Atom Probe Tomography of TiSiN Thin Films

David Engberg
This thesis concerns the wear resistant coating TiSiN and the development of the analysis technique atom probe tomography (APT) applied to this materials system. The technique delivers compositional information through time-of-flight mass spectrometry, with sub-nanometer precision in 3D for a small volume of the sample. It is thus a powerful technique for imaging the local distribution of elements in micro and nanostructures. To gain the full benefits of the technique for the materials system in question, I have developed a method that combines APT with isotopic substitution, here demonstrated by substitution of $^{15}$N with $^{15}$N. This alters the time-of-flight of ions with of one or more N and will thereby enable the differentiation of the otherwise inseparable isotopes $^{14}$N and $^{28}$Si. Signs of small-scale fluctuations in the data led the development of an algorithm needed to properly visualize these fluctuations. A method to identify the best sampling parameter for visualization of small-scale compositional fluctuations was added to an algorithm originally designed to find the best sampling parameters for measuring and visualizing strong compositional variations. With the identified sampling parameters, the nano-scale compositional fluctuations of Si in the metal/metalloid sub-lattice could be visualized. The existence and size of these fluctuations were corroborated by radial distribution functions, a technique independent of the previously determined sampling parameter. The radial distribution function algorithm was also developed further to ease in the interpretation. The number of curves could thereby be reduced by showing elements, rather than single and molecular ions (of which there were several different kinds). The improvement of the algorithm also allowed interpretation of signs regarding the stoichiometry of Si$_{ny}$. With a combination of analytical transmission electron microscopy and APT we show Si segregation on the nanometer scale in arc-deposited Ti$_{0.92}$Si$_{0.08}^{15}$N and Ti$_{0.81}$Si$_{0.19}^{15}$N thin films. APT composition maps and proximity histograms generated from Ti-rich domains show that the TiN contain at least $\sim$2 at. % Si for Ti$_{0.92}$Si$_{0.08}^{15}$N and $\sim$5 at. % Si for Ti$_{0.81}$Si$_{0.19}^{15}$N, thus confirming the formation of solid solutions. The formation of relatively pure Si$_{ny}$ domains in the Ti$_{0.81}$Si$_{0.19}^{15}$N films is tied to pockets between microstructured, columnar features in the film. Finer Si$_{ny}$ enrichments seen in APT possibly correspond to tissue layers around TiN crystallites, thus effectively hindering growth of TiN crystallites, causing TiN renucleation and thus explaining the featherlike nanostructure within the columns of these films.
This thesis is part of my doctoral studies in the field of materials science. It covers the most important parts of my research in the Thin Film Physics Division at the Department of Physics, Chemistry, and Biology (IFM) at Linköping University from May 2012 to December 2015. My work has been focused on the development and application of experimental and data treatment methods of the analysis technique atom probe tomography applied to TiSiN thin films grown by cathodic arc deposition. The key results of my studies are found in the appended Papers. The work has been conducted within Theme 2 of the VINN Excellence Center FunMat, in collaboration with Sandvik Coromant, Seco Tools and Ionbond Sweden.
LIST OF INCLUDED PUBLICATIONS

Paper 1

*Resolving Mass Spectral Overlap in Atom Probe Tomography by Isotopic Substitutions – Case of TiSi$^{15}$N*

Author’s contribution:
I participated in the deposition of the thin films and in running the atom probe. In addition, all the specimen preparation, characterization and data treatment were conducted by me, except the elastic recoil detection analysis and data treatment. Lastly, I wrote the Paper.

Submitted for publication.

Paper 2

*Solid Solution and Segregation Effects in Arc-Deposited Ti$_{1-x}$Si$_x$N Thin Films Resolved on the nanometer scale by $^{15}$N Isotopic Substitution in Atom Probe Tomography*

Author’s contribution:
As this Paper concerns the same samples and atom probe measurements as in Paper 1, I participated in the deposition of the thin films and the running of the atom probe for this Paper as well. I prepared specimens for all analyses and also did the atom probe data treatment. Lastly, I wrote the Paper.

In manuscript.
ACKNOWLEDGEMENTS

This work could not have been done without my supervisors. I would like to thank Lars Hultman for inspiring me to apply to this position, for believing in me, as well as his constant optimism and ways of turning things around to the better. My co-supervisor Magnus Odén also deserves praise for his unwavering support, good discussions, and for choosing the high road.

I thank my predecessor, former colleague, and friend Lars Johnson for introducing me to the project and the life as a graduate student, getting me started, and being patient with all my questions. I am very fortunate that you remained in the project even after leaving Linköping University.

Another person whom must be recognized for his support is Mattias Thuvander, my atom probe expert at Chalmers University of Technology. I always feel welcome when visiting your group in Gothenburg and I’m looking forward to continue our fruitful collaboration.

I would like to thank all the members of FunMat Theme 2 for nice discussions and helpful feedback. Special thanks are extended to Mats Johansson Jöesaar for helping me with the deposition system and film growth at SECO Tools in Fagersta.

My thank goes out to all friends and colleagues in the thin film physics division and the whole of IFM. Especially to Mathias Forsberg for being lost with me the first few weeks (months), as well as my current and former office mates, Lina Tengdelius, Olof Tengstrand and Amin Gharavi, for the respect we show each other and the nice work environment we have created together.

My friends and my family deserve my sincerest thanks for always being there and especially for your support when life dealt me a bad hand. The person I am today is as much a result of your strength and persistence, as my own.

Nevertheless, my greatest thank goes to my wife Linda, for standing me by all the long years we were apart and for the hard work and sacrifices she has made to make us a family. Every day you make my life better!
# TABLE OF CONTENTS

## Introduction 1

## 2 Materials 3

2.1 Bulk Materials 3

2.2 TiN 3

2.3 TiSiN 4

## 3 Thin Film Deposition 7

3.1 The Structure Zone Model 7

3.2 Hardening and Strengthening Mechanisms 9

3.3 Cathodic Arc Deposition 10

3.3.1 Modes of Operation 11

3.3.2 Reactive Cathodic Arc Deposition 11

3.3.3 Sample Rotation 11

## 4 Characterization 13

4.1 X-Ray Diffraction 13

4.1.1 Bragg-Brentano Setup 13

4.2 Electron Microscopy 14

4.2.1 Excitation Volume 14

4.2.2 Scanning Electron Microscopy 15

4.2.3 Transmission Electron Microscopy 16

4.3 Energy Dispersive X-ray Spectroscopy 16

4.4 Elastic Recoil Detection Analysis 17

4.5 Atom Probe Tomography 17

4.5.1 Specimen preparation 18

4.5.2 Detection efficiency 19

4.5.3 Delay-Line Detector and Multiple Events 20

4.5.4 Tip Reconstruction 21

4.5.5 Evaporation Behavior of Ceramics 22

4.5.6 Local Magnification Effects 23

4.5.7 Mass Spectrum and Ranging 24

4.5.8 Mass Resolution 25

## 5 APT Data Treatment 27

5.1 Voxel Size and Delocalization 27

5.1.1 Voxel Size for Visualization of Small-scale Fluctuations 28

5.2 Radial Distribution Function 29

5.3 Frequency Distribution Analysis 31

5.4 Surfaces and Proximity Histograms 32

## 6 Main Results & Contributions to the Field 33

## Bibliography 35

Papers 41

Paper 1 43

Paper 2 69
1. INTRODUCTION

Whether one can see them or not, thin films are found everywhere in modern society. They are the reason why steaks don’t stick in a modern frying pan; why so few reflections disturbs the farsighted reader; and why heat, in the form of infrared radiation, is not transmitted through a pane of window glass, whereas light in the visible part of the spectrum is.

This thesis focuses on the deposition and analysis of ceramic TiSiN thin films with industrial uses in cutting, drilling, and machining. The tools must be both hard and tough to survive in the harsh environment of such processes for any extended period of time; however, this combination is not commonly found, as hardness and brittleness goes together while toughness is associated with ductility. Nevertheless, by depositing a strong work piece material with a thin, hard coating, it is possible to combine these properties and significantly increase the lifetime and work temperature of the coated tool.

The films have been analyzed using a novel analysis technique called atom probe tomography (APT) that quite recently has been adapted for thin films and ceramic materials [1]. Due to mass spectral overlaps inherent to TiSiN, this technique has been combined with isotopic substitution within the deposited film; an uncommon method in this field but established in other analysis techniques [2-4]. As a consequence of this, parts of the thesis concern the development and evaluation of analysis procedures necessary when broadening the field to include ceramic materials in general and materials with mass spectral overlaps, such as TiSiN, in particular.
Several materials systems are part of, or relevant to, the investigations presented in this work. These materials systems are presented in this chapter. The motivation behind the choices of materials used is also given.

2.1 Bulk Materials

While thin films are the focus of my work and this thesis, these must be deposited on a substrate, and the choice of substrate can sometimes have significant effect on the characteristics of the film. In cutting tool applications, the bulk material must, of course, meet a number of demands unrelated to the film growth; strength, hardness, and thermal stability being some of the most important ones.

Common bulk materials for cutting tools are cemented carbides [5] and polycrystalline cubic boron nitride (PCBN) [6]. As cemented tungsten carbide (WC-Co) was the substrate of choice in Paper 1 and 2, it will briefly be described here. This composite material consists of small grains of WC surrounded by a binder phase consisting mostly of Co, but often includes small amounts of other metals. Its structure is quite similar to a brick wall, where WC plays the role of bricks while Co, which has good wetting properties [7], is the mortar keeping it together. The WC grains are randomly oriented, rather than positioned in the orderly fashion of bricks in a wall. The grains are very hard and brittle, while the binder phase provides toughness, which allows the tools to be deformed without immediate brittle failure. In addition to good wetting properties, Co makes the substrates magnetic, which allows simple sample mounting using magnets, but can create drift problems during specimen preparation with focused ion beam (FIB) unless first demagnetized.

2.2 TiN

TiN was first used as a decorative coating in the making of jewelry, because it has a golden yellow color. It was later adopted as a protective coating and is still popular due to its versatility. However, more recently engineered materials systems like TiAlN, TiCN and TiSiN have surpassed its cutting performance when processing selected groups of materials or in certain modes of operation.

No pure TiN layers were grown during this work, except for diffusion barriers between the substrate and the TiSiN films, but it can serve as a reference for the TiSiN film. TiN coatings for cutting tools are often grown by cathodic arc deposition. These coatings are generally dense and polycrystalline, with large columnar grains [8]. The crystals are fcc with the NaCl-structure (Fm̅3m space...
The lattice parameter is \( \sim 0.424 \) nm, with each unit cell containing 8 atoms, as shown in Fig. 1.

![The NaCl-structured unit cell of TiN.](image)

**Fig. 1.** The NaCl-structured unit cell of TiN.

### 2.3 TiSiN

In an attempt to increase the maximum work temperature of TiN by improving its oxidation resistance, oxide forming elements such as Al and Si were added to the mixture. The hypothesis was that when the coatings were exposed to air, a thin layer of oxide would be formed on the surface. As the diffusion coefficient of oxygen in these surface oxides is much lower than in the coating, atmospheric oxygen would effectively be prevented from diffusing further into the bulk, which would otherwise occur at high temperatures and be detrimental to the lifetime of the coating.

Even though a protective layer of SiO\(_2\) can be formed [9], it was not the main improvement gained from combining TiN with S. SiN turned out to be immiscible with TiN over a wide range of compositions [10], meaning that if supplied with enough energy, a solid solution of TiSiN will phase separate into TiN and SiN. The incorporation of a moderate amount (~2.5 at. % Si [11]) of an immiscible compound in TiN significantly altered the structure of the coatings, resulting in severely decreased grain size. Even so, the grains in TiSiN retain the crystal structure of TiN, i.e. NaCl (Fm\(\overline{3}\)m), possibly with small amounts of Si substituting Ti in the metal/metalloid sub-lattice, however without any significant change in the lattice parameter. It was found that these coatings excel at dry cutting and high temperatures.

Nanocomposite TiSiN is a structure with phases at, or close to, thermodynamic equilibrium, as opposed to the cathodic arc deposited coatings with metastable phases. More energy is supplied to the coatings during growth, in order to drive phase separation to the very limit, where the TiN and TiSi grains are very pure and surrounded by a Si\(_3\)N\(_4\) tissue phase [10]. The SiN is
grain refining also in the nanocomposite coatings, and the hardness reaches a maximum when the amount of Si in the coating corresponds to a monolayer around all TiN and TiSi grains; the exact amount is thus dependent on the grain size [12].

The micro and nanostructured TiSiN films investigated in this thesis are grown at low temperatures to allow formation of metastable phases, and not necessarily form nanocomposites with a Si$_3$N$_4$ matrix. The materials system was chosen because it is interesting from both scientific and commercial points of view. Details regarding the nanostructure needed to be determined in order to understand and model the growth in detail. This, in turn, could lead to better control of the cutting properties of deposited films.

APT was identified as a good technique for this task, had it not been for the mass spectral overlaps of Si and N, making them indistinguishable by the time-of-flight mass spectrometry used to identify ions in APT. By growing the films using $^{15}$N instead of $^{14}$N, the mass spectral overlap was largely avoided, as is thoroughly described in Paper 1, which enabled a detailed APT investigation of the film in Paper 2.
3. THIN FILM DEPOSITION

By tradition, techniques for making thin films are categorized into two major branches, chemical and physical vapor deposition (CVD and PVD, respectively). To be classified as a CVD process, the material must be deposited as a result of one or more chemical reactions, whereas the processes in PVD are purely physical.

Even though there are many different CVD and PVD techniques, each branch has several characteristics in common and the two branches are in many ways complementary to each other. CVD generally produce coatings of high quality, especially when complex geometries are to be coated. The deposition rate is generally slow, but this can be compensated by designing the reactor to provide a uniform gas flow, since that allows production of very large batches. However, all elements of the coating must be available in gas phase, which often means that toxic and environmentally unfriendly carrier gases are used. In addition, CVD coatings are generally grown at high temperatures, as many of the required chemical reactions have high activation energies.

PVD on the other hand, combines high deposition rate with low temperature. The absence of dangerous and environmentally unfriendly gases needed in many CVD processes makes PVD safer and less problematic to work with. All in all, this effectively reduces the cost of growing films compared to most CVD processes. Furthermore, it enables deposition onto sharp edges of a tool, which is a requirement for cutting tools. Lower temperature can also result in the deposition of metastable phases, which is the main reason why the thin films in this thesis have been made exclusively with PVD techniques.

This chapter starts with parts of the theory of thin film deposition relevant for this thesis, before the deposition techniques used to produce the investigated coatings are explained in more detail.

3.1 The Structure Zone Model

Common characteristics of polycrystalline thin films can often be described in general terms by the structure zone model (SZM). It reduces many practical parameters of film growth to a few parameters directly linked to the growth process. All versions of the SZM include the growth temperature on one axis [13, 14]; often normalized by the melting temperature of the deposited film (homologous temperature), but more recently also compensated for the potential energy of the arriving particles [15]. The other axis has changed during the years from substrate bias [13] and pressure [14] to the kinetic energy of the arriving particles [15].
As TiN growth with cathodic arc deposition is one example where the SZM can be successfully applied, it serves as a good starting point for describing how the growth is altered by the addition of Si. The SZM by Thornton [14] in Fig. 2 shows how the grain size and shape develop for different temperatures and pressures. In zone 1, the adatom mobility is low, which increases nucleation and results in small grains. With increasing energy, surface diffusion start to play a more important role, while grain boundary diffusion is still limited, resulting in the competitive growth that is characteristic to zone T [16]. With even more energy the grains grow into columns that often extend throughout the entire coating, which characterizes zone 2. In zone 3, the atom mobility is high enough to allow bulk diffusion and recrystallization, resulting in a dense, large grained structure.

TiN films for cutting applications are generally deposited in the transition zone (zone T). Barna and Adamik [16] extend the SZM by also taking impurities or co-deposited additives into account, which at least partly can be used to describe how the addition of Si affects the growth of TiN. The impurities may either be dissolved in the lattice or segregate to the surface and possibly disrupt structure forming phenomenon, reducing the grain size. Given the low solubility of SiN in TiN [10], this is believed to be the reason behind the grain refining effect of Si in TiSiN. This is discussed in more detail in Paper 2.

![The SZM as described by Thornton [14]. From [17] with permission.](image-url)
3.2 Hardening and Strengthening Mechanisms

There are many mechanisms to increase the strength and hardness of a material, some of which will be covered here. The plasticity of a material is highly related to the movement of dislocations. These dislocations are crystallographic defects that require a small amount of energy to move along the ordered structure of a lattice, but are hindered by disturbances of this order until sufficient energy is supplied. This energy can, e.g., be from gentle heating (so as to avoid recrystallization and defect annihilation) or mechanical work. However, the latter also creates new defects that hardens the material and is thus known as work hardening [18]; a mechanism uncommon in ceramics due to their brittleness.

The remaining mechanisms are focused on impeding the movement of dislocations, which can be done in several ways. By adding one or more alloying elements to a base material, dislocation movement can be impeded by the resulting imperfections of the lattice. The different sizes of the atoms in the alloy cause lattice strain, which increases the energy barrier for dislocation movement. Large alloying elements substitute the lattice atoms while small alloying elements are located in interstitial sites. If the solubility limit is reached, precipitates of another phase are formed and this is then called precipitation hardening or age hardening. Gentle heating, be it from deliberate annealing or use, may also be needed to allow diffusion of the atoms forming the precipitates. The precipitates slow dislocation movement and can sometimes stop crack evolution. If the precipitates are small they may strain the lattice to remain coherent, which can be compared to a single solute atom straining the lattice in solid solution strengthening. That dislocations require more energy to go through the strained regions close to coherent phase boundaries is called coherency strain hardening [19, 20].

When not coherent, or when the angular difference in the lattice orientation is significant, grain boundaries serve as strong breaches of the lattice order and thus hinder dislocation movement. When the difference in shear modulus of the two grains is significant, called a Koehler barrier, the dislocation movement across the grain boundary is impeded [21]. By decreasing the grain size, dislocations are likely to reach grain boundaries more often, thus slowing their movement in average. This method is known as Hall-Petch, or grain boundary, strengthening [22]. As dislocations move faster within the grain than between grains, there will be a pile-up of dislocations close to grain boundaries. The pile-up makes it easier for dislocations to cross into another grain. When the grain size decreases, fewer dislocations fit in the grain and the pile-up effect will decrease. Eventually, this will pin the dislocation in the grain, which increases the strengthening effect even more. Unfortunately, there is a limit to the Hall-Petch strengthening mechanism. If the grains become smaller than a critical grain size, typically 10 nm or less, they may start to move with respect to one
another [23]. This phenomenon, often called the inverse Hall-Petch effect, is caused by a deformation mechanism known as grain boundary sliding and should this mechanism be active, the effects of hindering dislocation movements will no longer be relevant in determining the material strength.

The thermodynamically most favorable state of SiN is the tetrahedrally coordinated $\text{Si}_3\text{N}_4$. It has, however, been shown that thin layers of SiN can be stabilized into a cubic-related phase by cubic TiN with coherent interfaces [24-26]. Such coherent interfaces would cause coherency strain due to the difference in the size of the unit cell of c-TiN and c-SiN. The coherency will slow dislocation movement, but when the layer becomes too thick, the coherency is lost.

With the addition of Si to TiN, the grain size is significantly decreased, which causes Hall-Petch strengthening, and the films exhibit an exceptionally high defect density [27]. Even though the grain size of the films analyzed in this thesis in some cases are very small, there will be a resistance against grain boundary sliding, since energy must first be used for breaking the coherency, effectively decreasing the critical grain size [28]. In addition to this, the shear modulus of TiN and SiN differ significantly, thereby adding Koehler barriers to the list of possible mechanisms influencing the properties of TiSiN.

### 3.3 Cathodic Arc Deposition

Cathodic arc deposition uses highly energetic arc discharges to remove material from a cathode. This material is then deposited on substrates placed in front of the cathode. The technique is also known as cathodic arc evaporation, which suggests that atoms evaporate from local, arc induced melts, but in reality a majority of the atoms are sublimated directly into an ionized state [29]. A negative bias is applied to the substrates to attract the positive ions, which will accelerate and impact the substrate at high speeds.

Cathodic arc was chosen as deposition technique for the films in Paper 1 and 2 because it has a high degree of ionization. This enables good control of the speed of impinging ions by the bias voltage. At high speeds, the intermixing between the substrate and the coating improves adhesion. At the same time, energy can be supplied to the coating without heating, which means that it is possible to tailor the structure of the growing films in accordance with the SZM without using high temperatures. This, in turn, is a prerequisite for growing the structures of interest in Paper 1 and 2, as these consist of non-equilibrium phases. Lastly, cathodic arc deposition is the work horse of the commercial cutting tool industry, as it is fast and thereby cheap. Since the films in Paper 1 and 2 are grown in an industrial system with similar parameters as those used commercially, any findings can be directly related to commercial products and quickly be put into practice. Through this, the research becomes more easily
available and useful to the society; one out of three major tasks assigned to all Swedish universities (the other two being research and teaching).

3.3.1 Modes of Operation
Cathodic arc deposition can be operated in direct current (DC) or pulsed mode. In DC mode, all cathodes supply material continuously and the coating composition is adjusted by varying placement of the substrates, the ratio of element in compound cathodes, or both. In pulsed mode, the coating composition is mainly adjusted by varying the pulse frequency of different pure cathodes, but compound cathodes are also possible to use. Other factors such as re-sputtering of deposited material will also affect the final composition, but is not used for regulatory purposes [30]. The growth of the films in Paper 1 and 2 was conducted in a manner that mimics the commercial growth of such films, which is done in DC mode, as the growth rate is high.

3.3.2 Reactive Cathodic Arc Deposition
When making ceramic coatings, alternate strategies for depositing non-metallic elements might be necessary. Cathodes with high electrical conductivity are necessary to sustain the arc. As ceramics are inferior conductors, it is most often not advisable to use compound cathodes. Pure non-metallic cathodes are possible with e.g. carbon, but a common technique is to supply the non-metallic element in gas phase when applicable. Metallic ions from the cathodes are subjected to a gas, e.g., N₂, and a ceramic hard coating is formed on the substrate through reactions with the gas, hence the name reactive cathodic arc deposition. As Paper 1 and 2 regarded nitride thin films, this technique was used for depositing the films of. However, there are, of course, problems associated with this method; e.g., that the gas will also react with the cathode. This process, called poisoning, causes the cathode deposition rate to decline [31] and will ultimately render the cathode unusable.

3.3.3 Sample Rotation
Cathodic arc deposition is a line of sight technique, meaning that ions from the cathode will be deposited more or less in front of the cathode. A common way to achieve homogeneous coverage is to mount the samples on a rotating stage. The simplest is 1-fold rotation, where the substrates are placed on a rotating drum. By, e.g., rotating several rotating drums in the chamber, 2-fold rotation is achieved while 3-fold rotation generally uses rotating substrate fixtures as well. Each additional rotation increases coating quality and enable more complex shapes to be homogenously covered, but at the cost of lower deposition rates.
As cathodic arc deposition is a line of sight technique and homogenous coverage is preferred in most cases, protective coatings are in general deposited using rotation around one or more axes. However, in Paper 1 and 2, stationary deposition was chosen because of a limited supply of $^{15}\text{N}$ combined with an interest primarily in the nanostructure, which should not be affected by inhomogeneous thickness. However, as the films in Paper 1 and 2 were grown without substrate rotation, they lack the compositional layering common in such films grown with single rotation. It has been shown that the sputtering yield during deposition varies with the angle of the incident ions to the surface normal [30]. Least Si is sputtered away from the film at normal incidence and it increases with higher gracing angle of incidence. This means that less Si should be sputtered away with the stationary setup compared to rotating. Thereby, a slightly higher Si:Ti ratio than what has been previously reported could possibly be achieved.
4. CHARACTERIZATION

To understand the relation between growth parameters and film properties, it is necessary to investigate the structure of the films. Because no technique can provide answers to all questions, several complementary techniques have been used in the works of this thesis. These are explained in this chapter.

4.1 X-Ray Diffraction

Light passing through an opening of the same size as the wavelength of the light will undergo diffraction, i.e. scatter in a predictable fashion. The lattice spacing of crystalline materials can be seen as a three-dimensional ordered constellation of openings through which light may pass. As the wavelength of light in the X-ray regime of the spectrum coincides with the spacing of the crystal lattice, diffraction will occur. The way in which the X-rays diffract can reveal important clues to the structure of the crystal. This technique is known as X-ray diffraction (XRD) and it has been used to evaluate films grown in this thesis.

4.1.1 Bragg-Brentano Setup

XRD and related techniques can be conducted in a number of different ways depending on the properties of the sample and what information is sought. The most common method, which has been used in this thesis, is with the Bragg-Brentano setup, also known as powder diffraction. In this setup, shown in Fig. 3, the X-Rays irradiate the material of interest while the angle of incidence and the detector angle are scanned, but always kept equal. At certain angles \( \theta \), constructive interference of X-rays reflected in crystal planes will occur, yielding an intensity peak in the detector.

![Fig. 3. The Bragg-Brentano, or powder diffraction, setup of the XRD system.](image-url)
With this setup, only planes perpendicular to the surface contribute to the peaks. For heavily textured samples, it is important to know how they are oriented in order to get diffraction, but when investigating powders or polycrystalline coatings, there will always be some grains oriented in such a way that their planes are parallel to the surface. Since only part of the sample contributes to the peaks, their intensity is decreased, but this can be counteracted by longer measurement times or wider slits; the latter may however also broaden the peaks, as the angular uncertainty increases. Because the lattice spacing varies between different crystals and in different directions of the same crystals, the angles $\theta$ at which intensity peaks occur can be used to identify the crystal structure of the sample. The different phases of the investigated materials are identified using reference spectra from an international database.

4.2 Electron Microscopy

Electron microscopy is a collection of different techniques that all uses accelerated electrons to characterize a sample. Because of this, they have much in common. One important benefit of electron microscopy compared to light microscopy is the resolution limit, which states that the size of the smallest resolvable object is on the order of the wavelength used to image it. The wavelengths of light visible to the human eye is in the range of 400-700 nm, which sets the absolute resolution limit. Even though it is possible to go beyond the wavelength range of the human eye by the use of detectors, the probability of light-matter interactions decreases rapidly with wavelength.

The wavelength of an electron depends on its speed, so by accelerating the electron using a bias, the wavelength can be tuned. Even though the probability of electron-matter interactions drops with decreasing wavelength, the change is not as rapid as the decrease of light-matter interactions. However, the resolution of electron microscopy is limited by the poor quality of electron lenses, compared to the optical counterpart, and the excitation volume.

4.2.1 Excitation Volume

An electron that interacts with a sample may scatter several times before its energy is lost, or it has found its way out of the sample. Statistically, the electrons interact with the atoms of a sample in a volume shaped like a teardrop hanging where the electron beam meets the sample. This teardrop shape, shown in Fig. 4, is called excitation volume and its size depends on the electron energy and the material in question. The teardrop shape is, however, only found if the sample is sufficiently thick.

As the electron interacts with matter, a number of processes can occur that may generate signals, e.g., secondary electrons, Auger electrons, and X-rays, which can be detected. Depending on at what depths these processes occur,
different signals may, or may not, reach the detector. The width of the excitation volume is of importance for the spatial resolution in electron microscopy and related techniques. This will be discussed further for each different technique below.

**Fig. 4.** A cross-section of the excitation volume in electron microscopy, i.e. the volume where electron-matter interactions are most likely to occur when a sample is hit by accelerated electrons. The detectable signals from different sample depths are also shown.

4.2.2 *Scanning Electron Microscopy*

A focused beam of accelerated electrons scanning over the surface of a sample is the working principle of the scanning electron microscope (SEM). At each position, the secondary electrons generated by the beam are accelerated toward a detector and the number of secondary electrons determines the brightness of the related pixel on the display. The secondary electron yield varies with the acceleration voltage, as well as the work function and atomic number of the surface atoms, but the local curvature of the surface has the most significant effect. Because of this, SEM is a good technique for getting an overview of the film surface.

The secondary electrons are generated throughout the excitation volume, but their energies are generally low, so most of the electrons reaching the detector have been generated fairly close to the surface. The width of this region is larger than the width of the electron beam, but smaller than the width of the entire excitation volume.
4.2.3 Transmission Electron Microscopy

In transmission electron microscopy (TEM), the electron beam is detected after passing through the sample. This means that the sample must be very thin, both to allow the electrons to pass through and to produce clear micrographs. Ideally, the sample should be thin enough that an electron scatters no more than once before leaving the sample, which corresponds to nanometer thicknesses. In addition, the likelihood of overlapping features that make micrographs hard to interpret becomes larger with increasing thickness. As the electrons scatter only once, the excitation volume is reduced from its drop-shape to a disc with approximately the width of the electron beam, which increases the spatial resolution compared to SEM. The low quality of the electron lenses is instead the limiting factor of the resolution in TEM.

The microscope can be used in different ways to investigate different aspects of the sample. In its most straightforward way, an image is formed in the focal plane and the contrast is generated by the difference in phase between the electrons. If an image instead is formed in the back focal plane, a 2-dimensional diffraction pattern appears. This pattern is closely related to the 1-dimensional diffraction pattern recorded with XRD, as crystals can diffract accelerated electrons in the same way as X-rays.

By inserting an aperture into the back focal plane, different features of the diffraction pattern can be selected, which influences the micrograph. If the center beam of unscattered electrons is chosen, dark areas will correspond to diffracting grains. This mode is called bright field. If another part of the diffraction pattern is selected, then the micrograph will be bright in areas contributing to that part of the diffraction pattern. Dark field TEM was used in Paper 1 to estimate the grain size of Ti$_{0.81}$Si$_{0.19}$N and in Paper 2 to show the difference in structure of Ti$_{0.92}$Si$_{0.08}$N and Ti$_{0.81}$Si$_{0.19}$N.

4.3 Energy Dispersive X-ray Spectroscopy

Energy dispersive X-ray spectroscopy (EDS, in some cases also abbreviated EDX or XEDS) is used in combination with SEM or TEM. It can identify what elements a region of a sample consists of by means of electron-matter interactions. Occasionally, when an accelerated electron collides with an atom, part of its kinetic energy is used to remove another electron from an inner shell of the atom. The atom is then in an excited state, but will eventually return to a relaxed state as an electron from an outer shell fills the inner. This relaxation process releases energy by emission of radiation (or of a valence electron, which is then known as an Auger electron). The wavelength of the radiation, typically found in the X-ray range of the spectrum, will be characteristic of the transition in question, i.e. the energy difference between the two shells, which in turn is element specific. By scanning the electron beam and detecting the wavelength of
the emitted light, the composition of a sample can be determined. As the technique relies on electron-matter interactions, it is most suitable for detecting elements with a high atomic number, as these have more electrons and are thus more likely to interact with the electrons of the beam. This difference decreases the certainty of quantitative analyses when the difference in atomic number is great, and is one reason why the N content measured by several techniques in Paper 1 differs the most.

The spatial resolution of EDS varies depending on the excitation volume. The characteristic X-rays measured interact weakly with matter, and will thus have no problem of reaching the detector, but they are only generated if the accelerated electrons have enough energy to knock out core electrons, thus excluding the edges of the drop-shaped volume. Unless the sample is very thin, such as in EDS conducted in scanning TEM, this part of the excitation volume will be significantly wider than the electron beam, which otherwise is the limiting factor of the spatial resolution of EDS. The resolution is thus too low to provide all the details of the fine grained TiSiN structure, but can be used as a complementary technique to determine the average composition in a large volume, as was done in Paper 1.

4.4 Elastic Recoil Detection Analysis

In time-of-flight energy elastic recoil detection analysis (ToF-E ERDA), accelerated heavy ions, $^{127}$I$^{8+}$ in the case of this thesis, hit an area of the sample from a pre-defined incidence angle. They have enough energy to remove atoms that are lighter than themselves from the sample through elastic collisions. At a specific angle, these recoiled atoms pass two timing gates and the time-of-flight between the gates of each atom is measured before it hits an energy sensitive detector. The time-of-flight relates to the mass of the recoiled atom while the energy relates to the depth from which it originated.

Plotting the energy as a function of time-of-flight (or vice versa) of all recoiled atoms gives banana-shaped curves that are ranged to a specific element or isotope. This enables depth profiling of the sample composition.

ERDA is used to provide a reference measurement of the TiSiN compositions in Paper 1, because EDS is generally less sensitive to N and other elements with low atomic number, while this is not the case for ERDA.

4.5 Atom Probe Tomography

The main principle of atom probe tomography is controlled demolition. A very small volume of the film is shaped into a tip with an apex diameter of the order of 100 nm. A DC voltage is applied between the tip and a counter-electrode in such a way that the atoms at the apex almost field evaporate. By also applying short
voltage or thermal pulses, the atoms can successively be pushed over the threshold to become free ions. The tip is kept at cryogenic temperatures in order to reduce surface diffusion on the apex and, in case of thermal pulsing, to reduce the risk of heat build-up. The DC voltage is regulated so that the evaporation rate is kept at fractions of a percent to increase the probability of generating no more than a single ion per pulse.

While the ion is still in the vicinity of the tip, electrons may tunnel back, thereby increasing the charge state of the ion in a process called post-ionization [32, 33]. Regardless, the ions are accelerated toward a multi-channel plate (MCP) by the electric field between the tip and the counter electrode. The MCP converts the ion signal into a burst of electrons, which then are detected. The detector measures the spatial x and y-coordinates as well as the time-of-flight, i.e. the time between the latest pulse and the detection event. A schematic image of the basic principles of the atom probe is shown in Fig. 5. From the time-of-flight, the elemental identity of the ion can be determined. Together with the spatial coordinates and the sequence in which the ions were field evaporated, a 3D reconstruction of parts of the tip, with elemental information, can be created by back-projection.

![Diagram of atom probe (APT) setup](image)

**Fig. 5.** The basic principles of laser-assisted atom probe tomography. The different components are not to scale.

### 4.5.1 Specimen preparation

As the analysis technique requires samples to be sharp tips, and the samples of interest are thin films, there are two possible ways to go; either to deposit sharp tips with film or to make sharp tips from already deposited films. The first option is poor because we cannot be certain that the growth is not affected by the choice of substrate or its geometry. WC-Co is not available as ready-made tips for APT and would be tedious to make. The best way is thus to remove small pieces of films grown on flat WC-Co substrates, attach them to ready-made Si truncated
tips using Pt or W deposition, and shape the pieces into tips with a diameter of about 100 nm using an annular milling pattern [34]. All of this can be done using FIB; an instrument that usually combines a SEM with a beam of Ga⁺ ions, or more recently Xe⁺ ions. Even though electrons primarily are used for imaging, and ions primarily for milling away material, both beams can be used for imaging and, with the addition of locally applied reactive gas, for depositing material.

The tips in Paper 1 and 2 were made using a Ga⁺ ion beam, which can result in implantation of Ga. This problem was minimized by protecting the tip from implantation of milling ions with a cap of Pt or W deposited before the pieces were removed from the rest of the film. What remains of the cap after annular milling can then be removed by milling with Ga⁺ ions with a lower acceleration voltage, effectively reducing the amount of implantation ions and the implantation depth.

4.5.2 Detection efficiency

One of the major benefits of APT is that the detection efficiency in most cases is uniform, i.e. not dependent on atomic number, ion charge state, etc. This means that there is no need for reference samples in quantitative composition analyses. High detection efficiency is important when small structures are to be investigated, as it influences how many atoms a feature must consist of to be distinguishable from noise. The detection efficiency is primarily influenced by obstacles in the ion flight path, such as the reflectron and the MCP of the detector.

The reflectron is essentially an electric field, generated by a set of curved grids, designed to alter the flight path of the ions in accordance with their speed [35]. Ions with high kinetic energy travel far into the reflectron before their trajectory is deflected toward the detector, while those with low energy travel a shorter distance. As the length of the flight path is adjusted to fit the ion speed, the differences in the time-of-flight for each ion kind decreases, resulting in higher mass resolution. However, ions created by thermal pulsing already have so low energy spread that the use of a reflectron is not necessary. Unfortunately it cannot be removed during thermal pulsing to increase the detection efficiency.

The MCP is shaped like a honeycomb, where each channel converts the ion into an electron signal, which is multiplied by secondary emission of electrons in order to create a detectable electric pulse. As the walls of the honeycomb cannot be made infinitely thin, some ions are bound to miss the active area of the MCP and remain undetected. About 60% of the MCP area is active in the instrument used in this thesis, but as it is also fitted with a reflectron, its detection efficiency is reduced to ~37%. This can be compared to the Cameca LEAP® 5000 XS (without reflectron) with a detection efficiency of ~80%, which is the highest of currently available instruments.
4.5.3 Delay-Line Detector and Multiple Events

The previous section regarded the detection efficiency of the instrument, however, the effective detection efficiency can be even lower depending on the detector design, what algorithms are used for analyzing the detector signal and the evaporation behavior of the sample. Imago LEAP® 3000X HR, the atom probe used in Paper 1 and 2, uses a delay-line detector [36]. In its basic form, it consists of two wires drawn back and forth to cover the detector in such a way that the wires are perpendicular to each other. When a signal from the MCP reaches the wires in a particular spot, it is divided and travel along each wire in both directions. By analyzing the difference in time when the signals reach the four wire ends, i.e. $\Delta t_x = t_{x2} - t_{x1}$ and $\Delta t_y = t_{y2} - t_{y1}$ using the notations from Fig. 5, the exact position can be calculated. Because of the detector design, a region around a detection site cannot detect another ion within a short period of time from a detection event. This region is called the dead zone and the period of time is called the dead time. They are related in such a way that the dead time decreases with distance from the detection event [37]. If a second ion hit the detector within the dead zone, at most one ion will be detected. When more than one ion reaches the detector in a pulse it is known as a multiple event, and with each additional ion the complexity in resolving which delay-line pulses correlate to a specific detector event increases significantly. To improve the detection capability for multiple events, and to be able to confirm the x and y-coordinates, an extra delay-line at 45° angle to the standard delay-lines has been added [38].

Multiple events can have several sources. The residual gas of the vacuum chamber constantly hits the detector. Should they arrive at the detector shortly after a voltage or thermal pulse on the sample, they will be recorded. Since such events are completely uncorrelated with the pulse, the time-of-flight and subsequent mass-to-charge state ratio will not depend on the ion mass or charge, and will thereby only contribute to the background noise. Should a residual gas atom hit the detector during the time-of-flight of a field evaporation ion, it will be considered a multiple event. Since these multiple events are random in character, they will not skew the compositional analysis; however, multiple events involving more than one field evaporated ion can produce more severe effects. Such effects can occur either by field evaporation of more than one ion in the same pulse, if the ions were in close proximity to each other on the tip, or through mid-flight dissociation of molecular ions.

With poor choice of DC voltage and pulse energy, multiple events can also occur when ions are field evaporated by the DC voltage between pulses. However, as these ions are uncorrelated to the pulse and will only contribute to the background noise, the main issue is the reduction of the relative abundance of that kind of atoms detected during pulsing, rather than the increase of multiple events.
4.5.4 Tip Reconstruction

After the tip has been field evaporated and the ions have been detected, the tip must also be reconstructed to make analyses possible. The volume of the reconstruction is more or less set by the number of detected ions, and the process serves to mold this volume into the correct shape. The reconstruction is based on the protocol by Bas et al. [39], with later improvements by e.g. Gault et al. [40], in which back-projection of each individual ion is used. The resulting magnification $M$ can be calculated by Eq. 1 below, where $L$ is the ion flight length, $R$ is the radius of curvature of the tip, and $\xi$ is an image correction factor that accounts for electro-static coupling of the field to the vacuum chamber.

$$M = \frac{L}{\xi R}$$

Knowing the magnification enables back projection of the x and y-coordinates of the detector event to the x and y-coordinates of the ion in the tip. The z-coordinate is calculated using the sequence in which the ions were evaporated, the volume per ion in the sample, the detection efficiency of the instrument and the estimated radius of curvature. The last can be calculated from the tip shank angle [41] or estimated from the evolution of the tip diameter in pre-analysis micrographs, as well as from the evolution of the applied DC voltage, $V$, using Eq. 2, where $k_f$ is a field factor that accounts for tip geometry and $F_e$ is the evaporation field.

$$R = \frac{V}{k_f F_e}$$

The product $k_f F_e$ can be found by measuring the end radius of curvature in SEM or TEM micrographs, given that the analysis is stopped before the tip ruptures from field induced stress [42, 43].

The reconstructions used in Paper 1 and 2 were aided by the tip diameter in pre-analysis SEM micrographs. In case the post-analysis tip was available for imaging, control reconstructions using the measured end radius of curvature were made as well.

The back-projection relies on the assumption that the tip assumes an equilibrium shape of a truncated cone with a hemispherical cap. This assumption works well for pure metals, e.g. Al and W [44], but has shown to be rough when analyzing samples containing different phases, as local magnification effects may distort the reconstruction [45]. Crystallographic features, such as poles or regions with regular lattice spacing, can be used to validate the reconstruction parameters, but only in samples where such features are visible [44].
4.5.5 Evaporation Behavior of Ceramics

The evaporation behavior of a material is influenced by its bond characteristics. Ceramics in general have bonds of covalent and ionic nature. As a result, the field evaporation of an atom or cluster from ceramics, such as TiSiN, is often followed by a burst of atoms and/or clusters from the same area. This process is called correlated evaporation and may reduce the accuracy of quantitative compositional analyses [46], because the APT detector is by design unable to detect ions that arrive too close in both space and time.

The algorithm assumes that field evaporation occurs from a hemispherical surface. With each ion detected, the z-coordinate is increased a minute distance, so that the z-coordinate has increased a length of the order of the nearest neighbor distance when an entire hemispherical layer has been detected. This works well with materials where field evaporation occurs evenly over the entire tip apex, but less so when correlated evaporation produces bursts of ions originating from a small part of the apex, which is the case for TiSiN. Fig. 6 shows a cross-section of an APT analysis of TiSiN, where each dot represents a detected Ti ion. Alternating dense and sparse crescents can be seen along the analysis direction (from right to left in Fig. 6). As a burst occurs in one area of the tip surface, it produces ions unnaturally close in the z-direction. This is how the dense regions of the cross-sectional image are created. In other parts of the tip, the z-coordinate is increased with few detected ions, producing the sparse regions. Because ions in protrusions have higher evaporation probability than in depressions, the shape will not deviate excessively far from the hemisphere assumed in the reconstruction, making the dense and sparse regions in cross-sectional images to appear as crescents.

![Fig. 6. A 5 nm thick cross-sectional image of the Ti ions in an TiSiN APT tip showing typical reconstruction algorithm artifacts with alternating dense and sparse crescents from correlated evaporation.](image)
4.5.6 Local Magnification Effects

As previously mentioned, the assumption of an equilibrium tip shape of a hemispherical cap on a truncated cone may be far from the truth when the analyzed sample contains several phases. If these phases require different field strengths to evaporate, the shape of the apex will no longer be hemispherical and the actual geometry of the tip surface may alter the trajectory of the field evaporated ions.

Take, e.g., a precipitate in a surrounding matrix phase. If the precipitate requires a higher field to evaporate than the matrix phase, it will largely persist while the matrix phase is field evaporated. Eventually, the precipitate will protrude from the otherwise hemispherical surface; a geometry that enhances the field in the precipitate and eventually causes it to field evaporate.

If the precipitate requires a lower evaporation field than the matrix phase, a flat surface, or possibly a depression, will be created instead, as the precipitate atoms have higher probability of field evaporation than the matrix atoms. Ions originating from a protrusion will be spread out in the x and y-plane while ions originating from a depression will be focused, as is shown schematically in Fig. 7.

Using the assumption of a constant hemispherical tip surface in the reconstruction will cause a high field precipitate to appear larger and less dense than it actually is, while a low field precipitate will appear smaller and denser. This phenomenon is often referred to as local magnification effects.

Even though the nanostructure in TiSiN is quite dissimilar from precipitates in a matrix, local magnification may still influence the size and composition measurements of the SiN	extsubscript{y} structures in Paper 2.

![Fig. 7. The influence of local magnification on the size of a precipitate (black atoms) in a matrix phase (white atoms) for different relative field strengths required for field evaporation.](image-url)
4.5.7 Mass Spectrum and Ranging

For each detected ion, the time between the field evaporating pulse and the detection event is recorded. This quantity is often referred to as the time-of-flight. Charged particles in an electric field in vacuum have potential energy described by $E_p = qU$, where $q$ is the particle charge and $U$ is the electric potential difference (voltage). When the particle has left the electric field, all potential energy has been converted to kinetic energy, $E_k = \frac{mv^2}{2}$, where $m$ is the particle mass and $v$ its speed, resulting in Eq. 3.

$$qU = \frac{mv^2}{2} \quad (3)$$

The speed after the acceleration can be described by $v = \frac{d}{t}$, where $d$ is the flight path length and $t$ is the time-of-flight. This results in Eq. 4, which can be rearranged to Eq. 5, where $q = ze$ has been applied to replace the particle charge with the unit-less charge state $z$ and the elementary charge $e$.

$$qU = \frac{md^2}{2t^2} \quad (4)$$

$$t = \frac{d}{\sqrt{2Ue}} \cdot \frac{1}{\sqrt{mz}} \quad (5)$$

Since the elementary charge, ion flight length and DC voltage are constant during the entire flight, the time-of-flight will only depend on the ion mass and charge. The mass is often given in Da and the unit-less charge state is used rather than the charge in C. Knowing the mass-to-charge state ratio is most often enough to identify the different ions when the main components of the sample are known. A histogram, known as the mass spectrum, of the number of ions within a time-of-flight (or mass-to-charge state ratio) range is plotted as function of the mass-to-charge state ratio. This plot has peaks for each kind of ion detected. It is used when manually assigning mass-to-charge state ranges to specific ions, a process called ranging.

In those cases when ions have similar mass-to-charge state ratios, partial or complete mass spectral overlaps will make ranging difficult. If the overlapping ions have several isotopes, these can be used, together with knowledge of the relative isotopic abundance in the sample, to estimate the ratio of the two overlapping peaks. Another option is to deconvolute the overlapping peaks [47] to aid in the ranging. Both these methods have the disadvantage of only correcting the bulk composition, while the ion positions remain unknown and are applied by chance.

If no more than one of the overlapping peaks is large, the damage from such ranging can often be ignored, but if that is not the case, local and global
compositions may be severely affected and analyses of ion distributions may not be reliable. To circumvent this, it may be possible to change the isotope distribution by making the sample with isotope-enriched materials. This is thoroughly covered in Paper 1 for the case of TiSiN, where overlaps of $^{14}\text{N}^+$ and $^{28}\text{Si}^{2+}$, among others, prevented quantitative measurements and distribution analyses by APT.

4.5.8 Mass Resolution
The shape of a mass spectrum peak is directly linked to the mass resolution. What determines the mass resolution is different depending on whether voltage or thermal pulses are causing the field evaporation. For voltage pulsing, it is highly dependent on the shape of the pulse, i.e. the applied voltage over time, but also the field evaporation rate [48], whereas the cooling characteristics of the tip and its shape are highly influential on the mass resolution in thermal pulsing. Even if the voltage pulse during which the ions are generated is very short, some ions will be generated as the pulse subsides. These ions will have a lower initial kinetic energy, which will influence their time-of-flight and reduce the mass resolution. The previously mentioned reflectron can compensate for this difference by adjusting the flight path length. Such time focusing improves the time-of-flight at the cost of detection efficiency, as some ions are bound to collide with the reflectron grids. Moreover, imperfections in the reflectron can introduce errors in the spatial positioning of the ions [36, 49].

Since laser-assisted APT can be applied to a wider range of materials and generally provides better mass resolution [50], it has been used in the work of this thesis and will therefore be considered from here on. The laser induced thermal pulse generates ions with negligible spread in kinetic energy, but atoms may be field evaporated as long as the tip remains heated. The time uncertainty caused by slow cooling cannot be corrected by energy compensating instruments.

The cooling time is related to properties of the tip as well as the heating laser. Samples with high thermal conductivity and long electron diffusion depth seem to benefit most from small laser spots and large shank angles [51]. Local heating from a small laser spot will quickly dissipate to the rest of the tip, lowering the temperature below the field evaporation threshold [52]. The heat dissipation is even quicker with a large shank angle as heat conduction along the tip increases with the cross-sectional area. Samples with low thermal conductivity and short electron diffusion depth will instead benefit most from a large tip diameter and short laser wavelength [51, Fig. 1 in 53]. In such samples, only the surface of the tip will be heated. The core remains cold and acts as a heat sink, effectively decreasing cooling times. A larger tip means a larger core and more efficient cooling, while a smaller wavelength reduces the thickness of the heated surface layer.
However, as large shank angles inevitably lead to an increase in diameter, the effects can be hard to differentiate and generalize to other materials systems than the one examined [54, 55]. The same applies to short wavelengths and small laser spots, since shorter wavelengths can be focused to even smaller spots [52]. As such, the proper parameter to change in order to increase mass resolution will depend on the material and geometry of the specific tip as well as the particular instrument used in the analysis.

As the electrical and thermal conduction of TiSiN are rather poor, the laser spot size and tip geometry are most important for improving the mass resolution, but the spot size is non-adjustable (assuming that it has been properly focused) in the Imago LEAP® 3000X HR instrument used in Paper 1 and 2. Changing to a newer model with a laser in the ultraviolet wavelength range would, however, enable smaller laser spot sizes, which potentially improves the mass resolution through faster cooling.
5. APT DATA TREATMENT

5.1 Voxel Size and Delocalization

When determining the local composition within an atom probe tip, one must first decide on how to sample the data. Local concentrations can only be calculated once a volume has been determined. There are usually two different approaches when employing such volumes, either constant number of atoms per volume or a constant shape and size. The choice depends on what kind of analysis is to be conducted. The former is e.g. used in the original algorithm for finding a proper block size by Hetherington et al. [56] mentioned in Paper 1 and the frequency distribution analysis (FDA) in Paper 2, as the data can be directly related to the binomial distribution when organized in this way. The latter choice is, e.g., made when visualizing the data. For simplicity, these volumes are in most cases cubes, or at least cuboids. As such, they can be considered the 3-dimensional version of the 2-dimensional pixel and thereby designated “voxels”, for volume pixels.

Sampling is, however, always related to sampling artifacts. Ions very close to the border between two or more voxels will only contribute to the voxel within which it is found. This can be tolerated as long as the surface-to-volume ratio is low, but with decreasing voxel size, the error grows.

One solution to minimize these sampling artifacts is to allow the ions to contribute partially to several voxels. The most common choice is to represent each ion with an ellipsoidal Gaussian function centered on the position of each ion. The function is normalized so that the volume integral of the entire ellipsoid returns exactly one ion. This smoothening of the ions is often called delocalization, as the ions become spatially delocalized within an ellipsoid delimited by the cut-off lengths.

In Paper 1 and 2, the cut-off is denoted the delocalization value and approximately corresponds to the $3\sigma$ value, as is standard in the IVAS software. Another standard of the software is to set the delocalization value to 3 times the voxel edge length in the x and y directions, and 1.5 times the voxel edge length in the z direction. This is because the resolution of the APT is considered best along the analysis direction, but due to the correlated evaporation behavior common in field evaporation of ceramics, this may not be the case.

As correlated evaporation may lead to unreasonably short interatomic distances in the z direction and the APT data in Paper 1 and 2 show signs of correlated evaporation, an isotropic delocalization has been chosen in both. The voxel size and delocalization parameter influence not only the visualization of the APT data, but also certain data treatment techniques, e.g., iso-concentration surfaces, and thus the related proximity histogram.
5.1.1 Voxel Size for Visualization of Small-Scale Fluctuations

An important part of Paper 1 involves finding the most meaningful voxel size in order to visualize the data in a correct way. A previously developed algorithm used for 1D APT data was first tested [56]. This algorithm was based on comparing the mean difference in composition of neighboring ion sequences \( \Delta C \) as a function of the sequence length, also known as the ion count. If \( \Delta C \) was found to be stable over a range of ion counts, the local composition would be independent of the ion count within that range. The smallest value within that range should then be chosen to maximize the spatial resolution. Unfortunately, this algorithm seemed to work only for strong compositional variations [56]. It had already been successfully tested for adaptation to 3D voxels, but all details of this conversion was not presented [57].

Unsurprisingly, the 3D version of the algorithm found no range in which \( \Delta C \) was stable for the small-scale compositional fluctuations of TiSiN. However, it turned out that the results could still be made useful if they were compared to the noise level for the specific voxel size. The noise level was simulated by re-arranging the mass-to-charge state ratio of the ions while keeping the average composition and ion positions the same, before the algorithm once again was applied to the APT data voxelized with different voxel edge lengths. This is shown in Fig. 5 of Paper 1. If the experimental curve would have had an obvious plateau of stable \( \Delta C \)-values for different number of ions per voxel (voxel sizes), the difference between the experimental curve and the randomized curve would have had a maximum at the rightmost end of the plateau, corresponding to larger voxels than the leftmost end.

It turned out that even though the experimental curve had no distinct plateau, there was a maximum in the plot of the difference between the experimental and the randomized curves. This is shown for six different datasets in Fig. 6 of Paper 1. The voxel size corresponding to this maximum would then be the best compromise between noise and excessive averaging. For practical reasons, the same voxel size of 0.8\(^3\) nm\(^3\) was used for all datasets even though the maximum were found for slightly larger voxels in a few datasets. This can, however, be motivated by the fact that the best choice of voxel size in the plot showing \( \Delta C \) (as a function of mean number of ions per voxel) would have been at the leftmost part of the plateau, since this increases the spatial resolution compared to the rightmost part. It is thus reasonable to choose a slightly lower value than the one indicated by the algorithm, so long as the change in mean compositional difference (\( \Delta C - \Delta C_{\text{random}} \)) is insignificant, i.e., the slope of the \( \Delta C \) difference curve is not too steep.

It is important to note that the voxel size determined by the algorithm is valid for the specific dataset, compositional fluctuation and delocalization parameter only, but if another dataset or delocalization setting is to be used or another compositional fluctuation than Si/(Ti+Si) is to be measured, new graphs can be generated to determine whether another voxel size is now the better choice.
However, in cases when two different compositional fluctuations are to be compared, it could be relevant to have the same voxel size for both, as long as the difference in most meaningful voxel size is not too large.

5.2 Radial Distribution Function

When short-range order, clustering behavior and corresponding length scales are of interest, as in Paper 2, the radial distribution function (RDF) is a powerful data treatment method to use. It shows how the local composition evolves at increasing distances from a particular kind of ion.

Before the analysis is started, one or several kinds of ions are selected to be centers. In the first step, spheres around each center ion, with radii set by the user, are defined and the relative concentration of all enclosed ions are calculated. In the next step, concentric shells around each sphere, with the same thickness as the sphere radius, are defined and the relative concentration of all enclosed ions is once again calculated. In the following steps, more and more concentric shells are defined and the concentration within each is calculated. Because the RDF samples the data in this way, it is not dependent on the user defined voxel size or delocalization setting.

![Diagram of Radial Distribution Function](image)

**Fig. 8.** The radial distribution function projected to 2D.

Fig. 8 shows an example of volumes defined by the radial distribution function projected to 2D. The concentration of a particular kind of ion as a function of distance from the center ions can then be plotted. Comparing curves generated in the same dataset with different center ions can sometimes be difficult, since the analysis will regard different parts of the sample. This is why minor
constituents generally are chosen as center ions, e.g., Cr when investigating Cr precipitates in Fe-Cr-Al alloys [58], or Si for enrichment of SiN in TiN-rich matrix, as is the case in Paper 2. There are, however, exceptions where RDFs of the matrix phase are used [59].

It is often useful to normalize each curve with the bulk concentration of that particular ion, as this makes deviations from homogeneity apparent. The enclosed volume increases with each consecutive shell, and thereby also the number of enclosed ions; thus, the compositional certainty is proportional to the distance from the center. However, the compositional certainty increases very quickly if there are many center ions.

When analyzing ternary ceramic coatings such as TiSiN with APT, a high number of different kinds of molecular ions are often present, because of the covalent nature of the bonds [60]. A high number of different kinds of molecular ions can make the RDF very difficult to interpret, since each molecular ion generates a curve that must be superimposed on the curves of its constituent ions. This is especially true in the case of RDF curves that have been normalized using the bulk composition of that particular ion, as these must be superimposed in accordance with their bulk composition.

To make interpretation of the RDFs in Paper 2 trivial, the molecular ions have been decomposed into single atoms (as in regular RDFs, the ion charge state is ignored). For all measured ions, the algorithm does not take into account the unlikelihood of finding several atoms at exactly the same position; instead they are simply allowed to occupy the same spatial coordinates. Potential negative effects of this would most likely only be found for very short distances from the center ions, as the increasing volume of each shell will increase averaging.

In case of molecular center ions with more than one of the intended single center ion, such as Si$_2$N in the Si-centered decomposed RDF, it was necessary to combine more than one RDF to decompose molecular ions into single. The molecular center ions posed a problem when decomposed, as the “extra” ions would all contribute to the concentration based on the lowest number of atoms. In this case, the ions were re-distributed into several bins using the Si nearest neighbor distribution measured in that particular APT dataset and their contributions were weighted with the detection efficiency, to account for the fact that they would have had a detection probability similar to the detection efficiency if they had evaporated as single ions.

In the case of Si center ions, which is the center ion used in Paper 2, this compensation for the “extra” ions generated by decomposing molecular ions containing Si had no significant effect on the interpretation of the RDF, but still reduced uncertainties regarding the effect of these “extra” ions.
5.3 Frequency Distribution Analysis

FDA is an example of an APT data treatment method where a block size is used rather than the voxel size. This means that the volume of each block is adjusted so that it contains exactly the chosen number of ions. A small number means that the compositional uncertainty within the block is larger, as the smallest compositional step is 1 at. % with 100 ions, 2 at. % with 50 ions, at. 10 at. % with 10 ions etc. On the other hand, the block volume (in combination with the density) influences the size scale of the analysis, as fluctuations much smaller than the block volumes are averaged.

By combining the FDAs with the ion density, it is possible to get an indication of what feature length scales are investigated. The approximate average ion density $\rho_{\text{ion}}$ in TiSiN have been calculated in Eq. 6, where $n_V$ is the number of ions in the volume $V$ and $\varepsilon_D$ is the detection efficiency.

$$\rho_{\text{ion}} = \frac{n_V \varepsilon_D}{V} = \frac{8 \times 0.37}{0.4243} \text{nm}^{-3} \approx 39 \text{ nm}^{-3} \quad (6)$$

In the calculation, the detection efficiency of 37% comes from the instrument design, while the volume and the number of ions originate from the TiN unit cell. The experimental ion density after reconstructing the material was $43\pm2$ nm$^{-3}$. As the spatial size of the blocks varies, only the average characteristic length $l_{\text{avg.}}$ of the FDA analysis can be calculated with Eq. 7 below, where $V_{\text{avg.}}$ is the average characteristic volume of the FDA with the block size $n_{\text{block}}$.

$$l_{\text{avg.}} = \sqrt[3]{\frac{V_{\text{avg.}}}{n_{\text{block}} \rho_{\text{ion}}}}$$

With $n_{\text{block}} = 100$ ions, the average characteristic length is $\sim1.33$ nm while $n_{\text{block}} = 50$ ions reduces it to $\sim1.05$ nm. Features smaller than this length will have limited influence on the results of the FDA, as they are averaged.

Using a fixed number of ions instead of a fixed volume is beneficial when comparing the experimentally measured concentration distribution with the expected distribution of a random solid solution of the same average composition. The random solid solution is created by re-distributing the mass-to-charge state ratio of the ions in the same way as previously described for the determination of the voxel size, i.e., keeping the average composition and ion positions the same. In a random solid solution, the concentration distribution follows the binomial distribution. By comparing the experimental distribution with the binomial distribution, it is possible to calculate the risk of mistaking the experimental data for being non-random, i.e. to erroneously discard the null hypothesis that the experimental distribution is random. FDAs were used in Paper 2 to statistically confirm the existence of non-randomness in the Si
distribution. This was interpreted as originating from SiN clustering seen with other techniques.

5.4 Surfaces and Proximity Histograms

Surfaces of constant concentration or density, so called iso-concentration and iso-density surfaces, can be generated from the voxelized data for a particular ion, group of ions, or ratio of ions, at a value specified by the analyst. This surface can then be used to enhance visualization, identify regions of divergent density from local magnification effects, define volumes for further analysis, or to find clusters, grain boundaries and other features with compositional or density differences. In Paper 2, an iso-density surface was used to identify the region in one APT dataset with lower density than the remaining tip.

An iso-concentration surface was also used in Paper 2 as the starting point of another analysis, the proximity histogram. It works similarly to the RDF, but the starting point is a surface, or a set of surfaces, rather than a particular kind of ion. Instead of measuring radially, the proximity histogram show how the concentration changes with distance from the surface in both directions. The error bars are directly related to the number of ions contributing to the measured composition.
6. MAIN RESULTS & CONTRIBUTIONS TO THE FIELD

The main scientific and technical advances of my work within the limits of this thesis are presented in this chapter. The impact of these advances is also discussed.

- I have shown that isotopic substitution can be used in combination with the analytical technique APT in general, and for TiSiN specifically. In the case of TiSiN I show that the isotopic substitution of natN with 15N enables quantitative APT with results in good agreement with ERDA. As is discussed in Paper 1, the method of isotopic substitution can be applied to a wide range of materials systems that are relevant in many different fields of research.

- Certain APT data analysis techniques require that the data must first be sampled into voxels, however, the parameters of such samplings are often arbitrarily chosen and based on the experience of the user. I have designed an algorithm for finding the most meaningful sampling parameter, which balances the compositional uncertainty of small sampling volumes with the averaging of small features when the sampling volume is large. This enables visualization of small-scale compositional fluctuations without relying on the experience of the user.

- I have shown that fluctuations of Si-content exist as domains in TiSiN films on the scale of ~5 nm, which confirm findings of analytical TEM. These domains are segregations, preferentially forming pockets between microstructural, columnar features in the film.

- In Ti_{0.81}Si_{0.19}N films, APT analyses show no strongly developed SiN_y tissue phase surrounding TiN crystallites. Thus, the films should be characterized as nanostructured rather than nanocomposite. This realization could not be made by analytical high-resolution TEM.

- My APT analysis further reveals a surprisingly homogenous distribution of Si in TiN. We conclude in Paper 2 that it is the presence of Si on growing TiN grains that repeatedly gives rise to renucleation and that the Si atoms are thus trapped in what becomes a mesh of TiN grains only a few nanometers large. These TiN crystallites are at an angle to each other because of an extreme density of dislocations, if not disclinations, caused by the incorporated Si and its preference for tetrahedral bonding to N over octahedral bonding, as in cubic-phase TiN. These findings have impact for understanding the superhardening and high thermal stability
of nanostructured TiSiN films because of alloy, defect and small grain size hardening effects. Since the Si is effectively in the lattice of the TiN phase, the thermal stability of the as-grown material becomes high. One can envision that otherwise expected grain coarsening and recrystallization is effectively hindered by local increase in Si concentration which is incompatible with growth of the TiN phase and that the development of SiN$_y$ tissue phases would act as diffusion barriers for growth.


43. Moy C., Ranzi G., Petersen T. C., Ringer S., Macroscopic electrical field distribution and field-induced surface stresses of needle-shaped field emitters, Ultramicroscopy 111 (2011) 397-404, DOI: 10.1016/j.ultramic.2011.01.024.


Paper 1

Resolving Mass Spectral Overlap in Atom Probe Tomography by Isotopic Substitutions – Case of TiSi$^{15}$N

Paper 2

Solid Solution and Segregation Effects in Arc-Deposited Ti$_{1-x}$Si$_x$N Thin Films Resolved on the nanometer scale by $^{15}$N Isotopic Substitution in Atom Probe Tomography
Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-122724