach, *Ultramicroscopy* **108**, 277 (2008).
[56] S. Rubino, P. Schattschneider, J. Rusz, J. Verbeeck, K. Leifer, *Journal of Physics D: Applied Physics* **43**, 474005 (2010).
[57] J. Rusz, J.-C. Idrobo, L. Wrang, *Phys. Rev. B* **94**, 144430 (2016).
[58] T. Thersleff, J. Rusz, B. Hjörvarsson, K. Leifer, *Phys. Rev. B* **94**, 134430 (2016).
[59] T. Thersleff, S. Muto, M. Werwinski, J. Spiegelberg, Y. Kvashnin, B. Hjörvarsson, O. Eriksson, J. Rusz, K. Leifer, *Scientific Reports* **7**, 44802 (2017).
[60] H. Lidbaum, J. Rusz, A. Liebig, B. Hjörvarsson, P. M. Oppeneer, E. Coronel, O. Eriksson, K. Leifer, *Phys. Rev. Lett.* **102**, 037201 (2009).
[61] Z. Wang, X. Zhong, R. Yu, Z. Cheng, J. Zhu, *Nature Communications* **4**, 1395 (2013).
[62] S. Muto, K. Tatsumi, J. Rusz, *Ultramicroscopy* **125**, 89 (2013).
[63] S. Muto, J. Rusz, K. Tatsumi, R. Adam, S. Arai, V. Kocevski, P. Oppeneer, D. Bürgler, C. Schneider, *Nature Communications* **5**, 3138 (2014).
[64] D. Song, G. Li, J. Cai, J. Zhu, *Scientific Reports* **6**, 18489 (2016).
[65] Z. Wang, X. Zhong, L. Jin, X. Chen, Y. Moritomo, J. Mayer, *Ultramicroscopy* **176**, 212 (2017), 70th Birthday of Robert Sinclair and 65th Birthday of Nestor J. Zaluzec PICO 2017 Fourth Conference on Frontiers of Aberration Corrected Electron Microscopy.
[66] J. Rusz, J.-C. Idrobo, S. Bhowmick, *Phys. Rev. Lett.* **113**, 145501 (2014).
[67] J. Rusz, J. C. Idrobo, *Phys. Rev. B* **93**, 104420 (2016).
[68] J. C. Idrobo, J. Rusz, J. Spiegelberg, M. A. McGuire, C. T. Symons, R. R. Vatsavai, C. Cantoni, A. R. Lupini, *Advanced Structural and Chemical Imaging* **2**, 1 (2016).
[69] M. Uchida, A. Tonomura, *Nature* **464**, 737 (2010).
[70] J. Verbeeck, H. Tian, P. Schattschneider, *Nature* **467**, 301 (2010).
[71] B. J. McMorran, A. Agrawal, I. M. Anderson, A. A. Herzing, H. J. Lezec, J. J. McClelland, J. Unguris, *Science* **331**, 192 (2011).
[72] J. Rusz, S. Bhowmick, *Phys. Rev. Lett.* **111**, 105504 (2013).
[73] P. Schattschneider, S. Löffler, M. Stöger-Pollach, J. Verbeeck, *Ultramicroscopy*

136, 81 (2014).
[74] D. Negi, J. C. Idrobo, J. Rusz, *Scientific Reports* **8**, 4019 (2018).
[75] D. Pohl, S. Schneider, J. Rusz, B. Rellinghaus, *Ultramicroscopy* **150**, 16 (2015).
[76] A. Béché, R. Juchtmans, J. Verbeeck, *Ultramicroscopy* (2016).
[77] D. Pohl, S. Schneider, P. Zeiger, J. Rusz, P. Tiemeijer, S. Lazar, K. Nielsch, B. Rellinghaus, *Scientific Reports* **7** (2017).
[78] J. Rusz, S. Muto, J. Spiegelberg, R. Adam, K. Tatsumi, D. E. Bürgler, P. M. Oppeneer, C. M. Schneider, *Nature Communications* **7** (2016).
[79] J. Rusz, J. Spiegelberg, S. Muto, T. Thersleff, M. Ohtsuka, K. Leifer, P. M. Oppeneer, *Phys. Rev. B* **95**, 174412 (2017).
[80] T. Schachinger, S. Löffler, A. Steiger-Thirsfeld, M. Stöger-Pollach, S. Schneider, D. Pohl, B. Rellinghaus, P. Schattschneider, *Ultramicroscopy* **179**, 15 (2017).
[81] Z. Wang, A. Tavabi, L. Jin, J. Rusz, D. Tyutyunnikov, H. Jiang, Y. Morimoto, J. Mayer, R. Dunin-Borkovski, R. Yu, J. Zhu, X. Zhong, *Nature Materials* **17**, 221 (2018).
[82] H. Kohl, H. Rose, volume 65 of *Advances in Electronics and Electron Physics*, pp. 173 – 227, Academic Press (1985).
[83] P. Schattschneider, M. Nelhiebel, H. Souchay, B. Jouffrey, *Micron* **31**, 333 (2000).
[84] A. Belianinov, R. Vasudevan, E. Strelcov, C. Steed, S. M. Yang, A. Tselev, S. Jesse, M. Biegalski, G. Shipman, C. Symons, A. Borisevich, R. Archibald, S. Kalinin, *Advanced Structural and Chemical Imaging* **1**, 6 (2015).
[85] S. Somnath, C. R. Smith, S. V. Kalinin, M. Chi, A. Borisevich, N. Cross, G. Duscher, S. Jesse, *Advanced Structural and Chemical Imaging* **4**, 3 (2018).
[86] W. Xu, J. LeBeau, *Ultramicroscopy* **188**, 59 (2018).
[87] K. Pearson, *Philosophical Magazine* **6** (1901).
[88] P. Trebbia, N. Bonnet, *Ultramicroscopy* **34**, 165 (1990).
[89] N. Bonnet, N. Brun, C. Colliex, *Ultramicroscopy* **77**, 97 (1999).
[90] P. G. Kotula, M. R. Keenan, J. R. Michael, *Microscopy and Microanalysis* **9**, 1 (2003).
[91] N. Borglund, P.-G. Åstrand, S. Csillag, *Microscopy and Microanalysis* **11**, 88 (2005).
[92] J. Spiegelberg, J. Rusz, *Ultramicroscopy* **172**, 40 (2017).
[93] H. Hotelling, *Journal of Educational Psychology* **24** (1933).
[94] G. H. Golub, C. Reinsch, *Numerische Mathematik* **14**, 403 (1970).
[95] K. Karhunen, *Ann. Acad. Sci. Fennicae* **34** (1946).
[96] M. Loéve, *Fonctions al eatoires du second ordre*, Gauthier - Villars (1948).
[97] T. Bouwmans, E. H. Zahzah, *Computer Vision and Image Understanding* **122**, 22 (2014).
[98] N. Halko, P. G. Martinsson, J. A. Tropp, *SIAM Review* **53**, 217 (2011).
[99] H. Zou, T. Hastie, R. Tibshirani, *Journal of Computational and Graphical Statistics* **15** (2006).
[100] C. Chatterjee, Z. Kang, V. P. Roychowdhury, *IEEE Transactions on Neural Networks* **11**, 338 (2000).
[101] M. Stéphane, in *A Wavelet Tour of Signal Processing (Third Edition)*, pp. 535 –

610, Academic Press, Boston (2009).
[102] R. Kilaas, *Journal of Microscopy* **190**, 45 (1998).
[103] R. Hovden, Y. Jiang, H. L. Xin, L. F. Kourkoutis, *Microscopy and Microanalysis* **21**, 436 (2015).
[104] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, *IEEE Transactions on Image Processing* **16**, 2080 (2007).
[105] A. Buades, B. Coll, J. M. Morel, *SIAM Review* **52**, 113 (2010).
[106] J. S. Charles Alban Deledalle, A. Dalalyan, in *Proceedings of the British Machine Vision Conference*, pp. 25.1–25.10, BMVA Press (2011), http://dx.doi.org/10.5244/C.25.25.
[107] M. Lebrun, M. Colom, A. Buades, J. M. Morel, *Acta Numerica* **21**, 475 (2012).
[108] N. Mevenkamp, P. Binev, W. Dahmen, P. M. Voyles, A. B. Yankovich, B. Berkels, *Advanced Structural and Chemical Imaging* **1**, 3 (2015).
[109] A. B. Yankovich, C. Zhang, A. Oh, T. J. A. Slater, F. Azough, R. Freer, S. J. Haigh, R. Willett, P. M. Voyles, *Nanotechnology* **27**, 364001 (2016).
[110] G. Bárcena-González, M. Guerrero-Lebrero, E. Guerrero, A. Yanez, D. Fernández-Reyes, D. González, P. Galindo, *Ultramicroscopy* **182**, 283 (2017).
[111] A. J. Craven, B. Sala, J. Bobynko, I. MacLaren, *Ultramicroscopy* **186**, 66 (2018).
[112] F. de la Pena, T. Ostasevicius, V. T. Fauske, P. Burdet, P. Jokubauskas, M. Nord, E. Prestat, M. Sarahan, K. E. MacArthur, D. N. Johnstone, J. Taillon, J. Caron, T. Furnival, A. Eljarrat, S. Mazzucco, V. Migunov, T. Aarholt, M. Walls, F. Winkler, B. Martineau, G. Donval, E. R. Hoglund, I. Alxneit, I. Hjorth, L. F. Zagonel, A. Garmannslund, C. Gohlke, I. Iyengar, H.-W. Chang, *hyperspy/hyperspy: HyperSpy 1.3* (2017).
[113] J. F. Cardoso, A. Souloumiac, *IEE Proceedings F - Radar and Signal Processing* **140**, 362 (1993).
[114] J.-F. Cardoso, *Neural Computation* **11**, 157 (1999).
[115] A. Hyvarinen, E. Oja, *Neural Networks* **13**, 411 (2000).
[116] F. Bach, M. Jordan, *Journal of Machine Learning Research* **3**, 1 (2002).
[117] E. G. Learned-Miller, J. W. F. III, *J. Mach. Learn. Res.* **4**, 1271 (2003).
[118] T. Blaschke, L. Wiskott, *IEEE Transactions on Signal Processing* **52**, 1250 (2004).
[119] S. Cruces, A. Cichocki, L. D. Lathauwer, in *2004 12th European Signal Processing Conference*, pp. 217–220 (2004).
[120] P. Comon, *Signal Processing* **36**, 287 (1994).
[121] A. Hyvarinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley (2001).
[122] J. F. Cardoso, *Proceedings of the IEEE* **86**, 2009 (1998).
[123] J. M. P. Nascimento, J. M. B. Dias, *IEEE Transactions on Geoscience and Remote Sensing* **43**, 175 (2005).
[124] N. Gillis, *Journal of MAchine Learning Research* **13**, 3349 (2012).
[125] G. Zhou, A. Cichocki, Q. Zhao, S. Xie, *IEEE Signal Processing Magazine* **31**, 54 (2014).
[126] K. Huang, N. D. Sidiropoulos, A. Swami, *IEEE Transactions on Signal Processing* **62**, 211 (2014).

[127] D. Lee, H. Seung, *Advances in Neural Information Processing Systems* **13** (2001).
[128] M. W. Berry, M. Browne, A. N. Langville, V. P. Paucac, R. Plemmons, *Computational Statistics and Data Analysis* **52**, 155 (2007).
[129] S. Jia, Y. Qian, *IEEE Transactions on Geoscience and Remote Sensing* **47**, 161 (2009).
[130] L. Miao, H. Qi, *IEEE Transactions on Geoscience and Remote Sensing* **45**, 765 (2007).
[131] J. Boardman, *Summaries 4th Annu. JPL Airborne Geoscience Workshop* **1**, 11 (1993).
[132] T. H. Chan, W. K. Ma, A. Ambikapathi, C. Y. Chi, *IEEE Transactions on Geoscience and Remote Sensing* **49**, 4177 (2011).
[133] M. E. Winter, *Proc. SPIE Conf. Imaging Spectrometry V* pp. 266–275 (1999).
[134] J. M. P. Nascimento, J. M. B. Dias, *IEEE Transactions on Geoscience and Remote Sensing* **43**, 898 (2005).
[135] M. D. Iordache, J. M. Bioucas-Dias, A. Plaza, *IEEE Transactions on Geoscience and Remote Sensing* **52**, 341 (2014).
[136] N. Dobigeon, N. Brun, *Ultramicroscopy* **120**, 25 (2012).
[137] X. Fu, K. Huang, B. Yang, W. K. Ma, N. D. Sidiropoulos, *IEEE Transactions on Signal Processing* **64**, 6254 (2016).
[138] T. H. Chan, C. Y. Chi, Y. M. Huang, W. K. Ma, *IEEE Transactions on Signal Processing* **57**, 4418 (2009).
[139] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, X. Li, *IEEE Transactions on Geoscience and Remote Sensing* **53**, 5067 (2015).
[140] N. Dobigeon, S. Moussaoui, M. Coulon, J. Y. Tourneret, A. O. Hero, *IEEE Transactions on Signal Processing* **57**, 4355 (2009).
[141] J. Li, J. M. Bioucas-Dias, A. Plaza, L. Liu, *IEEE Transactions on Geoscience and Remote Sensing* **54**, 6076 (2016).
[142] W. Naanaa, J.-M. Nuzillard, *Signal Processing* **130**, 254 (2017).
[143] X. Fu, K. Huang, N. D. Sidiropoulos, *IEEE Signal Processing Letters* **25**, 328 (2018).
[144] C. Boothroyd, M. Moreno, M. Duchamp, A. Kovács, N. Monge, G. Morales, C. Barbero, R. Dunin-Borkowski, *Ultramicroscopy* **145**, 66 (2014).
[145] T. G. Kolda, B. W. Bader, *SIAM Review* **51**, 455 (2009).
[146] A. Cichocki, D. Mandic, L. D. Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, H. A. PHAN, *IEEE Signal Processing Magazine* **32**, 145 (2015).
[147] N. D. Sidiropoulos, L. D. Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, C. Faloutsos, *IEEE Transactions on Signal Processing* **65**, 3551 (2017).
[148] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, B. Zhang, in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2949–2956 (2014).
[149] J. An, X. Zhang, L. C. Jiao, *IEEE Geoscience and Remote Sensing Letters* **13**, 1497 (2016).
[150] Y. Qian, F. Xiong, S. Zeng, J. Zhou, Y. Y. Tang, *IEEE Transactions on Geoscience and Remote Sensing* **55**, 1776 (2017).
[151] I. V. Oseledets, *SIAM Journal on Scientific Computing* **33**, 2295 (2011).
[152] S. Friedland, T. G. Kolda, L.-H. Lim, E. E. Tyrtyshnikov, J. Ballani,

L. Grasedyck, M. Kluge, *Linear Algebra and its Applications* **438**, 639 (2013).
[153] L. R. Tucker, *Psychometrika* **31** (1966).
[154] M. Ishteva, P.-A. Absil, S. V. Huffel, L. D. Lathauwer, *Chemometrics and Intelligent Laboratory Systems* **106**, 57 (2011).
[155] M. Mørup, L. K. Hansen, S. M. Arnfred, *Neural Comput.* **20**, 2112 (2008).
[156] M. P. Friedlander, K. Hatz, *Computational Optimization and Applications* **23**, 631 (2008).
[157] L. D. Lathauwer, B. D. Moor, J. Vandewalle, *SIAM Journal on Matrix Analysis and Applications* **21**, 1253 (2000).
[158] F. L. Hitchcock, *J. Math. Phys.* **6** (1927).
[159] J. D. Carroll, J.-J. Chang, *Psychometrika* **35** (1970).
[160] J. B. Kruskal, *Linear Algebra and its Applications* **18**, 95 (1977).
[161] I. Domanov, L. D. Lathauwer, *SIAM Journal on Matrix Analysis and Applications* **34**, 855 (2013).
[162] I. Domanov, L. D. Lathauwer, *SIAM Journal on Matrix Analysis and Applications* **34**, 876 (2013).
[163] L. D. Lathauwer, *SIAM Journal on Matrix Analysis and Applications* **30**, 1022 (2008).
[164] L. D. Lathauwer, *SIAM Journal on Matrix Analysis and Applications* **30**, 1033 (2008).
[165] L. D. Lathauwer, D. Nion, *SIAM Journal on Matrix Analysis and Applications* **30**, 1067 (2008).
[166] L. D. Lathauwer, *SIAM Journal on Matrix Analysis and Applications* **32**, 1451 (2011).
[167] G. Chabriel, M. Kleinsteuber, E. Moreau, H. Shen, P. Tichavsky, A. Yeredor, *IEEE Signal Processing Magazine* **31**, 34 (2014).
[168] X. Chen, Z. J. Wang, M. McKeown, *IEEE Signal Processing Magazine* **33**, 86 (2016).
[169] R. A. Harshman, *Mathematical and technical notes*, UCLA Working Papers in Phonetics (1972).
[170] H. A. L. Kiers, J. M. F. ten Berge, R. Bro, *Journal of Chemometrics* **13**, 275 (1999).
[171] E. Acar, M. A. Rasmussen, F. Savorani, T. Naes, R. Bro, *Chemometrics and Intelligent Laboratory Systems* **129**, 53 (2013), multiway and Multiset Methods.
[172] L. Sorber, *Data fusion - Tensor factorizations by complex optimization*, Ph.D. thesis, KU Leuven - Faculty of Engineering Science (2014).
[173] L. Sorber, M. V. Barel, L. D. Lathauwer, *IEEE Journal of Selected Topics in Signal Processing* **9**, 586 (2015).
[174] J. M. F. ten Berge, H. A. L. Kiers, *Psychometrika* **61**, 123 (1996).
[175] N. Faber, M. Meinders, P. Geladi, M. Sjöström, L. Buydens, G. Kateman, *Analytica Chimica Acta* **304**, 257 (1995).
[176] B. Nadler, *The Annals of Statistics* **36**, 2791 (2008).
[177] L. Zhuang, J. M. Bioucas-Dias, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **PP**, 1 (2018).
[178] J. Spiegelberg, *Advanced Signal Processing Tools for Noisy Hyperspectral Images*, Uppsala University (2016).

[179] M. Gu, S. C. Eisenstat, *SIAM Journal on Scientific Computing* **17**, 848 (1996).
[180] H. Cheng, Z. Gimbutas, P. G. Martinsson, V. Rokhlin, *SIAM Journal on Scientific Computing* **26**, 1389 (2005).
[181] L. Ling, *SIAM Journal on Scientific Computing* **38**, A1224 (2016).
[182] H. S. Anderson, J. Ilic-Helms, B. Rohrer, J. Wheeler, K. Larson, *Proc.SPIE* **8657**, 8657 (2013).
[183] A. Stevens, H. Yang, L. Carin, I. Arslan, N. D. Browning, *Microscopy* **63**, 41 (2014).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1672

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-348264

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2018