NTNU
Norwegian University of
Science and Technology

An X-ray Photoelectron Spectroscopy study of DyS$_x$ created by depositing Dy on MoS$_2$

For application in spintronics

Magnus Sebastian Christensen

An X-ray Photoelectron Spectroscopy study of DyS_x created by depositing Dy on MoS_2 - for application in spintronics

Magnus Sebastian Christensen
Supervisor: Justin Wells

June 11, 2017

Preface

This report is my master thesis of a Master of Science degree in Applied Physics at NTNU, carried out the spring of 2017, whose work is largely based on preliminary work done in my specialization project fall 2016. The master project is weighted one semester and has been carried out under the guidance of my supervisor, Associate Professor Justin Wells, and his research group at the Department of Physics at NTNU.

The project has relied on experiments that have been carried out at the synchrotron ASTRID2, at the Department of Physics and Astronomy at the University of Aarhus, as well as an X-ray photoelectron spectroscopy lab at the Department of Physics at NTNU. My work in this project includes planning and executing the experimental data acquisition and subsequent data analysis.

Most of the experiment and data analysis has been part of a collaboration with my fellow student, Stian Ruud Schikora, who was writing his own report on the same data set.

The assumed knowledge of the reader of this report is that of a master student in science, familiar with elementary quantum mechanics.

I would like to thank Justin Wells for shared insights and for including me in his team. I would also like to thank Postdoc Simon Cooil for all his time spent in the lab and answering questions.

If this report is read as a pdf, references to figures, equations, citations or tables are clickable links to the object referenced.

Magnus Sebastian Christensen
magnussc92@hotmail.com
June 2017
Trondheim

Abstract

There are many interesting materials under research for spintronics today, including the Transition Metal Dichalcogenides (TMDCs). They are particularly interesting because they collectively host materials which are semiconductors, superconductors and ferromagnets and combinations of these properties are useful for a myriad of spintronic applications. Since MoS_2 is a cheap and available semiconducting crystal, and Dy is the most magnetic of all elements, we attempt to create thin films of Dy on a MoS_2 substrate to see if we can create a periodic structure, and if so, if we get an interesting combination of the physical properties of MoS_2 and Dy.

By X-ray Photoelectron Spectroscopy (XPS), we find a chemical reaction takes place upon deposition of Dy on a -140 °C MoS_2 sample, and that subsequent annealing will, given a thick enough Dy layer, cause some DyS_x compound to move from the MoS_2-Dy interface to the Dy surface. The results indicate a S:Dy ratio of 2 when the Dy film is saturated with S.

The crystallinity of the created DyS_x should be investigated before its electrical and magnetic properties are determined for a thorough consideration of possible application to spintronic devices.

Sammendrag (Norwegian Abstract)

Mange interessante materialer forskes på for øyeblikket for potensiell bruk i spintronikk, inkludert de såkalte TMDC-materialene. Disse er spesielt interessante ettersom de inkluderer materialer som er halvledere, superledere og ferromagneter, og fordi kombinasjoner av disse egenskapene kan nyttes i et hav av ulike spintronikkomponenter. Fordi MoS_2 er en billig og tilgjengelig halvledende krystall, og fordi Dy er det mest magnetiske av alle grunnstoffer, forsøker vi å lage tynne filmer av Dy på et MoS_2-substrat for å sjekke om vi kan produsere en periodisk struktur, og hvis det er tilfellet, om vi får noen interessant kombinasjon av de fysiske egenskapene til MoS_2 og Dy.

Ved røntgen-fotoelektronspektroskopi (XPS) finner vi at en kjemisk reaksjon finner sted umiddelbart etter deponering av Dy på en -140 °C kald MoS_2-prøve, og at påfølgende oppvarming forårsaker, gitt en tykk nok Dy-film, noe DyS_x til å bevege seg fra MoS_2-Dy grenseflaten til Dy-overflaten. Resultatene indikerer et S:Dy-forhold på 2 når Dy-filmen er mettet med S.

Krystalliniteten av det produserte DyS_x bør undersøkes før dets elektriske og magnetiske egenskaper bestemmes for en grundig vurdering av mulige anvendelser innen spintronikkomponenter.

Contents

List of Tables

List of Figures

14

Abbreviations

ARPES   Angle Resolved Photoelectron Spectroscopy

FWHM   full width at half maximum

LEED   Low Energy Electron Diffraction

PES   PhotoElectron Spectroscopy

TMDC   Transition Metal Dichalcogenide

UHV   Ultra High Vacuum

XPS   X-ray Photoelectron Spectroscopy

1. Introduction

1.1 Background and motivation

Note: *Chapters 1, 2 and 3 are largely based on the preliminary work described in Ref. 1.*

In our world of electronics, there is always a pressure for the development of new electrical components with certain characteristics. Such characteristics could be faster signal processing speeds, wider signal transmission bandwidths (faster data transfer), lesser energy consumption, cheaper production and smaller component sizes, to mention just a few. With the advance of quantum mechanics, this pressure on development of electrical components has introduced economical incentive to research the field of spintronics.

"Spintronics" is a portmanteau of "spin transport electronics" and refers to the field of electronics expanded to also include effects caused by the spin (a quantum mechanical intrinsic property) of the electron, as opposed to just the electric charge of the electron which is used in electronics. The spin of the electron has been known since the advent of quantum mechanics in the first half of the 20th century and has been used in for instance imaging techniques (like magnetic resonance imaging in medicine) for many years already. In spintronics, however, the spin of the electron is both being manipulated and transported in controlled ways. A crucial point is that the spin manipulation and transport is combined with the charge transport of existing electronic components. This includes handling the behavior of spin in crystals like metals and semiconductors. Though several spin effects were predicted theoretically before they were confirmed experimentally, the first experimental confirmations started appearing around 1985 when Ref. 2 succeeded in injecting a spin-polarized (i.e. having a different amounts of spin up and spin down) current of electrons through an interface between a ferromagnetic and a paramagnetic metal.

Some years later a range of devices utilizing the spin of electrons has been conceived and experimentally investigated, some of which are summarized by Ref. 3 in 2001, and more thoroughly again in 2004 by Ref. 4. Common for these spintronic applications is that in order to fully realize experimentally the desired phenomena, it is necessary to understand the behavior of spin in crystals, which boils down to understanding the crystals

themselves. As described in Ref. 5, the spin polarization of electrons in crystals is closely related to different kinds of symmetries in the crystal. Therefore, to utilize the potential of the field of spintronics, an extensive knowledge of the crystal properties of different materials is called for. Furthermore, we need to understand the interplay between different kinds of materials. To make different spin manipulation devices interplay with traditional electronic devices, we must understand the effects of interfaces between materials which are combinations of ferromagnets, diamagnets, paramagnets, semiconductors, metals and superconductors.

A current hot topic of research for materials to use in spintronics is the group of materials known as Transition Metal Dichalcogenides (TMDCs). These materials are specifically interesting because they collectively host all the properties mentioned above: They are nonmagnetic in the bulk form but the broken inversion symmetry at the surface causes spin polarization[6][5], even of the bulk states[7]. Some TMDCs are even superconducting[8][9]. Adding impurities to TMDCs has been shown to create ferromagnetic and anti-ferromagnetic phase transitions[10], induce superconductivity[11] and change the superconducting temperature[12].

TMDCs constitute a class of compounds of the composition MX_2. M here is a transition metal like Mo, Ni and Zn and they are all similar in that they have interesting physics caused by the d-subshells (electron states with orbital angular momentum quantum number equal to 2) partially filled. X is a chalcogen like S, Se and Te, which is any element from the 16. group (column) of the periodic table, though usually not referring to O.

In this work we will consider the specific TMDC MoS_2. This choice is motivated by its being easily and cheaply accessible in good quality, while it has similar crystal structure and electronic properties to more exotic TMDCs and hence serves as a fitting stepping stone to future research. Inspired by the work of Ref. 13 from the spring of 2016, which investigated the effects depositing Fe on the surface of MoS_2 has on the electronic properties of MoS_2, we are here investigating the effects of depositing not Fe, but Dy, which is the most magnetic element in the periodic table. Preliminary work on such a Dy-MoS_2 system was carried out in the fall of 2016 in Ref. 1, with the intent of adding Dy as a magnetic impurity in the structure of MoS_2. However, it was found strong evidence of a chemical reaction taking place between Dy and S, forming a compound DyS_x, and that S is moving toward the surface during annealing. It is in light of those results that this work finds inspiration to investigate this reaction further.

Specifically, we wish to investigate at which temperatures the reaction takes place, and in what manner S moves to the surface.

1.2 Approach

A Dy-MoS$_2$ system is created by evaporating Dy close to the MoS$_2$, such that a film of Dy condenses on the MoS$_2$ surface. To keep the sample clean, this, and the subsequent measurements, are carried out in an Ultra High Vacuum (UHV) chamber. The measurement technique we use is X-ray Photoelectron Spectroscopy (XPS) which is chosen because it is adept at identifying chemical bonds and the amounts present of each element. It is also a very surface sensitive technique, i.e. measuring only the topmost nanometers of a sample, which in combination with the reactivity of Dy is why we need the cleanliness supplied by an UHV chamber.

In this project the XPS data is used to quantify the amount present (close to the surface) of different elements. These quantities are measured in four different systems. For both a thick (24 Å) and thin (6 Å) Dy film, we measure these amounts' dependence on both annealing temperature and photoelectron emission angle. That is, we measure a thin and thick film temperature series and a thin and thick film angle series.

The temperature series are measured after depositing Dy on cold MoS$_2$ and annealing in steps from around -100 ℃ to around 350 ℃, and its purpose is to show at which temperature the aforementioned chemical reaction between Dy and S happens.

The angle series is intended to show an angle dependence of the Dy-MoS$_2$ system, which is strongly connected to the way in which XPS is consecutively less able to measure the deeper layers of the sample. We use this attenuation of the signal from deeper layers to simulate a theoretical signal from a (layer) model. This theoretical signal depends on the relative concentrations of each element in each layer of the model. We (curve) fit the theoretical signal to the experimental one computationally by varying, and hence estimating, the relative concentrations. The goal is that these relative concentration estimates can show us how the different elements (S, Mo and Dy) are distributed with depth in the sample. For the thick film angle series we also wish to compare the aforementioned distribution in a Dy-MoS$_2$ system before and after annealing to investigate the hypothesis of Ref. 1 that S (in the form of DyS$_x$) moves toward the surface.

The source of X-rays for the thick film angle series is an X-ray gun, instead of a synchrotron which is the case for the rest of the measurements. This is because of limited time access to a synchrotron and because the X-ray gun is less surface sensitive and hence better to use on a thick film Dy-MoS$_2$ system than a thin film Dy-MoS$_2$ system. The synchrotron is the X-ray source of choice for two major reasons; It has a high X-ray intensity so the signal-to-noise ratio of the data becomes better, and it allows the user to chose the photon energy from a continuous spectrum which can be used among other things to make the measurements more surface sensitive. The latter is an advantage because it makes the angle series give more accurate information about the topmost layers of the sample.

2. Theory

In this chapter we present theory relevant to understanding the results and discussion of Chapter 4, starting by presenting essentials on MoS_2 and then proceeding to explain the general workings of XPS, how we model an XPS signal theoretically and how we fit the model to experiments with curve fitting. Details on the numerical treatment of the XPS data is presented in Chapter 3.

2.1 MoS$_2$

This section is based on Ref. 14. MoS_2 is a layered crystal of hexagonal structure as shown in Fig. 2.1. The three atomic layers S-Mo-S is called a monolayer, and as the figure shows, every other monolayer have the same orientation while the ones in between have the opposite. This is in fact only one of several structures - or polytypes - MoS_2 may form, denoted 2H, 3R and 1T. The one shown here is the 2H structure, which is the most abundant in nature and what we consider in this report. The atoms within a single monolayer are bound by covalent bonds while the monolayers are bound together only by weak van der Waals forces. Thus a crystal under stress usually shears between two monolayers. A photo of a real MoS_2 sample is shown in Fig. 2.2.

Figure 2.1: 2H-MoS$_2$ crystal structure with lattice parameters a and c. Taken from Ref. 13.

2.2 XPS in a nutshell

XPS is a technique under the umbrella term PhotoElectron Spectroscopy (PES). All PES techniques can in utmost simplicity be described as photons hitting a sample, kicking out electrons via the photoelectric effect, where the electron is called a *photoelectron*. In XPS the photons are X-rays and we measure the kinetic energy of the photoelectrons leaving

the sample in a given direction, i.e. measuring only one angle of emission. This angle is the angle between the velocity of the electron and the normal of the sample surface. These measurements could be done with any photon energy, but X-rays are suitable for looking at the electrons close to the atom core - *core electrons* - because the photon energy is large enough to excite a photoelectron out of the sample, but not so large that the kinetic energy of the photoelectron is a problem for the experimental equipment. An overview schematic of an XPS system is shown in Fig. 2.3, where the color of the light is meant to represent multichromatic X-rays, not visible light.

The main data XPS can give us is the binding energies of the core electrons present in the sample. This is a powerful technique for analyzing a sample for constituent elements and their chemical environment because the binding energies are element- and environment dependent. However, due to electron-electron interactions such as inelastic scattering, electrons are unlikely to be detected from deep in the sample, and hence XPS is a technique only able to measure the top few atomic layers of the surface. This means the measurements are easily disturbed by contaminants on the sample surface, which calls

Figure 2.2: Photo of MoS$_2$. Taken from Ref. 15.

for the sample to be contained in UHV during the measurements. Vacuum in the system is assumed in this section.

Reading information about the binding energy of an electron from a measured XPS spectrum and translating it into information about which elements are present and their chemical environments is not trivial because the photoelectron may experience a number of processes that change the kinetic energy we measure. Understanding the experimental setup of an XPS system and how it affects

Figure 2.3: XPS schematic. A photon source (a) produces multichromatic light (b). Monochromator/attenuator (c) selects a single frequency (d), which hits the sample (e). Photoelectrons (f) are created, and the ones ejected at emission angle θ (g) are detected by the electron energy analyzer (h). The output signal (i) is processed by some software (j) which produces the final output digital signal (k).

the measurements is therefore essential to a good data analysis. On this basis, this section contains a theoretical description of the basic physics going on in an XPS setup, from photon creation to final signal. In subsections 2.2.1-2.2.5 we look at the physics of each of the elements in Fig. 2.3 separately. Subsection 2.2.6 translates this physics into an expression for the measured X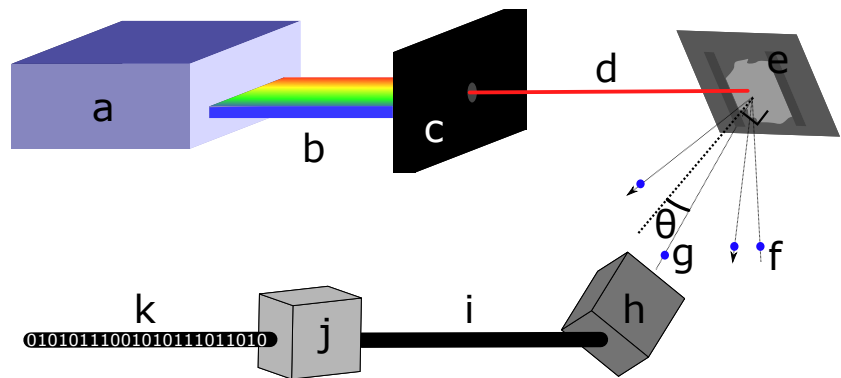PS spectrum, and subsection 2.2.7 makes some approximations in this expression to make it applicable to data analysis. Lastly, subsection 2.3 presents what a curve fitting model is and how we use it.

2.2.1 X-ray Source and monochromator

The X-ray source (a in Fig. 2.3) is usually an anode-cathode pair in an X-ray gun or a beam of accelerated electrons in a synchrotron. In order to discuss their properties, let us define an arbitrary photon energy as E_P, and the photon intensity out of the photon source as $I_{ps}(E_P)$ (b in Fig. 2.3).

An X-ray gun produces photons by accelerating electrons from a cathode onto an anode which then ejects photons in the

Figure 2.4: Example of X-ray gun intensity I_{ps} before and intensity I_{pm} after passing through an attenuator. Greek letters indicate the different relaxation processes that produce the X-rays. In this case, the attenuator makes the intensity of the $\alpha_{1,2}$ peak stronger relative to the other peaks.

X-ray range. In this case I_{ps} contains peaks at several E_P due to the different relaxations that may occur in the anode as shown in Fig. 2.4. Which E_P these peaks appear at, and their intensities, depend on the material of the anode and the electron acceleration voltage respectively. To get a more - but not perfectly - monochromatic X-ray beam, which is important to make the data easier to analyze, it can be passed through an attenuator (c in Fig. 2.3) that has a large absorption coefficient for the unwanted peaks, and a large transmission coefficient for the wanted peak. We term the intensity $I_{pm}(E_P)$ after the attenuator (d in Fig. 2.3) and an example of the effect of such an attenuator is also shown in Fig. 2.4. Due to I_{ps} being significantly intense only at a small number of photon energies (peaks), it is not possible to choose the energy of the peak of I_{pm} with any significant degree of freedom without a change of anode material. X-ray guns are therefore in practice only able to provide a limited discreet set of photon energies E_P, with as many elements in the set as we have X-ray gun anodes.

This brings us to the more advanced option, the synchrotron, and what advantage it has over the X-ray gun. The synchrotron creates the photons by accelerating electrons moving at velocities close to the speed of light. These emit radiation which is directed
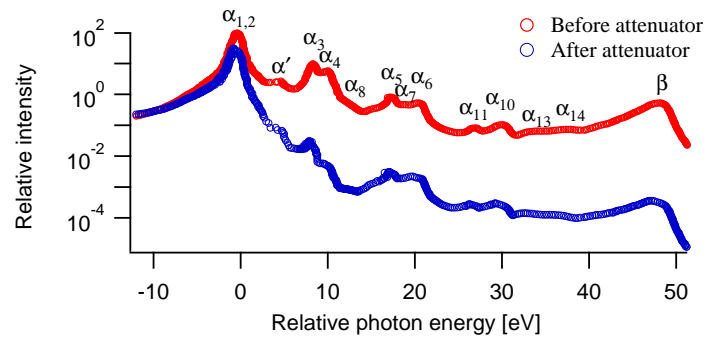
into the experimental setup by a set of focusing mirror lenses. In this setup, I_{ps} contains considerable contributions from a wide range of E_{P}, including the X-ray region, and is less characterized by peaks and more by a smooth curve, compared to an X-ray gun. An example of I_{ps} from a synchrotron is shown in Fig. 2.5. Note the wider range of energies and the smoother shape of the intensity compared to that of the X-ray gun in Fig. 2.4.

The photon distribution I_{ps} is in the case of the synchrotron guided into a monochromator which is not an attenuator like for the X-ray gun, but a device relying on diffraction. The monochromator works by letting the X-rays hit a diffraction grating. The different photon wavelengths will then constructively interfere at different diffraction angles. By selecting at which angle we place a slit that lets photons through to the sample - or by rotating the diffraction grating - we can then
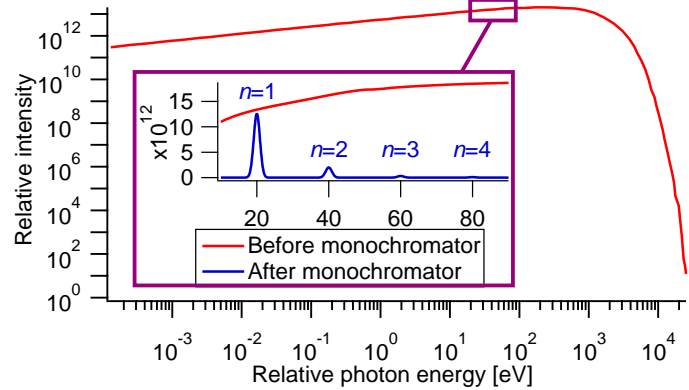
Figure 2.5: Example of synchrotron intensity I_{ps} before and intensity I_{pm} after passing through a monochromator. n annotates peaks at energies of n_{th} order photons in the case of selecting the monochromator to supply photons with energy $E_{\text{m}} = 20$ eV.

select the energy E_{m} of photons getting through the monochromator. I_{pm} then contains a narrow peak of photon energies close to E_{m}, as shown in the inset of Fig. 2.5. We exclude E_{m} as an argument of I_{pm} because it is constant for a given experimental setup. However, the criterion for constructive interference at the slit is what determines E_{m}, and this criterion will also be satisfied for higher energies, so in fact I_{pm} will contain finite-width peaks at all $E_{\text{P}} = n E_{\text{m}}$ for all positive nonzero integers n. We call the photons satisfying this criterion n_{th} order photons. Luckily, the peaks are weaker for increasing n as illustrated in the inset of Fig. 2.5. Further in this section, 1_{st} order photons are assumed.

Due to the broad energy distribution in I_{ps} (i.e. I_{ps} being significantly larger than zero for a broad range of photon energies), we can choose E_{m} from a continuous range of values without significantly reducing the photon flux reaching the sample, which is not feasible with the more discrete peaks from an X-ray gun. Later, in subsection 2.2.7, we will make assumptions about the shape of I_{pm}.

2.2.2 Electronic states - Notation and properties

Before we look at the photoionization (e in Fig. 2.3), we should define some notation and parameters describing the electronic states in an atom. The electronic (quantum mechanical) states of the electrons close to the atom cores in the sample - the core electrons - are described using notation like for the hydrogen atom. (The outer electrons have quite

different properties and are not relevant in this report.) We term such an inner electron state a *core level*, and denote it $\mathrm{A}nl_j$, where A is the element the electron belongs to. The parameters n, l and j are quantum mechanical quantities originating from the quantum mechanical solution of the hydrogen atom (neglecting electron-electron interactions). A detailed work on the hydrogen atom can be found in Ref. 16, and for now we just remind the reader that n, l and j represent the physical properties *energy*, *orbital angular momentum* and *total angular momentum*, respectively. The electronic states also have a dependence on the orientation of the angular momentum, which is represented by the quantum mechanical parameter m_j. In the absence of a magnetic field, which we assume to be case in this report, there are no physical properties of the system dependent on m_j. We therefore neglect m_j from the core level notation.

We should also note at this point the highly used notation for l, where each value of l is assigned a letter. The notation is shown in Table 2.1.

Table 2.1: Alternative notation for l-values.

l	0	1	2	3	4	5	>3
Alternative notation	s	p	d	f	g	h	Alphabetical

2.2.3 Photoionization and Scattering

This subsection describes the photoionization where the photoelectron is excited out of an atom in the sample (e in Fig. 2.3), and the scattering it may experience on its way out of the sample. We will look at the different physical properties that affect the probability of which kinetic energies we measure at which photon energies. We find an expression for this probability, which we will revisit in subsection 2.2.6 and translate into the measured XPS spectrum. We look only at electrons that scatter elastically after the photoionization process, though the curve fitting model discussed in subsection 2.3.1 will include a way to handle the neglected inelastic contribution to the XPS spectrum.

Energy conservation

Since we want to understand how the XPS spectrum is created, and XPS only measures the number of photoelectrons ejected from the sample as a function of the kinetic energy, we need to know what kinetic energy to expect for a given photon energy E_P. This relation is simply given by energy conservation, which for the elastic scattering case we are considering now, is

$$E_\mathrm{KAV}(E_\mathrm{P}, \mathrm{A}nl_j) = E_\mathrm{P} - E_\mathrm{B}(\mathrm{A}nl_j) - \phi_\mathrm{A}. \tag{2.1}$$

For a graphical representation of relevant energies we refer to Fig. 2.6. E_{KAV} is the kinetic energy of the ejected photoelectron when it is in the vacuum in the (electron energy) analyzer, which is where the kinetic energy actually is measured. The binding energy $E_{\mathrm{B}}(\mathrm{A}nl_j)$ is the energy required for the electron to get from its initial state to the Fermi level of the sample E_{FS}. E_{FS} is equal to the (electron energy) analyzer Fermi level E_{FA} because the sample and analyzer have been grounded together. ϕ_{A} is the analyzer work function, which is the energy required to get from the analyzer Fermi level E_{FA} to the vacuum in the analyzer. ϕ_{A} is excluded as argument of E_{KAV} since it is a constant of the setup.

We should note here that the kinetic energy we are interested in from this point onwards, is the kinetic energy we actually see in our XPS spectra. The measurement

Figure 2.6: **Energy diagram when the sample and analyzer are grounded together, showing how the kinetic energy $E_{\mathbf{KAV}}$ in the analyzer vacuum is different from the kinetic energy $E_{\mathbf{KSV}}$ in the vacuum just outside the sample. The sample work function $\phi_{\mathbf{S}}$ and the analyzer work function $\phi_{\mathbf{A}}$ are different, and $E_{\mathbf{KAV}}$ is measured with an offset $\phi_{\mathbf{A}}$ relative to the common Fermi level $E_{\mathbf{FS}} = E_{\mathbf{FA}}$. Adapted from Ref. 17.**

software compensates for the system dependency on ϕ_{A} by outputting a kinetic energy compensated for ϕ_{A}. Hence the kinetic energy of interest further is

$$E_{\mathrm{K}} = E_{\mathrm{KAV}} + \phi_{\mathrm{A}}, \tag{2.2}$$

which is the kinetic energy of the electron while still *in the sample*. Though this is not the actual kinetic energy of the electron as measured, it is a useful substitution because the energy conservation of Eq. (2.1) is simplified to

$$E_{\mathrm{K}}(E_{\mathrm{P}}, \mathrm{A}nl_j) = E_{\mathrm{P}} - E_{\mathrm{B}}(\mathrm{A}nl_j), \tag{2.3}$$

where the system independent (which makes for easier comparison with other work) binding energy E_{B} can be found by only knowing the photon energy E_{P}, which is known for each measurement.
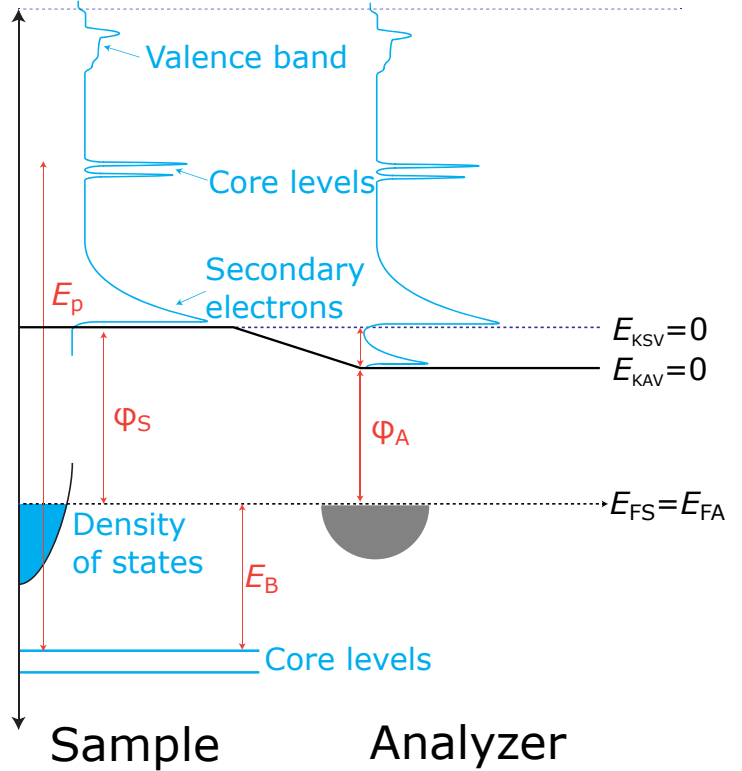
Photoionization cross-section

When a photon hits the sample, the probability that photoelectric effect takes place is proportional to the photoionization cross-section, which depends on the photon energy and the core level in question:

$$\sigma(E_{\mathrm{P}}, \mathrm{A}nl_j). \tag{2.4}$$

A quite comprehensive list of values for σ, showing the photon energy dependence, can be found in Ref. 18.

Multiplicity

The probability of a photoelectron being created from a given core level $\mathrm{A}nl_j$ is also proportional to the number of electrons each atom of element A has in that particular state. Since we have assumed no external magnetic field, the electron energy is the same for all states of different m_j and equal $\mathrm{A}nl_j$[16]. This means the probability of an electron being ejected from $\mathrm{A}nl_j$ is also proportional to g_{j} which is the multiplicity of $\mathrm{A}nl_j$. It is equal to the number of possible values of m_j and is given by[16]

$$g_{\mathrm{j}}(j) = 2j + 1. \tag{2.5}$$

Electron attenuation length

As we are under the assumption that photoelectrons undergo no inelastic scattering after the photoionization, we must consider the probability that a photoelectron indeed leaves the sample without suffering energy loss from inelastic scattering. This probability is exponentially damped with the distance the electron has to move to reach the sample surface:

$$e^{-\frac{d}{\lambda(E_{\mathrm{K}})\cos\theta}}. \tag{2.6}$$

Figure 2.7: Escape depth and emission angle. A core level electron (a) at a depth d in the sample experiences an escape length $\frac{d}{\cos\theta}$ when the excited electron (b) is emitted at emission angle θ.

Here the depth d is the shortest distance from the atom where the photoelectron is created, to the sample surface, i.e. measured in the direction of normal emission, and θ is the angle between normal emission and the photoelectrons direction, as shown in Fig. 2.7. This equation is the definition of the electron attenuation length $\lambda(E_{\mathrm{K}})$, which is dependent on the photoelectron's kinetic energy E_{K} in the sample. A subtle point here is that due to refraction at the sample surface, the emission angle θ is in fact not the same inside and
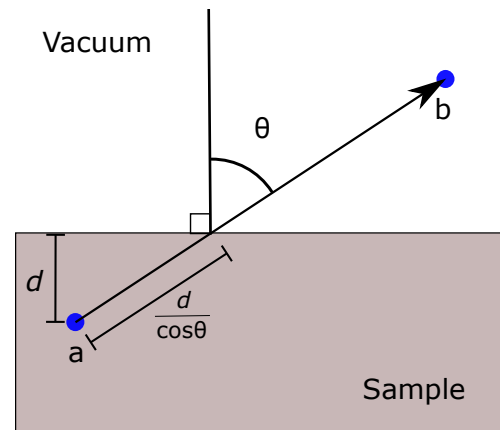
outside the sample. However, this effect is negligible if the kinetic energy change while exiting the sample is small, which we assume is the case in this work.

Core hole lifetime

This paragraph is based on Ref. 17. The last probability we will consider is a result of electron-electron interactions within the atom. A *core hole* is the hole left in the atom when a photoelectron is ejected from a core level, and its lifetime $\tau(\mathrm{A}nl_j)$ is a measure of how long it takes for it to be refilled by another nearby electron. In this process, the electron might come out of the atom with a different kinetic energy than that given by Eq. (2.3). Maybe surprisingly, the photoelectrons may not only come out with a little less energy, but also with more. One might think of it as the electrons around the core hole collapsing quickly enough after the photoelectron is ejected that there is some probability the photoelectron is receiving some extra push by the collapse. We define the probability of the electron having an extra kinetic energy ΔE_K instead of the originally expected E_K to be

$$p_\Delta(\Delta E_\mathrm{K}, \mathrm{A}nl_j). \tag{2.7}$$

We include the dependence on $\mathrm{A}nl_j$ because the core hole lifetime depends on $\mathrm{A}nl_j$. The shorter the lifetime, the faster the collapse, and the larger the energy change tends to be. We will later, in subsection 2.2.7, make assumptions about the shape of p_Δ.

Total probability

As we now head toward combining the above probabilities for a total probability, the curious reader might wonder why have not included other coordinates than depth in our probability. The reason is worth noting - we assume the sample is isotropic in the directions parallel to the monolayers of Fig. 2.1, at least in the area where the photons hit the sample.

Now, combining the above four probabilities, the total probability of photoelectrons leaving the sample with an extra energy ΔE_K from a core level $\mathrm{A}nl_j$ and depth d when hit with a photon of energy E_P is then

$$p_\mathrm{d}(E_\mathrm{p}, \Delta E_\mathrm{K}, \mathrm{A}nl_j, d) = N p_\Delta(\Delta E_\mathrm{K}, \mathrm{A}nl_j)\sigma(E_\mathrm{P}, \mathrm{A}nl_j)g_\mathrm{j}(j)e^{-\frac{d}{\lambda(E_\mathrm{P}-E_\mathrm{B}(\mathrm{A}nl_j)+\Delta E_\mathrm{K})\cos\theta}}. \tag{2.8}$$

The d subscript is to separate this probability from one introduced later where the dependence on d has been integrated away. N is a setup dependent constant and we have substituted E_K through Eq. (2.3). θ is excluded as argument of p_d because it is constant for any given setup, and completely independent of the other parameters of this chapter. In subsection 2.2.6 we will use Eq. (2.8) to find an expression for the XPS spectrum we measure.

2.2.4 Electron detection

The photoelectrons leaving the sample at the correct angle (g in Fig. 2.3) will reach the electron energy analyzer (h in Fig. 2.3). Here they are first decelerated by an electric potential, before they proceed into a hemisphere of two conducting plates at different electric potentials. Only the electrons with a given velocity will pass through the hemispheres without being absorbed by the walls. In the other end of the hemisphere each electron triggers an electron multiplier which hits a fluorescent screen that lights up. This light is mapped to a signal of one electron at the kinetic energy required (pass energy) for that particular deceleration voltage. By temporally scanning the signal for different deceleration voltages, one finds the XPS spectrum (i in Fig. 2.3).

2.2.5 Software processing

The specific setup of the system has a software (j in Fig. 2.3) that handles the parameter ϕ_{A}. In the case we consider, ϕ_{A} is added to the kinetic energy such that the final output XPS spectrum (k in Fig. 2.3) is the intensity as a function of kinetic energy E_{K} that the electron had in the sample. This is very useful because then the calculated binding energies will be system independent, and because the kinetic energy in the final output XPS spectrum is the one relevant for evaluating λ.

2.2.6 From probability to signal

We are now going to revisit the probability of Eq. (2.8) and translate it into the measured XPS spectrum. The final result should be an intensity as a function of only kinetic energy, and to match the output we are going to get from our software, we choose to formulate the intensity as a function of E_{K} and not any other of the previously introduced measures of kinetic energy. To achieve the desired result, the strategy is to weight the probability p_{d} with number of photons ($I_{\mathrm{pm}}(E_{\mathrm{P}})$) we have of a given energy (E_{P}), and then find the number of electrons we get at a given kinetic energy (E_{K}) by summing (in this report integration is taken to be a kind of sum) over all ranges of all parameters involved (d, ΔE_{K}, E_{P} and $\mathrm{A}nl_j$).

Summing over d

Since the depth parameter d only appears in the probability of Eq. (2.8) (i.e. not in I_{pm}), we (can) do the simplest first and sum p_{d} over all d before weighting it with I_{pm}. This gives the probability p of exciting a photoelectron in any depth:

$$p(E_{\mathrm{P}}, \Delta E_{\mathrm{K}}, \mathrm{A}nl_j) = \int_0^\infty p_{\mathrm{d}}(E_{\mathrm{p}}, \Delta E_{\mathrm{K}}, \mathrm{A}nl_j, d)\mathrm{d}d \tag{2.9}$$

$$= Np_\Delta(\Delta E_{\mathrm{K}}, \mathrm{A}nl_j)\sigma(E_{\mathrm{P}}, \mathrm{A}nl_j)g_{\mathrm{j}}(j)\lambda(E_{\mathrm{P}} - E_{\mathrm{B}}(\mathrm{A}nl_j) + \Delta E_{\mathrm{K}})\cos\theta. \tag{2.10}$$

Weighting and summing over $E_{\mathbf{P}}$ and $\Delta E_{\mathbf{K}}$

For simplicity we first introduce the intensity $I_{\mathrm{A}}(E_{\mathrm{K}}, \mathrm{A}nl_j)$ of photoelectrons that are coming from core level $\mathrm{A}nl_j$ and that have a kinetic energy E_{K}. Due to the extra kinetic energy ΔE_{K} that is introduced with probability p_Δ, the energy conservation, Eq. (2.3), becomes

$$E_{\mathrm{K}}(E_{\mathrm{P}}, \mathrm{A}nl_j, \Delta E_{\mathrm{K}}) = E_{\mathrm{P}} - E_{\mathrm{B}}(\mathrm{A}nl_j) + \Delta E_{\mathrm{K}}. \tag{2.11}$$

With this adjustment, we see that for a constant E_{K}, the contributions of photoelectrons to this same energy comes not only from having a photon energy E_{P} and a core hole lifetime-caused kinetic energy change ΔE_{K}, but also from any combination of photon energies $E_{\mathrm{P}} - E$ and kinetic energy changes $\Delta E_{\mathrm{K}} + E$. Summing the contributions from all E is then the same as summing contributions from all E_{P} and ΔE_{K}, and we have:

$$I_{\mathrm{A}}(E_{\mathrm{K}}, \mathrm{A}nl_j) = \int_{-\infty}^\infty I_{\mathrm{pm}}(E_{\mathrm{P}} - E)p(E_{\mathrm{P}} - E, \Delta E_{\mathrm{K}} + E, \mathrm{A}nl_j)\mathrm{d}E. \tag{2.12}$$

Summing over $\mathbf{A}nl_j$

Now, to get the total intensity $I(E_{\mathrm{K}})$ leaving the sample (g in Fig. 2.3), we sum over all $\mathrm{A}nl_j$ contained in our sample:

$$I(E_{\mathrm{K}}) = \sum_{\mathrm{A}nl_j} I_{\mathrm{A}}(E_{\mathrm{K}}, \mathrm{A}nl_j) \tag{2.13}$$

$$= \sum_{\mathrm{A}nl_j} \int_{-\infty}^\infty I_{\mathrm{pm}}(E_{\mathrm{P}} - E)p(E_{\mathrm{P}} - E, \Delta E_{\mathrm{K}} + E, \mathrm{A}nl_j)\mathrm{d}E. \tag{2.14}$$

2.2.7 Approximations

Before this is particularly useful or intuitive, we must rewrite the right side as a function of only E_{K} and $\mathrm{A}nl_j$. This should be possible since we already know that the integral is in fact summing over both E_{P} and ΔE_{K} - it is just a matter of substitution through Eq. (2.11) and shifting the integral. We start by writing out $p(E_{\mathrm{P}} - E, \Delta E_{\mathrm{K}} + E, \mathrm{A}nl_j)$, here called $P(E)$, fully through Eq. (2.10), neglecting for the moment to write the $\mathrm{A}nl_j$ and j

dependencies, and substituting out ΔE_{K} through Eq. (2.11):

$$P(E) = p(E_{\mathrm{P}} - E, \Delta E_{\mathrm{K}} + E, \mathrm{A}nl_j) \tag{2.15}$$

$$= Np_\Delta(\Delta E_{\mathrm{K}} + E)\sigma(E_{\mathrm{P}} - E)g_{\mathrm{j}}\lambda(E_{\mathrm{P}} - E - E_{\mathrm{B}} + \Delta E_{\mathrm{K}} + E)\cos\theta \tag{2.16}$$

$$= Np_\Delta(E_{\mathrm{K}} + E_{\mathrm{B}} + E - E_{\mathrm{P}})\sigma(E_{\mathrm{P}} - E)g_{\mathrm{j}}\lambda(E_{\mathrm{K}})\cos\theta. \tag{2.17}$$

At this point we are going to do two useful approximations, calling the approximated P for P'.

Approximation 1: Photoionization cross-section

The photoionization cross-section σ depends on the photon energy, so it will vary over photon energies contained in I_{pm}. We assume these variations to be small and approximate σ to be constantly equal to its value at the peak E_{M}:

$$\sigma(E_{\mathrm{P}}, \mathrm{A}nl_j) \approx \sigma(E_{\mathrm{M}}, \mathrm{A}nl_j) \equiv \sigma_{\mathrm{M}}(\mathrm{A}nl_j). \tag{2.18}$$

Since $E_{\mathrm{P}} - E$ in Eq. (2.14) runs over all the photon energies in I_{pm}, this is the same as approximating

$$\sigma(E_{\mathrm{P}} - E) \approx \sigma_{\mathrm{M}} \tag{2.19}$$

in Eq. (2.17). This approximation is in practice the statement that the photon energies let through the monochromator or attenuator, for a given core level, all have the same probability of exciting a photoelectron.

Approximation 2: Electron attenuation length

The second approximation we are going to do is in the electron attenuation length λ. It depends on the electron kinetic energy E_{K}, which for a given $\mathrm{A}nl_j$ will vary over a range both due to the range of photon energies E_{P}, and due to the range of core hole lifetime-caused kinetic energy changes ΔE_{K}. We approximate λ to the value it has when the photon energy is E_{M} and the kinetic energy change is 0:

$$\lambda(E_{\mathrm{K}}) \approx \lambda(E_{\mathrm{M}} - E_{\mathrm{B}}(\mathrm{A}nl_j)) \equiv \lambda_{\mathrm{M}}(\mathrm{A}nl_j). \tag{2.20}$$

This approximation is the statement that after the monochromator or attenuator, the combined effect of the presence of photon energies differing from E_{M} and the energy change ΔE_{K}, is not enough to change the kinetic energy of the electron so as to significantly change the probability of it scattering inelastically before leaving the sample.

Using approximations 1 and 2

Introducing these two approximations, Eq. (2.17) becomes

$$P'(E) = N g_{\mathrm{j}} \cos\theta \sigma_{\mathrm{M}} \lambda_{\mathrm{M}} p_{\Delta}(E_{\mathrm{K}} + E_{\mathrm{B}} + E - E_{\mathrm{P}}). \tag{2.21}$$

Inserting Eq. (2.21) into Eq. (2.14) and shifting the integration limits and integration parameter by $E_{\mathrm{P}} - E_{\mathrm{K}} - E_{\mathrm{B}}$ now gives

$$I'(E_{\mathrm{K}}) = N \cos\theta \sum_{\mathrm{A}nl_j} g_{\mathrm{j}} \sigma_{\mathrm{M}} \lambda_{\mathrm{M}} \int_{-\infty}^{\infty} I_{\mathrm{pm}}(E_{\mathrm{K}} + E_{\mathrm{B}} - x) p_{\Delta}(x) \mathrm{d}x, \tag{2.22}$$

where the prime indicates the expression has been approximated. We also recognize the integral as the convolution of $I_{\mathrm{pm}}(x)$ and $p_{\Delta}(x)$, evaluated at $x = E_{\mathrm{K}} + E_{\mathrm{B}}$. We are now going to do two final approximations.

Approximation 3: Gaussian shape of I_{pm}

We assume the intensity $I_{\mathrm{pm}}(E_{\mathrm{P}})$ out of the monochromator / attenuator to be a Gaussian distribution, as shown in Fig. 2.8, defined by a full width at half maximum (FWHM) w_{G}, a center E_{M} and an amplitude a_{G} which the distribution is normalized to:

Figure 2.8: Normalized and centered Gaussian distribution with full width at half maximum (FWHM) equal to w_{G}.

$$I_{\mathrm{pm}}(E_{\mathrm{P}}) = a_{\mathrm{G}} \frac{2\sqrt{\ln 2}}{w_{\mathrm{G}}\sqrt{\pi}} e^{-4\ln 2 \frac{[E_{\mathrm{P}} - E_{\mathrm{M}}]^2}{w_{\mathrm{G}}^2}}. \tag{2.23}$$

Again E_{M} has been excluded as argument since it is a constant of the setup. A comment on this approximation is made in section 4.6.
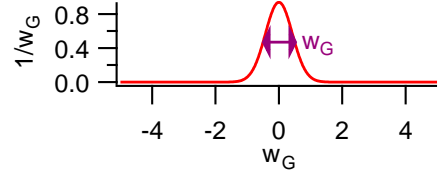
Approximation 4: Lorentzian shape of p_{Δ}

We assume the kinetic energy change probability $p_{\Delta}(\Delta E_{\mathrm{K}}, \mathrm{A}nl_j)$ to be a normalized Lorentzian distribution, as shown in Fig. 2.9, defined by a FWHM $w_{\mathrm{L}}(\mathrm{A}nl_j)$ and a center $\Delta E_{\mathrm{K}} = 0$:

Figure 2.9: Normalized and centered Lorentzian distribution with FWHM equal to w_{L}.

$$p_{\Delta}(\Delta E_{\mathrm{K}}, \mathrm{A}nl_j) = \frac{1}{\pi} \frac{2 w_{\mathrm{L}}(\mathrm{A}nl_j)}{4\Delta E_{\mathrm{K}}^2 + w_{\mathrm{L}}(\mathrm{A}nl_j)^2}. \tag{2.24}$$

A comment on this approximation is made in section 4.6.

Introducing the Voigt distribution

Before we introduce the last two approximations into Eq. (2.22), we will simplify the notation by introducing the Voigt distribution, which is defined as the convolution of the normalized and centered Gaussian and Lorentzian distributions:

$$V(y, w_\mathrm{G}, w_\mathrm{L}) = \frac{2\sqrt{\ln 2}}{w_\mathrm{G}\pi^{\frac{3}{2}}} \int_{-\infty}^{\infty} \frac{2w_\mathrm{L}}{4x^2 + w_\mathrm{L}^2} e^{-4\ln 2 \frac{[y-x]^2}{w_\mathrm{G}^2}}\, \mathrm{d}x. \qquad (2.25)$$

There is no analytical solution to the Voigt distribution, but as shown in Fig. 2.10, the smaller $\frac{w_\mathrm{L}}{w_\mathrm{G}}$ is, the more Gaussian-like the Voigt distribution is, while it looks more Lorentzian-like for larger values. The FWHM of the Voigt distribution, w_V, is a function of w_G and w_L.
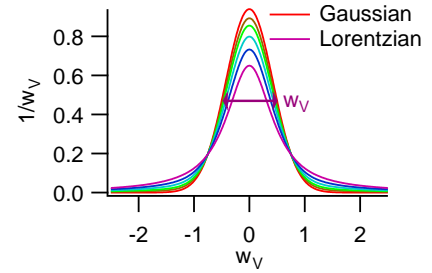

Figure 2.10: Normalized Voigt distribution with FWHM equal to $w_\mathbf{V}$, with $\frac{w_\mathbf{L}}{w_\mathbf{G}}$ ranging from 0 (entirely Gaussian) to much larger than 1 (entirely Lorentzian).

Using approximations 3 and 4

Inserting the two approximations Eqs. (2.23) and (2.24) into Eq. (2.22) and comparing with the Voigt distribution in Eq. (2.25), we see that the approximate intensity I' now becomes:

$$I''(E_\mathrm{ek}) = a_\mathrm{G} N \cos\theta \sum_{\mathrm{A}nl_j} g_\mathrm{j}\sigma_\mathrm{M}\lambda_\mathrm{M} V(E_\mathrm{K} - [E_\mathrm{M} - E_\mathrm{B}], w_\mathrm{G}, w_\mathrm{L}). \qquad (2.26)$$

The double primes remind us I has been approximated twice. Here all the dependencies on $\mathrm{A}nl_j$ have been made implicit. This is the final equation we will use as we now continue toward discussing curve fitting models and its applications. This equation shows the extremely useful result
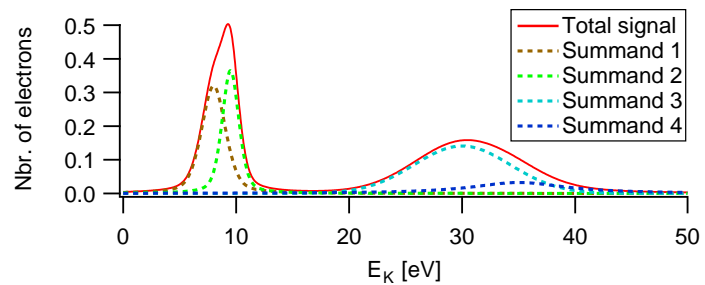

Figure 2.11: Example of signal composed of four separate peaks.

that each core level will contribute a Voigt-shaped peak to the total XPS spectrum, where the intensity, center and width of each peak depends on the core level. An example of a (hypothetical) spectrum containing four peaks is shown in Fig. 2.11.

2.3 Curve fitting

What we get from curve fitting

The whole idea behind curve fitting is to find unknown parameters of a system. We start by creating a theoretical model for a signal created by this system, where the signal depends in the unknown parameters we want to find. The theoretical model for the signal is used to recreate the experimental signal by letting the parameters of the theoretical signal vary until the two signals are as equal as possible. The parameters that satisfy this condition are taken to be true physical values of the system. A curve fitting model is then the combination of the set of *parameters* that vary, and the set of *constraints* that we impose on their variation.

Pitfalls

A big source of fallibility is the large number of degrees of freedom one easily gets in this kind of work. When including too many degrees of freedom, the theoretical signal might find several sets of parameters that recreate the experimental signal, all of which can not be correct. Which set we obtain will depend on the (usually numerical) method used to minimize the deviation between the theoretical and experimental signals. Another pitfall lies at the opposite end of the spectrum, where we try too hard to limit the degrees of freedom and end up imposing unphysical constraints, which also leads to wrong values. Good curve fitting is therefore relying on a large and correct set of physical constraints, which is often hard to come by.

2.3.1 Curve fitting: XPS Peaks

In this report we consider two models of curve fitting. The first model is presented in this subsection (the second model is presented in the next subsection) and is a curve fitting model that reproduces an XPS spectrum by using Eq. (2.26). We will refer to this kind of curve fitting as *peak fitting*.

Background

But before we go any further, we need to consider the inelastically scattered electrons - from now on also referred to as the *background* - that were neglected in deriving I''. This is because the expression for I'' alone not will be able to reproduce the experimental signal which includes inelastically scattered electrons, and hence parameters found by curve fitting would easily be pulled away from their true values in order to compensate for the lacking background term in the model. We will therefore

include a term in the curve fitting that is meant to sufficiently describe the contributions to the XPS spectrum from inelastic scattering. The model we will use for the background is an error function f_{erf} which is shown in Fig. 2.12 and defined as
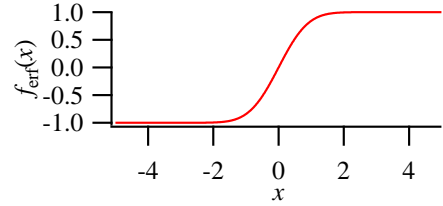
Figure 2.12: The error function $f_{\mathbf{erf}}$.

$$f_{\mathrm{erf}}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathrm{d}t. \qquad (2.27)$$

Each $\mathrm{A}nl_j$ gives a Voigt distribution in the XPS spectrum, and now each $\mathrm{A}nl_j$ will also include an error function with a center equal to that of the Voigt distribution, and a weight $-a_{\mathrm{erf}}$. The minus is to make positive a_{erf} correspond to the physically expected situation of having the background account for energy *loss*, not *gain*. The motivation for choosing the error function is that if the lower integration limit was $-\infty$, then $-f_{\mathrm{erf}}(a)$ would represent the total area of a (normalized and centered) Gaussian in the region $x > a$. Now, the change of integration limits is just the same as adding a constant. Hence, if we include an additive factor, a multiplicative amplitude, and a shift and scaling of x, then we can make the error function at x represent the total number of electrons from a Gaussian peak, with kinetic energy higher than x - this is exactly what we will do. In other words, we model the background at a given energy, from a given peak, to be proportional to the total number of electrons with higher energy, i.e. that can reach the energy in question through energy loss. We will for simplicity use the error function also for peaks that are Lorentzian-like to avoid too many degrees of freedom. An example of a Voigt peak with a background modeled as an error function is shown in Fig. 2.13.

In practice, when curve fitting an XPS spectrum, the spectrum will contain data from a limited range of E_{K}. Voigt distributions from $\mathrm{A}nl_j$ outside this range - i.e. $\mathrm{A}nl_j$ whose Voigt distributions are centered outside this range - are assumed to be zero inside the range, and their corresponding error functions are assumed to be constant. The total contribution from such constants of all the error functions that exist outside the range is modeled as a single constant, O, in the background. O will then effectively function as a constant offset which accounts for all electrons inelastically scattered from higher energies than those contained in the spectrum range. The constants from shifting the integrals of all error functions is also absorbed into O.
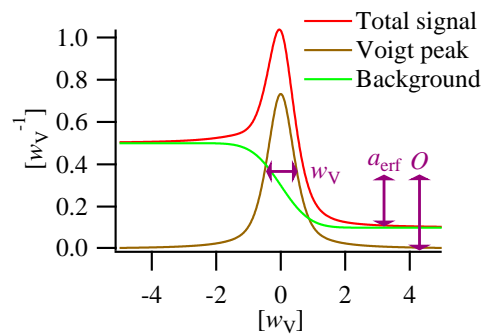

Figure 2.13: The Voigt distribution with background defined by Voigt width $w_{\mathbf{V}}$, amplitude $a_{\mathbf{erf}} = \frac{0.2}{w_{\mathbf{V}}}$ and offset $O = \frac{0.1}{w_{\mathbf{V}}}$.

Complete peak fitting model

Combining the two background terms introduced above, and adding them to I'', we get the general curve fitted signal I_{F} in the form that we will use it in this report:

$$I_{\mathrm{F}}(E_{\mathrm{K}}) = O + \sum_{\mathrm{A}nl_j} \left[a_{\mathrm{V}} V(E_{\mathrm{K}} - E_{\mathrm{c}}, w_{\mathrm{G}}, w_{\mathrm{L}}) - a_{\mathrm{erf}} f_{\mathrm{erf}} \left(\frac{E_{\mathrm{K}} - E_{\mathrm{c}}}{w_{\mathrm{V}}(w_{\mathrm{G}}, w_{\mathrm{L}})} \right) \right]. \qquad (2.28)$$

We have defined the Voigt amplitude a_{V} and the Voigt center E_{c} as

$$a_{\mathrm{V}} = a_{\mathrm{G}} N \cos \theta g_{\mathrm{j}} \sigma_{\mathrm{M}} \lambda_{\mathrm{M}}, \qquad (2.29)$$

$$E_{\mathrm{c}} = E_{\mathrm{M}} - E_{\mathrm{B}}. \qquad (2.30)$$

Eq. (2.28) is now formulated in terms of the parameters that are actually found through peak fitting; a global offset O, and a set containing a Voigt amplitude a_{V}, Voigt center E_{c}, error function amplitude a_{erf}, Gaussian width w_{G} and Lorentzian width w_{L} for each core level. An example of a (hypothetical) XPS spectrum of four Voigt peaks and their error function backgrounds is shown in Fig. 2.14.
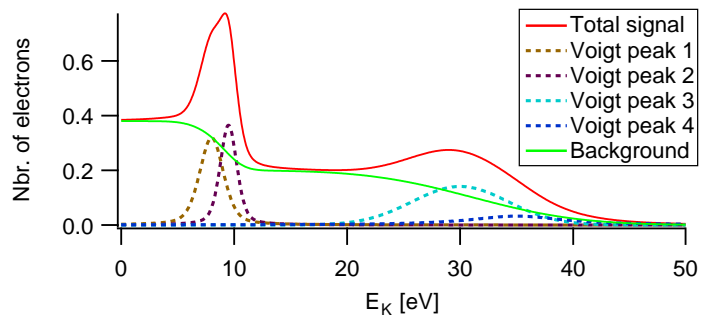

Figure 2.14: Example of signal composed of four separate peaks and their corresponding backgrounds.

Constraints

The shape of Eq. (2.28) implies there are $5n + 1$ parameters to vary, or *degrees of freedom*, when there are n core levels included in the model. Two reduce the number of degrees of freedom, and hence increase the reliability of parameters calculated by peak fitting, we impose constraints on the $5n + 1$ parameters. We can divide these constraints into three types as follows.

The first type of constraints are restrictions on the range of values the five parameters of one core level are allowed to vary over. For physical reasons we restrict a_{V}, E_{c}, a_{erf}, w_{G} and w_{L} to be positive.

The second type of constraints are ones relating one or several of the parameters of one core level to the parameters of another core level. One example could be to constrain the Gaussian widths of two different peaks (core levels) to be equal because the photon source, which introduces the Gaussian distribution in the first place, is the same in both

cases.

The third type of constraint is to limit the number of Anl_j we include in the sum of Eq. (2.28). This is usually done by first measuring an XPS spectrum over a large range of E_K - a wide-scan or broad-scan - and compare table values of core level binding energies (including a reasonable margin of error) with the peaks in our spectrum to identify which elements are present in the sample. Most XPS setups have software that can identify the elements for us. Then, we look up which core levels the present elements have, which have binding energies (again within a reasonable margin of error) such that they would appear in the XPS spectrum we are peak fitting. The sum over Anl_j is now restricted to the core levels satisfying the above criteria.

The numerical details of implementing the peak fitting model is presented in sections 3.5 and 3.6.

2.3.2 Curve fitting: Layer Model

There is one more curve fitting model used in this report, which we will call a *layer model* and a curve fit produced by this model we call a *layer fit*. Before we explain which experimental result we want to reproduce with this model, it is useful to explain the model itself.

We construct a model of the sample by dividing it into N_L layers, where layer nbr. i from the surface, starting at 0, has thickness t_i. In each layer i there is a total number of atoms $N_{i,a}$, and the fraction of these who are of element A is denoted Q_{iA}. All Q_{iA} are the parameters that are being varied in the layer fit. The number of different elements present is N_A, meaning Q is a N_L-by-N_A matrix. Normalization requires all elements in a layer to make up every atom in the layer: $\sum_A Q_{iA} = 1$.

Each atom is modeled to eject the same intensity of photoelectrons, I_a, (which is reasonable because the experimental data we curve fit later will be corrected for factors changing this, including the photoionization cross-section). The intensity from layer i is then attenuated with the thickness $\sum_{k=-1}^{k=i-1} t_k$ of all the above layers combined, according to the attenuation in Eq. (2.6). Note the summation starts at $k = -1$, and we also define $t_{-1} = 0$, such that the sum is defined and equal to zero for $i = 0$, which means the topmost layer has no attenuation. The total intensity I_A from one element but all layers and atoms will then be

$$I_A = \sum_i I_a N_{i,a} Q_{iA} e^{-\frac{\sum_{k=-1}^{k=i-1} t_k}{\lambda_A \cos\theta}}. \tag{2.31}$$

We have added the subscript A to λ to remind us that the different elements may have different values for λ because their photoelectrons may have different kinetic energies. This fact can not be compensated for in the same manner as the intensity from each atom when handling the experimental data, because it is not just a matter of a multiplicative

factor. Hence we include the element dependence of λ in this model.

We will now normalize I_A to what we call the relative intensity R_A as

$$R_A = \frac{I_A}{\sum_A I_A}. \tag{2.32}$$

This is useful because it removes some experimental factors when the experimental data we layer fit is normalized in the same way. Inserting Eq. (2.31) into Eq. (2.32) we get

$$R_A(\theta) = \frac{\sum_i N_{i,\mathrm{a}} Q_{iA} e^{-\frac{\sum_{k=-1}^{k=i-1} t_k}{\lambda_A \cos\theta}}}{\sum_{A,i} N_{i,\mathrm{a}} Q_{iA} e^{-\frac{\sum_{k=-1}^{k=i-1} t_k}{\lambda_A \cos\theta}}}. \tag{2.33}$$

The θ dependence of R_A has been made explicit because R_A as a function of θ is what we will find experimentally and reproduce with this layer fit. The strength of this model is that the relatively easily measured (non-invasively) θ dependence can be used to find Q, which contains explicit information about the depth distribution of each element in the sample. We should point out that any curve fitting requires some sort of deviation between experimental and theoretical (model) signal to be minimized, and in this case, there is one R_A for each core level, but the definition of one R_A includes all elements of Q so all R_A cannot be fit independently. Rather, a compromise must be made by layer fitting every R_A simultaneously, minimizing the (possibly weighted) sum of deviations for all R_A.

In the case of modeling every layer to have the same thickness $t_{i\neq-1} = t_0$, and assuming each layer to contain the same number of atoms, we get the simplified version, R_A':

$$R_A'(\theta) = \frac{\sum_i Q_{iA} e^{-\frac{(i-1)t_0}{\lambda_A \cos\theta}}}{\sum_{A,i} Q_{iA} e^{-\frac{(i-1)t_0}{\lambda_A \cos\theta}}}. \tag{2.34}$$

The numerical details of implementing the layer fit is described in section 3.8.

2.4 Deposition thickness

The probability of an electron escaping the sample is assumed to be exponentially damped, as stated in Eq. (2.6). This implies the intensity I_0 of a core level of the substrate before deposition, and the intensity I_t after being covered by a layer of thickness t, are related by

$$t = \lambda(E_K) \cos\theta \ln\frac{I_0}{I_t}, \tag{2.35}$$

which can easily be seen by summing the intensity from all depths larger than 0, and all depths larger than t, and solving for t. This can be used to calculate the deposition thickness from the attenuation of an intensity.

3. Method

3.1 Equipment

3.1.1 Facilities

Most of the experimental data presented in this report is collected at the MATLINE beamline of the ASTRID2 synchrotron at the University of Aarhus, Denmark. A picture of the lab is shown in Fig. 3.1. The remainder of the data is collected in an XPS lab (with an X-ray gun as X-ray source) at the Department of Physics at NTNU.
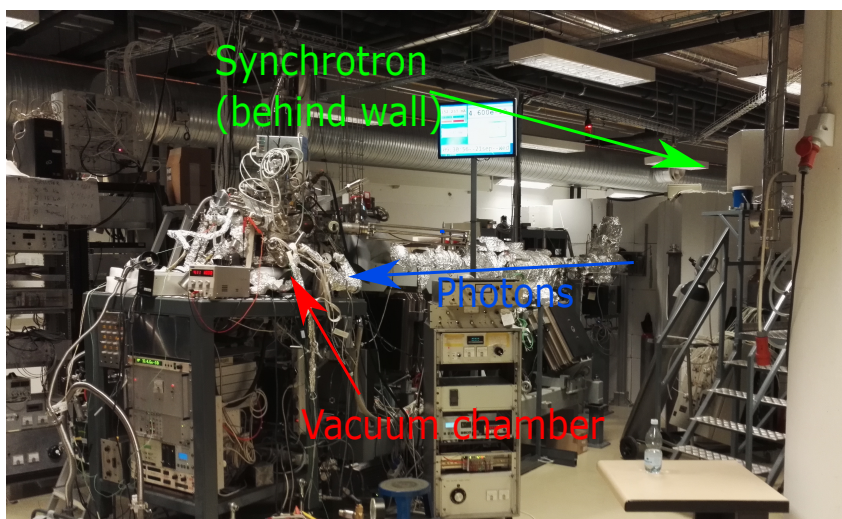

Figure 3.1: Photograph of MATLINE. The X-rays enter from the right (blue) and are guided into the vacuum chamber to the left (red). ASTRID2 is behind the concrete wall to the right (green).

3.1.2 Sample preparation

XPS is a surface sensitive technique because the photoelectrons generated beneath the surface have a low probability of leaving the sample elastically. This is due to the small electron attenuation length (sometimes inaccurately called the electron inelastic mean free path), which is usually in the nanometer scale[17]. This means that small amounts of contamination on the surface is enough to upset the measurements. To minimize the amount of contam-
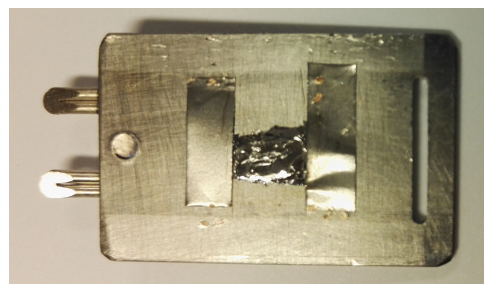

Figure 3.2: Photo of a sample plate with a sample of MoS_2.

inants, several preparation techniques are used.

First, the sample (smaller than 7 mm x 7 mm x 0.2 mm) surface is cleaved off to leave a clean new surface as late as possible before conducting the measurements. This is done by utilizing the weakness between some of the crystal planes in the sample, in attaching scotch tape to the surface and pulling the tape off, bringing a thin layer of the surface with it. After cleaving, the sample is attached to a sample plate, as shown in Fig. 3.2, and inserted into an UHV chamber, as shown in Fig. 3.3, for the remainder of the experiment. This reduces the rate at which new contaminants reach the sample.

3.1.3   Dy evaporator

Dy deposition is performed with a home-made Dy evaporator, as shown in Figs. 3.3 and 3.4. The way the evaporator works is as follows.

A piece of Ta foil is wrapped into a cylinder and welded shut. Inside are pieces of Dy and the cylinder ends are bent to close them. A hole is made in the cylinder for the Dy to be able to escape in the desired direction, and a current is applied through the
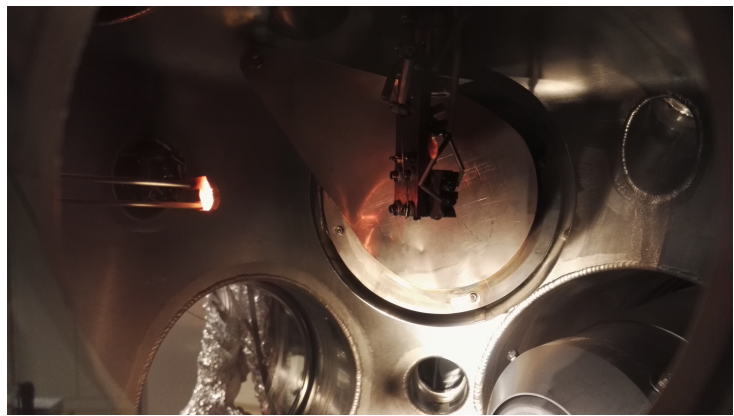

Figure 3.3: Photo of the vacuum chamber inside. To the left the Dy evaporator is glowing, and in the center an arm is holding the sample turned away from the evaporator.

Ta, which gets hot enough to evaporate the enclosed Dy. The deposition is also done at UHV.

3.1.4   Work function

The sample and electron energy analyzer are connected to the same potential with a wire so the sample Fermi level and analyzer Fermi level are aligned, $E_{\text{FS}} = E_{\text{FA}}$, such that Eq. (2.1) holds. The resulting kinetic energy is shifted by ϕ_{A}, such that the peaks from any core level appear at a kinetic energy E_{K} as assumed in deriving Eq. (2.28).

3.1.5   Samples

All measurements are performed on natural 2H-MoS$_2$ crystals from the same mine - Kingsgate mine in New South

Figure 3.4: Photo of a home-made Dy evaporator. Taken from Ref. 19.

Wales, Australia. In total, five different samples are used - one for calculating Dy deposition rate, and one for each of the four measurement series.

3.2   Degassing

Degassing is to remove impurities from a sample by heating it so the impurities evaporate and leave the sample. This is done before deposition so that the subsequent post-deposition annealing will cause less impurities to be released from the bulk MoS_2 into the measured surface layers. This is important because the presence of impurities in the measured layer increases the probability of unwanted chemical reactions or catalytic effects. The choice of temperature is a compromise of having a low enough temperature to not destroy the sample, and having a large enough temperature to remove impurities and in not too much time.

Every sample is degassed before measurement by heating it to 350 ℃ and keeping it there until preliminary measurements show low amounts of carbon and oxygen. Temperature increase is done in steps to avoid pressure increases above 10^{-8} mbar. This works because every time the sample is heated to previously unreached temperatures, new gases are released, so letting the released gasses at one temperature be pumped away before releasing new gases by further heating helps keep the pressure consistently low, which is important to keep the sample as clean as possible. The Dy evaporator is degassed before Dy deposition by heating it to slightly below the evaporation temperature of Dy for 6 to 8 hours.

3.3   Dy Deposition

Deposition procedure

A home-made evaporator like the one described in subsection 3.1.3 is inserted into an UHV chamber with the MoS_2 sample such that the Dy-filled Ta cylinder of the evaporator has a hole pointing towards the surface of the MoS_2 sample. We rotate the sample plate such that it is between the evaporator and the sample while we heat the evaporator by applying a current of 9 A and 3.2 V through the Ta cylinder. The purpose of doing this is to release impurities from the evaporator before we do the actual deposition so the Dy remains as clean as possible. This current is reached by increasing it in steps to avoid the pressure increasing above 10^{-8} mbar due to too rapid heating as discussed above for the degassing of the samples. After approximately 2 minutes at 9 A, the pressure stops increasing from evaporated Dy because the Dy sticks to the vacuum chamber walls and counters the pressure increase by reacting with other gas particles. At this point the sample plate is turned so the sample is facing the evaporator until the desired deposition

time has passed, and the sample is turned away and the current through the evaporator is turned off.

Finding deposition rate

A test sample is used to deposit Dy as described above, for 10 minutes, and the deposition thickness is calculated with Eq. (2.35). The intensities used in this calculation are intensities of the Mo3d core level(s) which are found by curve fitting the Mo3d core level with the peak fitting model of Eq. (2.28) as described below, in section 3.5 and Fig. 3.7. One chemical environment is used for Mo3d before deposition, and two chemical environments after deposition (since we know from Ref. 1 there is a chemical reaction taking place which introduces a second pair of Mo3d peaks). The value used for λ is the same as given in section 3.8. We find a deposition (growth) rate of $\frac{6\,\text{Å}}{5\,\text{minutes}}$.

Temperature series

By running liquid N through a tube thermally connected to the sample (by direct metal contact), the samples used for the temperature series are cooled to approximately -140 °C before deposition of Dy. Then the sample for the thin film temperature series is deposited on for 5 minutes to give a 6 Å Dy layer, while the sample for the thick film temperature series is deposited on for 20 minutes to give a 24 Å Dy layer.

Angle series

The samples used for the angle series are deposited on at room temperature. The sample for the thin film angle series is deposited on for 5 minutes to give a 6 Å Dy layer.

The sample for the thick film angle series is deposited on a bit differently than the cases mentioned above because a different evaporator (and lab) is used. In this case, Dy deposition is carried out over several separate deposition events, until the deposited layer is calculated to be 20 Å to 30 Å. For simplicity we refer to this also as a 24 Å thick-film. The distinction is not important here, and the difference is comparable to the error in the deposition thickness anyway.

3.4 Measurements

General comments

Saying a measurement is done of Anl_j in this report means the E_K range of the spectrum is covering the expected (by table values like for instance those found in Ref. 20) location of the Anl_j peak, and neglecting j from the notation means there only is one possible j value or that we are aiming to measure both j values - or both spin values.

3.4.1 Thick film temperature series

XPS measurements in the thick film temperature series are carried out at the MATLINE synchrotron beamline. The temperature series consists of measurements of core levels directly after Dy deposition on a cold (approximately -140 °C) sample, and after annealing the sample to 300 °C in approximately 50 °C steps, including measurements after each annealing step.

In this measurement series, the core levels measured mainly include S2p. Measurements of C1s and O1s are done roughly (not as good data quality but quicker), just to keep track of impurities, and the Mo3d core level is measured to confirm there is no measurable Mo, which is expected because the Dy film is too thick compared to λ and the hypothesis we are testing is that S, and only S, moves (significantly) toward the Dy surface during annealing.

The S2p core level is chosen for quantification of S because it consists of two separate sharp peaks, which therefore are easy to peak fit, and easy to use to identify chemical changes. These measurements are done at a photon energy $E_P = 230$ eV to maximize the photoionization cross-section according to the values of Ref. 18, and to thereby increase the signal-to-noise ratio. Lower photon energies would increase the cross-section further but is avoided to keep the kinetic energies from getting too low, which would have both the undesired effects of decreasing the surface sensitivity (by increasing λ) and making the electron energy analyzer work outside its optimal energy region.

Mo3d is chosen for quantification of Mo because it, like S2p consist of two well-defined quite sharp and separated peaks, with the same befits as S2p. The S2s core level is in the same region but separated enough that Mo3d should be relatively easy to peak fit. It is measured with $E_P = 400$ eV, C1s with $E_P = 370$ eV and O1s with $E_P = 610$ eV, all of them for the same reasons as for the choice of E_P for S2p.

3.4.2 Thin film temperature series

The measurements in the thin film temperature series are the same as that of the thick film temperature series. The only difference is that in this case the Dy film is thinner and we therefore expect measurable signal from Mo, and hence measure Mo3d with better data quality than for the thick film temperature series.

3.4.3 Thick film angle series

XPS measurements in the thick film angle series are carried out at an XPS lab at NTNU, using an X-ray gun with $E_P = 1253.6$ eV. The core levels measured include S2p, Mo3d and Dy5p to quantify the amounts of S, Mo and Dy, while O1s and C1s are measured to check the purity of the sample. The Mo is expected to be measurable for this thick film, unlike for the thick film temperature series, because the photon energy in this case gives kinetic

energies which result in larger λ and hence less surface sensitivity. The measurement of the aforementioned core levels is carried out for a series of photoelectron emission angles (θ). This angle series is measured straight after Dy deposition and after post-deposition annealing to 300 °C, with the aim to see how the depth dependencies of the Dy and S concentrations change during annealing.

3.4.4 Thin film angle series

XPS measurements in the thick film angle series are carried out at the MATLINE synchrotron beamline. The core levels measured are S2p, Mo3d, C1s and O1s for the same reasons as for the thick film temperature series, but the Dy quantification is done by measuring Dy4d instead of Dy5p. This is because the photon energies available at this synchrotron are too small to excite the Dy5p core level.

The photon energies used for S2p, Mo3d, C1s and O1s are the same as for the other measurements at MATLINE. The Dy4d and S2p core levels have peaks that lie in the same region of binding energies, so when we measure one we also measure the other. For S2p measured at $E_\mathrm{P} = 230$ eV, the photoionization cross-section of Dy4d is so small compared to that of S2p, that the spectrum with good approximation only includes S2p. However, when we want to quantify the amounts of Dy by measuring Dy4d, we chose $E_\mathrm{P} = 450$ eV to increase the Dy4d cross-section relative to the S2p cross-section, but the best we can do is make them comparable to each other. This means measurements we will refer to as measurements of Dy4d show the same region of binding energies as S2p, but S2p contains only S2p, while Dy4d contains both and as much of Dy4d as possible.

The aforementioned core levels are measured at a series of photoelectron emission angles (θ) straight after Dy deposition.

3.5 Curve fitting core level spectra

For all four measurement series (thin/thick temperature/angle) we are interested in finding the total intensity from one or several core levels. For the temperature series we simply wish to see if certain core levels change intensity with annealing, and for the angle series we wish to use the intensities to find relative intensities that we layer fit with Eq. (2.33). In any case, the intensity from a peak in an XPS spectrum is the sum of all data points in that peak, which is approximately equal to the area under the peak if we correct for the energy step between each data point. But since the peaks of different core levels are easier to separate from each-other than their backgrounds, we use the peak fitting model of Eq. (2.28) to find areas of individual Voigt peaks and ascribe the core level intensities to these peaks instead of both peak and background. Further in this section, we describe how we peak fit each spectrum, i.e. which core levels we include in each fit and which constraints we impose.

In general for $l = \mathrm{p}$ and $l = \mathrm{d}$ core levels we impose a constraint on the Voigt amplitudes and error function amplitudes as

$$\frac{a_{\mathrm{V}}(\mathrm{A}nl_j)}{a_{\mathrm{V}}(\mathrm{A}nl_i)} = \frac{a_{\mathrm{erf}}(\mathrm{A}nl_j)}{a_{\mathrm{erf}}(\mathrm{A}nl_i)} = \frac{g_{\mathrm{j}}(j)}{g_{\mathrm{j}}(i)} = \frac{2j+1}{2i+1}. \tag{3.1}$$

This is because the different j core levels have the same probability of photoionization, but a different, and known, number of states, so we know how intense one peak should be compared to the other.

Thick film temperature series

In the thick film temperature series we peak fit the S2p core level as shown in Fig. 3.5. We have imposed the constraint of Eq. 3.1 in the form

$$\frac{a_{\mathrm{V}}(\mathrm{S2p}_{\frac{3}{2}})}{a_{\mathrm{V}}(\mathrm{S2p}_{\frac{1}{2}})} = \frac{a_{\mathrm{erf}}(\mathrm{S2p}_{\frac{3}{2}})}{a_{\mathrm{erf}}(\mathrm{S2p}_{\frac{1}{2}})} = \frac{g_{\mathrm{j}}(\frac{3}{2})}{g_{\mathrm{j}}(\frac{1}{2})} = 2. \tag{3.2}$$

Moreover, since the $S2p_{\frac{1}{2}}$ and $S2p_{\frac{3}{2}}$ core levels have the same core hole lifetime and are photoionized by the same distribution of photon energies, the Gaussian and Lorentzian widths are constrained to be independent of j:

$$w_{\mathrm{G}}(\mathrm{S2p}_{\frac{3}{2}}) = w_{\mathrm{G}}(\mathrm{S2p}_{\frac{1}{2}}) \tag{3.3}$$

$$w_{\mathrm{L}}(\mathrm{S2p}_{\frac{3}{2}}) = w_{\mathrm{L}}(\mathrm{S2p}_{\frac{1}{2}}). \tag{3.4}$$

Thin film temperature series

The S2p data in the thin film temperature series are peak fitted with the model shown in Fig. 3.6. This model consists of two separate models of the kind used for the S2p core level in the thick film temperature series (Fig. 3.5). There is one pair of S2p peaks for one chemical environment, and one pair for another chemical environment. The S2p peak pair of one single chemical environment is constrained in
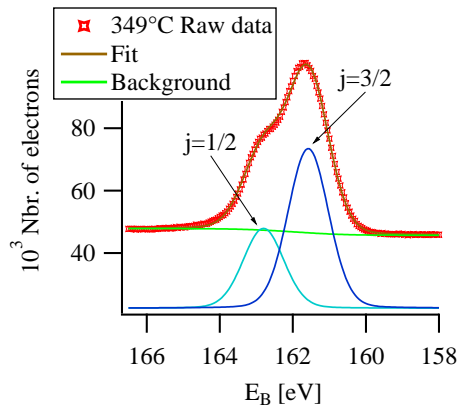

Figure 3.5: Peak fitting model for S2p in thick film temperature series. The model includes S2p$_{\frac{1}{2}}$ and S2p$_{\frac{3}{2}}$. The fit is the sum of the background and the two marked Voigt peaks.
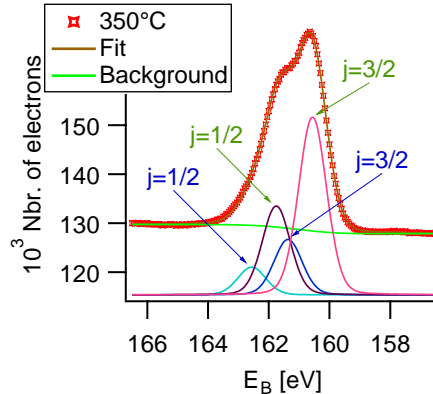

Figure 3.6: Peak fitting model for S2p in thin film temperature series. The model includes one pair of peaks consisting of S2p$_{\frac{1}{2}}$ and S2p$_{\frac{3}{2}}$ from one chemical environment, and a second pair from a second chemical environment. The fit is the sum of the background and the four Voigt peaks marked with green arrows for one chemical environment and blue arrows for a different chemical environment.

just the same manners as for the thick film temperature series. In addition, we restrict the S2p peak pair from the second chemical environment to be identical in every manner to the first chemical environment except for two degrees of freedom: the second pair is shifted (both the $j = \frac{3}{2}$ and $j = \frac{1}{2}$ peaks and their backgrounds by the same amount) relative to the first pair, and the second pair is reduced (or increased) by a factor (the $j = \frac{3}{2}$ and $j = \frac{1}{2}$ peaks and their backgrounds by the same factor) relative to the first pair.

The Mo3d data in the thin film temperature series are peak fitted as shown in Fig. 3.7. This model includes two chemical environments of Mo, with both a $j = \frac{3}{2}$ peak and a $j = \frac{5}{2}$ peak from each chemical environment. The constraints imposed on these four peaks are the same as for the double S2p pair described in the above paragraph, except for the ratio of Eq. (3.1) being $\frac{3}{2}$ instead of 2, and one extra degree of freedom: in the Mo3d core level, the $j = \frac{3}{2}$ peak (and background) is allowed a different Lorentzian width than the $j = \frac{5}{2}$ peak (and background) because they may have different core hole lifetimes (which is typical d-core levels. This width is still the same for the two peaks of different chemical environments and equal j. The Mo3d peak fit model also includes the S2s peak which is modeled as single Voigt distribution with no constraints except that
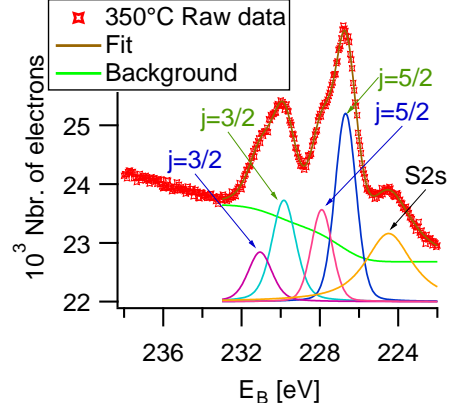

Figure 3.7: Peak fitting model for Mo3d in thin film temperature series. The model includes one pair of peaks consisting of Mo3d$_{\frac{3}{2}}$ and Mo3d$_{\frac{5}{2}}$ from one chemical environment, and a second pair from a second chemical environment, as well as a single S2s$_{\frac{1}{2}}$ peak. The fit is the sum of the background, the four Voigt peaks marked with green arrows for one chemical environment and blue arrows for a different chemical environment, and the S2s Voigt peak.
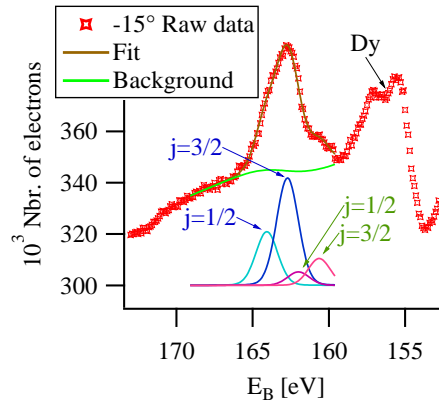

Figure 3.8: Peak fitting model for S2p in thick film angle series. The model includes one pair of peaks consisting of S2p$_{\frac{1}{2}}$ and S2p$_{\frac{3}{2}}$ from one chemical environment, and a second pair from a second chemical environment. The fit is the sum of the background and the four Voigt peaks marked with green arrows for one chemical environment and blue arrows for a different chemical environment. The background includes a linear term.

its Gaussian width is the same as for the Mo3d peaks because they are excited by the same photon energy distribution.

Thick film angle series

In the thick film angle series, the Dy5p core level straight after Dy deposition is peak fitted like the S2p core level in the thick film temperature series (Fig. 3.5), while after post-deposition annealing it is peak fitted with an extra chemical environment like the S2p peak fit in the thin film temperature series (Fig. 3.6).

The S2p core level is peak fitted as shown in Fig. 3.8, both straight after Dy deposition and after post-deposition annealing. This model is like the double S2p model of Fig. 3.6, with one exception: at the photon energy used, S2p and Dy4d are comparable in intensity and they lie in the same region of binding energies, so the background has been modeled not only to include an offset and the error functions (backgrounds) of each S2p peak, but also a linear background term approximating the Dy peaks that lie directly beneath the S2p peaks.

Thin film angle series

In the thin film angle series, the S2p data are peak fitted with double chemical environments like the S2p in the thin film temperature series in Fig. 3.6.

The Dy4d is peak fitted with one model of double S2p like in the above paragraph, which peak fits the S2p part of the Dy4d spectrum. In addition, the Dy4d peak fitting model includes two peaks with no constraints, which model the two most distinct Dy peaks. The model is shown in Fig. 3.9.
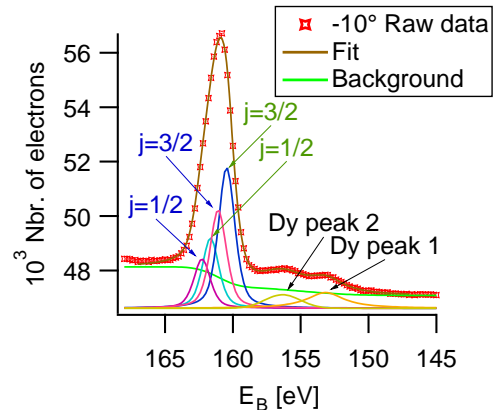

Figure 3.9: Peak fitting model for Dy4d in thin film angle series. The model includes one pair of peaks consisting of S2p$_{\frac{1}{2}}$ and S2p$_{\frac{3}{2}}$ from one chemical environment, a second pair from a second chemical environment, and two Dy peaks. The fit is the sum of the background and the six Voigt peaks marked with green arrows for one S2p chemical environment, blue arrows for a different S2p chemical environment and black for Dy.

3.6 Computational implementation of peak fitting

3.6.1 Software

The peak fitting model of Eq. (2.28) is implemented, and the variation of parameters performed, in the software *Igor Pro 6.37*. The parameters of the model are estimated by finding a least-squares solution for the deviation between the peak fit and the experimental spectrum, using built-in functionality based on the Levenberg-Marquardt method.

3.6.2 The pseudo-Voigt approximation

Due to the Voigt distribution not having an analytical expression, it must be calculated numerically. Brute force computation of the integral in the Voigt distribution can be time-consuming, but there are often efficient algorithms and approximations used when the Voigt distribution is a built-in part of analytical software or packages to programming languages. Such is also the case for our software of choice, *Igor Pro 6.37*: there exists a built-in Voigt distribution, and a look at its source code reveals it is an elegant approximation. However, it also reveals the developer only included it as a test project and that its implementation could change any time. Furthermore, the specific implementation lacks good documentation and is cumbersome to work with when we want to control the Gaussian and Lorentzian widths independently. We therefore implement it numerically ourselves, using an approximation documented in Ref. 21 - the *pseudo-Voigt approximation*, which has the advantage of faster computation times since it has an analytical expression. This approximation is made by replacing the Voigt distribution - which is a convolution of the normalized and centered Gaussian and Lorentzian distributions $G(x, w_\mathrm{G})$ and $L(x, w_\mathrm{L})$ - with the pseudo-Voigt distribution V_PV, which is a weighted sum of G and L where they have a common FWHM called the pseudo-Voigt width w_PV. For a Lorentzian weight μ and Gaussian weight $1 - \mu$, the normalization is conserved and we have:

$$V_\mathrm{PV}(x, w_\mathrm{G}, w_\mathrm{L}) = (1 - \mu)G(x, w_\mathrm{PV}) + \mu L(x, w_\mathrm{PV}). \tag{3.5}$$

Since V_PS approximates a function V which is only characterized by w_G and w_L, V_PS must also be uniquely defined by w_G and w_L. This means μ and w_PV also can be functions of no other parameters than w_G and w_L. In the implementation used is this report, on the basis of Ref. 21, w_PV is implemented as

$$w_\mathrm{PV} = \left[w_\mathrm{G}^5 + 2.69269 w_\mathrm{G}^4 w_\mathrm{L} + 2.42843 w_\mathrm{G}^3 w_\mathrm{L}^2 \right.$$
$$\left. + 4.47163 w_\mathrm{G}^2 w_\mathrm{L}^3 + 0.07842 w_\mathrm{G} w_\mathrm{L}^4 + w_\mathrm{L}^5 \right]^{\frac{1}{5}}, \tag{3.6}$$

and μ is implemented as

$$\mu = 1.36603\left(\frac{w_\mathrm{L}}{w_\mathrm{PV}}\right) - 0.47719\left(\frac{w_\mathrm{L}}{w_\mathrm{PV}}\right)^2 + 0.11116\left(\frac{w_\mathrm{L}}{w_\mathrm{PV}}\right)^3. \tag{3.7}$$

3.6.3 Initial guesses

When using the implementation of the peak fitting model to fit experimental data, the set of parameter values found to minimize deviation, can be lured into a local minimum instead of a global one, if the initial guess for the variation procedure is far from the global minimum. To reduce the probability of that happening, we manually adjust parameters

to make the initial guess resemble the experimental data before starting the parameter variation (the peak fitting). These initial guesses are made by first setting the centers, intensities and widths roughly correct, then running the peak fitting. For the temperature and angle series, this is done at one temperature or angle, and the parameters found from peak fitting this temperature or angle is used as initial guesses for the other temperatures or angles.

3.7 Finding relative intensities

In order to find relative concentrations by using Eq. (2.33), we need to convert our experimental data to the same kind of relative intensity that Eq. (2.33) produces. To achieve that, we start with the peak fitted angle series data. For each core level and angle, we have one peak fitted spectrum which consists of a background and several Voigt peaks. The total area of all the Voigt peaks from that core level is mapped to that core levels total intensity by scaling with the energy step of the experimental data. This total intensity has been affected by the core levels photoionization cross-section, which we correct for by dividing the intensity with the cross-section. Now we have (an intensity proportional to) the intensity as it would have been had each atom emitted the same intensity of photoelectrons. We then divide this corrected intensity, and the ones from different elements/core levels, by the sum of this intensity and the ones from different elements/core levels. What we are left with is the relative intensity we wanted, and this is what we peak fit with the layer fitting model. This last division cancels out all experimental factors that, for a given angle, are the same for all core levels, including the photon flux.

3.8 Computational implementation of the layer model

This section explains how we use Eq. (2.34) to curve fit the experimental relative intensities described in section 3.7, which we will call $R_{\exp}(\theta)$.

We use Eq. (2.34) with layer thicknesses $t_0 = 6$ Å to coincide with the monolayer thickness of MoS$_2$ ($\frac{c}{2}$ in Fig. 2.1), and the number of layers N_{L} is set to 5 because signal from deeper than that does not affect the theoretical signal $R'_{\mathrm{A}}(\theta)$ beyond the uncertainties of the experimental data $R_{\exp}(\theta)$. The values for λ are taken from the NIST database[22]; $\lambda = 3.5$ Å (for S) at a kinetic energy of 62 eV, $\lambda = 4$ Å (for Mo) at a kinetic energy of 172 eV, $\lambda = 7$ Å (for Dy) at a kinetic energy of 295 eV and $\lambda = 15$ Å (for Dy and S) at a kinetic energy from 1025 eV to 1225 eV. We should note that these values are highly uncertain, with an uncertainty in the order of a factor of 2 in either direction (i.e. from 50% to 200%). With the above values as input, the relative concentrations Q_{iA} are found numerically by the method described below.

The relative concentrations Q_{iA} are found numerically by changing them such that the error (defined below) between $R'_A(\theta)$ and $R_{\exp}(\theta)$ is minimized. The method used to accomplish this in this work is a user-written "brute force" algorithm written in the software *Igor Pro 6.37*.

The error which is minimized by this algorithm is a sum of two terms; a deviation term and an entropy term:

$$\alpha \sum_{\theta, A} \frac{\left[R'_A(\theta) - R_{\exp}(\theta)\right]^2}{R'_A(\theta)} + \sum_{i, A} \left[\frac{Q_{iA}}{\sum_i Q_{iA}}\right] \ln \left[\frac{Q_{iA}}{\sum_i Q_{iA}}\right]. \tag{3.8}$$

The range of θ is the set of θ that the angle series contain measurements at. The left term is the deviation term, which is simply the sum of squared deviations between the theoretical and the experimental relative intensities, relative to the theoretical one. There is a scaling factor α to control how much this term is prioritized compared to the right term. Note that the deviation term (without the α) is what is commonly called χ^2 in peak fitting. Its numerator is the square of the deviation between the model (theoretical) and experimental signal, while the denominator is the square of the uncertainty in the number of electrons (the uncertainty in the number of electrons, R, is $\sqrt{R}$). This means that if the deviation term (without the α) is less than the number of data points (elements in the sum), then the fit is deviating less from the experimental result than the uncertainty of the experimental result itself, and hence we are using the fit to extract data from the system which has a significant chance of being incorrect (over-fitting). This is why the entropy term is introduced.

The entropy term is the term to the right, but this term is *minus* the entropy (which comes with a minus of its own). The argument to the natural logarithm can be interpreted as the probability that a random atom of element A is located in layer i. This probability distribution is what we take the entropy of, and the right term is simply this entropy summed up for all elements. The opposite sign ensures minimizing the error corresponds to *minimizing* the deviation in compromise with *maximizing* the entropy. Physically, maximum entropy is achieved when each element has a constant concentration through all layers, meaning any layer is equally probable as the location of a random atom of that element. This is in accordance with the viewpoint that maximum entropy corresponds to maximum disorder, or maximum mixing.

The reason for including the entropy term is to avoid the Q_{iA} to vary unphysically much from layer to layer in attempting to minimize the deviation term (over-fitting as mentioned above). The factor α is there solely to allow us to control the extent at which we wish to safe-guard against over-fitting. The effect of a large α is that the curve fit is only concerned with making the theoretical relative signals coincide with the experimental relative signals, i.e. a pure χ^2 optimization. A small α means the curve fit is only concerned with maximizing the entropy of the relative concentrations,

making the fit reproduce the signal of a uniform sample. An α not too large and not too small gives a compromise between the two effects, but what constitutes large and small depends on the situation, for instance on how many data points (angles) the relative signal plot has, compared to the number of points (layers in the layer model) the relative concentration has. In any case, α should be small enough that the deviation term is larger than approximately the number of data points to avoid over-fitting.

To minimize the error described above, the relative concentrations Q_{iA} are varied one by one. First, one Q_{iA} is changed temporarily by 0.1 both up and down, under the constraint that it does not exceed 1 or go below 0, and if any change reduces the error, then Q_{iA} is changed permanently to that value that reduced the error (the most), otherwise the change is retracted. After changing Q_{iA} by 0.1 (or to 0 or 1), the rest of the Q_{iA} in the same layer are changed to re-satisfy normalization to 1, but in such a way that the initially changed Q_{iA} is not changed by the re-normalization. After this set of operations, the single Q_{iA} has been changed either exactly once or not at all, and the same operations are then performed on the rest of the Q_{iA} in the same layer, one at a time. Once this is done, all of the above is repeated for the next layers one layer at a time. After all of the above, every Q_{iA} has had the opportunity to reduce the error by changing once by 0.1 (or changing by less than 0.1, to 1 or 0). All of the above is repeated until the error stops decreasing, at which point the change is reduced from 0.1 to 0.01 and the entire process is repeated until again no reduction in error is made. At this point, we have found all Q_{iA} to the second decimal, satisfying minimizing the error. No better precision is attempted because the resulting changes in the curve fit of the relative intensities would be much less than the variance/fluctuation in the experimental data.

4. Results and Discussion

This chapter presents and discusses measurement results from four different cases; a thin (6 Å) and thick (24 Å) deposition of Dy on MoS_2, and for each of them both a temperature series and an angle series. We will also discuss the validity and consequences of important assumptions and approximations and possible improvements.

4.1 Thick film temperature series

Results from the thick film temperature series are shown in Fig. 4.1. Whereas Ref. 1 shows clear evidence of S moving toward the Dy surface when annealing from room temperature to between 300 °C and 500 °C, we here expand the temperature range and resolution as shown in the inset of Fig. 4.1. This temperature dependent area plot clearly shows successful oppression of S migration for temperatures up to approximately room temperature. Though not shown in this data set, the work of Ref. 1 shows the area increase (per temperature) of S between 300 °C and 500 °C is less than the slope of our temperature series at 300 °C. This indicates there is some temperature of S saturation in the Dy layer, in the first few hundred °C above our temperature series.

The Mo3d measurements show no measurable signal during this temperature series, which strongly indicates S is migrating to the surface *without* Mo, in agree-
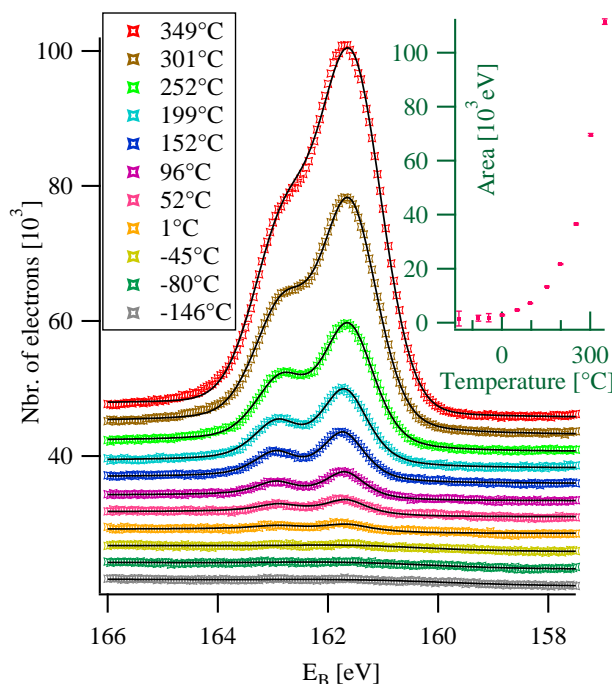

Figure 4.1: **Thick film temperature series S2p raw data, peak fits and areas. All raw data graphs may have an arbitrary vertical offset. The temperature indicated is the highest temperature annealed to after Dy deposition. Fitted results in black. The inset shows the area under the peaks at each temperature, with error bars as ± one standard deviation.**

ment with Ref. 1 and its assertion that S is reacting with Dy to create DyS_x.

Above 300 °C the valley between the two S2p peaks starts to fill up, showing a broadening of the peaks. This points to a new chemical environment appearing for S, and could indicate the DyS_x is being destroyed or stoichiometrically changed by the increased temperature.

The peak fits of the S2p core levels, also shown in Fig. 4.1, are in excellent agreement with the experimental data, indicating the peak fitting model described in section 3.5 is adequate in accuracy for quantifying this core level. The small error bars in the area plot further supports this. The larger error (bars) at lower temperatures is probably caused by the Voigt peak height being so close to zero that the width of the Voigt peak is free to (incorrectly) become very large without the peak fitting becoming significantly worse. When this happens, increasing the area of the Voigt peak(s) can be done without significantly increasing the height. An intuitive justification is that an infinitely wide peak with an infinite area is indistinguishable from a zero-area peak and a constant offset. It is therefore quite probable that peak fitting very flat spectra will be quite susceptible to sliding into local minima where the zero-intensity peak has a larger degree of freedom in varying its intensity.

There initially is no signal from S, so most likely all S that contributes to the data in Fig. 4.1 has moved out of the MoS_2 and into the Dy layer. Since the peak fitting model, which contains only one chemical environment for S, seems reasonably well fitting, it is reasonable to guess that S only appears as (one kind of) DyS_x up to the broadening at 300 °C.

4.2 Thin film temperature series

The raw data, peak fits and areas of the S2p core level for the thin film temperature series are shown in Fig. 4.2. We clearly see signal directly after deposition (i.e. at -134 °C) so we know we see some S from the chemical environment of clean MoS_2 (because
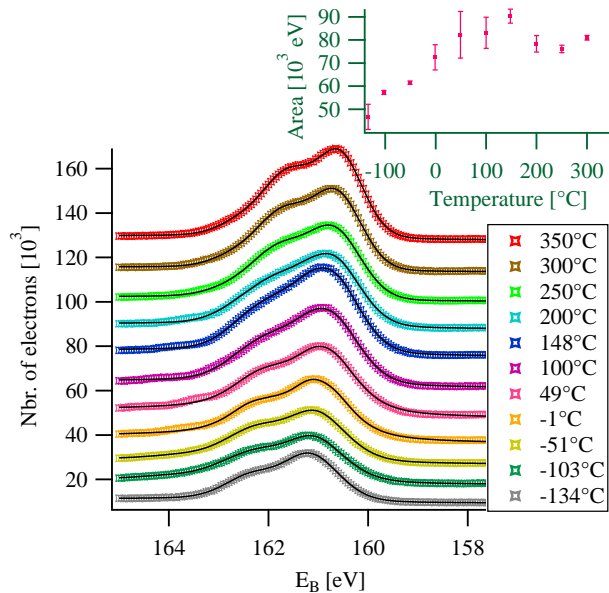

Figure 4.2: Thin film temperature series S2p raw data, peak fits and areas. All raw data graphs may have an arbitrary vertical offset. The temperature indicated is the highest temperature annealed to after Dy deposition. Fitted results in black. The inset shows the area under the peaks at each temperature, with error bars as ± one standard deviation.

the Dy layer is not thick enough to attenuate it completely). Hypothetically this could

be due to S moving to the surface directly upon deposition, but we find this unlikely since S2p is completely attenuated for the same situation in the thick film temperature series (Fig. 4.2). Since we also know from Ref. 1 there is a chemical reaction between S and Dy immediately upon deposition at room temperature, we expect a second chemical environment for S at least at room temperature and above. Furthermore, since this thin film temperature series shows (at all temperatures) less of a valley between the two peaks than that exhibited for the thick film temperature series, we can be fairly certain there are more chemical environments in this series (even for low temperatures) than there is for the thick film temperature series. This means there is probably a chemical reaction immediately upon dosing also for this cold deposition, and that there are two chemical environments of S in this entire temperature series.

The assertion that there are two chemical environments is supported by the excellent agreement between the experimental results and the peak fitting model (described in section 3.5) which contains two chemical S2p environments.

The shifting shape and peak centers Fig. 4.2 shows for increasing temperature can be explained by annealing causing one of the chemical environments to increase in intensity while the other one decreases. The systematic shift to lower binding energies for higher temperatures shows the lower-binding-energy chemical environment is increasing with temperature, relative to the higher-binding-energy chemical environment. From the area increase up to room temperature (and then stagnation-like behavior), this indicates the lower-binding-energy chemical environment up to room temperature is increasing more than the higher-binding-energy chemical environment is decreasing, while above room temperature the increase and decrease are more or less similar in magnitude. The natural interpretation of this is that up to room temperature, DyS_x is being created by pulling S slightly toward surface, from the MoS_2 topmost layers and into the Dy layer, while above room temperature the thin film is more or less saturated with S and the only change is DyS_x moving slightly towards the surface within the thin film. The statement that the thin Dy film gets saturated with S is supported by the fact that the thick film temperature series, which has more Dy available, reaches S2p areas (at high temperatures) much larger than the ones reached here, even though the thick film does not show signs of saturation at the temperatures investigated here.

From the above observations we could note that the thick film has the capacity to absorb much S, compared to this thin film which is almost saturated immediately upon deposition. Given an even thinner layer (say half the thickness since the area is roughly doubling during annealing) might be roughly the threshold thickness of reaching saturation with cold deposition even without annealing, while the film thickness used here is roughly ideal for saturation immediately upon dosing at room temperature.

The peak fitting model's excellent agreement with the experimental data indicate two chemicals environments is reasonable. But even if it is not an accurate description of the

physics of the system, the fact that the fit is so good indicates the area will be accurate even if the number of chemical environments is not. We should note that when modeling four peaks so close together as is the case here, there is a lot of room for areas and widths of individual peaks to vary while still keeping the peak fit quite good. We believe this is the reason for the large area uncertainties at some temperatures. We suggest it is very small at 350 °C because this is the spectrum whose fit is used as initial guess for the rest of the temperatures. All other temperatures then get an error bar which is very dependent on the initial guess compared to the raw data, giving a quite varying error bar for different temperatures.

Let us now move on to the Mo3d results for the thin film temperature series, shown in Fig. 4.3. Unlike the S2p core level where the presence of a second chemical environment directly after deposition (at -134 °C) is not beyond doubt, we here see an extra pair of peaks which removes all doubt of a chemical reaction taking place immediately upon deposition, even for cold deposition as performed here. This is also supported by the excellent fit at all temperatures, using a model with two chemical environments of Mo3d.

The lower-binding-energy chemical environment increases relative to the higher-binding-energy chemical environment starting around room temperature, so there definitely is some temperature dependence of the Mo state. The area-temperature plot shows increasing total area, so the natural question to ask is if Mo is also moving to the surface. However, we know from the thick film's lack of measurable Mo after annealing, that this movement cannot be significant. Moreover, the S2p results in this
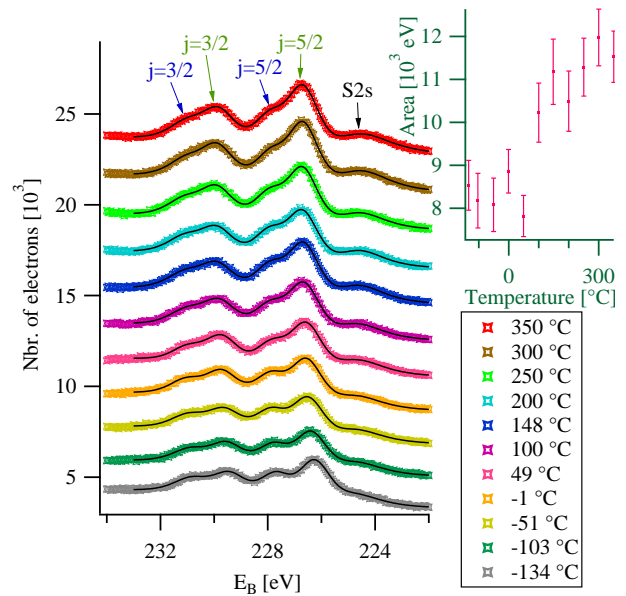

Figure 4.3: Thin film temperature series Mo3d raw data, peak fits and areas. All raw data graphs may have an arbitrary vertical offset. The temperature indicated is the highest temperature annealed to after Dy deposition. Annotations show one chemical environment for Mo3d (green), another environment (blue) and S2s (black). Fitted results in black. The inset shows the area under the peaks at each temperature, with error bars as ± one standard deviation.

thin film shows increase in S signal at some temperatures, and both Mo and S cannot both move to the surface at same time unless something else entirely simultaneously moves down into the sample. This means either Dy is intercalating, or Mo is not really moving but is increasing in intensity by some other mechanism. We discard the possibility of Mo moving toward the surface while S does not, on the grounds that the thick film sample

56

shows indisputable evidence of S migration toward the surface. Hence we are left with the possibility of Dy intercalating, or Mo increasing in intensity without migrating, for instance due to some reorganization which increases the amount of constructive interference of electrons in the measurable layers.

The excellent agreement between peak fits and experimental data suggests the model of two Mo3d chemical environments and one S2s is reasonable. We would expect there two be two chemical environments of S2s, and there probably is, but the model works fine with just one. This is not unreasonable, because the split between the two S2s chemical environments can be small enough that the fit can compensate sufficiently by adjusting the Lorentzian and Gaussian widths. As mentioned for the S2p fit in this thin film temperature series, there is some compensation capacity that allows one peak area to decrease while another increases. We postulate that this freedom is more restricted when the peak separation is larger, which is why the error bars of the Mo3d areas are less spectrum (temperature) dependent than the ones for S2p, and why the Mo3d error bars are smaller (relative to the respective area) than the largest of the S2p error bars.

4.3 Thick film angle series

Dy5p raw data

Results from the thick film angle series Dy5p core level are shown in Fig. 4.4. There is no obvious angle dependence of the Dy5p core level neither before nor after annealing. However, there are some signs of a second chemical environment appearing after annealing, in the slight broadening of the lower-binding-energy shoulder of the $j = \frac{3}{2}$ peak, and the increase of the $j = \frac{3}{2}$ peak relative to the $j = \frac{1}{2}$ peak while the $j = \frac{1}{2}$ peak remains fairly constant in (peak-to-background) height.

This is supported by the good fit of the peak fit to the experimental data, both be-
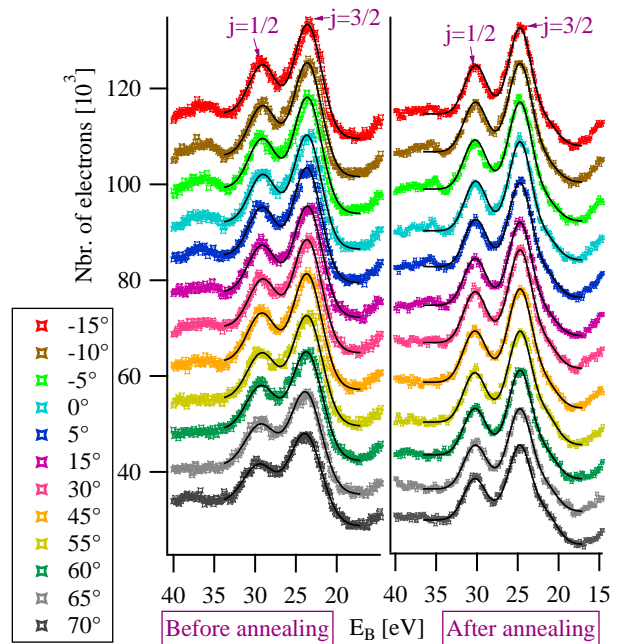

Figure 4.4: Thick film angle series Dy5p raw data and peak fits before and after annealing. All graphs may have an arbitrary vertical offset. The angle indicated is the photoelectron emission angle θ. Annotations indicate Dy5p$_{\frac{1}{2}}$ and Dy5p$_{\frac{3}{2}}$. Peak fits are shown in black.

fore and after annealing, where the peak fitting model includes only one chemical environment before annealing, and two chemical environments after. The one-chemical-environment model has been tried and found inadequate after annealing. The second

chemical environment appears after annealing of a thick film, where we know from the thick film temperature series that S migrates to the surface, so a reasonable explanation is that the new chemical environment is from DyS_x while the original one is clean Dy.

The fact that the second chemical environment Dy peak pair is not disappearing or increasing for large or small angles indicates the S is distributed throughout the entire Dy layer and not just localized at the Dy-MoS_2 interface.

The peak fits agree quite well with the experimental data, so we argue, as for the S2p core level of the thin film, that the areas we extract from the peak fitting models are accurate even if the number of chemical environments should be incorrect.

S2p raw data

Moving on to the S2p core level of the thick film angle series, the results are shown in Fig. 4.5. Since these measurements are from an X-ray gun and not a synchrotron, the photon energy used provides relatively low photoionization cross-sections for both S2p and Dy4d, whose cross-sections then are comparable in magnitude. Therefore, the S2p peaks lie on top of non-negligible Dy signal.

When comparing, before annealing, the Dy peak height with the S2p peak height, it is obvious by eye that there is some angle dependence which is different for the two elements, showing that if S is distributed (only as DyS_x) throughout the entire Dy layer as

Figure 4.5: **Thick film angle series S2p raw data and peak fits before and after annealing. All graphs may have an arbitrary vertical offset. The angle indicated is the photoelectron emission angle θ. Annotations indicate $S2p_{\frac{1}{2}}$ and $S2p_{\frac{3}{2}}$. Peak fits are shown in black.**

suggested from the Dy5p core level, then it is not a uniform distribution in the Dy layer. After deposition, this dependence seems washed out, except for a slight increase in S2p relative to the Dy peaks around normal emission ($\theta = 0$). This is indicative of the S distribution becoming more or less uniform in the Dy layer, and that some S is measured in MoS_2 beneath the Dy layer for emission angles close to normal emission. We know the thick film temperature series shows no Mo signal through roughly the same Dy layer thickness, but in this case we have 3 to 4 times as large electron attenuation length so we expect less surface sensitivity in this series. The increased S signal close to normal
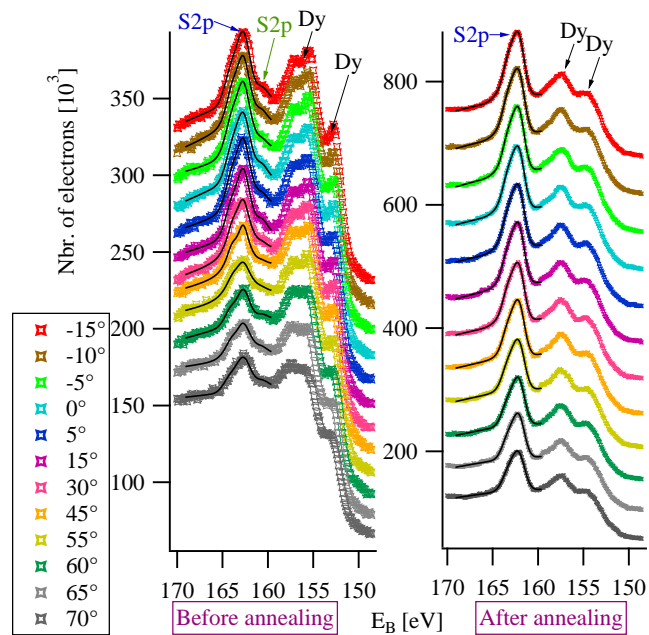
emission could also be explained by a S-free layer at the topmost part of the Dy layer, which is more dominant in off-normal angle signal.

In addition to the weaker angle dependence after annealing, we also see that the S2p intensity is systematically increased relative to the Dy peaks for all angles, which can be explained by S moving toward the Dy surface as we already know from the thick film temperature series happens in the annealing temperature range we anneal over in this thick film angle series (from room temperature to 300 ℃).

We can also see from Fig. 4.5 that the peak-to-background of the Dy peaks are roughly the same before and after annealing, while the S peaks increase. This shows that any significant intercalation of Dy is unlikely between room temperature and 300 ℃.

Another thing we can see happens during annealing is that the Dy peaks become broadened. This can be explained by annealing causing S migration toward the surface as we know happens from the thick film temperature series, but only as long as not all Dy has reacted with S for if that was the case then the Dy peaks should resolve themselves to narrower widths again, with shifted centers. This broadening is therefore in accordance with the thick film temperature series discussion in that the Dy is not saturated by S at 300 ℃.

The peak fits are in good agreement the experimental data, and as discussed for previous core levels, the number of chemical environments are not very relevant for the accuracy of the areas as long as the fit is good. However, in the case of the thick film angle series S2p, the underlying Dy peak(s) pose quite a pickle. The tail of inelastically scattered electrons from the annotated Dy peaks in Fig. 4.5, as well as some peaks of electrons that have suffered energy losses by other mechanisms like plasmon resonance, make the total background beneath the S2p peaks (i.e. everything except the S2p Voigt peaks and backgrounds) quite far from constant. Modeling such a complex background is beyond the scope of this work, so the problem has been "solved" by including a linear term in the background of the S2p peak fit. This linear term has slope that increases for larger binding energies, which is nonphysical as it corresponds to electron energy gain, not loss, as it traverses the sample. We allow it nonetheless because it is not supposed to model only a background but also peaks from Dy.

The reason this might affect our areas is that the linear background means neglecting any Dy peaks which are significantly non-linear in the region of the S2p peak fit. Now, since the S2p peak fitting model in this case includes two close-together chemical environments (one was tried and found insufficient), the fit can compensate quite well for the (possibly) neglected Dy without significantly worsening the fit. However, as the S2p and Dy peaks do not vary very much relative to each other in this angle series, the relative error in estimating the S2p area is possibly a quite constant factor. Even if it is non-constant, the consequence of the approximation is a systematic *under*-estimation of the S2p area.

Layer fit

The layer fit results from the thick film angle series are shown in Fig. 4.6. The relative intensities show a decrease of Dy5p and increase of S2p upon annealing, in accordance with the above discussion of the raw data.

The theoretical relative intensities are created with the layer fitting algorithm described in section 3.8. Like the case for previously discussed core levels modeled with two chemical environments, the error bars are probably larger than the actual uncertainty in the area. The fits agree well with the experimental data, and similarly well for all fits, although the experimental results themselves are quite fluctuating. This means that parameters of the peak fit that rely on the fine

Figure 4.6: Layer fit of thick film angle series. Experimental relative intensities of S2p, Mo3d and Dy4d are shown with error bars as ± one standard deviation. Theoretical (fitted) relative intensities shown correspond to the relative concentrations shown to the right.

structure of the theoretical relative intensities probably will be quite unreliable. However, the general qualitative behavior of which core level is more intense, and whether the intensity increases or decreases with angle is still reliable.

The relative concentrations in Fig. 4.6 are the concentrations that create the theoretical (layer fitted) relative intensities in the same figure. The relative concentration of S before annealing is found to be close zero in the topmost 12 Å of the Dy layer, but close to 60% in the next 18 Å. This is in support the claim from the discussion of the (thick film angle series) raw data that there is S in the entire Dy layer, except possibly the topmost section which we now see evidence is the case.

We also see, with annealing, an increase in the S concentration relative to the Dy concentration, in accordance with the same observation in the raw data and relative intensities. The increase is found to be in the ballpark of 0.2 in the topmost 12 Å, and much less in deeper layers. The S:Dy ratio is close to, but not quite, 2 close to the Dy-MoS_2 interface, and less than 0.5 close to the Dy surface. The roughly equal ratio in the three deepest 0.6 nm sections of the layer model could indicate there is a saturation of S when there is DyS_2. We should keep in mind the possible under-estimation of the S2p
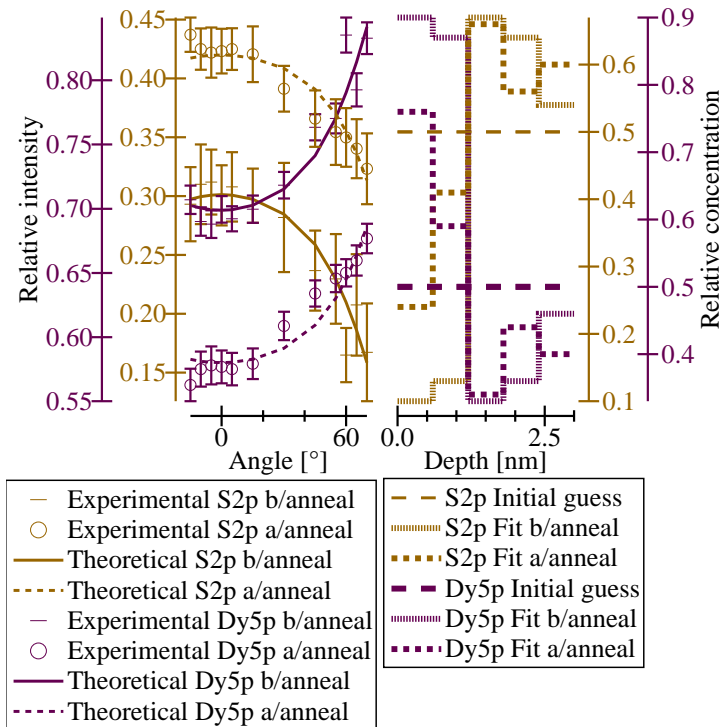
area which would mean the ratio is really somewhat higher.

We should note that the layer fit has been made with a large α, i.e. almost completely neglecting the entropy term of the minimized error, because the resulting relative concentrations are not obviously unphysical. This corresponds to a pure χ^2 optimization, with a value of χ^2 of less than 1 per data point. This means we are over-fitting in the sense described in section 3.8, but in this case decreasing α does not solve the problem without making the fit obviously incorrect. Hence the deviation-entropy compromise turns out not to help in this case, so instead we just bear in mind the resulting uncertainty in the peak fitting parameters caused by over-fitting. Most important here is probably the uncertainty of relative concentrations in deeper layers of the sample model.

4.4   Thin film angle series

Raw data and peak fits

The raw data and peak fits from the thin film angle series are shown in Fig. 4.7. Due to photon flux and the share of detected electrons changing when the sample is rotated and moved between each angle, there is also a scaling of the peak heights in our angle series which is potentially different for each angle. This is probably the reason for most of the attenuation at extreme angles. Therefore, we look only at the shape of the peaks when we discuss them now, and we look only at the intensi-
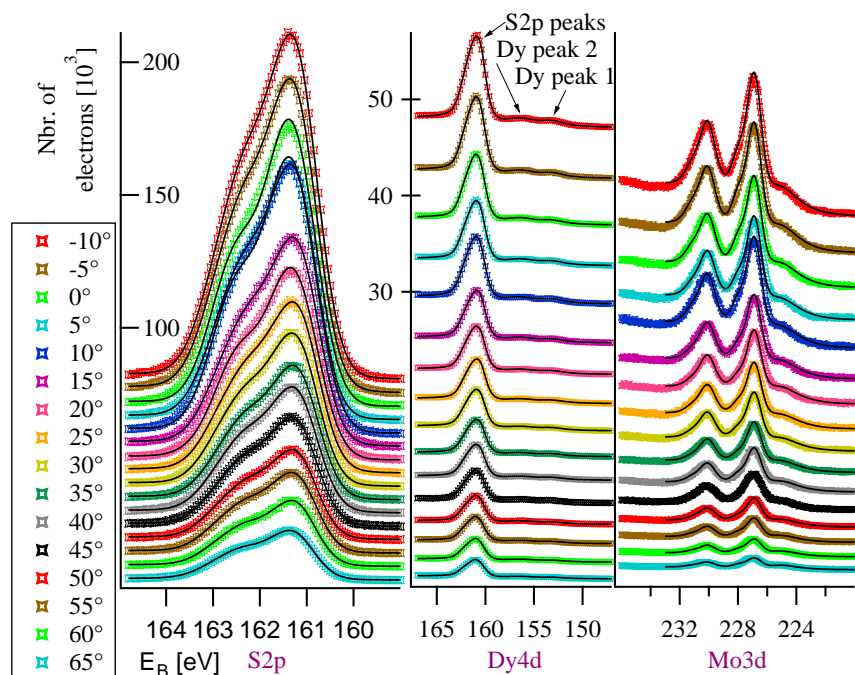

Figure 4.7: **Thin film angle series S2p, Dy4d and Mo3d raw data and peak fits. All graphs may have an arbitrary vertical offset. The angle indicated is the photoelectron emission angle θ. Peak fits are shown in black. Annotations indicate S and Dy peaks included in the peak fitting model.**

ties (or peak heights) *relative* to other core levels at the *same angle* in the layer fit results presented below.

The S2p core level shows no apparent angle dependence of its shape, suggesting the binding energy change of S2p when S is reacting with Dy, and the amount of S reacting

with Dy, both are too small to show their effects. We know from the thin film temperature series discussed above, and from Ref. 1, that there has been a reaction at this stage, but the two chemical environments might have peaks that vary very little in binding energy. The presence of two chemical environments is also supported by the excellent agreement between the experimental data and the fit using a model with two chemical environments.

The Dy4d core level is much less intense than the S2p core level in the same region, so in comparison, the Dy4d only looks like two blips on the background. A figure more suitably scaled to see the Dy4d is the peak fitting model of Fig. 3.9. The peak fits agree very well with experiments, but the S2p core level is modeled with two chemical environments so it might be able to compensate for potentially missing Dy peaks directly beneath the S2p peaks.

The Mo3d core level shows good agreement with the peak fits which include two chemical environments, as we already know are present.

Layer fit

Fig. 4.8 shows the layer fit of the thin film angle series. An α corresponding to pure χ^2 optimization has been used for the same reasons as for the thick film angle series layer fit.

The relative intensities show increasing S and Dy with more extreme angles, indicating these elements are present in higher concentrations closer to the surface than they are deeper in the sample. Similarly, the decreasing Mo relative intensity with extreme angles show the topmost sample layers are Mo depleted relative to the deeper ones.

The larger error bars on the S relative intensity compared to the

Figure 4.8: **Layer fit of thin film angle series. Experimental relative intensities of S2p, Mo3d and Dy4d are shown with error bars as ± one standard deviation. Theoretical (fitted) relative intensities shown correspond to the relative concentrations shown to the right.**

other ones are presumably caused by the model of S2p containing two chemical environments closer to each other than those of Mo3d, hence the compensation capabilities of the model to wrong intensities are larger. The Dy peaks then have the smallest error bars because the Dy4d model only includes two peaks which are fairly well separated.

Since the photon energy used on each core level is different, and we do not precisely
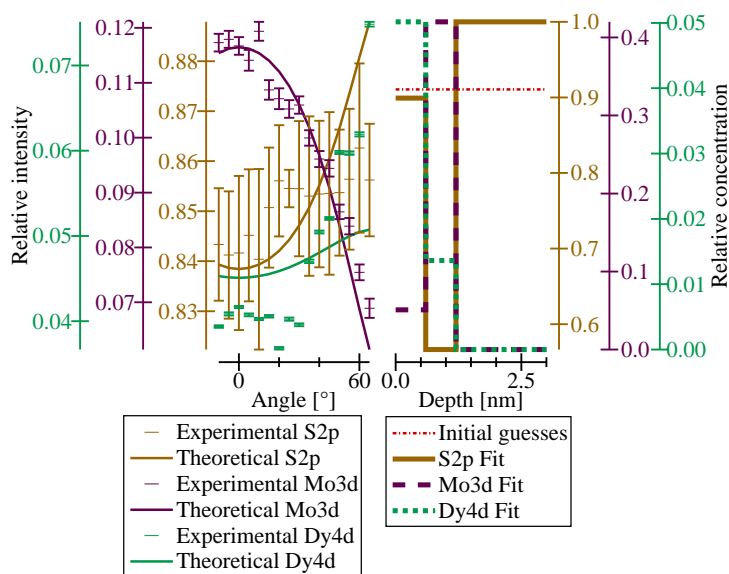
know the photon flux at each photon energy, there is an unknown scaling of all (three) relative intensities (at the same angle) relative to each other. There is also approximately such a scaling effect caused by the systematic over- or under-estimation of each area due to the peak fitting models being wrong. The area errors in the Mo and S2p relative intensities are presumably small, but the Dy4d area is based only on the two most intense and resolved peaks of a wide and unresolved set of complicated peaks and backgrounds both directly beneath the S2p core level and further towards higher binding energies. This is not believed to significantly affect the S2p model accuracy because S2p is so much more intense than the Dy4d, but the effect of neglecting these peaks in the Dy4d peak fitting model probably has significant effects on the estimated Dy quantity. Most likely, the Dy relative intensity is strongly under-estimated. The difficulty in correcting this by including the neglected peaks in the peak fitting model is the lacking theoretical understanding of the processes involved in emission of d-core level electrons, and the particulars of Dy in particular. More experimental knowledge of Dy is required as input in a peak fitting model if the neglected peaks are to be included, because the overlap of different peaks makes the curve fitting algorithm highly susceptible to sliding into local minima of significantly incorrect peak fitting parameters.

Due to the unknown scaling factors mentioned above, we cannot rely on the experimental results' slope with angle to tell us to which elements are more angle dependent and which are less. Furthermore, the theoretical (layer fitted) results show a significant deviation from the experimental Dy results. This means the model used is unable to explain the experimental results, and hence the model cannot be adequately correct. A better agreement between the theoretical and experimental results can possibly be achieved by finding the correct relative scaling factors and correct for them when handling the data. We note that such scaling factors are not an issue for the thick film angle series because the photon flux in that case is independent of core level, and the peak fitting model used to quantify Dy is a p-core level which probably estimates the Dy quantity much better than the d-core level used here.

The intuitive situation of the sample is that the Dy is contained in the topmost part of the measurable sample layers, so it follows that the Dy relative intensity is expected to increase rapidly for extreme angles, as is indeed the case of the experimental data. Adding to this, a correction to the under-estimation of the Dy intensity relative to the other intensities would imply an even more increasing Dy relative intensity for extreme angles. Due to normalization this would have to be an increase in slope at the expense of the slope of either the Mo or S relative intensities. As the S relative intensity fit seems to be overshooting at extreme angles, we suggest some of the deviation in both the Dy and S fit can be explained by systematic under-estimation of the Dy quantity relative to the S quantity, i.e. a shortcoming of the peak fitting model (or alternatively, the pre-curve-fitting data processing).

The relative concentrations found from the layer fit shows Dy exclusively contained in the topmost 12 Å of the sample, where it is less than 10% of the S quantity. In light of the suggested DyS_2 S saturation ratio suggested for the thick film angle series, this also supports the above assumption of the Dy quantity being under-estimated. The S relative concentration is found to be 1 in the deeper layers which might both be a sign of systematic scaling factors as discussed above, and an effect of the larger surface sensitivity in this angle series compared to that of the thick film angle series, making the deepest 12 Å of the layer model not very important for the outcome of the layer fit. In any case, also with unknown scaling factors, the larger S concentrations at shallower depths in the topmost 18 Å of the sample indicate some S migration into the Dy layer, which is reasonable in light of our previous knowledge that the S-Dy reaction happens immediately upon deposition. The Mo relative concentration supports this by being lower in the topmost layers (i.e. in the Dy layer) and then increasing in deeper layers. It goes to zero at the two deepest layers, and it does not quite reach zero at the topmost layers, both effects which we again attribute to incorrect scaling of the elements, because there is no compelling reason to believe these parts of the results in light of previously discussed results. As for the thick film angle series we do have reason to believe the qualitative observations of increasing or decreasing relative concentrations in the top 18 Å of the sample because it corresponds to the experimental relative intensities and because this kind of behavior is insensitive to distortion by incorrect scaling factors.

4.5   Peak fitting model: Areas of core level peaks

Sources of error and signs of validity

Peak fitting is in this work only performed to find the total intensities of different core levels of certain elements, which is then used to calculate the amount of measured atoms of the element in question. The error bars introduced into these intensities (or areas) are solely based on the precision of the intensities of each peak in the peak fitting model which contribute to the intensity. As discussed previously, this uncertainty might be quite larger than the real uncertainty in the total intensity of the core levels because we are adding uncertainties of intensities in individual peaks when the compensation capability of a peak fit to an increase in one peaks intensity might be almost exclusively relying on another peaks intensity decreasing. Hence much of the uncertainty would be canceled out in a more rigorous handling of the error bars. In addition to this, there is also an important source of uncertainty in the peak fitting models which is not included in error bars in this work. This is a systematic error caused by a peak fitting model of a given core level not correctly representing the physics of the situation, for instance the wrong number of peaks or the wrong constraints on a peak fit. The results presented above make a self-consistent set of observations that support the peak fitting models used except for the

thin film angle series S2p and Dy4d relative intensities which exhibit significant deviation from experimental results. As discussed for the thin film angle series, we believe there is a systematic unknown scaling factor of each core level's area due to varying photon flux with photon energy, as well as a factor caused by under-estimating the Dy quantity with too small a part of the true Dy4d core level and all its photoelectron emission processes.

Error consequences

An important consequence of this model short-coming is that the calculated relative intensities of each element is scaled relative to the other elements' relative intensities. In extreme cases this might even change the sign of the slope of a relative intensity as a function of angle, meaning the interpretation will be an increasing relative concentration with increasing/decreasing angle, instead of a decreasing relative concentration.

Improvement suggestions

To correct for this error, we can suggest several approaches. The error caused by under-estimating the Dy amount can be corrected by a more rigorous treatment of the Dy4d core level. Experimental information could be used to guess which additional peaks to expect from Dy4d, and then to include in a peak fitting model, but finding sufficient experimental knowledge to accurately describe the entirety of the Dy4d core level photoelectron emission peaks is, to the best of the author's knowledge, not possible without new research on the topic.

An alternative approach is to use a photon source of sufficient energy to excite for instance the Dy5p core level as done in this work for the thick film temperature series. This core level is much easier to quantify through peak fitting than the Dy4d core level, especially with the presence of S.

The effects caused by varying photon flux can be solved by using a photon source where the relative photon flux difference between core levels is known, for instance by knowing the flux of the used synchrotron at different photon energies, or by using an X-ray gun or synchrotron at only one photon energy, where the flux then becomes independent of the core level.

4.6   Gaussian I_{pm} and Lorentzian p_Δ

We assumed in subsection 2.2.7 that the photon intensity I_{pm} is Gaussian, that the core hole lifetime-caused kinetic energy change probability p_Δ is Lorentzian, and we demonstrated that these result in a Voigtian XPS spectrum. These are truths with modifications which warrant justification. A photon source consisting of a synchrotron and a monochromator gives an approximately Gaussian I_{pm}, but for the X-ray gun with an attenuator,

a more Lorentizian-like I_{pm} is produced. Furthermore, the convolution of I_{pm} and p_Δ should strictly be convoluted with a third function, namely a Gaussian-like broadening effect caused by the finite resolution of the electron energy analyzer.

The consequences of neglecting this in deriving the peak fitting model are presumably negligible because all three convoluted functions are approximately Voigtian and a convolution of Voigtian peaks is in turn approximately Voigtian.

4.7 Peak fitting model: Background

Sources of error and signs of validity

The background of every peak in any peak fitting model used in this report contains the error function to account for inelastically scattered electrons. When calculating intensities (or areas) of each peak or set of peaks, these electrons are excluded. Ideally we would want to find the intensity of all photoionized electrons because that is what represents the true amount of an element present. This exclusion does of course affect the interpretation of how much of a given element is measured. But this effect should be dependent only on the electron attenuation length (and escape length), through the probability of inelastic scattering occurring. Hence this point becomes moot when considering relative changes as in the temperature series. For the angle series, this becomes a scaling factor like that discussed above, which is different for different kinetic energies measured, and hence in general different for different core levels.

Error consequences

The consequences of such an error is the same as for the similar effect caused by the aforementioned peak fitting model short-coming in including the wrong number of peaks.

Improvement suggestions

To correct for this error it could be worth a try to restrict the amplitude of every error function to represent the probability of inelastic scattering. A challenge with this is the dependence of this probability on the escape length of the photoelectron, but perhaps an iterative procedure could be used in which one first estimates the layered structure of the sample and then uses this to restrict the background amplitude which again is fed into finding the layered structure until some equilibrium is found.

4.8   Peak fitting model: Variational procedure

Sources of error and signs of validity

As supported by obvious fluctuations (noise) in above area and intensity plots, there is more uncertainty to account for than those introduced by the peak fitting model. We suggest most of this is caused by the variational procedure itself. When a peak fit is performed using a given initial guess the exact resulting peak fit parameters depend on the initial guess.

Error consequences

The result of this is that each area calculated gets some random-like uncertainty.

Improvement suggestions

To remedy this error, a suggestion could be to adapt initial guesses to each situation by not using the same set of initial guess parameters for every spectrum, but rather let the solution of one spectrum peak fit be the initial guess of the next spectrum peak fit, where the order of the spectra is such that the change in peak fitting parameters are minimal between each spectrum. Another suggestion is to peak fit once as is done in this report, then analyze not just the areas' variation through the spectra, but also the widths and centers. Then the widths and centers can be investigated to see if their behavior can be curve fitted in a reasonable way, and then these parameters could be constrained to this curve fit and used as input in a new peak fit, possibly repeated iteratively.

5. Conclusion

5.1 Summary

Directly after deposition of a thin film (6 Å) Dy on the surface of MoS_2, we see evidence of a chemical reaction taking place involving Dy and S, even when the deposition is done on a cold sample (approximately -140 ℃). The angle series and layer model of a thin film shows discrepancies between theory and experiment, but still strongly suggests S mixing with the Dy layer and the formation of DyS_x.

For the thick film (24 Å), both the temperature and angle series show the DyS_x immediately upon (cold) deposition is constrained to the Dy-MoS_2 interface.

When annealing the thin film sample, the temperature series shows creation of more DyS_x with higher annealing temperature, and a saturation of S in the Dy layer somewhere below 350 ℃.

For the thick film sample, the temperature series shows DyS_x migration toward the Dy surface with increasing temperature, starting around room temperature and not showing signs of S saturation at 350 ℃. When reaching 300 ℃, there are signs of chemical reactions taking place, possibly changing the S:Dy ratio in DyS_x. The angle series indicates that annealing causes increased DyS_x creation starting at the Dy-MoS_2 interface and spreading out towards the Dy surface while S saturates at a S:Dy ratio close to 2, i.e. as DyS_2.

All in all this work successfully demonstrates the intended results by shedding light on the temperature dependence and layered structure of a Dy-MoS_2 system where DyS_x is created by depositing Dy on MoS_2 and annealing the sample.

5.2 Suggested new work

This work demonstrates some parameters of creation of DyS_x by depositing Dy on MoS_2, so the natural follow-up investigation towards possible use in spintronics, is to find out which DyS_x compound we are making, and which crystalline properties it has, or alternatively how to make it crystalline. After such an investigation, it could be possible to start considering specific spintronic applications.

We suggest starting the way forward by saturating a thick film of Dy with S in a Dy-

MoS$_2$ system, by annealing to 250 °C for a few hours (and avoid the possible stoichiometric change at 300 °C). This DyS$_x$ layer is then possibly DyS$_2$, and its crystalline properties could be investigated by Low Energy Electron Diffraction (LEED) or other diffraction techniques. If it is not sufficiently periodic, further annealing might be attempted to increase the crystallinity. In light of the work done here, there is a risk this might destroy the DyS$_2$, but there is also a possibility that another Dy:S ratio, with another crystal structure, might emerge. Once sufficient crystallinity is achieved, a spin-resolved Angle Resolved Photoelectron Spectroscopy (ARPES) measurement can be used to map the valence- and conduction bands. This would give direct information about the possible role of DyS$_2$ (or other Dy-S compounds) as part of a spintronic device.

5.3 Future potential

With the strong magnetic properties of Dy, it is quite possible that DyS$_x$ would contribute new possibilities to the current pool of possible materials for spintronic devices. This contribution could subsequently be broadened by expanding the DyS$_x$ properties through doping, substitution or spatial restrictions. Even further expansion could be attempted through copying the DyS$_x$ creation procedure to other TMDCs than MoS$_2$; a TMDC containing S, but with a different crystal structure, to get a different crystal structure of DyS$_x$ - or with another chalcogen than S to get another Dy compound entirely, maybe even a superconducting one. Finding possible uses of different Dy compounds once made, relies on uncovering their electric and magnetic properties.

The above hypothesized Dy compounds are just a part of the conceivable uses of the work done here. Some of the suggestions for further work might not work as expected, but there are also the uses that we have not conceived, which can only thoroughly be investigated once satisfactory crystallinity has been achieved and proven.

Bibliography

[1] Magnus Sebastian Christensen. Studying the effects of dosing MoS2 with Dy for application in spintronics. Technical report, NTNU, Trondheim, 2016.

[2] Mark Johnson and R. H. Silsbee. Interfacial charge-spin coupling: Injection and detection of spin magnetization in metals. *Physical Review Letters*, 55(17):1790–1793, 1985.

[3] D. M. Wolf, S. A. and Awschalom, D. D. and Buhrman, R. A. and Daughton, J. M. and von Molnár, S. and Roukes, M. L. and Chtchelkanova, A. Y. and Treger. Spintronics: a spin-based electronics vision for the future. *Science (New York, N.Y.)*, 294(5546):1488–95, 2001.

[4] Igor Žutić, Jaroslav Fabian, and S. Das Sarma. Spintronics: Fundamentals and applications. *Reviews of Modern Physics*, 76(2):323–410, 2004.

[5] Xiuwen Zhang, Qihang Liu, Jun-Wei Luo, Arthur J. Freeman, and Alex Zunger. Hidden spin polarization in inversion-symmetric bulk crystals. *Nature Physics*, 10(5):387–393, 2014.

[6] L. Bawden, S. P. Cooil, F. Mazzola, J. M. Riley, L. J. Collins-McIntyre, V. Sunko, K. Hunvik, M. Leandersson, C. M. Polley, T. Balasubramanian, T. K. Kim, M. Hoesch, J. W. Wells, G. Balakrishnan, M. S. Bahramy, and P. D. C. King. Spin-valley locking in the normal state of a transition-metal dichalocogenide superconductor. *Nature Communications*, 7(May):7, 2016.

[7] J M Riley, F Mazzola, M Dendzik, M Michiardi, T Takayama, L Bawden, C Granerod, M Leandersson, T Balasubramanian, M Hoesch, T K Kim, H Takagi, W Meevasana, Ph Hofmann, M S Bahramy, J W Wells, and P D C King. Direct observation of spin-polarized bulk bands in an inversion-symmetric semiconductor. *Nature Physics*, 10(11):835–839, 2014.

[8] D. J. Rahn, S. Hellmann, M. Kalläne, C. Sohrt, T. K. Kim, L. Kipp, and K. Rossnagel. Gaps and kinks in the electronic structure of the superconductor 2H-NbSe 2 from angle-resolved photoemission at 1 K. *Physical Review B - Condensed Matter and Materials Physics*, 85(22), 2012.

[9] Xiaoxiang Xi, Zefang Wang, Weiwei Zhao, Ju-Hyun Park, Kam Tuen Law, Helmuth Berger, László Forró, Jie Shan, and Kin Fai Mak. Ising pairing in superconducting NbSe2 atomic layers. *Nature Physics*, 12(November):1–6, 2015.

[10] Chin Mo Kim, Chul Sung Kim, Sung Baek Kim, and Eun Joo Hahn. Magnetic properties of Fe-doped 2H-TaS2. *Journal of the Korean Physical Society*, 60(1):79–82, 2012.

[11] L. J. Li, W. J. Lu, X. D. Zhu, L. S. Ling, Z. Qu, and Y. P. Sun. Fe-doping–induced superconductivity in the charge-density-wave system 1T-TaS 2. *Europhys. Lett.*, 97:67005, 2012.

[12] L. J. Li, X. D. Zhu, Y. P. Sun, H. C. Lei, B. S. Wang, S. B. Zhang, X. B. Zhu, Z. R. Yang, and W. H. Song. Superconductivity of Ni-doping 2H-TaS2. *Physica C: Superconductivity and its Applications*, 470(5-6):313–317, 2010.

[13] Kristoffer William Hunvik. *A Spectroscopic Study of in situ Magnetic Doping of Novel Layered Materials for Spintronics*. Master thesis, NTNU, 2016.

[14] Intek Song, Chibeom Park, and Hee Cheul Choi. Synthesis and properties of molybdenum disulphide: from bulk to atomic layers. *RSC Adv.*, 5(10):7495–7514, 2015.

[15] 2D Semiconductors. http://www.2dsemiconductors.com/molybdenum-disulfide-mos2-small-10x10mm/ Accessed 29-11-2016.

[16] P. C. Hemmer. *Kvantemekanikk*. tapir akademisk forlag, Trondheim, 5 edition, 2005.

[17] S. P. Cooil. *Controlling the Epitaxial Growth of Graphene On Diamond Surfaces*. Phd, Aberystwyth University, 2014.

[18] Atomic Calculation of Photoionization Cross-Sections and Asymmetry Parameters https://vuo.elettra.eu/services/elements/WebElements.html Accessed 12.12.2016.

[19] Magnus Sebastian Christensen, Ann Julie Utne Holt, and Mari Kristine Kristensdotter Røed. Dy dosed MoS2 - A LEED and XPS study (Working title). Technical report, NTNU, Trondheim, 2016.

[20] Gwyn P. Williams. X-Ray Data Booklet http://xdb.lbl.gov/Section1/Sec_1-1.html Accessed 10.12.2016.

[21] T. Ida, M. Ando, and H. Toraya. Extended pseudo-Voigt function for approximating the Voigt profile. *Journal of Applied Crystallography*, 33(6):1311–1316, 2000.

[22] NIST. NIST XPS Database https://srdata.nist.gov/xps/Default.aspx Accessed 11.12.2016.