Lukas Etzlinger, BSc

Simulation of Reaction and Diffusion of Defects in Silicon

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Advanced Materials Science

submitted to

Graz University of Technology

Supervisor
Dipl.-Phys. Dr.rer.nat. Karin Zojer

Co-Supervisor
Univ.Prof. Ph.D. Peter Hadley

Institute Of Solid State Physics

Graz, January 2018

EIDESSTATTLICHE ERKLÄRUNG

AFFIDAVIT

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit/Diplomarbeit/Dissertation identisch.

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis/diploma thesis/doctoral dissertation.

Datum / Date

Unterschrift / Signature

Contents

1 Introduction **7**

2 Background **8**

2.1 Silicon . . . 8
2.2 Defects In Silicon . . . 9
2.2.1 Point Defects . . . 9
2.3 Semiconductor Doping . . . 11
2.4 Manufacturing a Silicon Wafer . . . 12
2.4.1 Polycrystalline Silicon . . . 12
2.4.2 The Czochralski Process . . . 13
2.4.3 The Float Zone Method . . . 13
2.4.4 Epitaxial wafers . . . 14
2.5 The Ion Implantation Process . . . 14
2.6 Stopping and Range of Ions in Matter (SRIM) . . . 17

3 The Simulation Setup **18**

3.1 Aims of the Simulation . . . 18
3.2 Restrictions . . . 18
3.3 Assumptions and Simplifications . . . 19

4 Implementation **21**

4.1 Mathematical Background . . . 21
4.1.1 The Finite Difference Scheme . . . 25
4.1.2 The Thomas Algorithm . . . 27
4.2 The Simulation Structure . . . 30
4.2.1 The Reaction and Dissociation Matrix . . . 30
4.3 Code Optimization and Increasing Performance . . . 33
4.3.1 The Adaptive Time Step . . . 33
4.3.2 Evaluation of Diffusion, Reaction and Dissociation . . . 34
4.4 Parallelization . . . 34

5 Results **35**

5.1 General Overview . . . 35
5.2 Influence of Process Temperature . . . 43
5.3 Influence of Implantation Current . . . 49
5.4 Influence of Initial Carbon Concentration . . . 56
5.5 Influence of Initial Oxygen Concentration . . . 59
5.6 Simulation Speed and Specifications . . . 59

6 Problems and Issues **67**

7 Summary and Outlook **69**

Abstract

A code has been implemented that monitors the diffusion, reaction and dissociation of 86 different defects in a silicon wafer during the process of implanting hydrogen protons into the sample. The main aim is to extract concentration profiles along the depth of the wafer of all the present defects. A secondary aim is to make the simulation fast enough to perform implantation steps with a higher workload, i.e., subsequent annealing steps. Process parameters like the process temperature, the implantation current and the initial concentrations of the impurities carbon and oxygen have been varied in several simulations to study their influences on the sample. The mathematical-numerical problem at hand is a system of coupled time-dependent equations that consider diffusion, dissociation, and reactions of all defects. To treat the evolution in time, the implicit Euler scheme was employed. The resulting equations were then partially solved by using the Thomas algorithm. The entire algorithm was parallelized.

Due to the amount of different partially unknown parameters and the complexity of the problem it is hard to predict quantitatively accurate profiles. However, overall trends, in particular with respect to groups of defects, can be extracted. General trends show, that increasing the process temperature leads to a promotion of more complex defects at the expense of smaller, basic defects. An increase in implantation current reduces the duration of the process proportionally. This impedes the formation of defects, that need a longer time to build up a significant concentration and favours smaller defects, that are often present from the beginning in vast amounts. A higher initial concentration of carbon tendentially favours the formation of small carbon-related defects as well as compound carbon defects ($\mathrm{C_iC_S}$), but impedes the formation of carbon defects with additional impurities. Increasing the initial concentration of oxygen leads to an increase in concentration of $\mathrm{V_xO_y}$- and $\mathrm{I_xO_y}$-complexes, while the concentrations of oxygen-related defects with impurities are disadvantaged. An increase in simulation speed could be achieved. However computationally more intense simulations are still not viable for the industry.

Next steps in this field will include the removal of physical simplifications to get more accurate results as well as usage of more computational resources combined with a more efficient code to further increase the simulation speed.

Kurzzusammenfassung

Es wurde ein Programmiercode implementiert, der die Diffusion, Reaktion und Dissoziation von 86 verschiedenen Defekten in einem Siliziumwafer während der Protonenimplantation von Wasserstoff überwacht. Das Hauptziel ist es, Konzentrationsprofile aller Defekte entlang der Tiefe des Wafers zu ermitteln. Ein weiters Ziel besteht in der Beschleunigung der Simulation, um sowohl aufwändigere Protonenimplantationen als auch Annealing-Prozesse in angemessener Zeit simulieren zu können. Die Prozessparameter, wie zum Beispiel Prozesstemperatur, Implantationsstrom und anfängliche Konzentrationen von Kohlenstoff und Sauerstoff wurden in verschiedenen Simulationen variiert um deren Einfluss auf die Probe analysieren zu können. Das mathematisch-numerische Problem ist ein System von gekoppelten, zeitabhängigen Gleichungen, die die Diffusion, die Dissociation, und die Reaktion von allen Defekten berücksichtigt. Um diese Gleichungen zu behandeln wurde die implizite Euler-Mehode angewandt. Die daraus resultierenden Gleichungen wurden teilweise mit dem Thomas-Algorithmus gelöst. Der gesamte Algorithmus wurde mit parallelisierter Programmierung durchlaufen.
Durch die unterschiedlichen, teilweise unbekannten Parameter und die Komplexität des Problems ist es nicht möglich quantitativ korrekte Profile vorrauszusagen. Es können jedoch grundsätzliche Trends, besonders im Hinblick auf Defektgruppen, extrahiert werden. Allgemeine Trends zeigen, dass die Erhöhung der Temperatur die Formation von komplexeren Defekten auf Kosten von simpleren und kleineren Defekten begünstigt. Eine Erhöhung des Implantationsstroms führt zu einer proportionalen Reduktion der Dauer des Implantationsrozesses. Dies behindert die Entstehung von Defekten, die eine längere Zeit brauchen um signifikante Konzentrationan aufzubauen, und begünstigt deshalb kleinere Defekte, die sich von Beginn an in großer Zahl im Wafer befinden. Erhöht man die anfängliche Konzentration von Kohlenstoff im Wafer, bilden sich tendentiell mehr kleine Kohlenstoffdefekte aus, während die Bildung von Kohlenstoffdefekten mit zusätzlichen Verunreinigungen eingedämmt wird. Eine Erhöhung des Sauerstoffgehalts im Wafer führt zu einer vermehrten Bildung von V_xO_y- und I_xO_y-Komplexen, auf Kosten der Konzentrationen von sauerstoffhältigen Defekten mit weiteren Verunreinigungen. Die Simulationsgeschwindigkeit konnte zwar erhöht werden, die Simulation von rechenintensiveren Protonenimplantationen und Annealings ist jedoch noch nicht im zeitlich brauchbaren Rahmen.
Die nächsten Schritte in diesem Gebiet inkludieren das Entfernen von physikalischen Vereinfachungen, um genauere Resultate zu erhalten, und die Verwendung größerer Rechenleistung in Kombination mit einem effizienteren Simulationscode, um die Simulationsgeschwindigkeit weiter zu verbessern.

Acknowledgement

First of all I would like to thank my supervisors Dr. Karin Zojer and Prof. Peter Hadley for their support during this thesis and for taking their time to discuss emerging problems and answer any of my questions, be it about the simulations or the writing process.
I want to thank Martin Faccinelli for enduring numerous long discussions about the topic, for his smart ideas and not to forget for his relentless optimism. A big thank you also goes to my colleague Julia Hassler, for sharing ideas and opinions and for making the time at the university much nicer. Moreover I would like to thank Markus Krammer, Anton Fernandez, Georgii Krikun and Natalia Bedoya for furthe discussions and inputs to this thesis.
I am very grateful to my parents Daniela and Gerald as well my sister Maren for supporting me my entire life and especially so during my studies and for being there for me.
Finally, I want to thank Sophie Adlgasser for all her love, support and also her patience, especially in the last couple of months during writing this thesis.

This work has been performed in the project EPPL, co-funded by grants from Austria, Germany, The Netherlands, Italy, France, Portugal- ENIAC member States and the ENIAC Joint Undertaking.

1 Introduction

In proton implantation, hydrogen atoms are implanted into silicon. Depending on their energy, they travel a certain distance in the host material, before they have no energy left and reside at a certain depth in the wafer. On their way, they knock silicon atoms out of place, thus forming vacant sites (vacancies) and interstitial silicon atoms (interstitials). These are three of the seven point defects, that will be present in the sample. The remaining four are carbon and oxygen (from the manufacturing process) and boron and/or phosphorus from the semiconductor doping. These seven point defects may diffuse in the sample, react with each other and dissociate again. Upon reaction, they form defect complexes, compounds of single point defects, that have different properties than their beginning state (e.g. ability to diffuse).

The aim of this work is to simulate the diffusion, reaction and dissociation processes in order to extract a concentration profile of all the defects along the depth of the wafer. Knowing where to find which defect in the sample leads to a better understanding of the proton implantation process. From this, electronic devices can be created in a much more purposeful way, because even though the process of proton implantation is already employed in semiconductor device manufacturing, the defect complexes and their behaviour are still not fully understood.

One of the crucial factors of this simulation is its speed. Because of the many defect complexes that have to be considered, the computational effort is very high and it takes a long time to obtain results. The simulation time is strongly dependent on simulation parameters (e.g. temperature), and more often than not a slight change in one of them has the consequence of a drastic increase of duration. So, not only is it the aim of this work to properly simulate the given model, but to make the simulation work as fast as possible, so it becomes usable for predicting the outcome of an implantation process with certain parameters.

The simulation utilizes the finite difference scheme, more specifically the implicit Euler method. Compared to the Crank-Nicolson method it offers higher speed in exchange for less precision. This loss of precision is negligible compared to the many assumptions and simplifications we had to make in order to get the simulation running at a satisfying speed in the first place. Among these assumptions for instance is the fact, that many of the characteristic values of the defect complexes (such as binding and dissociation energies) are guessed, because for many defects values cannot be found in the literature so far.

2 Background

This section gives an introduction to important background topics for this thesis. As it deals with the formation and motion of defects in a silicon wafer, it is important to know about silicon and its properties. Further, different approaches of making a silicon wafer are presented. Another section explains the basic defects found in the silicon crystal structure, and goes more into detail about defect complexes. Next, the process of proton implantation and important parameters thereof will be discussed. Lastly, this section gives a brief explanation of SRIM simulations which were also used in this thesis.

2.1 Silicon

About 90% of the earth's crust is composed of silicates, which makes silicon the second most abundant material on earth after oxygen [1]. For most applications silicon is not or only barely purified. Silicates are often used in construction as a part of concrete and for a wide variety of glasses and ceramics. Elemental silicon is often used in metallurgical processes involving steel and aluminium. Only a small part of highly purified silicon is used for today's modern electronics as integrated circuits, where silicon is still the most dominant and important material.
In the periodic table silicon is a group IV element. Similar to other elements from that group (e.g. carbon) it has 4 electrons in its outer shell. This gives silicon the electron configuration $[\mathrm{Ne}]3\mathrm{s}^2\mathrm{p}^2$, with two unpaired electrons in the 3p-shell. Just like in carbon, the outer s-orbitals and the p-orbitals combine to form four energetically equivalent orbitals, resulting in four possible covalent bonds (sp^3-hybridization) in the solid. In standard conditions it exhibits a face-centred diamond-cubic structure, with two atoms in the base sitting at positions (0, 0, 0) and $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ as depicted in figure 1. Its lattice constant is 543 nm which amounts to a concentration of silicon atoms of about 1.5×10^{28} m^{-3} [2] [3].

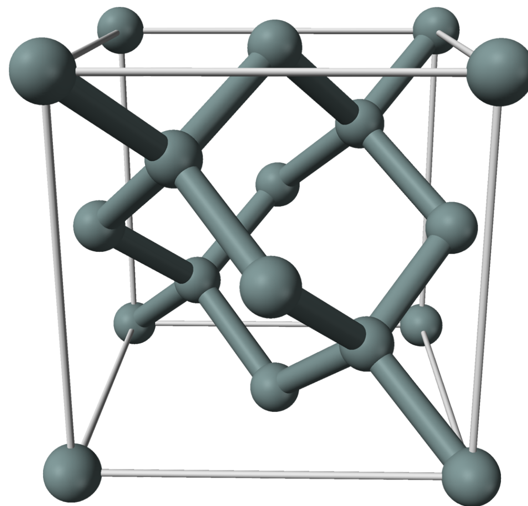


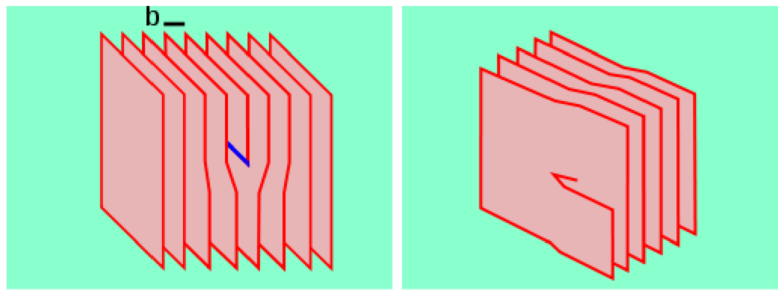
Figure 1: Unit Cell of silicon. Picture retrieved from `https://upload.wikimedia.org/wikipedia/commons/f/f1/Silicon-unit-cell-3D-balls.png` on 17. November 2017.

2.2 Defects In Silicon

This section will give an overview over the different types of defects in silicon crystals and how they emerge, as well as how they influence the material they are embedded in. Most of the information in this section is taken from the books "Materialwissenschaft und Werkstofftechnik: Physikalische Grundlagen" by Günther Gottstein [4] and "Intrinsic Point Defects, Impurities, and Their Diffusion in Silicon" by Peter Pichler [5].
The structure of silicon discussed in the previous section refers to a perfect crystal structure. This is an ideal case and never happens in reality. The crystal lattice always contains defects which can be of varying type and size. The concentration of these defects varies over time, as they can move around and form new defects or more complex clusters of defects. They alter the properties of the material, so some of them are introduced intentionally. Depending on the kind and number of defects introduced one can modify various properties. However these modifications can also have negative effects on the device and its performance.
The most popular way of distinguishing between the defect types is by their expansion, or dimensionality. Point defects technically have no expansion in space and thus are zero-dimensional. These are the most important defects in this work, because they are able to diffuse around in the host lattice and combine with other point defects to form bigger defect complexes. There is, however, no tangible limit, as to when such a cluster of zero-dimensional point defects has an expansion in space. The defects and defect complexes that are featured in the simulations are always treated like point defects, regardless of their actual size. They are more extensively explained in section 2.2.1. Dislocations are one dimensional crystal defects and occur in one of two variations (or a combination of both which is most common in real materials). Both types are also depicted in figure 2. To better explain these defects, it is convenient to imagine the atoms in the crystal as planes which are stacked on top of each other. The first type of one-dimensional defect is a screw dislocation. It is formed by "cutting" the crystalline material along a plane halfway through and then moving one of the halves over by one lattice constant. The lattice planes now roughly resemble the shape of a very flat spiral staircase. The Burgers vector is often used to describe the direction and the magnitude of the dislocation. In case of the screw dislocation, the Burgers vector is parallel to the dislocation line (which, in analogy, would be the central "pillar" of the spiral staircase). Edge dislocations can be described by putting an extra half plane midway into the crystal lattice. The end of the half plane is now the line dislocation, but the Burgers vector points in the direction perpendicular to the introduced half plane, which is also the characteristic of the edge dislocation.
Examples of two-dimensional defects are grain boundaries, phase boundaries and stacking faults. The latter is a mismatch in the sequence of the lattice planes. Grain boundaries are the interfaces between two crystallites that are oriented differently, while phase boundaries mark the interface between two different phases. Lastly there can be voids or precipitates of other materials in the crystal structure, which represent the three-dimensional defects.

2.2.1 Point Defects

Point defects can be subcategorised into intrinsic and extrinsic point defects. Intrinsic point defects consist of only the host material (or its absence). If an atom is missing from a regular lattice site it is called a "vacancy", often just denoted as "V". They can be formed when the crystal grows from the



(a) An edge dislocation. Picture retrieved from `https://upload.wikimedia.org/wikipedia/commons/5/52/Dislocation_screw_e.svg` on 17. November 2017.

(b) A screw dislocation. Picture retrieved from `https://de.wikipedia.org/wiki/Versetzung_(Materialwissenschaft)#/media/File:Dislocation_edge_d2.svg` on 17. November 2017.

Figure 2: Schematic drawings of edge and screw dislocation. For the edge dislocation, the Burgers vector is perpendicular to the dislocation line. The Burgers vector in a screw dislocation is parallel to the dislocation line.

melt because of lattice vibrations (phonons), or by a local arrangement of the atoms. A common model to describe this local arrangement is the movement of silicon atoms near the surface to the surface. It leaves behind a vacancy, which is moved further to the inside of the crystal, as other neighbouring silicon atoms try to fill in the gap, thus reducing the total energy. Vacancies may also combine with each other to form new defects or even bigger defect complexes that contain several vacancies (e.g. divacancies).

The presented vacancy sources however are not of much importance in this thesis, as it deals with already processed silicon wafers, where the crystal growth is long over. Another source of vacancies, the most important one for this work, is bombardment by ions. By shooting high-energy ions into the crystal, the ions collide with the lattice atoms. If their energy is higher than the bonding energy between the silicon atoms, they are knocked out of their place, generating vacancies, as well as self-interstitials (the second of the intrinsic point defects). One ion normally has enough energy, to generate multiple vacancy-interstitial pairs, before it settles in the lattice.

As mentioned, self-interstitials are also point defects. They are silicon atoms displaced of their normal position and sit between the lattice atoms on an interstitial lattice site, hence their name. They are often abbreviated with "I". It seems obvious, that upon generating a vacancy, the atom we cut out has somewhere to go. In the case of ion bombardment it is often the case, that the atoms reside on an interstitial place. This combination of vacancies and interstitials is called a Frenkel-pair or Frenkel-defect.

Interstitials do not reside anywhere in the lattice, but rather have some defined locations where they can be placed. Every crystal lattice has voids between the regular lattice atoms. Depending on the kind of lattice as well as the kind of atoms that it is composed of, the voids can be of different number and size. Silicon has a face-centred diamond-cubic structure, which is fairly open. In its unit cell, there are 5 possibilities for an interstitial atom to rest. In order to make some room for vacancies and interstitials, the crystal lattice has to deform around these point defects. Thus it is often the case, that 2 silicon atoms share one regular lattice spot (but neither of them occupies the spot).

Another kind of point defect are extrinsic point defects. As the name suggests, they are formed when atoms are involved that differ from the host material. These defects play a major role in the

semiconductor industry, as they are often deliberately introduced in order to change the electronic properties of the material. Elements like phosphorus or boron are added in exact amounts to achieve a desired functionality of the material. This process is called doping and will be further explained in section 2.3.
Just like self-interstitials, impurity atoms may occupy an interstitial lattice site. Depending on their size, they will fit more or less easily into the designated space, again giving rise to a stress in the crystal lattice around that defect. Another way of placing foreign atoms into the lattice is to put them in a vacancy (or rather: on a regular lattice site). This is a substitutional impurity, as it substitutes an atom from the host material. To know, whether we are talking about an interstitial or a substitutional defect, it is customary to take the element symbol of the impurity and add a subscript "S" if it is on a substitutional place, or a subscript "I", if it is on an interstitial place (e.g. P_S is substitutional phosphorus and P_I is interstitial phosphorus).
Again, these extrinsic point defects can combine either with each other or with intrinsic point defects to form new and bigger defects and defect complexes. This thesis deals with seven basic "point defects" that have only one constituent. Hydrogen, silicon vacancies and silicon interstitials are present due to the implantation process (explained in section 2.5), thus, their amount depends on the implantation parameters. Also, these are the only defects that are continuously added during the implantation step, because the hydrogen is added gradually, defined by the implantation current and the implantation dose. More about the implantation process can be found in section 2.5. Oxygen and carbon are contaminations from the wafer manufacturing process. Depending on which manufacturing process was chosen (more on the different methods in sections 2.4), their concentrations may vary. Their concentration has to be constant, as no more oxygen or carbon is able to enter or leave the sample (apart from minor out-diffusion from the sample, which is neglected in the simulation). Boron and phosphorus are intentionally added as dopants in precise amounts during wafer preparation. Just like oxygen and carbon, their total amount hast to stay constant, as no more is added later and they cannot leave the sample. These seven point defects will react with each other, to form defects and defect complexes. Every formed defect is treated as a defect with its own properties. These properties include values for diffusion (diffusion constant and activation energy), reactions with each other (information about which defect reacts with which and a certain radius they have to be in) and dissociation (attempt frequency and binding energy). In total, this simulation monitors 86 defects (the seven basic point defects included) with up to eight constituents. It is important to note, that not all of these defects are able to diffuse, or dissociate. Especially larger defect complexes are immobile and cannot move through the lattice due to their size.

2.3 Semiconductor Doping

Silicon is classified as a semiconducting material, with an electric conductivity σ of about 4×10^{-4} S/m [6]. As a comparison, metals have electric conductivities of $> 10^6$, while insulators show values of typically $< 10^{-8}$ [7]. Opposite to metals, (where valence and conduction and overlap and the Fermi energy lies in either one of them), semiconducting materials have an energy gap between the bands. In the case of silicon, this gap amounts to about 1.11 eV [3]. The position of the Fermi energy and thus the electrical properties of the material can be somewhat adjusted, by introducing foreign atoms to the silicon crystal lattice.
As already mentioned, silicon shows sp^3-hybridization and thus has 4 valence electrons. These valence

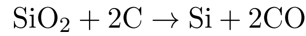
electrons form covalent bonds with neighbouring valence electrons of the other atoms and thus form the silicon crystal. One way of doping silicon, is exchanging a small number (0,01 up to 1 ppm) of silicon atoms with elements that have a different number of valence electrons. This process is called doping.
Elements from group V in the periodic table have an additional valence electron compared to silicon. A popular candidate for doping from group V is phosphorus. If phosphorus atoms take places of silicon atoms in the crystal lattice, they will form the same covalent bond. However, phosphorus atoms have an additional electron that does not participate in the bonding. It can move freely through the sample and has the ability to carry current, thus increasing the electric conductivity of the material. Since the phosphorus atoms surrender their electron, they are called (electron) donors. The phosphorus atoms themselves are positively charged. The process of adding donors into the semiconducting sample is called n-type-doping. The Fermi energy is raised and lies just below the valence band.
P-type doping works the other way around. By introducing elements that have only three valence electrons, there will be a constant shortage of electrons in the sample, as not all the covalent bonds are satisfied. Elements are typically taken from group III, with the most popular candidate being boron. In semiconductor physics, the absence of an electron is often called "hole" and is treated just like an electron, with the opposite charge (+e instead of -e, where e is the elementary charge). Electrons can now move around and carry current, due to a shortage of electrons introduced by the boron atoms. Inversely, one could also treat this as a movement of holes that carry current, just in the opposite direction of electrons. Because the boron atoms essentially take away electrons from the sample, they are called electron acceptors, or just acceptors. The boron atoms will be negatively charged, due to the additional electrons that partake in the covalent bonds. This p-type doping also moves the Fermi energy, but this time down towards the conduction band. By combining p- and n-type materials as well as varying their concentrations it is possible to manufacture many different devices which are the fundamentals of modern electronics, like diodes, transistors or solar cells.

2.4 Manufacturing a Silicon Wafer

Silicon wafers represent the substrate, the base material of a semiconducting device. They are discs with a diameter ranging between 50.8 mm (2 inch) and 300 mm (11.8 inch) and a thickness between 275 mm and 775 mm. The diameter of wafers has steadily increased over the years, as this increases throughput and effectiveness in the manufacturing process by reducing silicon waste. Therefore, wafers with a diameter of 450 mm have already been proposed and made, but are not yet used to mass-produce semiconducting devices. The discs are cut from a highly pure, almost perfect single-crystalline silicon ingot. Depending on the process used to manufacture the ingot, different impurities are present in different concentrations.

2.4.1 Polycrystalline Silicon

Silicon is the second most abundant element in the earth's crust. It is however only present in impure and oxidised forms. To achieve what is called metallurgical grade silicon (elemental silicon with a purity of at least 98%) silicon dioxide (SiO_2) is reacted with carbon in an electric arc furnace at about 1800 °C [1] according to the following reaction [8]:



$$SiO_2 + 2C \rightarrow Si + 2CO$$

Metallurgical grade silicon is mainly used in the manufacture of steel. Only a small fraction is used in the Siemens process to produce electronic grade polycrystalline silicone, with a purity between 9N-11N (which stands for ”nine nines” and ”eleven nines”, so a concentration of 99.9999999% and 99.999999999% respectively). The Siemens process is a chemical vapour deposition method. Trichlorosilane gas ($HSiCl_3$) is blown over monocrystalline silicon rods at a temperature of about 1150 °C. Silicon crystallites are grown directly on the surface of the pure silicon rods in the chamber. Polycrystalline silicon is then either used as a conducting gate material in MOSFETs, or it is further processed to achieve monocrystalline silicon, which is needed for wafer production.

2.4.2 The Czochralski Process

The highly pure polycrystalline silicon is put in a crucible usually made of quartz (SiO_2) and melted at about 1420 °C. To prevent contamination of the silicon, the process is normally performed under a protective and inert atmosphere such as argon. From above the crucible a silicon seed crystal is lowered until it barely touches the melt. The seed crystal is then slowly rotated in one direction, the crucible in the other. Upon slow retraction of the seed (while maintaining rotation) the liquid silicon will cool and solidify, continuing the perfect crystal structure given by the seed [1]. The rate of retraction is very slow to avoid the development of defects. In order to make n-type or p-type semiconductors, impurity dopants such as phosphorus or boron are added to the melt in precise amounts. Because of the high temperatures, the walls of the crucibles will be dissolving into the melt during the growth process of the ingot, giving rise to a contamination with oxygen. The concentration of oxygen in the sample is typically in the range of 10^{23} m^{-3} [9]. It is not always unwanted to have oxygen in the sample, as it can have positive benefits, such as trapping unwanted impurities. Another prominent impurity in Czochralski grown single crystals is carbon. It originates from contaminations of the crucible and has concentrations of about 10^{22} m^{-3} [10]. The resulting ingots have diameters of up to 300 mm and lengths of up to several meters.

2.4.3 The Float Zone Method

The float zone method (or more generally zone melting) includes melting a polycrystalline silicon ingot with heating coils, where a high frequency alternating current passes through. The process is executed either in vacuum or an inert gas that may contain small concentration of a dopant. Slowly, the coil moves down the ingot, melting the silicon where the coil surrounds it. That way, small regions of the material are molten, while other parts are solid. This is important, because impurities will diffuse to a liquid region if confronted with a solid/liquid interface. Volatile impurities such as silicon monoxide are gassed out during the process. By sweeping the coil towards the end of the ingot, most of the impurities will be found at the end of the ingot. This end is cut off and the process is repeated if a higher purity is needed. The oxygen and carbon concentrations of Fz-Si is typically lower than that of Cz-Si, because of the absence of a crucible. Typical values are $< 10^{22}$ m^{-3} for oxygen and $< 5 \times 10^{21}$ m^{-3} for carbon [11]. Dopants are added to the melt before the process starts. Since only a small portion of the ingot is molten, Fz-Si is prone to deviations in doping concentrations, which leads

to variations in some physical properties such as resistivity. A further disadvantage is the restriction of the ingot size. Due to the surface tension of the molten material, silicon ingots cannot have a diameter greater than 150 mm in order to be processed [1].

2.4.4 Epitaxial wafers

Epitaxial wafers are wafers where one or more layers of crystalline material is grown epitaxially on the wafer surface, to generate a layered semiconductor with desired properties. The base material is usually a polished Cz-wafer. On top of that, layers are grown which exhibit a better quality than the underlying Cz-wafer, resulting in much higher purity and low concentrations of unwanted impurities such as oxyagen and carbon. [11]. One of the possibilities to grow epitaxial layers is to subject the monocrystalline wafers to a flow of silicon tetrachloride ($SiCl_4$), hydrogen and a suited dopant [1]. This is done at temperatures below the melting temperature. The crystal structure of the epitaxial layer is determined by the crystal structure of the underlying substrate. In case of homoepitaxy which is used for the epitaxial silicon wafers, the epitaxial layer and the substrate are of the same material. The crystal structure can be continued without any internal stresses or defects. This approach is used to fabricate wafers with a smoother surface (which is also more pure) and to generate layers which exhibit different doping levels. In heteroepitaxy the epitaxial layer and the substrate are different with regard to crystal structure. With heteroepitaxy it is possible to generate thin crystalline layers of materials, that otherwise do not exhibit such a crystal structure.

2.5 The Ion Implantation Process

In ion implantation ions are extracted from a source and accelerated towards the target material with the aim of implanting those ions into the target. That way impurity atoms are introduced which serve as doping for the semiconductor. The ion source will generate the kind ions needed for the process. The ions are then extracted from the source and accelerated electrostatically until they reach the implantation energy. This energy varies from a few 10 keV to a few MeV, depending on application and desired outcome. The ions in the beam are then separated by their energy. At the end of this energy separation, a slit lets ions of a certain energy pass, so that the beam is then constituted of monoenergetic ions. After this, a magnetic deflection system is in place to divert the beam to a certain location on the target. The ion beam is focussed on one point on the sample for a specific time (the so called dwelling time), before it is being deflected to the next point. By doing so, a small area on the target is illuminated. The dwelling time is determined by the scanning frequency. Its inverse, the scanning time, tells how long it takes the beam to sweep through all the points and return to its original position. A schematic of the whole process is shown in figure 3.

The ions penetrating the sample interact with the atoms and electrons of the target material and lose energy in the process. The interactions can be separated into electronic deceleration and nuclear deceleration. The former is the dominant contribution, if the ion has a low atomic number and/or a high energy. This is the case for this thesis, as it deals with the implantation of ionised hydrogen atoms into monocrystalline silicon. It can be assumed, that the energy loss of the ions due to interactions with the electrons happens continuously because of the high density of electrons in silicon as well as

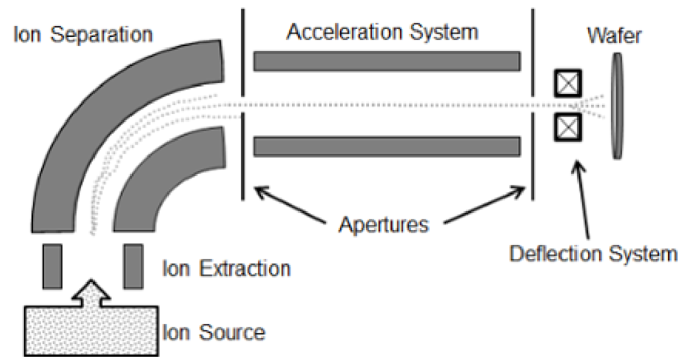


Figure 3: Schematic of the ion implantation process.

the high range of the Coulomb interaction between ion and electron. The electronic deceleration can again be separated into several cases. (i) The incoming ions may interact with free electrons in the substrate, which represents an elastic collision accompanied with an energy loss of the ion. (ii) Another option is the collision between ion and a bound shell electron of the atom. In this case the atom may be excited, if the ion has sufficient energy. Since this excitation is based on an interaction between ion and electron, it is an electronic deceleration, even if it is the atom that is being excited. The atom is ionised, if the energy of the ion exceeds the atom's ionization energy. In case of a semiconductor this leads to the generation of an electron-hole pair. (iii) Lastly, the ion itself might be further ionised by its own kinetic energy within the first few atomic layers after the initial impact. Upon deceleration inside the material, electrons are collected again [12].

In nuclear deceleration the ions show Coulomb interaction with the atoms of the crystal lattice. During the interaction, energy is transferred from the ion to the lattice atom. How much energy is transferred depends on the masses of the incoming ion and the lattice atom. If the transferred energy exceeds the Frenkel energy E_F, a Frenkel defect is formed. E_F is constituted of the binding energy of the lattice atoms, which has to be overcome to separate them and a small displacement energy, which is necessary to dislocate the silicon atoms and move it to an interstitial lattice site. If the energy is lower than that, the ions will generate phonons instead. The energy of the ions decreases with each one of these interactions. Eventually their energy is reduced to zero, at which point they will reside at a certain depth in the material, called the implantation depth or the projected range. This projected range depends on the implantation energy of the ions. A higher energy translates to a greater penetration depth, but also more damage in the crystal lattice. In materials with a crystal structure an effect called "channeling" may occur. Since atoms in a crystalline material are ordered (apart from the crystal defects), there are certain directions or "columns" in which no atom can be found. If an ion enters such a column, it experiences far less interactions with nuclei and electrons. The deceleration is much slower and the projected range is increased [12].

The simulation presented in this thesis uses ionised hydrogen atoms as the ions. The target material is a silicon wafer. Since hydrogen atoms exhibit a low mass, they are able to penetrate deeper into the substrate. Also, the spread along the beam is much narrower, which allows for a more precise "placement" of the hydrogen atoms. Figure 4 shows the distribution of hydrogen atoms, vacancies and interstitials along the depth of a silicon wafer. Hydrogen shows a distinct peak, which is the projected range, where most of the hydrogen atoms are embedded. The concentration of vacancies and interstitial atoms peak here as well. This means that most of the defects are introduced at the projected range.

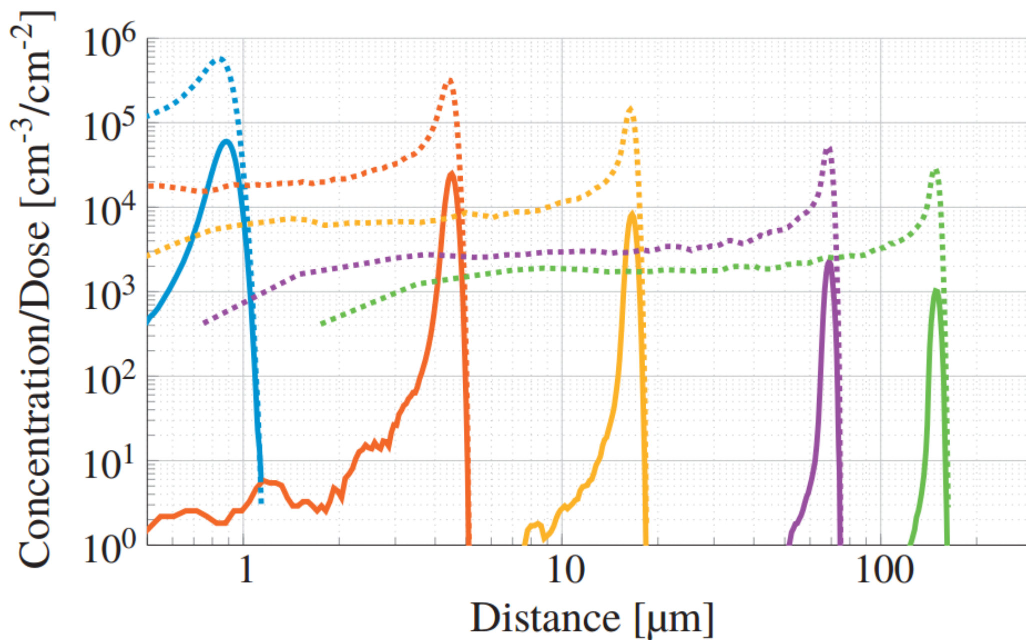


Figure 4: Implantation profiles for different implantation energies (100 keV, 400 keV, 1 MeV, 2.5 MeV, 4 MeV, from left to right) simulated with SRIM. The profiles are normalized to the proton dose. Dotted lines represent the concentrations vacancies and interstitial atoms, solid lines represent the concentration of hydrogen. Image taken from [13].

Implanting protons into silicon has different purposes based on the radiation dose used. Low implantation doses (up to about 10^{19} m^{-2}) are used to change the charge carrier lifetime in silicon. The implanted hydrogen atoms cause defects in the silicon substrate (most prominently divacancies and vacancy-oxygen complexes). These introduce energy levels in the band gap of silicon that ease the recombination of the charge carriers. The proton implantation process has benefits over other options like diffusing platinum into silicon. The implantation depth of the protons can be controlled by changing the implantation energy accordingly (figure 4). This allows for a precise positioning of regions with a reduced charge carrier lifetime. At the same time one needs significantly less energy to achieve this, since protons have the highest penetration depth for a certain energy of all available ions. Moreover, the damage done to substrate regions that should not be affected is low in comparison [12]. If medium doses are used, hydrogen related donors are formed which lead to an n-type doping of the substrate. Again, proton implantation is a viable alternative to other doping processes due to the aforementioned high penetration depth. If the implantation dose exceeds about 10^{22} m^{-2} the concentration of hydrogen and lattice defects near the projected range rises. This gives rise to plate-like crystal defects that are oriented parallel to the surface. By applying an annealing step at several hundred Kelvin after the implantation these plate-like defects grow. Ultimately it is possible that the layer between surface and projected range is separated from the bulk material due to extension of the plate-like defects. This effect is utilized in the Smart-Cut process which is used to generate thin slices of silicon [12].

The implantation current (or the related current density) specifies how many protons are implanted per area and time. Increasing the implantation current has two main effects. Firstly the amount of protons per area is increased, which leads to more damage to the irradiated region. It also gives rise to higher concentration of defects that involve hydrogen, as it is present in much higher amounts. Secondly, the time it takes for a certain implantation dose to be implanted will be lowered, as more protons are implanted in the same time. This impedes diffusion to a certain degree, as there is less

time for defects to diffuse. It is this reason why bigger defect complexes have decreased concentration, or maybe do not form at all.
It is also possible to vary the process temperature of the implantation. The simulations were done at temperature ranging from $50 - 150$ °C. The temperature greatly affects then diffusion, reaction and dissociation of the defects.

2.6 Stopping and Range of Ions in Matter (SRIM)

Stopping and Range of Ions in Matter (SRIM) are computer programs that calculate the interaction of ions when shot into one or several layers of a target material. The programs are based on the binary collision approximation (BCA). Here the ions travel through the target material and experience collisions with the lattice atoms. Between the collisions it is assumed that the ions travel on a straight path and are only influenced by electronic stopping power. A collision is dealt with by solving the scattering integral between the two colliding particles. The result yields the energy loss of the ion as well as the scattering angle. From that it is possible to calculate the new energy of the particle, which is then used as a basis for the next collision. This method itself is only able to "track" the ions and their path, but not the recoils they produce. Interstitial atoms generated by the collisions can have sufficient energy to damage the crystal lattice as well and generate more crystal defects (collision cascade). To account for this case and get more exact results it is necessary to also calculate the time integral of the collision, which yields the duration of the collision [14] [15].
The input parameters for the programs are the type of ion and its initial energy (implantation energy) as well as the material of the target, which of course implies its physical properties, such as density. The most interesting information for this thesis is about the distribution of the ions in the sample in all three dimensions (e.g. penetration depth) and the concentration of vacancies considering collision cascades. The output of the SRIM simulations were provided in lists, containing the concentrations of hydrogen and vacancies depending in the depth in the target material.

3 The Simulation Setup

Before going into (mathematical) details of the simulation, it is necessary to explain the general framework of the simulation. This section covers the main motivation of the simulation, and what the aims and expected outcomes are. It is also necessary to make a few assumptions and simplifications, as otherwise the simulation will become too complex and the computation time would be far too long. It is important to find a good compromise of simplification and complexity. Some of the simplifications are made to decrease computation and/or implementation time. Those that decrease implementation time usually do not impact the result of the simulation very much, as otherwise it would be important and necessary to make an effort to implement the complex way. Some simplifications are however made because the complex way is not yet accessible, due to missing parameters and values. Apart from simplifications there are also some restrictions that need to be considered.

3.1 Aims of the Simulation

It is the aim of this work to gather information about the distribution of the different defects in the silicon wafer, depending on the initial parameters. Those parameters are settings of the implantation process and the properties of the wafer, i.e. its contaminations from the manufacturing process as explained in section 2.4. The output of the simulation are graphs that depict the defect concentration versus the depth of the silicon wafer. To be more precise, the simulation needs to be able to properly simulate the implantation process with all its parameters. This is achieved with the help of files obtained from previously mentioned SRIM-simulations. Simultaneously the simulation is required to compute how the different defects diffuse inside the material, dissociate and react with each other.
The second objective was to reduce computation time of the simulation to a minimum. The simulation has already been written using programming language Python, which is not particularly suited for the task at hand. Simulations with medium workload had a runtime of up to several months, depending on the initial parameters, which is not feasible for the industry. Thus, efforts were made to rewrite the existing code to another, more suited programming language. The simulation for this thesis was written in FORTRAN 95. Computation time was further decreased by decreasing the amount of unnecessary mathematical operations as well as employing parallel computing.

3.2 Restrictions

Some physical restrictions were applied when modelling the silicon wafer to describe it numerically. First of all it has to be mentioned that the wafer is not modelled in three spatial dimensions, but only one, namely its depth. It is assumed that the small deviations in concentration caused by the sweeping motion of the proton beam can be neglected compared to other existing deviations originating from other simplifications and assumptions. The result of the simulations are graphs that depict the defect concentration of one ore more defects versus the depth of the wafer. The second restriction which plays a key role in the simulation is the fact that at no point in space or time negative values for the defect concentration may appear, as they do not make physical sense. The reason why this is important lies in the methodological errors induced by numerical methods used to evaluate the concentration values. These methods never yield an exact solution but only approximate the solution. The quality of the approximation depends on the problem at hand and the numerical parameters

chosen by the experimentalist. Due to these errors it is almost unavoidable that at some point during the simulation negative concentration values appear. Section 4.3 explains how they emerge and how they are dealt with in this simulation. Lastly the simulation does not take into account defects leaving the wafer by diffusion anywhere on the sample. The impact of this contribution is very small, while the computational effort to include it as well as to monitor the defects that way is rather high which is why it was neglected.

3.3 Assumptions and Simplifications

One of the (temporary) simplifications include the use of an effective diffusion for every defect. Usually, every defect can be found in several charges states. These charged states have slightly different properties (e.g. diffusion constants) and thus behave differently in the silicon wafer. However it is difficult to obtain knowledge of how these charged states are distributed, which is why the effective diffusion was introduced, which acts as an average value over all the charged states of a defect. To acquire the distribution of the defect states depending on the current chemical potential it would be necessary to perform DFT calculations for all the separate defects. A problem with introducing charged defect states however is the performance of the simulation. While the results would be much more accurate, the computational effort would rise greatly, as one effectively has to deal with many more defects than before (considering the charged states of a defect as a single defect).
The chemical potential of the system is set to be constant during the whole implantation process and assumed to be centered in the band gap of silicon. It is however never used in any of the calculations performed. If the chemical potential would be introduced, one could determine the distribution of the charged defect states (presuming the corresponding data from DFT calculations were available), which in turn would remove the above mentioned simplification of the effective diffusion. Also, the introduction of the charged defect states would subject the chemical potential to a constant change in its value. This means that the value of the chemical potential has to be calculated as well as other quantities associated with it (e.g. the capture radius). This again slows the simulation down, with the upside of having a more realistic model and more accurate results.
The implantation process itself has also been simplified in the simulation. Usually the proton beam is moved over a specific sample area in a scanning fashion. After scanning over that area once it returns back to the initial position and starts again. This movement of the proton beam has not been regarded in the simulation. Instead, the simulation will not implant anything until one scanning cycle is completed. The time it takes for this to happen is given by the inverse scan frequency. After that, a dose equivalent to the elapsed time is implanted all at once.
Next, many values of defect properties are either not known or guessed. Important physical properties for the defects include the diffusion constant, the activation energy for the diffusion, the binding energy of a defect, its attempt frequency as well as the capture radius of one reaction between two reactants. The values used for the simulation were found in literature search. If there is for example knowledge of one defect being able to diffuse but no value was found in the literature, a value for the diffusion constant was simply guessed, based on values of other, maybe related defects. This of course greatly impacts the accuracy of the results and is one of the biggest issue of the simulation as it is. Moreover, it is possible that defects are formed in a real experiment, that are not taken into account by the simulation.
Some assumptions were also made about the way defects react and dissociate. For the reaction it was

assumed that only second order reactions occur. This only allows reactions of the form $A + B \rightarrow i$, where two defects react to form one product. The reverse is true for the dissociations, where we assume that one defect may only dissociate to form two products. Moreover, one dissociating defect might have multiple possible dissociation results, i.e. one defect does not always dissociate into the same two defect species. Moreover, there is a probability, that in the real case reactions and dissociations happen, that are not regarded by the simulation due to lack of exact knowledge.

4 Implementation

After giving introductions both to the (physical) background as well as the aims and restrictions of our model it is time to look at the mathematics behind the simulation and shed some light on the simulation structure. First of all, this section will explain the mathematical necessities used to tackle the problem at hand. It is then explained, (i) how the finite difference method is used to combine the model of the silicon wafer with the underlying mathematical framework (ii) and what methods are employed to get a solution from this system. Afterwards, the simulation structure is explained. It is shown, how data and parameters are stored and handled, as well as what the simulation code does in more detail. Finally this section explains measures that are included in the simulation that improve its performance, for example including an adaptive time step and using parallelization.

4.1 Mathematical Background

As already mentioned, the simulation aims to gather information about the concentration of defects along the depth of the wafer. From a mathematical point of view one needs to determine how the defect concentrations change over time during the implantation process when starting with certain initial conditions (initial impurity concentrations in the wafer, implantation parameters, temperature, etc...) . The change of concentration of one defect called i over time is expressed as follows:

$$\frac{dc_i}{dt} \tag{4.1}$$

For the sake of simplicity, we now look only at one specific point in the wafer. As already mentioned, the different mechanisms and "actions" that defects can perform in the wafer are diffusion, reactions and dissociations. Each of the presented mechanisms changes the concentration of the defects in every single point of the wafer. Hence, we need to know how these mechanism change the concentrations of the corresponding defects over time. Once we gather all the contributions we can simply combine them to one main equation that is able to describe the whole system and simultaneously takes care of the applied simplifications and restrictions that we made in section 3.

Diffusion

First of all, some defects are able to diffuse in the sample. Diffusion of one defect occurs if regions with different concentrations of that defect are present. The diffusion mechanism tries to compensate for the concentration gradient by moving the defects accordingly.

$$J = -D\frac{\partial c}{\partial x} \tag{4.2}$$

Fick's first law (equation 4.2, in one dimension) states, that the particle flux is proportional to the concentration gradient opposite to the direction of the diffusion, which means that the diffusion goes from a region of high concentration to low concentration. The proportionality constant is called diffusion coefficient (often abbreviated with D) and depends strongly on the nature of the diffusing particle and the medium. Thus, the diffusion coefficient for every defect will be different in the

simulation. The diffusion coefficient contains parameters that are specific to every single defect, as shown in equation 4.3:

$$D_i = D_{0,i}\ exp\left(\frac{E_{A_i}}{k_B T}\right) \tag{4.3}$$

$D_{0,i}$ is the diffusion constant (of defect i) and $E_{A,i}$ is the activation energy needed for that defect to start diffusing. Not every defect in the sample is actually able to diffuse. Especially bigger defects and defect complexes cannot move through the silicon crystal lattice due to their size, and are considered immobile. Correspondingly, their values for $D_{0,i}$ and $E_{A,i}$ are zero. The variables k_B and T in equation 4.3 are the Boltzmann constant and the temperature of the implantation process respectively. The lower the activation energy for the diffusion and the higher the process temperature, the higher the contribution of the diffusion.
Combining equation 4.2 with the continuity equation 4.4 (the mass of the defects while diffusing has to be constant, as no defect are allowed to leave the sample)

$$\frac{\partial c}{\partial t} = -\frac{\partial J}{\partial x} \tag{4.4}$$

gives Fick's second law, which is also called the diffusion law (in one dimension):

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} = D\Delta c \tag{4.5}$$

The change in concentration due to diffusion is the first term in the main equation for the system.

Reaction

It was already mentioned, that only second order reactions are regarded in the simulation. That means that two defects will react to form one product (a new defect). To derive the term that describes the reaction, let us look at one possible reaction, where defect A reacts with defect B to form defect i. The indices used in the upcoming equations strictly refer to the different defects involved in the processes. For example, sub-index a states, that the variable attached to it is a property of defect A.
There are two factors, that determine how much of defect i will be produced. Firstly, it is dependent of the concentrations of the partaking defects A and B. The higher the concentrations of either one, the higher the concentration of the produced defect i. The other factor is the reaction rate k (unit $\left[\frac{\mathrm{m}^3}{\mathrm{s}}\right]$), which is a measure of how fast the reaction takes place. The reaction rate (here as an example the reaction rate of defect A and B) is described by the following relation:

$$k_{ab} = 2\pi r_{c,ab}(D_a + D_b) \tag{4.6}$$

$r_{c,ab}$ is the so-called capture radius (of reaction $A+B \rightarrow i$). Only if the reaction partner of A is within the capture radius, a reaction between the two is possible. This mechanism prevents two defects from reacting with each other, if they are very far apart from each other (e.g. on either end of the sample), which would not be realistic. For the reaction to be possible, the reaction partners have to

be somewhat close to each other. Further we can find the diffusion constants of both reaction partners in the relation for the reaction rate. This implies that no reaction will take place, if both defects are immobile. In order to react, defects have to move close to each other. Hence, if they are not able to move in the first place, no reaction will take place.
For the change of defect concentration over time we need to multiply the concentration of the reacting defects and multiply this with the according reaction rate:

$$\frac{\partial c_i}{\partial t} = k_{ab} c_a c_b \tag{4.7}$$

The simulation deals with 86 defects. Theoretically, all of them are able to react with each other (and themselves), so it is obvious to assume, that there is more than one reaction which yields defect i as a product. Accordingly it is necessary to sum over all the reactions where i is a product, which extends the previous relation to

$$\frac{\partial c_i}{\partial t} = \sum_{a,b \neq i} k_{ab} c_a c_b \tag{4.8}$$

In this sum, indices a and b cannot be i, because it is already the product; the reaction would be impossible. This contribution of the reaction to the main equation is still not done however, because we have not yet considered, that defect i (whose concentration value we are currently observing) can also be one of the reaction partners. In this case the concentration of defect i would be decreased instead of increased. In analogy to the previous case, we have to again consider that there is more than one reaction, where i is partaking to form a new defect:

$$\frac{\partial c_i}{\partial t} = \sum_{a,b \neq i} k_{ab} c_a c_b - \sum_{a \neq i} k_{ai} c_a c_i \tag{4.9}$$

Dissociation

The new defects that were generated by the reaction may also dissociate again into smaller defects they are composed of. Similar to the reactions, some assumptions were made for the dissociations. Most importantly, defects will only dissociate into two smaller defects and there is usually more than one combination of defects they can dissociate into.
The mathematical terms that describe the dissociations are very similar to the ones of the reactions. Again we observe the change in concentration of defect i. The amount of products generated upon the dissociation of a defect depends on its concentration as well as the dissociation rate, which is given by the following relation:

$$d_a = f_{A,a}\ exp\left(\frac{E_{B,a}}{k_B T}\right) \tag{4.10}$$

Here, variable $f_{A,a}$ is the attempt frequency, which states how many times a second a defect tries to break its bond and dissociate. $E_{B,a}$ is the binding energy which has to be overcome in order to break the bond. The contribution of one dissociation is given by

$$\frac{\partial c_i}{\partial t} = d_a c_a \tag{4.11}$$

Just like for the reactions, it is necessary to sum over every reaction that yields defect i as one of the two (or even both) products, because then the concentration of i is increased:

$$\frac{\partial c_i}{\partial t} = \sum_a d_a c_a \tag{4.12}$$

Finally it is necessary to consider, that defect i might dissociate again. In this case, the concentration of i is lowered. Further, we only have to consider this term once, since there is only this one case, where the defect concentration of i is decreased by dissociation. The whole contribution of the dissociation towards the main equation thus looks as follows:

$$\frac{\partial c_i}{\partial t} = \sum_a d_a c_a - d_i c_i \tag{4.13}$$

Implantation

Hydrogen vacancies and interstitial are not only formed by reactions and dissociations, they are mainly introduced during the implantation process. Thus, their total amount in the wafer constantly increases over time. The concentration however may be locally decreased, if one of the above mentioned processes dominates the contribution of the implantation. The contribution of the implantation towards the main equation is considered by the following relation

$$\frac{\partial c_i}{\partial t} = c_i^g \tag{4.14}$$

where c_i^g is the concentration increase of defect i per time due to the implantation.

The Main Equation

Combining all the contributions from the previous sections, one ends up with the main equation that mathematically fully describes the system:

$$\frac{dc_i}{dt} = D_i \Delta c_i + \sum_{a,b \neq i} k_{ab} c_a c_b - \sum_{a \neq i} k_{ai} c_a c_i + \sum_{a \neq i} d_a c_a - d_i c_i + c_i^g \tag{4.15}$$

Equation 4.15 shows the change of concentration over time for one of the 86 defects, for a small step in time (dt). To achieve knowledge of how all the defects are distributed in the wafer after the implantation process is over, it is necessary to evaluate equation 4.15 for every different defect, on every "point" in the wafer sample (more details about those "points" in section 4.1.1) and until the implantation process is over (i.e. until it has been calculated for enough time steps dt and the implantation dose is reached).

When computing the results, it is important that the simulation code is written in a way not to

perform reactions twice or not at all. A short example will illustrate two common problems when evaluating the contributions: Let us look again at a reaction of type $A + B \rightarrow i$. While adding the obtained concentrations to defect i, it is important not to forget to subtract the same amounts from defects A and B, as they are transformed to make up defect i. Moreover, it is important to discard the reaction $B + A \rightarrow i$ until the next evaluation of that contribution, as that is the same reaction as $A + B \rightarrow i$. The code needs to be set up in way that these things are prevented, and still guarantee a fast computation time. A part of section 4.3 addresses this issue.
Formula 4.15 only gives information about the change of defect concentration over time, but we currently have no information about how the defects are distributed in the wafer or how the wafer is shaped. To incorporate this kind of information, we have to map the equation onto our sample, as will be described in the next section.

4.1.1 The Finite Difference Scheme

It is now important to map equation 4.15 to the underlying silicon wafer. This is done by discretization. The sample to be modelled is divided into cells that are represented by single points, the so called grid points. Physical properties and quantities (in this case the defect concentration) can then be defined on those grid points. Sometimes physical quantities need to be defined between grid points (e.g. current density in simulations for charge carriers in a semiconductor). It is up to the experimentalist to choose how many grid points should describe the sample and to find a good compromise. More grid points improve resolution, but also increase computation time (often in a non-linear fashion, i.e. doubling the amount of grid points often increases computation time by a factor of >2). One also has to take into account different materials and their interfaces at which the physical properties change. One example would be the interface between a semiconductor and an oxide material in a MOSFET. In our case, the only "interface" is between the wafer and the surrounding atmosphere. One of the simplifications states, that no out-diffusion from the wafer to the atmosphere is allowed, which is considered in the upcoming steps.
The wafer we are trying to simulate has three dimensions. However, as mentioned earlier, we use a simplification, in which we assume that the defect concentrations only change along the depth of the wafer, not in the other two dimensions. Thus only the depth of the wafer needs to be discretized, which corresponds to simulating a one dimensional silicon rod where the defects can diffuse, react and dissociate.

The Numerical Methods

The most difficult part about discretizing equation 4.15 is dealing with the second derivative in space in the diffusion term. This makes the equation an ordinary differential equation that needs to be solved numerically. This can be done using one of three closely related numerical methods which are presented in this section.
The aim is to get a solution to the differential equation only from knowing the initial values, namely the defect concentrations. In order to make the explanation of those methods easier, we imagine having to calculate a curve that satisfies a given differential equation (for example equation 4.15) and whose initial value we know. With the help of the differential equation and the initial value as a starting point, we can determine the slope of the curve in that point. From the slope, we can further

determine the tangent line. We can now move along the tangent line and set the next point of our function, that tries to approximate the real curve. The longer the distance between starting point and next point (henceforth called "step size"), the more the next point will deviate from the true position of the curve. Thus, if we only use a very small step size, the result will be much more accurate, but it takes much longer to move forward in time. From this point on it is again possible to determine the slope and the tangent line and determine the next point of the function, and so forth. This method is called "forward Euler" or "explicit Euler" method. It is obvious, that the step size is a crucial quantity, when trying to determine solutions to differential equations this way. Large step sizes yield fast results which have a larger numerical error, while small step sizes will decrease the error but take longer to compute. The explicit Euler method is the simplest one, but also has issues regarding stability. The solution might start to oscillate, depending on how the step size was chosen.
The second method is not prone to such instabilities, but is more complex. The "implicit Euler" method is using the same principle as the explicit Euler method with a small difference. The slope and tangent line is not taken from the initial point, but from the point that is going to be calculated next. This yields an implicit equation that needs to be solved in every step, hence the name of the method. As already mentioned, this method is much more stable than the explicit Euler method, but usually involves greater computational effort, due to its implicit nature.
The third possible method represents a middle ground between the previous two. The Crank-Nicolson method can be understood as an average between the explicit and the implicit method. Strictly speaking however it is not exactly an average between the two, as it still contains the implicit dependency from the implicit Euler method. The benefit of this method are improved accuracy at the cost of computation speed, as essentially both previous methods have to be calculated.
Because computation time is an issue for this simulation, it was decided to use the implicit Euler method, due to its good combination of stability speed and numerical errors. Moreover, many of the simplifications and the fact that some of the initial values are not known and just "good guesses". The uncertainty caused by those issues is assumed to be far greater than the difference of numerical error between implicit Euler method and Crank-Nicolson method.

The Discretization

To fully discretize equation 4.15 in space and time every derivative needs to be rewritten to a finite difference. The time derivative on the left side of the equation is thus transformed to

$$\frac{dc(x,t)}{dt} \rightarrow \frac{c_x^{t+1} - c_x^t}{\Delta t} \tag{4.16}$$

This introduces the step size of this simulation. While the step size in the examples provided in the last section had no specific dimension or physical meaning, the step size in this simulation represents the elapsed time of the implantation process. Thus, the step size is also called "time step" in this thesis. The concentration variables c in equation 4.16 have 2 indices: index t is the time coordinate. A value of t represents the "present", index $t+1$ represents a value one step ahead in the "future", i.e. the next value to be computed. A good choice of the value of the time step Δt is crucial to the outcome and performance of the simulation. This is why an adaptive time step was introduced, which is explained more thoroughly in section 4.3.1. Index x represents the coordinate in space, more precisely which grid point is addressed. $x = 0$ is the surface (top) of the wafer that is subjected to the

proton implantation, while $x = N$ is the bottom of the wafer (with N the total amount of grid points used, to describe the wafer depth).
By applying the implicit Euler method to the diffusion term on the right hand side of the main equation 4.15 we get

$$\Delta c = \frac{d^2c}{dx^2} \rightarrow \frac{c_{x+1}^{t+1} - 2c_x^{t+1} + c_{x-1}^{t+1}}{(\Delta x)^2} \tag{4.17}$$

where Δx is the distance between the grid points. It is defined in the following way, where d is the total simulated depth of the wafer

$$\Delta x = \frac{d}{N-1} \tag{4.18}$$

Applying equations 4.16 and 4.17 to equation 4.15 yields (indices previously used to describe the defect type are now top left)

$$\frac{{}^ic_x^{t+1} - {}^ic_x^t}{\Delta t} = D_i \frac{{}^ic_{x+1}^{t+1} - 2\,{}^ic_x^{t+1} + {}^ic_{x-1}^{t+1}}{(\Delta x)^2} + \sum_{a,b \neq i} {}^{ab}k\,{}^ac_x^t\,{}^bc_x^t - \sum_{a \neq i} {}^{ai}k\,{}^ac_x^t\,{}^ic_x^t + \sum_{a \neq i} {}^ad\,{}^ac_x^t - {}^id\,{}^ic_x^t + {}^i_gc_x^t \tag{4.19}$$

Solving the equation for the unknown, future values (with index $t+1$), omitting the indices $_{i,a,b}$ for the sake of readability and using the substitution $r = \frac{D_i \Delta t}{(\Delta x)^2}$ one obtains

$$-rc_{x+1}^{t+1} + (1+2r)c_x^{t+1} - rc_{x-1}^{t+1} = c_x^t + \Delta t \left[\sum_{a,b \neq i} k_{ab}\,{}^ac_x^t\,{}^bc_x^t - \sum_{a \neq i} k_{ai}\,{}^ac_x^t c_x^t + \sum_{a \neq i} d_a\,{}^ac_x^t - d_i c_x^t + {}_gc_x^t \right] \tag{4.20}$$

To advance the system further in time by one time step, this equation has to be evaluated for every grid point and every defect. Every term can be computed independently (causing increments or decrements in the concentration for that defect at that grid point) which is summed up as soon as all parts are computed. This opens up the opportunity to parallelize the computation and let the simulation run on a computer cluster which significantly reduces the computation time.

4.1.2 The Thomas Algorithm

In equation 4.20 every quantity on the right side is known, while the concentration values on the left side are not known and have to be calculated. This can be done by using the Thomas algorithm. To make application easier, values for the concentrations of the defects are stored in matrices. One column will describe one defect type (e.g. hydrogen), while the rows represent the grid points in the sample. If we look at only one defect and the whole depth at a time (i.e. only at one column of the concentration matrix) we can rewrite the concentrations as vectors which contain N values, namely the concentration values for that one defect on every grid point. Vectors that hold information about current/present concentration values (noted with index t) are named d. Thus it is possible to calculate

every term on the right side of the equal sign in equation 4.15 and call it d, because the terms only contain concentration values of the present. Vectors that hold the future/updated concentration values (noted with index $t+1$) are named x. By doing all this and using the substitutions shown in 4.22, we manage to cast equation 4.20 into the following matrix equation:

$$\underbrace{\begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & a_3 & \ddots & \ddots & \\ & & \ddots & \ddots & c_N \\ & & & a_N & b_N \end{pmatrix}}_{=M} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{N-1} \\ d_N \end{pmatrix} \tag{4.21}$$

with

$$b_{1,\dots,N} = 1 + 2r, \quad a_{1,\dots,N} = c_{1,\dots,N} = -r, \quad r = \frac{D_i \Delta t}{(\Delta x)^2} \tag{4.22}$$

The coefficient matrix M is tridiagonal. All other values of M are 0. This equation will now be solved using the Thomas algorithm. In a first sweep all the a_i's are eliminated. A subsequent second sweep then yields the solution of the x_i's by a backward substitution.

Written out, the first line of the matrix equation is

$$b_1 x_1 + c_1 x_2 = d_1 \tag{4.23}$$

Dividing by b_1 and using $\gamma_1 = \frac{c_1}{b_1}$ and $\delta_1 = \frac{d_1}{b_1}$ one obtains

$$x_1 + \gamma_1 x_2 = \delta_1 \tag{4.24}$$

or in matrix equation form

$$\begin{pmatrix} 1 & \gamma_1 & & & \\ a_2 & b_2 & c_2 & & \\ & a_3 & \ddots & \ddots & \\ & & \ddots & \ddots & c_N \\ & & & a_N & b_N \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} = \begin{pmatrix} \delta_1 \\ d_2 \\ \vdots \\ d_{N-1} \\ d_N \end{pmatrix} \tag{4.25}$$

The second row reads

$$a_2 x_1 + b_2 x_2 + c_2 x_3 = d_2 \tag{4.26}$$

Multiplying the first row with a_2 (4.28) and subtracting from the second row (4.27) eliminates the first term and we get a new expression for the second row, without a coefficient for x_1 (4.29)

$$a_2x_1 + b_2x_2 + c_2x_3 = d_2 \tag{4.27}$$

$$a_2x_1 + \gamma_1 a_2 x_2 = a_2\delta_1 \tag{4.28}$$

$$(b_2 - \gamma_1 a_2)x_2 + c_2x_3 = d_2 - a_2\delta_1 \tag{4.29}$$

Division by $(b_2 - \gamma_1 a_2)$ yields

$$x_2 + \frac{c_2}{b_2 - a_2\gamma_1}x_3 = \frac{d_2 - a_2\delta_1}{b_2 - a_2\gamma_1} \tag{4.30}$$

$$x_2 + \gamma_2 x_3 = \delta_2, \quad \gamma_2 = \frac{c_2}{b_2 - a_2\gamma_1}, \quad \delta_2 = \frac{d_2 - a_2\delta_1}{b_2 - a_2\gamma_1} \tag{4.31}$$

After using the substitutions in 4.31, the matrix equation looks as follows:

$$\begin{pmatrix} 1 & \gamma_1 & & & \\ 0 & 1 & \gamma_2 & & \\ & a_3 & \ddots & \ddots & \\ & & \ddots & \ddots & c_N \\ & & & a_N & b_N \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ d_{N-1} \\ d_N \end{pmatrix} \tag{4.32}$$

The procedure from the second row is now repeated until the very end. This can be expressed by the following rules for the substitutions:

$$\gamma_i = \begin{cases} \dfrac{c_i}{b_i}, & \text{if } i = 1 \\ \dfrac{c_i}{b_i - a_i\gamma_{i-1}}, & \text{if } i = 2, ..., N \end{cases} \tag{4.33}$$

$$\delta_i = \begin{cases} \dfrac{d_i}{b_i}, & \text{if } i = 1 \\ \dfrac{d_i - a_i\delta_{i-1}}{b_i - a_i\gamma_{i-1}}, & \text{if } i = 2, ..., N \end{cases} \tag{4.34}$$

After performing these steps down to the last row, the matrix equation has the following shape

$$\begin{pmatrix} 1 & \gamma_1 & & & \\ 0 & 1 & \gamma_2 & & \\ & 0 & \ddots & \ddots & \\ & & \ddots & 1 & \gamma_{N-1} \\ & & & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{N-1} \\ \delta_N \end{pmatrix} \tag{4.35}$$

at which point the last row already represents the solution of the last element in x:

$$x_N = \delta_N \tag{4.36}$$

This makes it possible to solve the second to last row of the matrix equation. This is repeated backwards until the whole vector of x is solved, which includes computing all the values for δ and γ as well. This algorithm can also be applied if the Crank-Nicolson method is used, as the only differences are the values of vector d. This calls for slightly different (and more complex) substitution parameters γ and δ, but the main principle stays the same.

4.2 The Simulation Structure

First versions of the simulations were written in python by Martin Faccinelli. Since python is not optimized for a number crunching task like this, a first step was to rewrite the simulation in a more potent programming language. The computational aspects of the simulation were programmed using FORTRAN 95, while Python was used to prepare the data for FORTRAN 95 and to generate the output.
The foundation for the simulation is one file, that contains all the (known) information about all the defects we want to consider. This file was not made to be usable directly by FORTRAN code, which is why python is used to generate further files in which the information is stored in a more useful way. This is used to deal with the reactions and the dissociations during the implantation process. The simulation itself consists of basically two parts. The task of the first part is to import and prepare the data for the second part of the simulation, which is its core functionality. The second part consists of a loop that calculates and updates the defect concentrations in the wafer, as the implantation process advances in time. This section will explain both parts more thoroughly and will go into detail about how the available information was set up and processed.

4.2.1 The Reaction and Dissociation Matrix

The most straight forward way of determining the contribution of every possible reaction towards every defect would be to generate a matrix (henceforth called the reaction matrix Mr), whose rows and columns depict the defects. To make this task simpler, every defect is given an id number (from 0 to 85, because array indices in most programming languages start at 0 and there are 86 defects). To determine the outcome (if any) of a reaction of defect A (with id number x) with defect B (with id number y), one can quickly refer to element $Mr(x, y)$ and find the id number of the product, or a different entry if there is no reaction. The case of "no reaction" can have two reasons: First, it is possible, that two defects simply do not react in real life to form a product. If a reaction is, in principal, possible, but both defects are considered immobile and unable to diffuse, there can't be a reaction, as the reaction rate is 0.
If one implements a reaction matrix like this, its dimensions would be 86 $\times$ 86. When checking every possible reaction however it would be wrong to check every element in Mr, as the elements $Mr(x, y)$ and $Mr(y, x)$ refer to the same reaction! Thus, only the elements in the main diagonal and below (or above) have to be considered. This way, only 3741 elements in the matrix have to be checked for reaction, every time the concentrations are updated.

A much more efficient way to handle the reactions is to generate a matrix that contains information only about those defects, that are really able to react. The rows of such a matrix (we will still call it Mr) contain the information about one reaction. The first column then holds the id number of the first reaction participant, the second columns the id number of the second participant and the third column denotes the id number of the product. That way the computation is cut down to the 207 reaction that take place, and for every update only 207×3 elements have to be checked (just for the reaction). During the simulation this can be further reduced, by checking if the concentration of either reaction participant is 0, in which case the reaction cannot take place.

The exact same procedure is done to deal with the dissociations in the system. A dissociation matrix Md with three columns is generated, where the first column is the dissociating defect, and the other two columns are the product of the dissociation. When generating Md it is checked if a defect has values for the attempt frequency and/or the binding energy. If yes, this defect is able to dissociate and is added to the matrix. Otherwise it would have been necessary to check all 86 defects for dissociation in every update, even though only 34 of them actually dissociate. These matrices can also be split up by rows, calculated independently and later by added together again. This is very important, as it opens up the possibility to employ parallel computation to further reduce the computation time.

Storing the Defect Concentration

The defect concentrations are of the main interest in this simulation. It is necessary to store the concentration of all 86 defects on every grid point. Further it should be possible to easily retrieve and update those values. The way this information was organized was by putting it into a matrix. The rows of this matrix represent the grid points of the modelled silicon wafer while the columns hold the concentration values for the different defect types. The concentration values are frequently updated and thus need to be overwritten, which can easily lead to erroneous values if done incorrectly. This is why a system with two concentration matrices was employed. One holds "old" concentration values and thus serves as a backup if necessary. Further, these are the values from which the update concentration values are calculated. The updated values are stored in the second concentration matrix, which will then be subjected to several checks, whether or not the calculations were successful. If they were, the first matrix is overwritten with the values from the second one and the process is repeated again (more on that process can be found in the next section).

The Main Loop

After the import and preparation of all kinds of data has been finished, the simulation will execute its core part, the main loop, which is responsible for calculating and updating the concentration values of all the defects. Figure 5 shows a flowchart of this part of the code, which is executed until the implantation process is over, i.e. until a certain implantation dose is reached.
Inside the main loop of the simulation, it is first checked if an implantation step has to be done. This check monitors the time passed since the last implantation step. If enough time has elapsed (i.e. the proton beam has finished scanning over the entire target area on the sample once) another implantation step is performed. This is done by adding a certain amount to the concentrations of hydrogen, vacancies and interstitials (the only defects that are generated due to implantation). SRIM-simulations

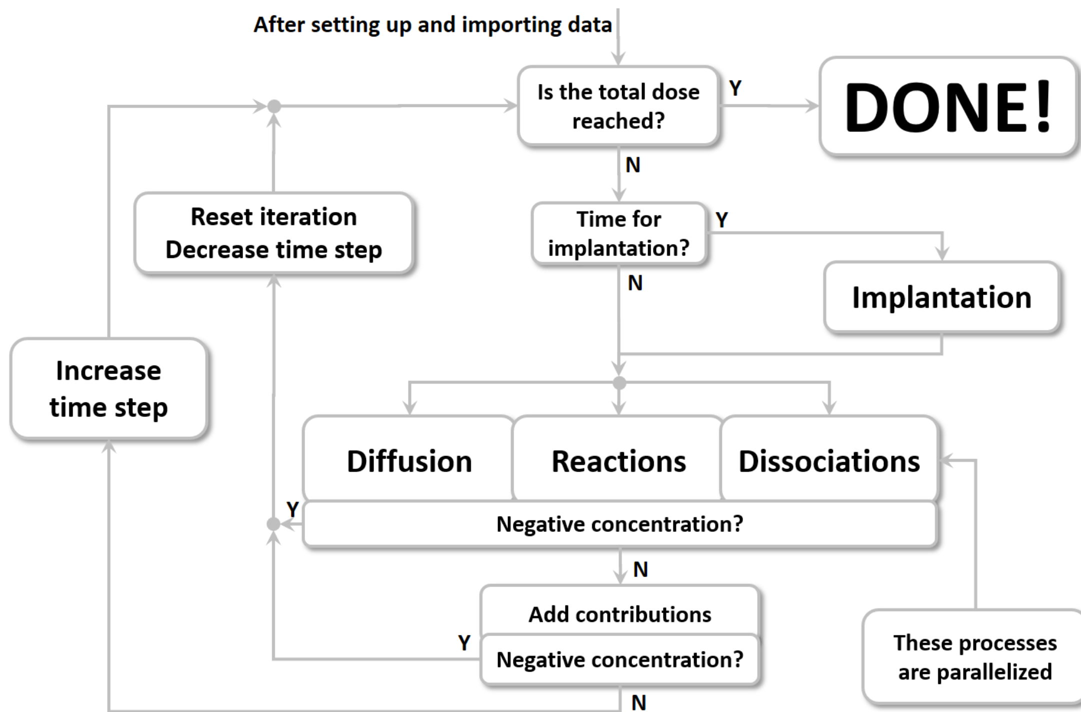


Figure 5: Flowchart of the main loop of the simulation.

at specific energies determine how much has to be added to the corresponding concentrations. It is one of the tasks of the first part of the simulation to correctly import the files from the SRIM-simulation. If the scanning process has not progressed enough, this step is simply skipped.
After the implantation itself has been dealt with, the program computes the different contributions by defect diffusion, reactions and dissociations, as described above. Here, one loop loops through the rows of the reaction and dissociation matrix (described in section 4.2.1) and simultaneously addresses the concerning defects with their identification number. A second loop inside the first loop then loops through all the grid points. It is possible to treat and compute these three parts independently, albeit this again being a small deviation from the real case. The program calculates updated concentrations by the three different parts from the same set of concentrations. In reality though, as soon as something diffuses, reacts or dissociates, the concentrations for the next process would be different, leading to different results which are however negligible compared to other simplifications and assumption already made. It is this part of the simulation where parallel computing is used. Section 4.4 will treat this in more detail.
After computing these contributions, they are added to the "old" concentration values (still independently of each other). The program now checks, if there are negative concentrations somewhere in the sample (originating from either diffusion, reactions, or dissociations). If there are, the last step is reset and done again, with a lower time step (more on the time step in this section). This is repeated until no negative concentrations appear anymore. The next step is to add all of the contributions together and to add this sum to the previous concentration values. Again, if there are negative concentration values appearing, the diffusion, reaction and dissociation is computed again with a lower time step, until no more negative concentration values appear. If this step is successful, the time step is increased and the outer loop starts over again, until the total implantation dose is reached.

The time step is the most crucial part of the simulation. The smaller the step size (here: the time step) the smaller the numerical error. In turn it takes longer for the simulation to finish if the time

step is small. However, one of the explicit goals of this thesis is to make present versions of the simulations faster, since they already took several weeks, depending on the parameters even months to finish. To speed up the simulations it was decided to keep the time step as high as possible. This in turn gives rise to larger numerical errors, as previously mentioned. It is however safe to say, that all the simplifications and assumptions that were made to perform these calculations outweigh this numerical error.
Moreover, said time step is not held constant. As the simulation time progresses, more defects are introduced into the sample and new defects are generated by reactions, leading to a higher computational effort as the simulation progresses. As a consequence, a constant time step would at some point yield negative concentration values. To counteract this, the simulation uses an adaptive time step, that increases by a certain factor if a loop iteration was successful (= no negative concentration on any grid point for any defect generated), and decreases it if the loop iteration was unsuccessful. A good choice of increasing and decreasing factors greatly impacts the simulation speed (more on this in section 4.3).

4.3 Code Optimization and Increasing Performance

Since computational performance is one of the main objectives in this work, great efforts have been made to reduce unnecessary calculations. Partly, this was achieved by restructuring the given data beforehand, as already explained in section 4.2.1. On the other hand an adaptive time step was implemented (but not perfected) to let the simulation progress as fast as possible. Further, the code was structured in a way that the use of parallel programming could be included. These measures have decreased the simulation speed compared to earlier versions of the simulation. A brief discussion about the results concerning speed is found in section 5.6.

4.3.1 The Adaptive Time Step

The larger the time step is chosen, the higher the numerical error will be. In extreme cases this will result in negative concentration values on some grid points, which are physically impossible and have to be avoided. If negative concentration values are encountered, the simulation has to be either continued with a lower time step or has to be restarted with a lower initial time step. Usually in such simulations the time step should be as small as possible, so that the obtained solutions are as close to the real solution as possible. The downside however is the longer computation time. Since computation times are already in the time frame of weeks to months, depending on the initial parameters, an even longer duration is not viable, in particular not for industrial purposes. Instead of choosing either a time step that is too high or too low, an adaptive time step was implemented. It starts out with some initial value that is lowered by a certain factor once negative concentrations have been detected at some point during updating the concentration values. If no negative concentrations were produced, the time step is increased again, in order to drive the simulation forward at the highest possible speed. By doing so, the simulation is kept at the very edge of physically relevant values.
Whenever a loop iteration was successful (i.e. no negative concentrations were generated), the time step is multiplied with a factor Δt_+ which is greater than 1. On the other hand if negative concentrations are present after a loop iteration the time step is multiplied by a factor Δt_- smaller than 1. The performance of the simulation drastically changes depending on the values chosen for Δt_+ and Δt_-.

If Δt_+ is too high (> 1.1) the time step will be too high on the next loop iteration, causing negative concentrations, which in turn calls for Δt_- to reduce the time step again. If at the same time Δt_- is set too high as well (> 0.8, but smaller than 1), it takes more than one reduction to bring the time step down to a level, where no negative concentrations are generated anymore. In extreme cases about 75% of the performed iterations are discarded in such a scenario, leading to a massive drop in overall simulation speed. If both Δt_+ and Δt_- are set too low, the simulation will not produce negative concentrations, but will not run at the highest possible time step either. Even though this represents a better case than both factors being too high, it still leads to a significant time loss. "Good" values for both Δt_+ and Δt_- were not determined mathematically, but by trial and error. The simulations that were performed and analysed for this thesis used $\Delta t_+ = 1.01$ and $\Delta t_- = 0.6$.

4.3.2 Evaluation of Diffusion, Reaction and Dissociation

While the simulation deals with 86 different types of defects, by far not all of them do diffuse or dissociate and not every defect reacts with every other defect. Looking at equation (4.15) it seems obvious that corresponding terms do not need to be evaluated. Accordingly, the simulation first checks, which defects diffuse, react and dissociate and then only evaluates the terms for those, leaving out all the rest. Further it is possible not to evaluate the left over terms, if the concentration on that grid point or the corresponding reaction rates are 0. To achieve this, it is however necessary to prepare the data in a specific manner, as already explained in section 4.2.1. Because of this, it was possible to reduce the amount of contributions by diffusion from 86 to 15, by dissociations from 86 to 34 and by reactions from over 3700 to only 207. Moreover, every single calculation of diffusion, dissociation and reaction includes the calculation for that value on every grid point, which is a great saving of computation time, considering the simulations were run with 201 points.

4.4 Parallelization

In parallel computing a complex and sophisticated computation is broken down to smaller and simpler tasks which are then computed simultaneously instead of sequentially. This simulation was run on a computer cluster which is composed of a set of connected computers that work together on a task. The workload to be computed simultaneously should be evenly distributed across the single entities in the computer cluster to avoid unnecessary idle time (if one part of the workload is finished significantly faster than the other). The results obtained from every single node are combined whenever the computation is over.

In order to increase to speed of the computation, the underlying code has been changed to allow the use of this kind of parallel computing to some extent. It is not possible to parallelize the whole simulation code (including importing data and so forth), as some parts are required to run sequentially; a simultaneous computation would be qualitatively wrong. Parts that can be computed at the same time are the evaluation of the contribution of diffusion, reaction and dissociation for each time step. Due to the unbalanced workload (only 14 defects diffuse, while there are 207 reactions to take care of), the subroutines that take care of each mechanism were subdivided again, to form approximately equal chunks of workload. That way it was possible, that every small task roughly needed the same time to compute, which made distribution of the task to the different nodes easier and more effective.

5 Results

The results obtained from the performed simulations are presented in this section and its subsections. In total 81 simulations with different parameters were performed. The following parameters were changed between three values to perform a parameter study: process temperature, implantation current, and the initial concentrations of carbon and oxygen. The implantation energy was held constant at 400 keV, as this basically only shifts the projected range deeper inside the wafer. Further, the implantation dose was held constant at a value of 10^{19} m^{-2}. As explained in section 2.5, this dose can be used to introduce hydrogen related thermal donors which cause n-type doping.
Changing the temperature of the implantation process has a bigger effect, because it changes the diffusion coefficients (and thus the diffusion and reaction of defects) as well as the dissociation rates. With increasing temperatures, defects will diffuse faster and the reactions and dissociations will yield more products. Changing the initial impurity concentration affects those defects and defects complexes that contain impurities. Further it is important to take a look at the simulation time and how it changes when the process parameters are altered.

All of the varying parameters were given a set of three different values, that will be called "high", "medium" and "low". Table 1 shows the values for those parameters associated with those keywords. Some simulations featuring a carbon concentration of 5×10^{21} m^{-3} have been identified as being erroneous. Because of these failed simulations, only two options are regarded for the carbon concentration, as depicted in table 1. Table 3 shows the simulation numbers and their associated parameters. Table 4 shows all the defects present in the simulation and their parameters. Both table 3 and table 4 can be found at the very end of this section.

Table 1: Values for the varying parameters and the keywords they are referred by in the text.

	"Low"	**"Medium"**	**"High"**
Temperature [K]	323	373	423
Implantation Current [μA]	50	100	200
Carbon Concentration [m^{-3}]	10^{22}	-	2×10^{22}
Oxygen Concentration [m^{-3}]	5×10^{22}	10^{23}	2×10^{23}

5.1 General Overview

This section gives a general overview of the results by showing some basic graphs that have been extracted from one of the simulations. All the parameters of this simulation were set to "low" (see table 1). Subsequent sections will show the change in defect concentrations when one ore more parameters are varied.
The sample depth can be roughly divided into three different regions. The irradiated region ranges from the surface down to about 4.5 microns. This is the region affected by the implantation of the protons and contains vacancies and interstitials. However most protons have sufficient energy to penetrate deeper into the sample than just this irradiated region. The highest concentration of protons, vacancies and interstitials is found in the second region, the projected range or implantation

depth. It is a quite small region at about 4.5 microns and can be easily identified by either a peak or a valley for many defect concentrations. It is this region, where most of the reactions and dissociations take place, as the defect concentrations are some orders of magnitude higher than in the irradiated region. Finally, in the regions beyond 5 microns the concentration of the implanted point defects rapidly decreases.
The only defects initially present in this region are the impurity atoms (oxygen, substitutional carbon, substitutional phosphorus and substitutional boron) with a certain concentration that resembles the manufacturing process. Other defects first have to diffuse into this regions before they can react. Depending on the temperature and the properties of the single defects, the diffusion can be rather slow. Correspondingly, the presented graphs will only show the concentrations of defects up until 6 microns, even though the simulated depth ranges down to 10 microns.

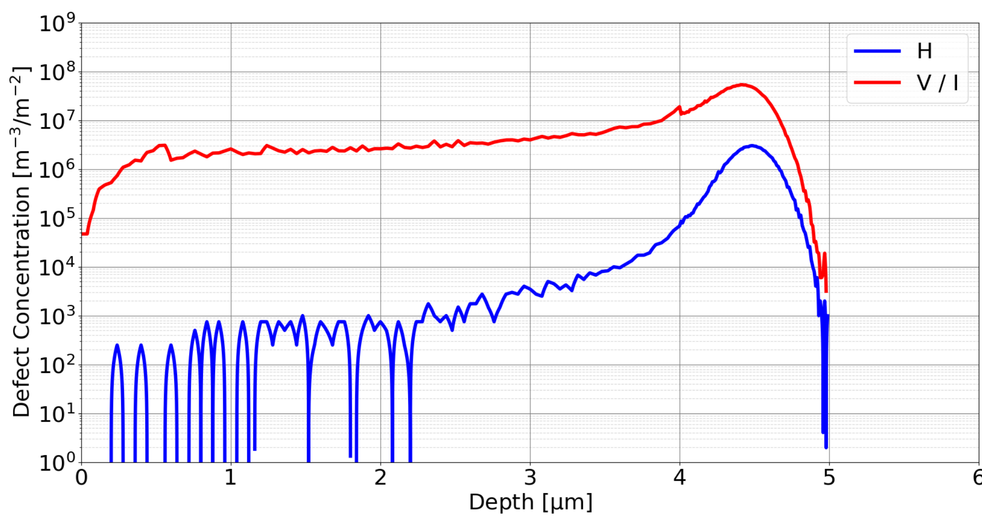


Figure 6: Semi-logarithmic plot of the concentrations of hydrogen (blue) and vacancies and interstitials (red) normalized to the proton dose simulated with SRIM. The implantation energy is 400 keV.

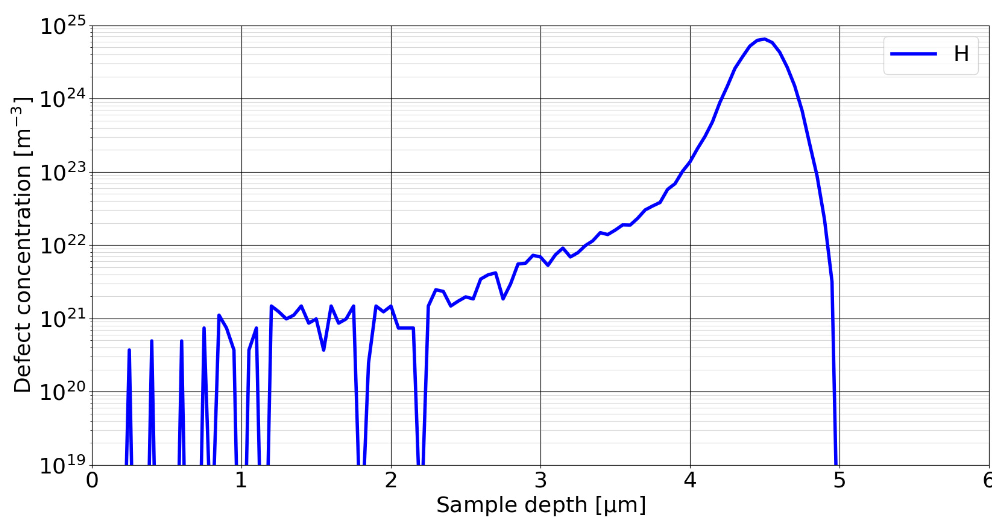


Figure 7: Semi-logarithmic plot of the concentration of hydrogen in the wafer after implanting a dose of 10^{19} m^{-2}. Note that the scale on the y-axis has changed from the previous graph.

Figure 6 shows the data that is given by the SRIM files, for an implantation energy of 400 keV. The blue curve shows the concentration of hydrogen. Up to a little over 2 microns, there are a lot

of discontinuities, because at some depths the concentration was calculated to be zero, i.e. these values cannot be plotted in a semi-logarithmic plot. The concentration of vacancies/interstitials is much higher than the concentration of hydrogen, because every hydrogen is able to knock several silicon atoms out of their place in the crystal lattice. Both curves peak at the implantation depth (i.e. projected range) at about 4.5 microns deep. Some hydrogen defects exceed the projected range and reach 5 microns, but their concentration drops rather quickly after the implantation depth. The regions deeper down in the wafer are not directly affected by the implantation. Hydrogen, vacancies and interstitials only reach these regions by diffusion; then they are able to react with the impurities, which are evenly spread out along the depth of the wafer. These concentration profiles of protons, vacancies and interstitials are repeatedly implanted into the wafer, until the total dose is reached.

Hydrogen. As a consequence, the concentration of hydrogen in the sample after implanting a dose of 10^{19} m^{-2} assumes a distribution that is shown in figure 7. Some of the discontinuities in the irradiated region have become smaller or even vanished because of the diffusion of hydrogen. Another difference is the scale on the y-axis. As the profile from figure 6 is implanted repeatedly, the concentration of hydrogen is obviously much higher, namely by the factor of the implantation dose. However the numbers do not completely add up. The peak concentration of 3×10^6 m^{-3} in figure 6 should give rise to a total 3×10^{25} m^{-3} in figure 7, but is actually a little less. This discrepancy occurs because hydrogen reacts and forms other defects, such as 'H_2', which decreases the concentration of 'H'.

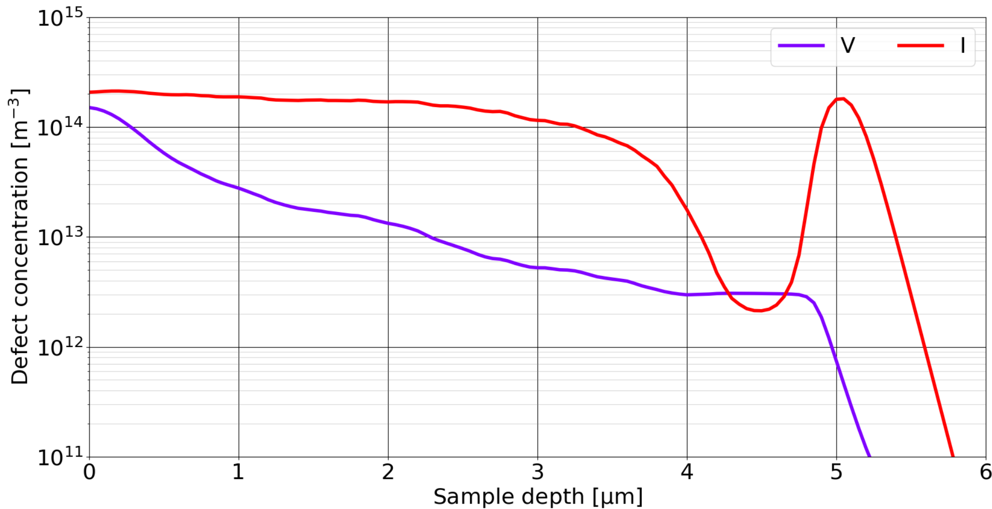


Figure 8: Semi-logarithmic plot of the concentration of vacancies and interstitials in the wafer after implanting a dose of 10^{19} m^{-2}. Again the scale on the y-axis had to be adapted.

Vacancies and interstitials. While the hydrogen profile after the implantation resembles the profile of the SRIM simulations, the concentration profiles of vacancies and interstitials look very different (figure 8); their concentrations are several orders of magnitude lower than the one of hydrogen. The data from the SRIM simulation suggests, however, that there should be 100-1000 times more vacancies and interstitials than hydrogen. It has to be mentioned, that these curves are merely 'snapshots' of concentrations that change very quickly. As soon as new vacancies and interstitials are introduced with the concentration shown in figure 6, the total concentration of V and I will shortly increase and quickly react away again. The reason for the generally low concentrations is the fact, that vacancies and interstitial mostly react with each other, to form a defect 'VI'. This combination of a vacancy and an interstitial atom is just a silicon sitting at a regular lattice site. One can view the "defect"

'VI' as the healing process of the crystal from the radiation damage. 'VI' is also unable to dissociate , so this reaction is irreversible and the concentration cannot be decreased anymore. Thus this defect shows the highest concentration of all. At its peak it reaches a concentration of 5×10^{26} m^{-3}, which is about 3% of the density of silicon. Figure 9 shows the evolution of 'defect' 'VI' as the implantation process progresses. Every curve represents 10% of the total process. It can be seen, that very high concentrations are already present from very early on. The increase in concentration becomes gradually slower as the implantation process progresses.

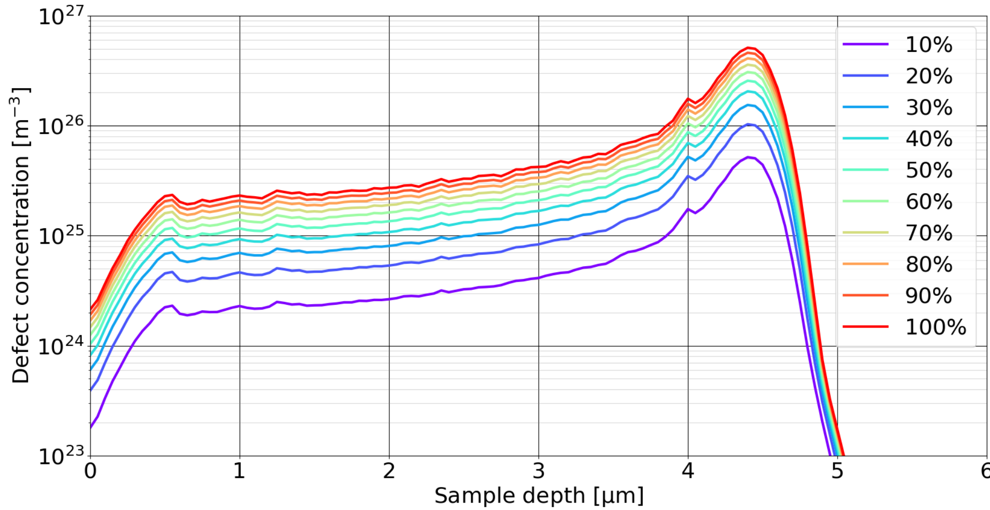


Figure 9: Semi-logarithmic plot of the concentration of 'VI' in the wafer after implanting a dose of 10^{19} m^{-2}.

The concentration differences from 0-4 microns in figure 8 might originate from the different amount of defects, in which vacancies and interstitials participate. While 33 defects contain one or more vacancies, only 22 are made up of one or more interstitial silicon atoms. Thus, vacancies have more possibilities to react and show a lower concentration. The silicon interstitial concentration then drops below the vacancy concentration at the implantation depth. This is most likely due to the following: interstitial complexes occur in seven distinct species ('I_2'-'I_8'), where pure vacancy complexes only range from 'V_2'-'V_6'. The maximal concentrations of these complexes lie between $(2.8-3.2)\times 10^{23}$ m^{-3} for 'I_2'-'I_8' and between $(2.5-2.7) \times 10^{23}$ m^{-3} for 'V_2'-'V_6'. Since the interstitial complexes are comprised of two more defects with about the same concentration, the concentration of interstitial silicon atoms decreases in that region. Even though vacancies are able react to defects where interstitials have no counterpart to, these defects have a much lower concentration (factor of minimum 1000), which does not visibly affect the concentration.

Main impurities. All the main impurities (oxygen, substitutional carbon, substitutional boron and substitutional phosphorus) have a constant concentration throughout the wafer depth before the implantation process starts. Carbon, boron, and phosphorus can also be on interstitial lattice sites, when a substitutional impurity and a silicon interstitial atom switch places. This reaction can also be reversed: an interstitial impurity becomes a substitutional impurity when meeting a vacancy. The concentrations of the impurities are depicted in figure 10. In the regions deeper than 5 microns the initial concentration is unchanged, because there are no reaction partners present yet. At the projected range the concentration of the main impurities drops because this is the region where most defects are

present. Closer to the surface the concentrations rise again, because the amount of reaction partners decreases again. But it is still lower than deeper in the sample, because of the introduced defects.

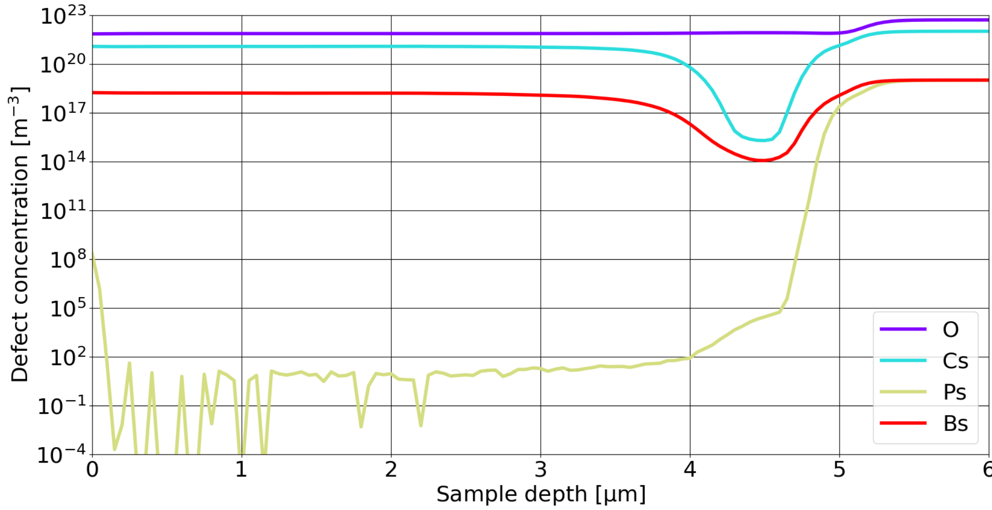


Figure 10: Semi-logarithmic plot of the concentration of the impurity defects in the wafer after implanting a dose of 10^{19} m^{-2}.

Substitutional and interstitial impurity defects of phosphorus, boron, and carbon always show roughly the same concentrations in the irradiated region, even though typically one of them has more reaction partners than the other. Interstitial defects are first formed when substitutional defects react with an interstitial by switching places. This happens very fast, usually within the first 5% of the whole implantation process. From then onward, the interstitial defects react to more complex defects, while at the same time new interstitial defects are produced from the substitutional defects. Thus some kind of equilibrium is generated between the two, that can be observed for 'P_S'/'P_i', 'B_S'/'B_i' and 'C_S'/'C_i'.

Phosphorus. The concentration values closer to the surface are basically zero, almost all the substitutional phosphorus reacts away. The reason why only phosphorus and not also boron or carbon exhibit this, is because equivalent defects are missing for carbon and boron. Defects 'V_2P_S' and 'P_iI' consume almost 100% of all the phosphorus in the irradiated region. 'VP_S' only has a very low concentration because despite being generated in vast amounts, it immediately reacts to form 'V_2P_S'. Since a lot of vacancies and interstitials are present in the irradiated region during the implantation (as they are steadily implanted), a lot of these aforementioned phosphorus containing defects are formed, thus consuming basically all of the 'P_S'. Figure 11 shows the concentrations of the most abundant defects that contain phosphorus.

Carbon. The concentration of 'C_S' shows a bigger decrease at the projected range than 'B_S' because of defects 'C_iOI' and 'C_iOI_2' that do not have equivalents for 'B_S'. Since 'B_S' is missing those two defects, more 'B_S' and 'B_i' remains at the projected range. Figure 12 shows the dominating defects that contain carbon. In the projected range, 'C_iOI' and 'C_iOI_2' dominate with a concentration three times higher than 'C_i' and 'C_S'. At the projected range, 'C_iH' takes the lead, and almost reaches the concentration level of the initial concentration of substitutional carbon before the beginning of the process. The discontinuities of 'C_iH' in the irradiated region again originate from the numerical

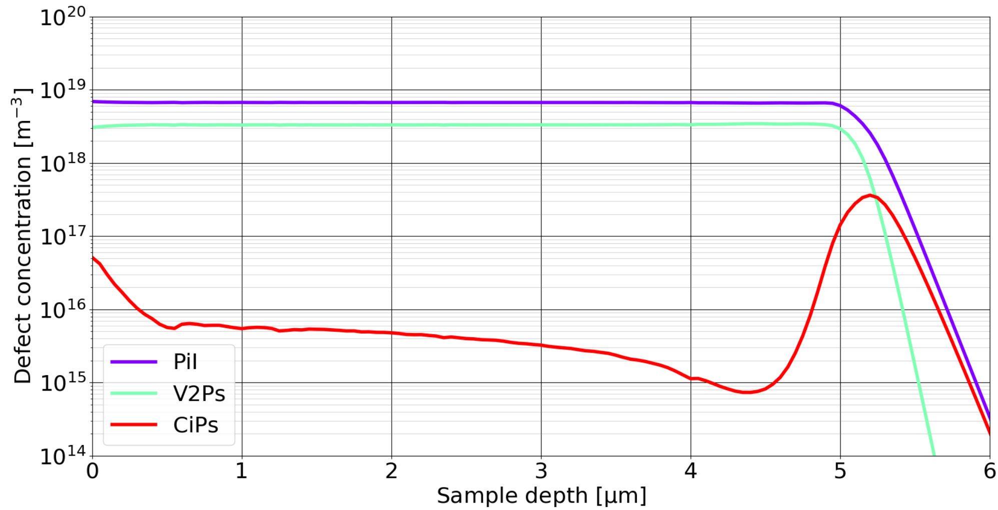


Figure 11: Concentrations of the most abundant phosphorus-related defects. While 'VP_S' is very frequently formed, it does not show in this figure, because almost all of it immediately reacts away and forms 'V_2P_S'.

discretization. 'C_iC_SI' and '$C_iC_SI_2$' show the same trends as 'C_iOI' and 'C_iOI_2', but have a lower concentration by about an order of magnitude, because oxygen is a more abundant reaction partner than 'C_i' or 'C_S'.

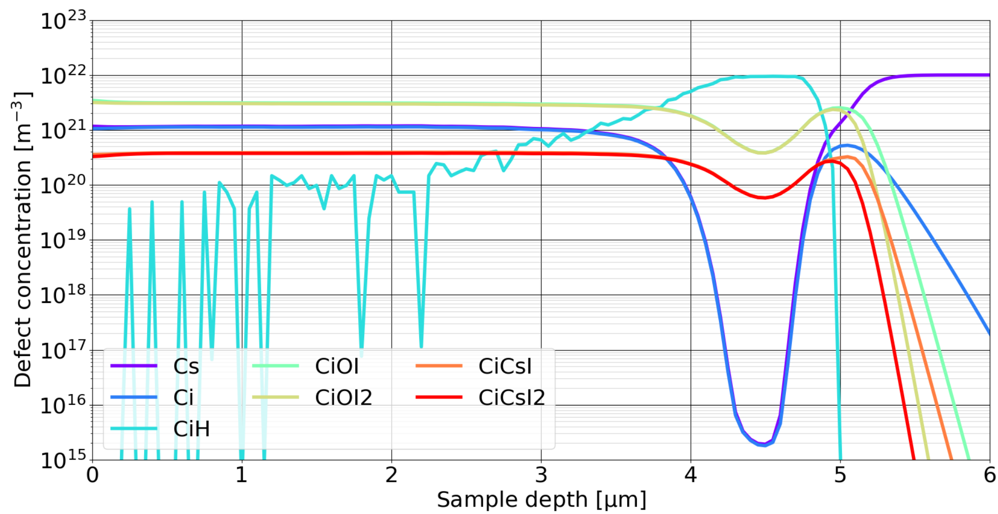


Figure 12: Concentrations of carbon-related defects with highest concentrations at a temperature of 323 K.

Oxygen. As shown in figure 13, the concentration of oxygen in the irradiated region is reduced to about 15-16% of the initial concentration. The oxygen mostly reacts to 'V_3O' (16%), 'V_2O' (15%), 'VO' (15%), 'IO' (14%) and 'I_2O' (13%). The percentages are calculated from the initial concentration of oxygen (5×10^{22} m^{-3}). 'C_iOI' and 'C_iOI_2' are also present with about 6% of the initial oxygen concentration. At the projected range these relations barely change. The concentrations of 'C_iOI' and 'C_iOI_2' decrease slightly at the implantation depth, probably due to the lack of 'C_i'. Also in the implantation depth at rather low concentrations one can see members of the 'VOH'-family.

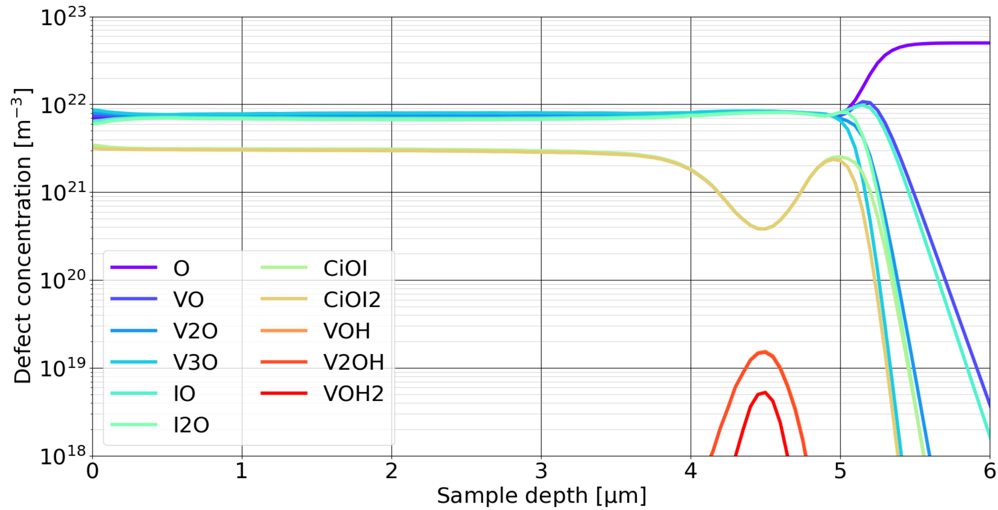


Figure 13: Concentrations of oxygen-related defects with highest concentrations at a temperature of 323 K.

Boron. About half of the boron in the irradiated region reacts with oxygen to form 'B_iO' (green line in figure 14). Just like for carbon, there is an equilibrium between interstitial and substitutional boron ('B_i' and 'B_S') who still make up for about 16% of boron each. Roughly a fifth of the boron can be found in 'B_iC_S' (18%). 'B_iB_S' is also present, but can be disregarded due to its very low concentration (almost four orders of magnitude below the concentration of 'B_iO'). All of these mentioned defects show a decrease at the projected range. This decrease is more pronounced (about four orders of magnitude) for both 'B_i' and 'B_S', as they are essential constituents for other, more complex defects. Similar to oxygen and also carbon, the concentration of boron defects containing hydrogen rises in the projected range. Most importantly, defect 'B_iH' has the highest concentration there, with values that even exceed the initial concentration of 'B_S'. This can only explained by a chain of events that includes reaction from 'B_S' to 'B_i' and simultaneous diffusion.

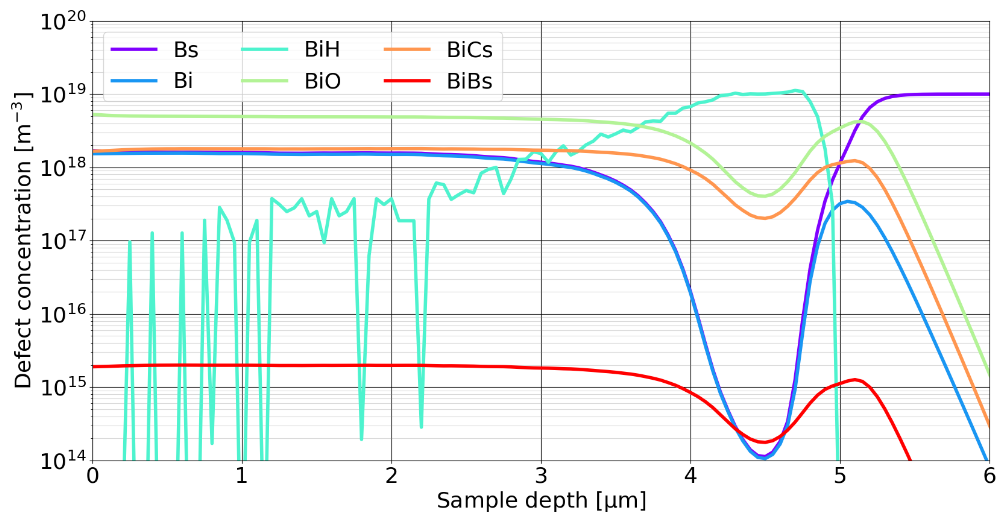


Figure 14: Concentrations of boron-related defects with highest concentrations at a temperature of 323 K.

Table 2: Different defect families.

Defect Family	**Family Members**
B_i - complexes	B_i, B_iH, B_iH_2, B_iH_3, B_iB_S, B_iC_S, $B_{i2}B_S$, B_iO
C_i - complexes	C_i, C_iC_S, C_iC_SI, $C_iC_SI_2$, C_iO, C_iOI, C_iOI_2, C_iOH, C_iH, C_iP_S
I - cluster	I - I_8
V - cluster	V - V_6
OH - complexes	OH, OH_2
IO - complexes	IO, I_2O
VO - complexes	VO - VO_3, V_2O - V_2O_3, V_3O - V_3O_3
VH - complexes	VH - VH_6, V_2H - V_2H_6
VOH - complexes	VOH, VO_2H, VOH_2

Defect Families. Figure 15 shows the concentration of various defect families. A defect family is composed of defects that share a common defect (e.g. 'B_i' or a group of defects (e.g. 'V_nO_m')). A full list of the defect families and their members can be found in table 2. In the irradiated region, vacancies and interstitials, who have not already recombined to form VI, rather bind with oxygen to form 'VO'- and 'IO'-complexes instead of forming clusters (i.e. I - I_8 or V - V_6). At the implantation depth the concentrations of pure clusters of vacancies and interstitials exceed the concentration of 'VO'- and 'IO'-complexes by about an order of magnitude. The 'VH'-family shows the largest concentration at the implantation depth. However, their concentration drops sharply and becomes negligible at 5 microns. The concentration of other defect families starts to decrease at 5 microns; this reduction occurs at a slower rate than for 'VH'. A possible explanation for the rapid decay of 'VH' is the slow diffusion of hydrogen at 323 K due to which it is not available in regions beyond the implantation depth. Both the 'B_i'- and 'C_i'-families show a very constant concentration in the irradiated region. Their concentration roughly equals the initial concentration of the respective substitutional defect of 10^{22} m^{-3} for 'C_S' and 10^{19} m^{-3} for 'B_S'. At about 4.7 microns the concentrations even slightly exceed these values (due to a combination of reaction and diffusion). In both cases, the defects 'C_iH' and 'B_iH' are the dominating defects in their families in the irradiated region. The concentration of 'OH'-complexes is very low because it only features 2 members. These members very frequently react with vacancies to become members of the 'VOH'-family. The 'VOH'-family also features a low concentration, partly because it is a small family with only 3 defects. Moreover, the 'VOH'-family loses concentration to the less complex 'VH'-family, which features the highest concentrations in the projected range. Between defects that belong to category 'VH_m' and 'V_nH_m' a pairwise equilibrium can be found: concentrations of, for example, 'VH_2' and 'V_2H_2' are almost exactly the same. This is true for the bigger complexes as well ('VH_3' and 'V_2H_3', etc...), but their concentration is lower than the one of smaller complexes. This is because the "predecessors" first have to be formed by reactions.

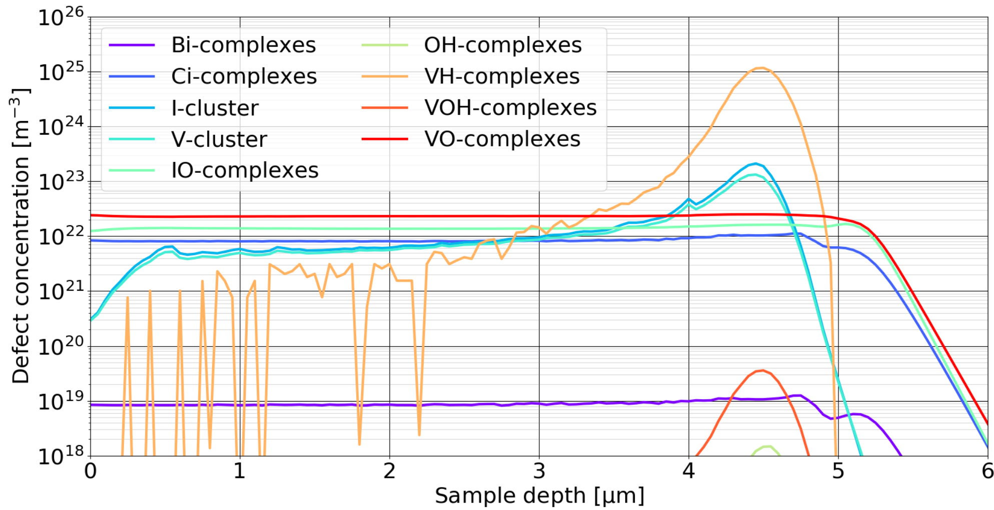


Figure 15: Comparison of the concentration of different defect families. The concentration of every member of the family has been summed to form the concentration of the defect family.

5.2 Influence of Process Temperature

Changing the temperature influences both the diffusion coefficient and the dissociation rates (equations 4.3 and 4.10). Due to being in the denominator in a negative exponential function, rising temperatures will drastically increase both quantities. Further, the reaction rates are affected, because they are made up of the diffusion coefficients. Last chapter was dealing with low temperatures of 323 K. Other simulations were performed at temperatures of 373 K and 423 K. This chapter focusses on the comparison of the concentration of several defects at different temperatures.

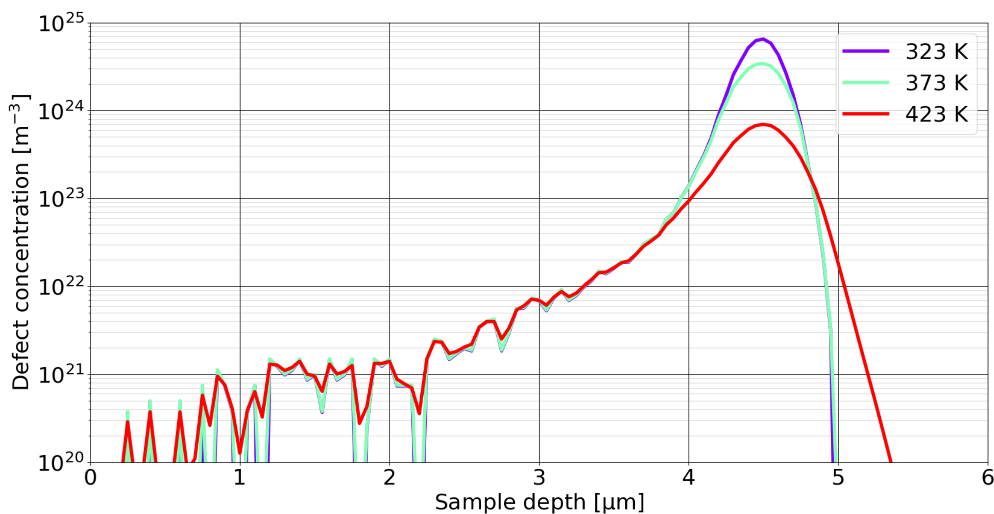


Figure 16: Concentration of hydrogen in the wafer at different temperatures.

Hydrogen. Figure 16 shows the concentration profiles of hydrogen at 323 K (purple curve), 373 K (cyan curve) and 423 K (red curve). Again, the discontinuities up until about 2 microns stem from the process of discretizing the sample into the 201 grid points. It is possible to see that at higher temperatures these discontinuities are less pronounced and the curve becomes smoother. The diffusion equalises the differing concentrations and is faster for higher temperatures. Further differences can be

noticed at the projected range at 4.5 microns and beyond. The peak concentration decreases for higher temperatures. The difference between the purple and the cyan curve are mainly due to the increased reaction rates. Even though the diffusion is already stronger (which can be seen by comparing the slope of both curves after the peak), it is basically negligible. Increasing the temperature even more leads to a further decreases of maximum concentration. This is a compound effect of the occurring reactions and the diffusion. Due to the higher temperature, even more hydrogen reacts away. On the other hand, the increased diffusion is now clearly visible by the slope of the red curve from 5 microns onwards. Integrating over these curves in order to determine the total amounts of defect 'H' in the whole simulated sample reveals that the number is reduced by about 83% when increasing the temperature from 323 K to 423 K.

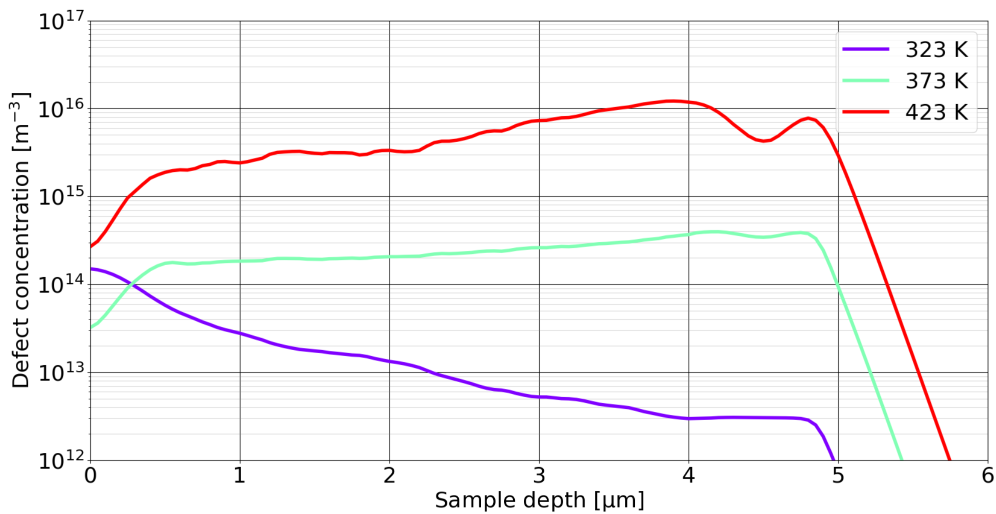


Figure 17: Concentration of vacancies in the wafer at different temperatures. The concentration values are rather low, compared to the ones of other defects. This reduces the impact of vacancies on the properties of the wafer.

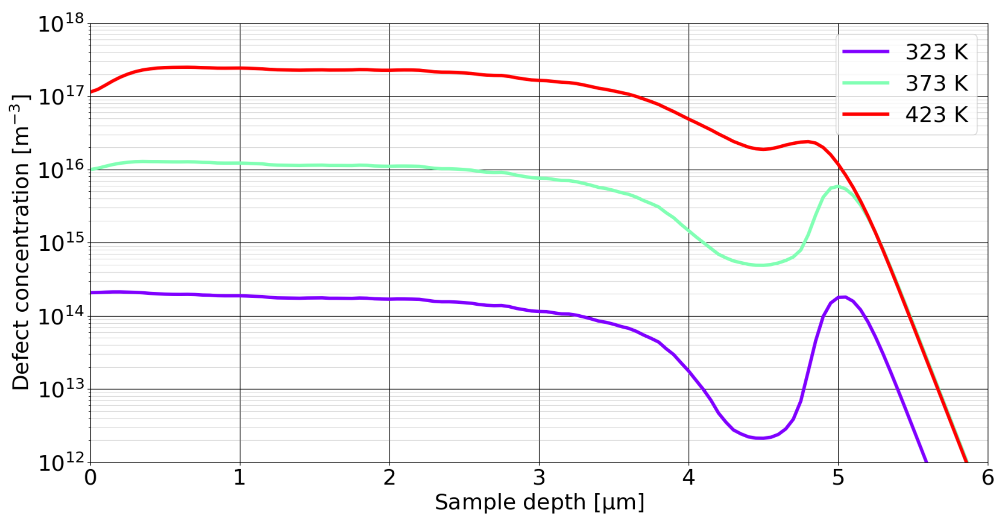


Figure 18: Concentration of interstitials in the wafer at different temperatures.

Vacancies and interstitials. A significant impact of temperature can be seen when looking at the concentrations of vacancies (figure 17) and interstitials (figure 18). Their concentrations increase by 1-2 orders of magnitude for an increase in temperature of 50 K. Since their concentrations are rather

low compared to other defects, it is hard to tell where this increase in concentration comes from; if the excess of vacancies would, for example, originate from a dissociation of 'VH', the corresponding reduction in the concentration of 'VH' would not be noticeable, as its concentration at the implantation depth is still about 100 million times higher than the one of 'V' for 423 K. Moreover vacancies can be part of 33 different defects. Since these 33 defects continually transform into each other, a lot of interplay happens and it is hard to pinpoint a single origin for the increase in concentrations of vacancies and interstitials at higher temperatures. Further, the concentrations of both vacancies and interstitials are not very stable, since they are both part of many reactions and dissociations. The profiles shown in figures 17 and 18 are merely snapshots of quite rapidly changing concentrations.

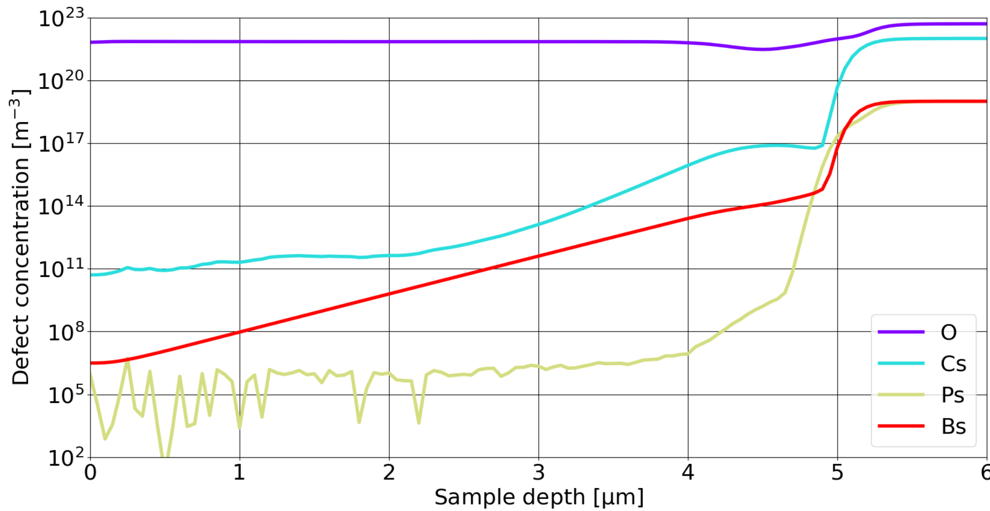


Figure 19: Concentrations of the main impurities at a temperature of 423 K.

Comparing the concentrations of the impurities in one graph makes it confusing and unreadable. Figure 19 shows those defect concentrations at a high temperature (compared to figure 10, which shows them at a low temperature). To make comparing these concentrations easier, they are treated separately in upcoming graphs, that independently show and discuss the concentrations of oxygen, carbon, boron and phosphorus. Straight away it can be said, that major differences can be seen for both carbon and boron, as they now show the same behaviour as phosphorus at lower temperatures.

Oxygen. The concentrations of the dominating oxygen-related defects do not change so much in the irradiated region (figure 20). 'V', 'VO', 'V_2O', 'V_3O' and 'IO' show a lower concentration only at the projected range. This is due to the appearance of defects 'VOH', 'V_2OH' and 'VOH_2 who now dominate in the implantation depth. Higher temperatures appear to promote the formation of defects that belong to the 'VOH'-family, because the 'VOH'-related concentration rises by about three orders of magnitude as seen in figure 13 (which is also explained later in this section, figure 25). Defects 'C_iOI' and 'C_iOI_2' gained some concentration due to the higher reaction rates, caused by the higher temperatures. The dip at the projected range at 4.5 microns at 323 K is replaced by a quite homogeneous concentration up until about 5 microns. 'I_2O' loses concentration throughout the sample, especially from the implantation depth onwards. This is probably due to the fact that 'I_2O' is able to dissociate. On top of a possible dissociation it can be converted to 'C_iOI_2' by reacting with 'C_i' and thus lose more concentration. It can further be assumed, that the pair of 'C_iOI' and 'C_iOI_2' is in an equilibrium as their concentrations are almost alike.

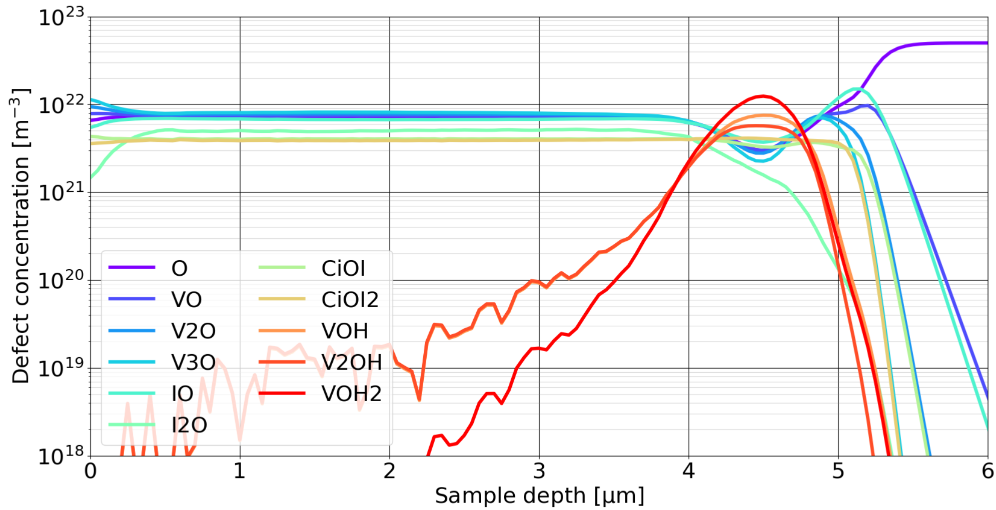


Figure 20: Concentrations of the dominating oxygen-related defects in the wafer at 423 K.

Carbon. 'C_S' now shows a decrease in concentration, but is far from depleted (figure 21). For increasing temperatures, substitutional and interstitial carbon are more and more expelled from the irradiated region. As a consequence the concentration of defects 'C_iC_SI', '$C_iC_SI_2$', 'C_iOI' and 'C_iOI_2' increase slightly. Furthermore, the dip in concentration at the implantation depth in figure 12 is gone and all four defects show a homogeneous concentration up until 5 microns. 'C_iH' dissociates at a higher temperature and thus shows a decreased concentration, especially in the irradiated region. The peak concentration is reduced by about an order of magnitude.

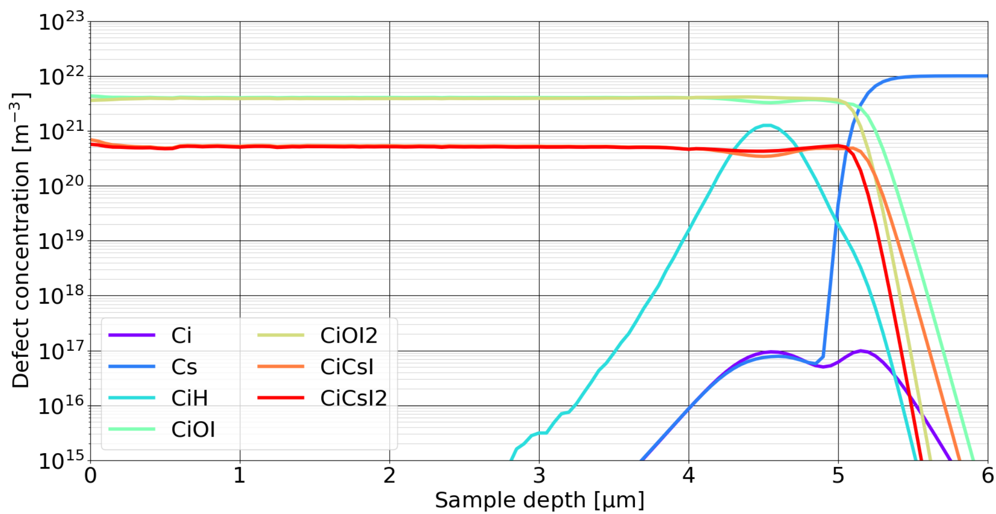


Figure 21: Concentrations of the dominating carbon-related defects in the wafer at 423 K.

Boron. At 423 K about 90% of all the boron in the irradiated region is present in defect 'B_iO'. About 10% can be found in 'B_iC_S' (figure 22). Almost every other defect containing boron loses concentration in every region of the sample (compared to figure 14). Defects 'B_i' and 'B_S' lose about ten orders of magnitude in the irradiated region. While 'B_S' resumes its initial concentration shortly after the projected range, the concentration of 'B_i' is also decreased in that region for elevated temperatures. The concentration of 'B_iH' also features a substantial decrease in every region of the sample for the same reasons as 'C_iH'. At the projected range, 'B_iH' only plays an inferior role at higher temperatures.

It can be assumed, that hydrogen will rather bind to other defects than boron (probably with members of the VOH- and VH-family), and that the remaining boron is solely used to form 'B_iO'.

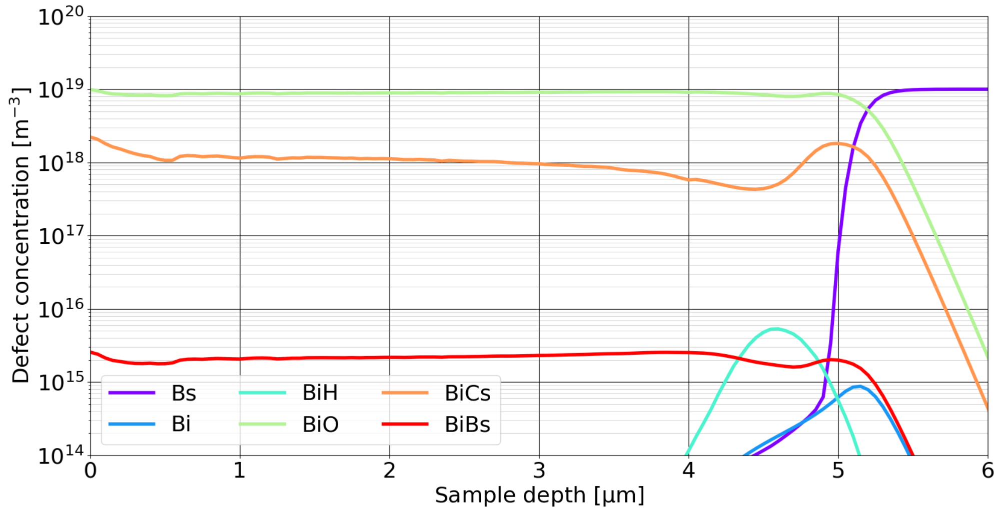


Figure 22: Concentrations of boron-related defects with highest concentrations at a temperature of 423 K.

Phosphorus. Phosphorus features one major change when increasing the temperature to 423 K (figure 23). Comparing to figure 11, the concentration of 'C_iP_S' is increased by two orders of magnitude and contains roughly 6% of all the phosphorus in the irradiated region. This gain in concentration comes from a slightly decreased amount of 'P_iI', which still holds about two thirds of the total amount of phosphorus. Defect 'V_2P_S' remains unchanged. A comparison of the different main impurities for a high temperature can be found in figure 19.

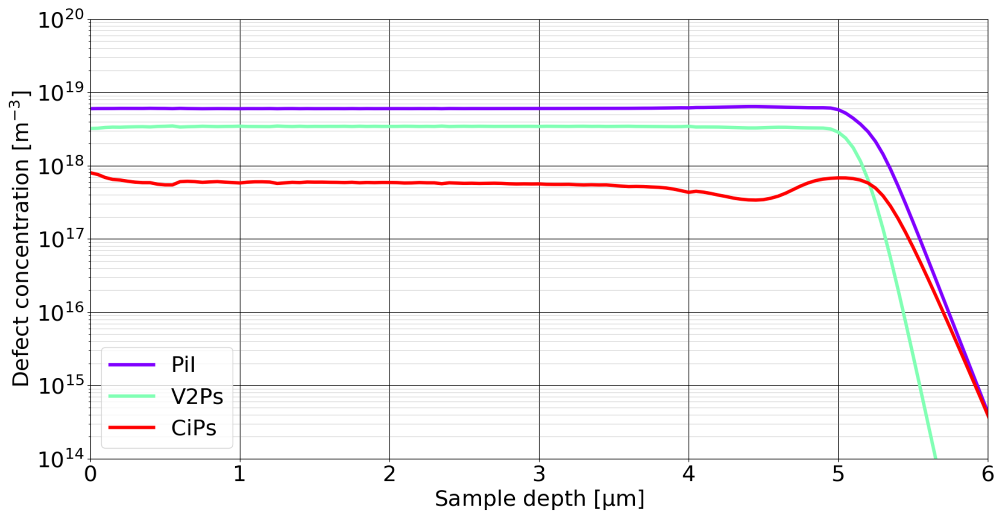


Figure 23: Concentrations of phosphorus-related defects with highest concentrations at a temperature of 423 K.

Defect families. Figure 24 shows the concentrations of the defect families listed in table 2 at a temperature of 423 K (the concentration profiles of the same defects at 323 K can be found in figure 15) . The concentration of B_i- and C_i-complexes slightly increase towards the initial value of the concentration of their substitutional counterparts B_S and B_S.

Vacancy- and interstitial clusters remain unchanged in the irradiated region. They show slight changes in the projected range. Here, the concentration of vacancy-clusters forms a shoulder after the peak at about 4.75 microns. This is probably due to the diffusion of vacancies in this region. The concentration of interstitial-clusters more than triples at the implantation depth. The peak becomes wider, which is due to the diffusion of defects 'I' and 'I_2'. It is assumed that the increase in concentration originates from a decrease in concentration of 'IO'-complexes.
The concentration of 'VO'-complexes is decreased by about 80 % in the implantation depth at 4.5 microns, but stays the same everywhere else. This loss in concentration can be explained by the increase in concentration of 'VOH'-complexes by about 3 orders of magnitude in both the irradiated region and the implantation depth. Due to the increased diffusion, the peak of the 'VOH'-complexes is broadened. 'OH'-complexes show the same trend as 'VOH'-complexes with the only difference being an overall lower concentration by about 1 to 2 orders of magnitude.
The concentration of 'VH'-complexes is halved at the projected range, since other defects containing H show an increased concentration in the same region. 'VOH'-, 'VOH'- and 'VOH'-complexes now show a less steep drop in concentration at 5 microns and deeper. Since the higher temperature causes a faster diffusion, hydrogen atoms can now move deeper into the wafer.

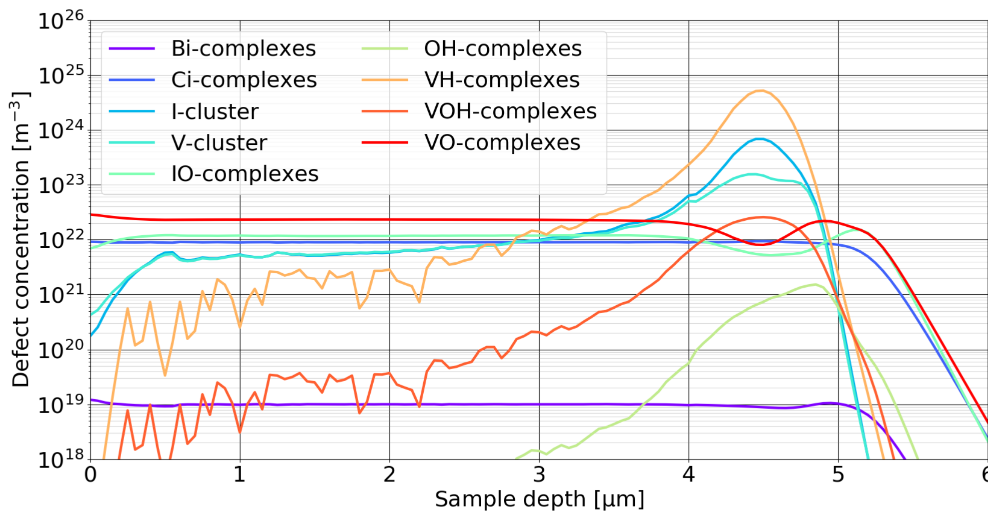


Figure 24: Concentrations of different defect families at a temperature of 423 K.

VH_n**-defects.** Figure 25 shows the concentration of defects 'VH'-'VH_6' at a low temperature of 323 K. In comparison, figure 26 shows the same defects at 423 K. At the projected range, all complex defects except 'VH' increase their concentration by several orders of magnitude. 'VH_6' even is the most abundant defect in that region of all the defects, neglecting 'VI' as this is essentially the regular crystal lattice. The same behaviour can be observed for related defect "families" such as 'V_2H'-'V_2H_6'. The total concentration of the defects shown in figure 26 is decreased compared to figure 25. This is because the concentration of VH is decreased by a factor of about 10. The concentration gain of the other defects does not compensate for this loss, which is why the overall concentration of VH-complexes decreases when a higher temperature is used.

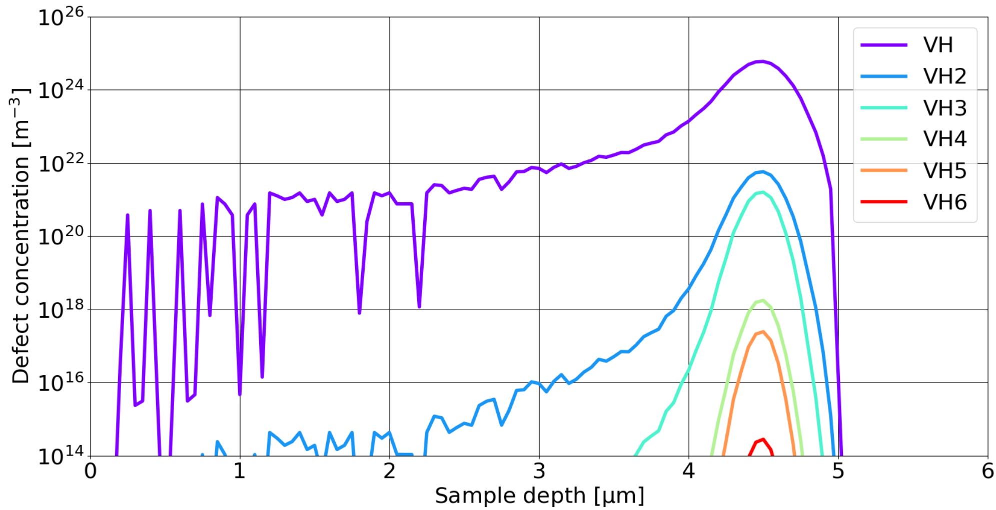


Figure 25: Comparison of defects 'VH'-'VH_6' at a temperature of 323 K.

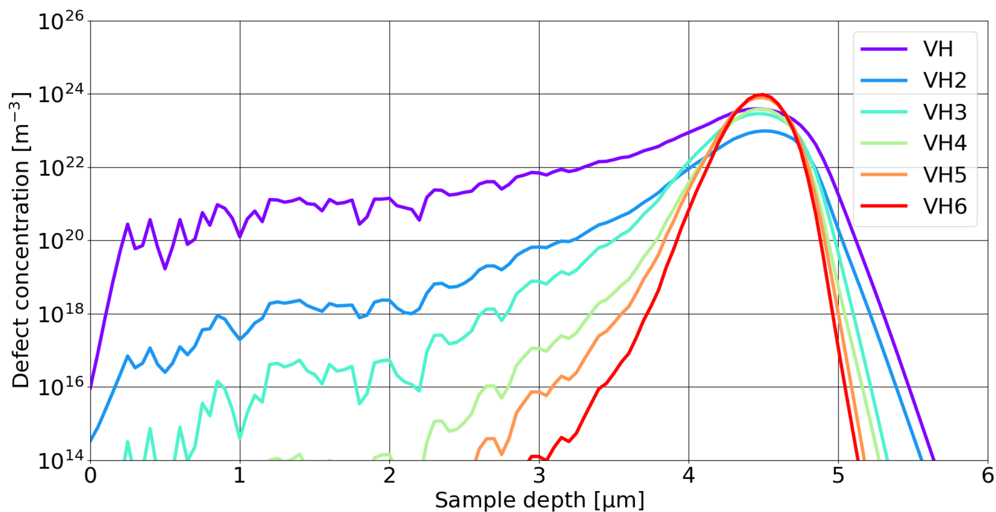


Figure 26: Comparison of defects 'VH'-'VH_6' at a temperature of 423 K. More complex defects clusters drastically increase their concentration at the projected range.

5.3 Influence of Implantation Current

Increasing the implantation current increases the amount of hydrogen atoms shot at the wafer in a certain period of time. The duration of the process and also the simulation time will be decreased, as the target dose will be reached in a shorter time. The change in duration follows an indirect proportionality, where doubling the implantation current will halve the duration of the process. The enhanced current will also impact the main processes occurring in the wafer, leading to a different outcome. For each concentration profile shown, where the implantation current is first varied at a low process temperature. That way, the influence of changing the implantation current is decoupled from temperature-related effects. Then, the implantation is varied at a high process temperature.

Hydrogen. Figure 27 shows the concentration of hydrogen at a temperature of 323 K for three different implantation currents. The concentration profile for hydrogen basically remains unchanged, even though the same amount of hydrogen was introduced at different rates. This implies, that

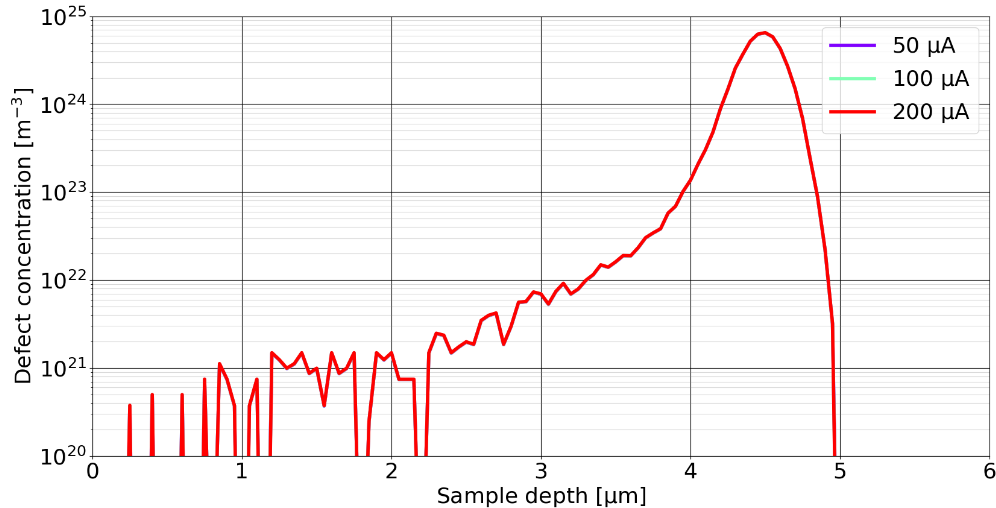


Figure 27: Concentrations of hydrogen H in the wafer depending on different implantation currents at a temperature of 323 K. Due to the speed of reactions and dissociation no differences can be seen. The temperature is too low for the diffusion to take any noticeable effect.

the change in duration caused by differing implantation currents has no effect on the reactions and dissociations, as they happen much faster. The diffusion is also not affected, as for hydrogen it is only noticeable at higher temperatures (see also figure 16). Using a higher temperature of 423 K differences are finally visible (figure 28). For low currents, the discontinuities are less pronounced. This is solely due to the influence of the implantation current, as all curves were recorded at the same temperature. For lower currents, the concentrations are subject to diffusion for a longer period of time, resulting in a smoother graph. This can also be seen in the region beyond the projected range, where the decrease of concentration is slower for lower currents due to an increased simulation time. At the projected range a lower concentration of hydrogen can be found for lower currents, as there is more time to form larger H-containing complexes. The peak concentrations of hydrogen for higher temperature (regardless of current) are almost an order of magnitude lower than for lower temperatures. This reduction can be explained by the increase in the reaction rates due to a higher temperature.

Vacancies and interstitials. Again, it is very hard to determine and predict the behaviour of vacancies and interstitials. Especially for low temperatures the concentration of vacancies and interstitials vary substantially during the implantation process, as can be seen in figure 29. This graph shows the concentration of vacancies over time at low temperatures and a medium implantation current. The concentrations vary by almost 2 orders of magnitude for every 10% of the implantation process. There seem to be three areas of concentration, where the lines reside in the irradiated: one at between 10^{14} m^{-3} and 10^{15} m^{-3}, one between 10^{16} m^{-3} and 10^{17} m^{-3} and finally one at about 10^{19} m^{-3}. The lines all reach a common value of about 10^{13} m^{-3} at the projected range, and subsequently spread out to various concentrations differing by several orders of magnitude. Thus it is not meaningful to compare such data sets with each other. This behaviour is not yet completely explained. It is known that the concentration of vacancies and interstitial fluctuate quite a bit, due to the constant generation, diffusion, and reaction. It is however not evident, whether these fluctuations alone are single-handedly responsible for this peculiar time evolution of the concentration of vacancies and interstitials. A parameter artefact can be excluded to be the cause, as simulations with the same code, but different parameters (e.g. high temperatures and a high implantation current) produce consistent

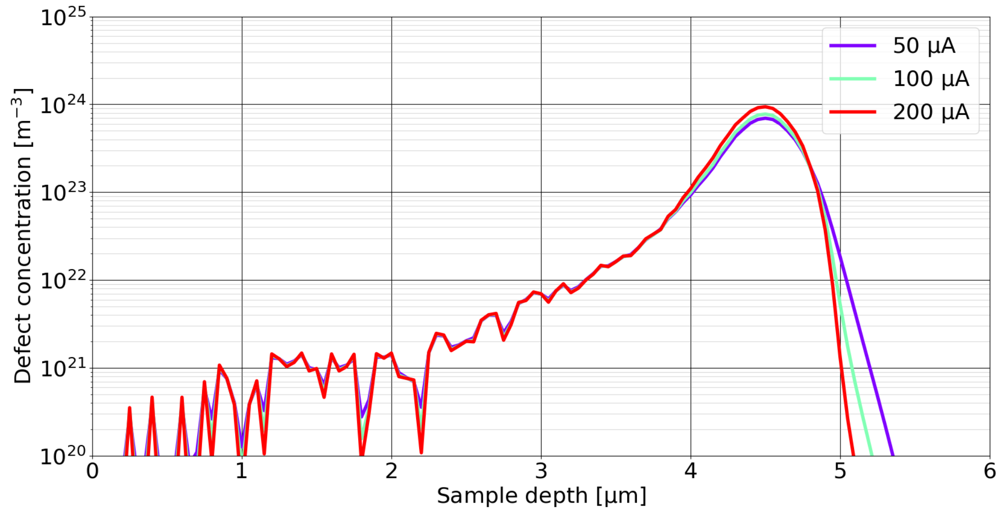


Figure 28: Concentrations of hydrogen H in the wafer depending on different implantation currents at a temperature of 423 K. Diffusion is visible in the irradiated region and beyond the implantation depth. Effects of reactions and dissociations can be seen at the projected range.

graphs. The absolute concentration of vacancies and interstitials in the sample is however very low (about ten orders of magnitude below the concentration of the most common defects), so that we refrain from further discussing.

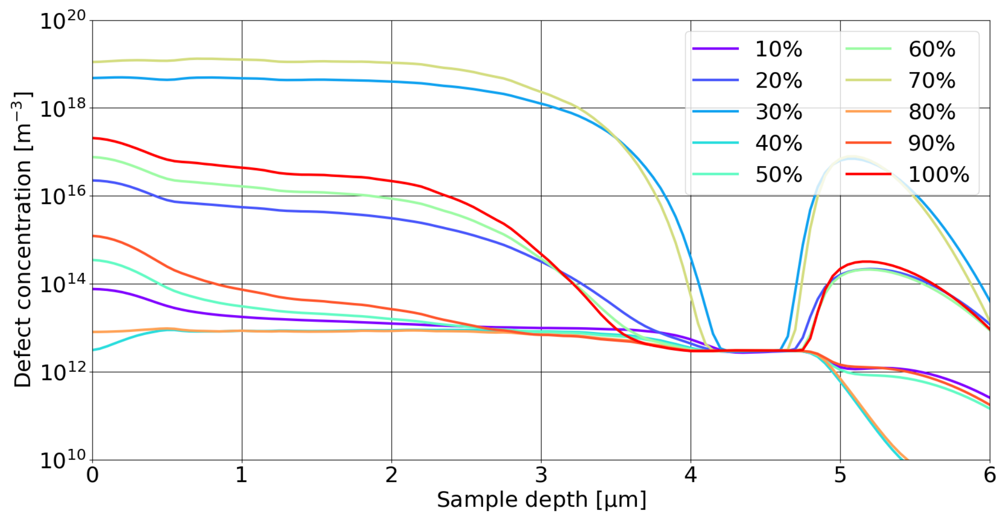


Figure 29: Concentrations of vacancies as time elapses, at a temperature of 323 K and medium implantation currents. The concentration fluctuates by a few orders of magnitude, and does so in an unpredictable, seemingly non-systematic fashion.

Main impurities. The main difference for the main impurities between high and low currents (at a low temperature) are the concentrations around the implantation depth (figure 30). Carbon and Boron show dips in concentration, that are more pronounced and wider for a lower implantation current. Their concentration in the irradiated region is lower for a lower current, as more time elapses and more defects react away. Figure 31 shows the concentration of the main impurities at 423 K and 200 μA. This has to be compared with the graphs in figure 19, where the implantation current is 50 μA at the same temperature. For lower currents, basically all of the impurities lose concentration throughout the whole sample, due to more elapsed time.

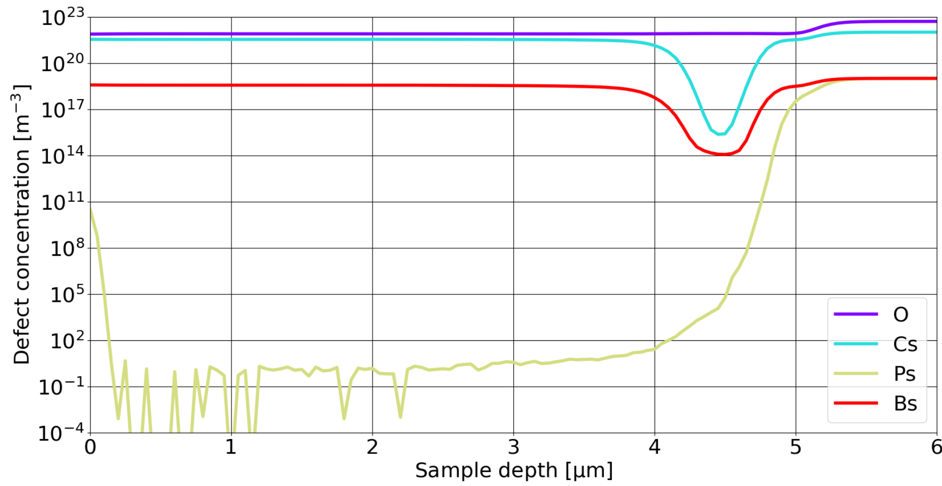


Figure 30: Concentrations of the main impurity atoms at a temperature of 323 K and an implantation current of 200 μA. Differences compared to figure 10 (323 K and 50 μA) can only be seen for carbon and boron, who show an overall slightly increased concentration and narrower valleys at the implantation depth.

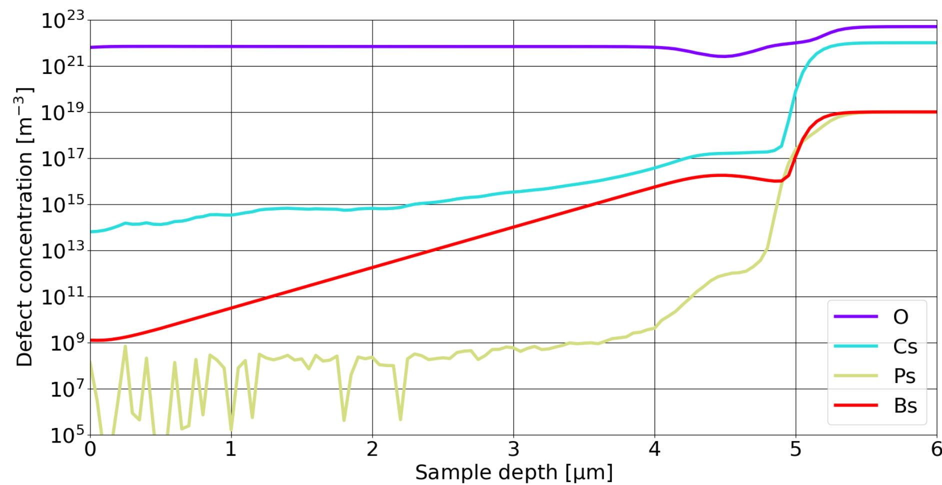


Figure 31: Concentrations of the main impurity atoms at a temperature of 423 K and an implantation current of 200 μA.

Oxygen. Compound and more complex oxygen-related defects experience a decrease in concentration for higher currents (figure 32). This visibly affects 'C_iOI' and 'C_iOI_2' as well as defects from the 'VOH'-family whose concentration is cut in half. More basic defects that hold the majority of oxygen, such as 'V_xO' and 'I_xO' show an increase of about 10% when higher currents are applied. A comparison of different currents at a high temperature shows only minor differences. As already shown in figure 20, defects of the 'VOH'-family dominate in the projected range at 423 K. While the peak concentration does not change between different implantation currents, the width of the peak becomes wider for lower currents. This widening is due to the diffusion of basic defects (e.g. 'V' and 'O') and subsequent reactions to larger complexes, that cannot diffuse anymore. The peak concentrations barely change, even though the temperature seemingly supports the formation of defects belonging to the 'VOH'-family, and the implantation process takes four times longer at a low current compared to a high current. Thus it is assumed that a maximum concentration, at least for defects belonging to the 'VOH'-family, exists.

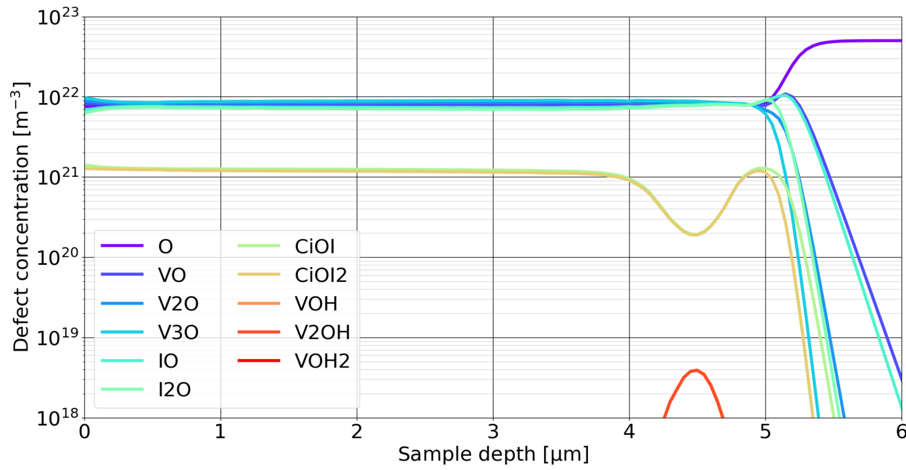


Figure 32: Concentrations of the most abundant defect containing oxygen at 323 K and an implantation current of 200 μA. A comparison with figure 13 shows lower concentrations for defects 'C_iOI' and 'C_iOI$_2$', because less time has elapsed and thus less of those defects is formed. The dominant defect species only show minor changes in concentration.

Carbon. For carbon-related defects, the implantation current at low temperatures determines the defects that dominate in the irradiated region (figure 33). At low currents, the dominating defects are 'C_iOI' and 'C_iOI$_2$' because they had enough time to form from 'C_i' and 'C_S'. At higher currents 'C_i' and 'C_S' show three times the concentration of 'C_iOI' and 'C_iOI$_2$' in the irradiated region, due to the short process time. High implantation currents are responsible for narrower peaks and dips at the projected range. Because of the low temperatures and the accompanying slow diffusion and reactions, 'C_iH' is almost unaffected by the change of the implantation current.

Looking at the differences in implantation current at a process temperature of 423 K (instead of 323 K), no major differences in the concentration profiles of carbon-related defects can be seen. In the irradiated region the concentration of 'C_iH' increases when a higher current is applied. The peak concentration of 'C_iH' is slightly increased for higher currents. This suggests, that a combination of temperature and concentration, the dissociation of 'C_iH' starts to be the dominating mechanism for that defect, causing its concentration to decrease.

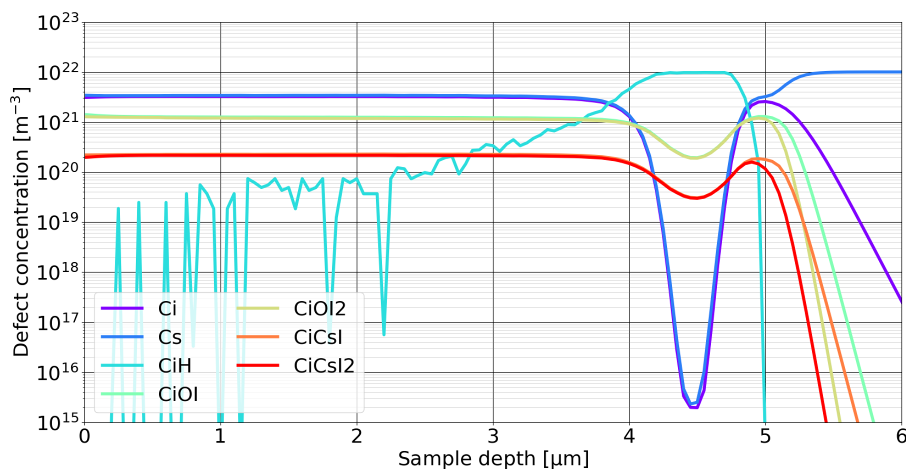


Figure 33: Concentrations of the most abundant defect containing carbon at a temperature of 323 K and an implantation current of 200 μA.

Boron. The leading defects containing boron show the same behaviour as their carbon-related counterparts (figure 34). For high currents of 200 μA, 'B_i' and 'B_S' show the highest concentrations in the irradiated region. With decreasing the implantation current, they increasingly transform to defects 'B_iO' and 'B_iC_S'. At low currents of 50 μA the concentration of 'B_iO' and 'B_iC_S' exceed the concentration of 'B_i' and 'B_S'. Defects containing hydrogen ('B_iH' in this case) are basically unaffected by changes in current and temperature. At higher temperatures, 'B_iH', 'B_i' and 'B_S' are expelled from the projected range with their concentration dropping by several orders of magnitude (figure 35). The dominating defect species at 423 K, i.e., 'B_iO' and 'B_iC_S' barely change in concentration.

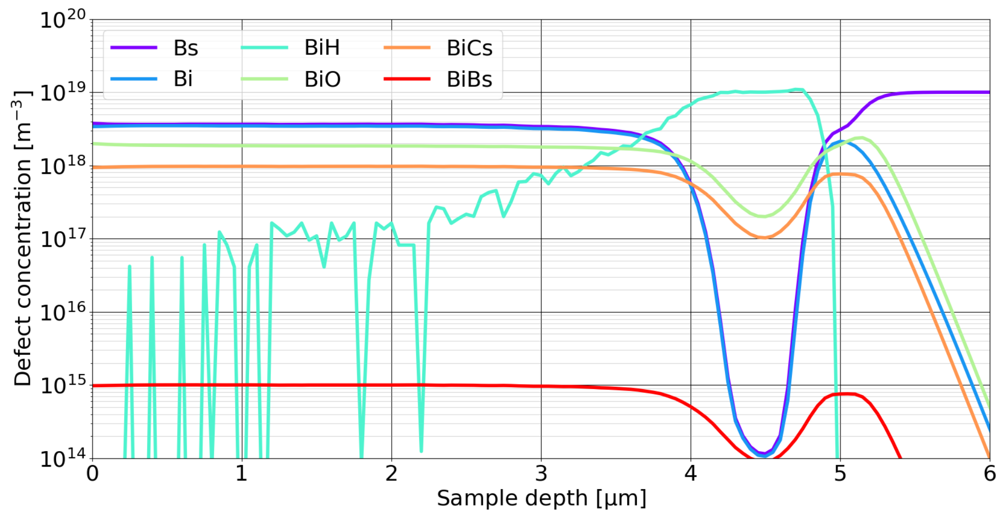


Figure 34: Concentrations of the most abundant defect containing boron at a temperature of 323 K and an implantation current of 200 μA.

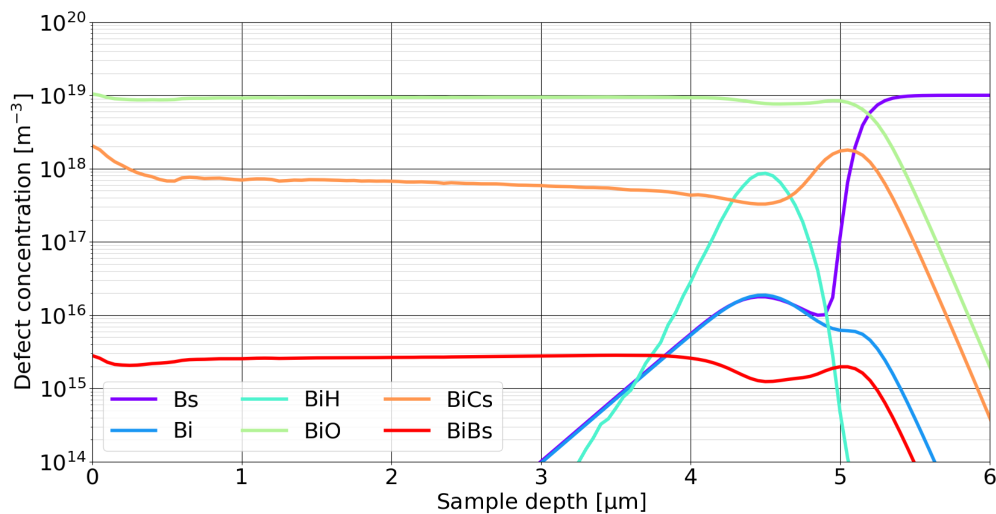


Figure 35: Concentrations of the most abundant defect containing boron at a temperature of 423 K and an implantation current of 200 μA.

Phosphorus. The implantation current does not significantly alter the concentration of phosphorus-related defects. The only difference that can be seen (for both low and high temperatures) is the concentration of defect 'C_iP_S' which slightly increases its concentration at lower currents. However, its concentration is 2-3 orders of magnitude below the one of the dominating defects species ('P_iI' and 'V_2P_s'), which is why it only plays an inferior role.

Defect families. Looking at the concentration of the different defect families some minor changes happen. Figure 36 (low temperature and high implantation current) shows an increase of interstitial- and vacancy-clusters by a factor of about 5. Defect families 'VH', 'VO' and 'IO' do not show a change at all. Boron and Carbon-complexes only increase their concentration in the irradiated region, so that they are now close to the initial concentrations of 'B_S' and 'C_S' respectively. 'VOH'- and 'OH'-complexes show a decreased concentration in the projected range. It seems that, initially, vacancies and interstitial rather form clusters and react with themselves. As time progresses 'V' and 'I' probably react away in a pairwise fashion: vacancy-clusters react with single interstitials and vice versa and irreversibly form 'VI'. In both cases, the concentration of the clusters is decreased over time. This also explains, why defect complexes like 'VO' and 'IO' are seemingly not affected by this decrease in concentration, despite having lower concentrations than the clusters. 'VOH'- and 'OH'-complexes show a lower concentration, as they need enough time and a high enough temperature to form.
The trends described in the previous paragraph can be spotted if one compares the concentration profiles at different implantation currents at a higher temperature of 423 K (see figure 37). Vacancy- and interstitial-clusters show an increased concentration for higher currents, but only in the irradiated region. The concentration of 'OH'-complexes increases. The concentration of 'VOH'-complexes does not change, probably since it is reasonably close to the initial concentration of oxygen and 'IO' in the sample. It might be the case, that dissociations compensate the increase of the concentration by reaction, thus holding the concentration of 'VOH'-complexes at a constant position.

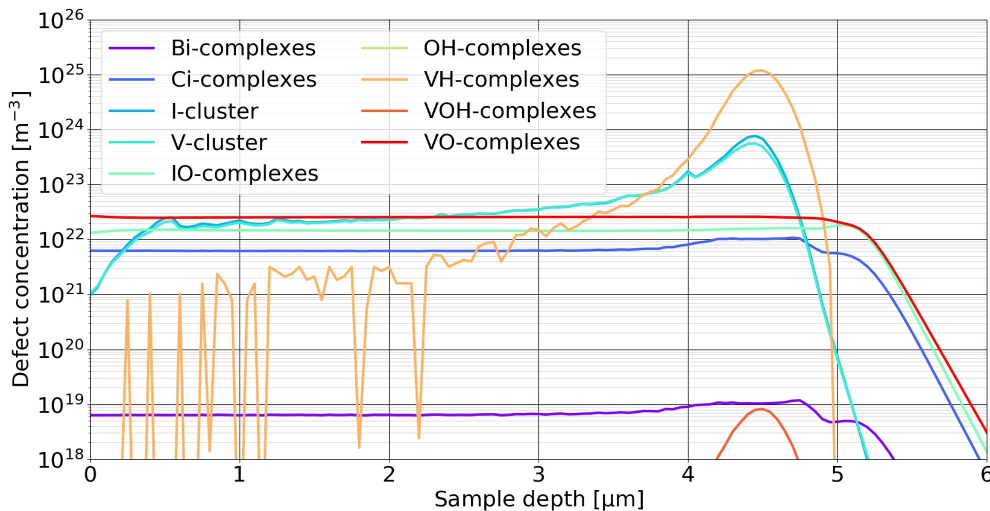


Figure 36: Concentrations of defect families mentioned in table 2 at a temperature of 323 K and an implantation current of 200 μA.

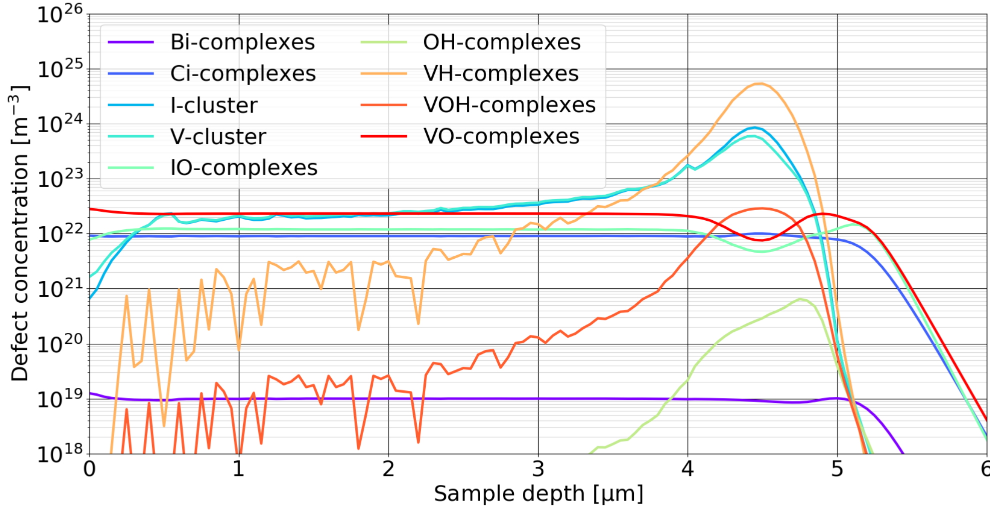


Figure 37: Concentration of defect families mentioned in table 2 at a temperature of 423 K and an implantation current of 200 μA.

5.4 Influence of Initial Carbon Concentration

As already mentioned in the beginning of this section, some simulations that feature a carbon concentration of 5×10^{21} m^{-3} have been identified as being erroneous. Thus it was decided not to take those into account and to only compare simulations that were performed at 10^{22} m^{-3} and 2×10^{22} m^{-3}. Since the main focus lies on different concentrations of carbon in the sample, carbon-related defects are prioritised in this section, but the main impurity atoms ('H', 'V', 'I', 'O', 'B', and 'P') are also examined.
Simulations will now be compared in pairs. One simulation will have a low concentration of carbon (10^{22} m^{-3}), the other one a high concentration (2×10^{22} m^{-3}). Between two simulations, other process parameters (process temperature, implantation current, and oxygen concentration) will not change. A ratio between the concentration values on every grid point is calculated. For carbon-related defects, this ratio is expected have a value of two, since the initial carbon concentration has been doubled. Other investigated defects, that do not contain carbon should exhibit a value of 1 for their ratio. Ratios below these values indicate, that increasing the initial carbon concentration negatively affects corresponding defects, as their concentration is lower than expected. The same holds in the opposite case, where defects are regarded to be positively affected, when their ratio exceeds the corresponding value. To cover all possible combinations of process parameters, 4 pairs of simulations have been investigated for every defect (temperatures of 423 K and 423 K, implantation currents of 50 μA and 200 μA).
Consequently, eight simulations are analysed for every investigated defect. This makes the use of graphs cumbersome and complicated. To give a clearer picture of this analysis, figure 38 shows an example graph of the analysis performed on the investigated defects.

Figure 38 shows four different lines, which represent the values of the previously mentioned ratios for every grid point (i.e. throughout the depth of the wafer) and every simulation pair. The investigated defect in this case was 'C_S'. It can be seen, that the values are very close to a value of two in the whole sample for low temperatures, regardless of current (red and purple curve). Higher temperatures however seem to favour the presence of 'C_S', because the ratios increase to about 2.5 for both high

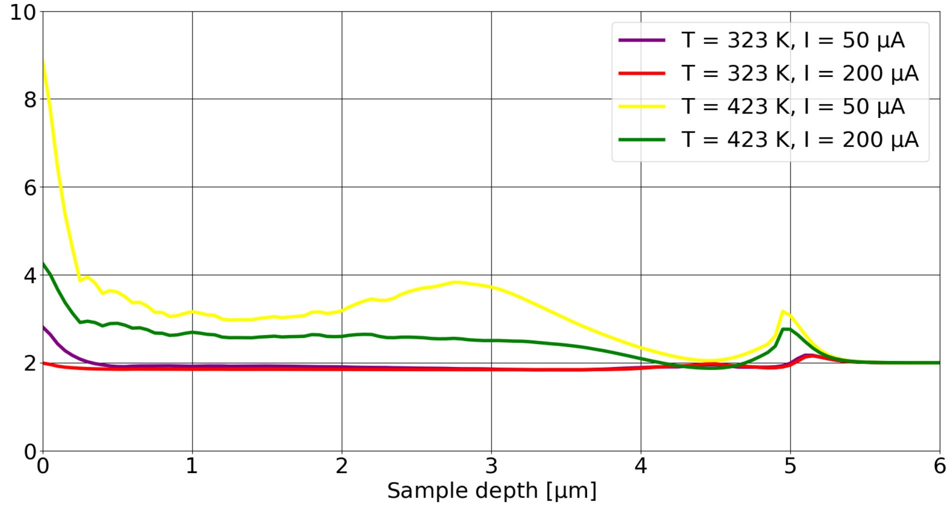


Figure 38: Ratio values for defect 'C_S'. Values higher than two indicate that more concentration is present than expected and vice versa.

temperatures and currents (green curve) and almost reach a value of 4 in the irradiated region for high temperatures and low currents (yellow curve). At the penetration depth, the ratios of all simulation pairs reach a value of two, before spreading apart again at about 5 microns. After this region, they converge to a value of two again, since this is the region, where defects only reach after a long time of diffusion.

Near the sample surface the ratios are elevated, which is believed to be an artefact of the simulation. As the defect concentration of 'C_S' is very low in this region, already tiny differences in concentration between the two compared simulations can lead to very high or low ratios. This artefact appears for almost every defect and is remedied as one goes deeper into the irradiated region, where the ratio values resume in a more constant fashion.

At the projected range basically all the examined defects show expected changes/ratios in their concentrations. Carbon-related defects show a ratio-value of approximately two, meaning that their concentration doubles when the initial carbon concentration is doubled. Defects not containing carbon show a ratio-value of about one. Unexpected behaviour can be observed in the irradiated region as well as at higher process temperatures. Many of the carbon-related defects show an increase in concentration that is higher than the expected factor of 2. This means, that raising the initial carbon concentration and increasing the temperature benefits those defects more than other defects. However this behaviour is often not spotted for all combinations of parameters. Most defects show this increase only at high temperatures (regardless of implantation current), where the ratio-values typically lie between 2.5 and 4 (meaning up to four times the concentration). Defects that show this trend are less complex and smaller, like 'C_i', 'C_S', 'C_iH' and 'C_SH'. The compound carbon defects 'C_iC_SI', '$C_iC_SI_2$' and 'C_iC_S' show this increase at every combination of implantation current and process temperature. Carbon-related defects that contain oxygen generally show a concentration which is lower than expected, when the initial carbon concentration is raised. Among them are defects 'C_iOI', 'C_iOI_2' and 'B_iC_S'. Their concentrations are lower in the irradiated region regardless of the combination of parameters. Typical ratio-values for those defects lie between 1.65 and 1.85. 'C_iO' and 'C_iOH', also show this decrease but only at low temperatures. At higher temperatures their ratios are close to two, indicating not much of a change.

Consequently, the concentration of oxygen, interstitials and vacancies (as well as compound defects like 'I_2O') show a decreased concentration as well. Even though carbon-related defects with oxygen do not show the increase in concentration by a factor of 2 (like one would expect), their concentration is still raised, causing the aforementioned defects to show a lower concentration as a result. Hydrogen does not seem to influence the carbon concentrations so much (and vice versa), as its concentration remains pretty much unchanged throughout the sample and regardless of parameters. Interestingly, 'B_i' shows more than twice the expected concentration in the irradiated region at elevated temperatures. A possible explanation for this is the lower than expected concentration of 'B_iC_S'.

5.5 Influence of Initial Oxygen Concentration

The differences in concentration are analysed the same way as for carbon in the previous section. Again, simulations will be compared in pairs, with one having a high concentration of oxygen (2×10^{23} m^{-3}) and one having a low concentration of oxygen (5×10^{22} m^{-3}). The process temperature, implantation currents and carbon contents will have the same value between two compared simulations. Consequently, the ratios between the two concentrations found in the simulations are expected as follows: if an oxygen-related defect is investigated, the ratio is expected to have a value of four, because the high initial concentration is four times higher than the low initial concentration. If oxygen-free defects are compared, the ratio is expected to be one. Deviations from these values again mean, that the increase of oxygen concentration in the sample either supports or hinders the formation of the defects. The investigation will primarily focus on oxygen-related defects that show a high concentration as well as the main impurities ('H', 'V', 'I', 'O', 'B', and 'P').

Oxygen-related defects combined with impurities as well as impurities on their own generally show a lower concentration than expected, regardless of process temperature or implantation current. This is because the concentrations of the other impurities are not raised to the extent the concentration of oxygen is enlarged. The concentration of non-oxygen impurities are limiting factors in those reactions. Their ratios are close to or slightly above one, indicating that their concentration barely changes or increases, despite the presence of oxygen. This can be observed for example for defects 'B_iO', 'C_iO', 'C_iOI' and 'C_iOI_2'. There is now comparatively much oxygen in the wafer, as the impurities cannot react with it. VO- and IO-complexes show higher concentrations than expected. Lower implantation currents and higher temperatures favour the generation of the VO- and IO-complexes even more. These deviations from the expected case are not very substantial and less distinct than in the case for changing carbon contents. On average, concentrations deviate by about ±5% from the expected results. Only defect 'OH' stood out by having between five and twenty times the expected concentration. The reasons for the favourable formation of 'OH' are the increased concentration of oxygen, and the abundance of hydrogen everywhere in the sample up to the projected range. This, in turn boosts the concentration of the 'VOH'-family, that also show a slightly higher concentration than expected. Defect 'OH_2' even shows a decrease in concentration, compared to lower initial oxygen concentrations. The concentration of defects 'H' and 'H_2' remain almost unchanged.
Almost all defects show a lower concentration after the projected range, when the initial concentration of oxygen is higher. Since more oxygen is present, reaction partners are more abundant for diffusing defects, whose motion is impeded by the fact that reactions convert them to defects that probably do not diffuse anymore. Hence, the concentration after the implantation depth is lower by about one order of magnitude for every micron deeper in the wafer.

5.6 Simulation Speed and Specifications

When looking at the simulation speed, three things are of interest that are tied to each other: the time step of the simulation (Δt), the amount of iterations performed per second, and the ratio between the simulated (process) time and the elapsed real time. The time step represents the amount of seconds the process is propelled forward in every iteration. Its value is rather high in the beginning (up to several ten milliseconds) but quickly decreases to values of 10^{-5} s and below (see figure 39), depending on the implantation parameters. The time step (Δt) is lowered if the computations yield negative

concentration values anywhere in the sample and is increased otherwise. The factors by which Δt is increased or decreased have a huge impact on the speed of the simulation. This has been more thoroughly explained in section 4.3.1.

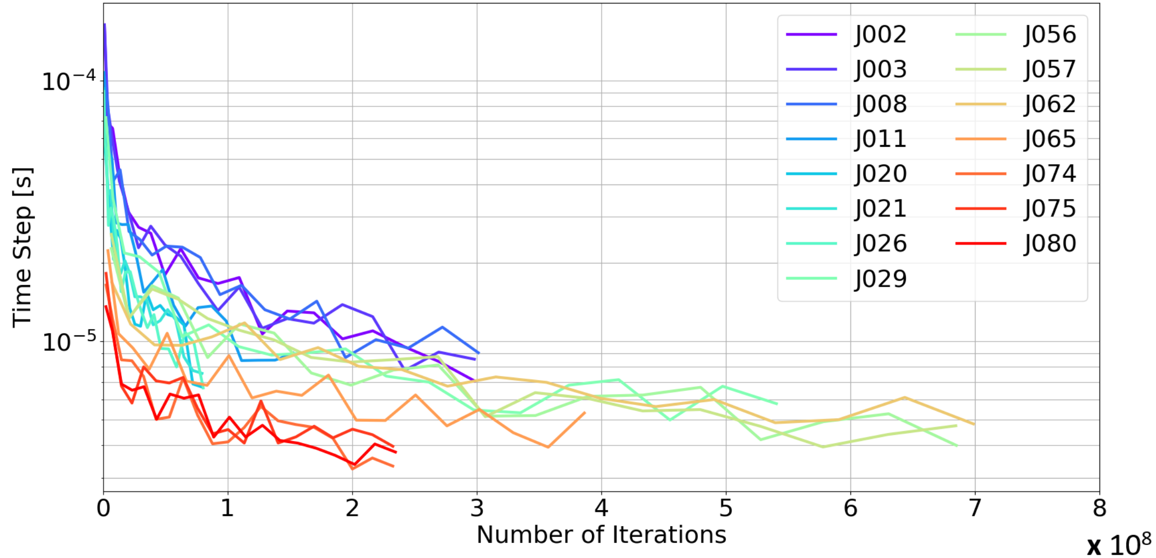


Figure 39: Values for the time steps of the investigated simulations only. An overview of the simulation numbers and their parameters can be found in table 3. Note that the values on the x-axis are multiplied with 10^8.

If the amount of iterations per second is increased, the simulation speed increases as well. The number of iterations per second is again influenced by the simulation parameters, but also on how the code is parallelized. Separating the code into many pieces and distributing it to different nodes for computation saves computational time, because results are determined in a simultaneous fashion. However this does not hold for an arbitrary number of separations, i.e. there is a limit to the amount of separate parts, where parallelization shows a decrease in computation time. Looking at averaged values for iterations per second for all the performed simulations it can be observed that they split up in to roughly two values or regions (see also figure 40): Some simulations run between 955-970 iterations per second, while others only perform at 685-700 iterations per second. This behaviour cannot be explained or tied to one specific parameter. There is the tendency towards more iterations per second when going for higher temperatures, however the temperature is not solely responsible for the increase in iteration speed. A factor, that was not measured, is the amount of "bad" iterations, i.e. iterations where negative concentrations occurred, which have to be redone. They occur more frequently if the factors that increase/decrease Δt are chosen in a bad fashion. These factors were not changed for different simulations. There is a possibility, that different sets of parameters require slightly different factors in order to run optimized. The value of "iteration per second" includes those "bad" iterations and they might occur only for those simulations, where the factors were not fitting perfectly. This might be one of the reason, the values are split up like that.

Lastly, the ratio between the simulated process time and the elapsed real time is calculated and is given in $[\frac{\mathrm{ms}}{\mathrm{s}}]$. Values are again higher in the beginning, with up to 200 $\frac{\mathrm{ms}}{\mathrm{s}}$, but they decline quickly to values of 10 $\frac{\mathrm{ms}}{\mathrm{s}}$ and below. To increase the speed of the simulation, output files have only been created every 5% in the progress of the simulation, resulting in a total of 20 data points for every

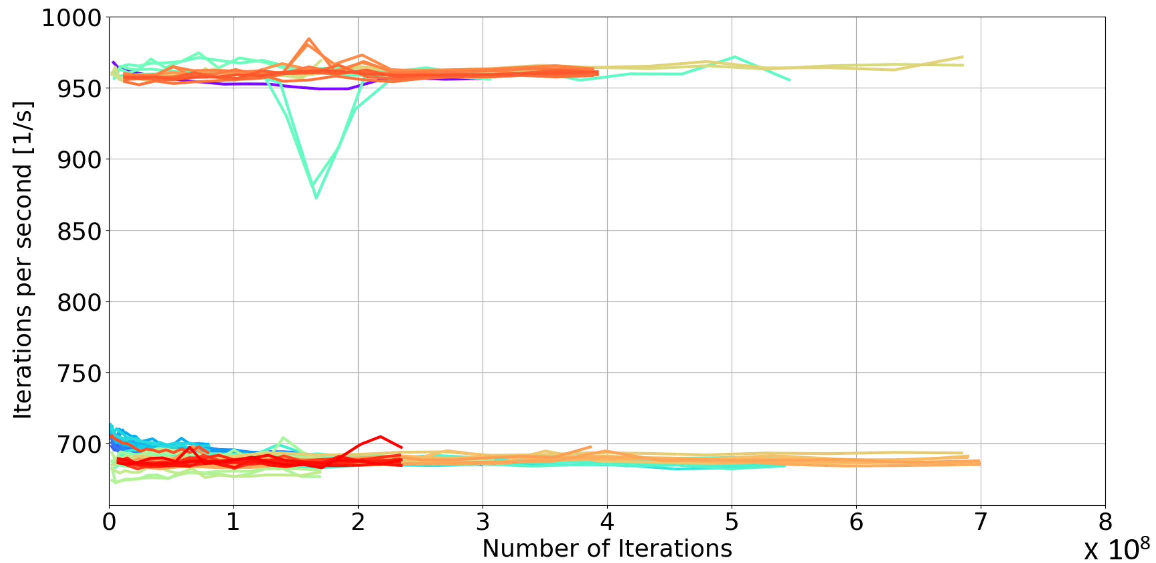


Figure 40: Values for iterations per second over the amount of total iterations for all the performed simulations in this thesis. Note: The values on the x-axis are multiplied with 10^8. Further some curves are longer than others, because not all simulations take an equal amount of iterations to finish.

simulations. Unfortunately, some of the details do not show with only so little data points (like the working of the adaptive time step, which increases and decreases the time step). General tendencies can however be observed without a problem.

From all the simulations that were performed for this thesis, the fastest one took about 31 hours to be completed (the implantation parameters were: low process temperature, high implantation current, low oxygen concentration, low carbon concentration). The slowest simulation took approximately 283 hours (11 days, 19 hours; the implantation parameters were: high process temperature, low implantation current, high oxygen concentration, low carbon concentration). The parameters used are considered to be of medium workload; simulations of greater interest include higher temperatures and higher end doses, which both increase the simulation time further. If the end dose is doubled, the simulation time is at least doubled, because the simulation speed decreases the longer the simulation runs.

From all this data, a few considerations can be made. The aim is to increase simulation speed, because simulation times of weeks (and months, if necessary annealing steps should be simulated as well) are not viable for the industry. There are some possibilities to increase the simulation speed:

1. Increasing the time step

Increasing the time step, increases the amount of time the simulation proceeds forwards in every iteration. Unfortunately, there is an upper limit to the time step, over which the results lose their physical meaning (negative concentration values). Below this limit, the results are still qualitatively wrong. A "correct" result could only be achieved by using a time step infinitely small, as the numerical approximation would then be as small as possible. As the correctness of the results should always be prioritized over the time in which the results are obtained, the time step should be kept as low as possible, while still guaranteeing fast results.

<u>2. Increasing the number of iterations per second</u>

Since the time step should rather be lowered than elevated, another option is to increase the amount of loop iteration per unit of time, to propel the simulation forwards quicker. This can be achieved by using a suited language for the task at hand and by writing efficient and fast code. Among the most popular programming languages, FORTRAN is one of the favourites for number-crunching tasks like this, but it can be argued, that using C might even benefit the speed of the simulation by small amounts. It was however beyond the scope of this work to thoroughly investigate the difference in speed between the two.

Much time was spent making the code more efficient and to perform as little as possible calculations as possible. This consists of many different steps. To begin with, it is important to set up the initial data in a away, that it can be processed fast and efficiently. Further it is necessary to employ smart code structures, that use as little computational time as possible. Finally, the use of parallel computing is inevitable to get the fastest simulation times. Not only is it important to apply parallel computing, but also to optimize it. The simulations performed for this thesis needed up to 7×10^8 iterations to finish, so saving only a fraction of a second on every iteration will have a great impact on the simulation speed.

One great possibility to increase the amount of iterations per second, is to increase the ratio between "good" and "bad" iterations. In figure 40, the number of iterations pro second only includes "good" iterations (i.e. no negative concentration was detected and the time step could be increased), but gives no information about how often an iteration was redone, to prevent negative concentration. Depending on the corresponding factors, that increase and decrease the time step, the ratio between "good" and "bad" iterations can be even higher than 75%, in favour of the bad iterations. By testing, these factors were chosen, such that this ratio is already as low as possible ($< 25\%$). Fully eliminating the "bad" iterations would increase the amount of iterations per second by 33%, given that 25% of the iterations were previously bad.

One possibility to achieve this, would be to completely remove the concept of the adaptive time step, and use a fixed one with a very low value instead. This in turn will have negative effects on the simulation speed, with the benefit of having much more accurate results. It is possible to perform some approximate calculations with the information we have, to see if this approach would be viable in the future:

In order to use a fixed time step, we have to use the lowest time step that occurred during the single simulations (or a lower one), in order to produce viable results. This makes the simulation much slower in the beginning, where otherwise the time step would be much higher. Towards the end however, the differences are only very small. By calculating the average time step throughout the whole simulation and dividing by the minimum time step, we can approximate a ratio, by which the simulation with an adaptive time step is faster than without it. These values lie between 1.4 and 4.4, for all the performed simulations. Additionally, there are no "bad" iterations anymore, as they are only caused by a time step chosen too high. Thus we can assume an increase in simulation speed of 33% at best, based on a ratio between "good" and "bad" iterations of 75% in favour of "good" iterations. The fastest simulation would then be done in about 54 hours (compared to 31 hours before) and the slowest would take 425 hours (compared to 283 hours before). Contrary to the objective, this approach increased the simulation time. However, we can expect these results to be more accurate, because the time step was not held as high. It has yet again to be decided, if the additional simulation time is worth the more accurate results.

Table 3: Simulation numbers and associated parameters of investigated simulations.

Simulation Number	**Temperature [K]**	**Implantation Current [μA]**	**Oxygen Concentration [m^{-3}]**	**Carbon Concentration [m^{-3}]**
J002	323	50	5×10^{22}	10^{22}
J003	323	50	5×10^{22}	2×10^{22}
J008	323	50	2×10^{23}	10^{22}
J011	323	100	5×10^{22}	10^{22}
J020	323	200	5×10^{22}	10^{22}
J021	323	200	5×10^{22}	2×10^{22}
J026	323	200	2×10^{23}	10^{22}
J029	373	50	5×10^{22}	10^{22}
J056	423	50	5×10^{22}	10^{22}
J057	423	50	5×10^{22}	2×10^{22}
J062	423	50	2×10^{23}	10^{22}
J065	423	100	5×10^{22}	10^{22}
J074	423	200	5×10^{22}	10^{22}
J075	423	200	5×10^{22}	2×10^{22}
J080	423	200	2×10^{23}	10^{22}

Table 4: Overview over all the present defects, their parameters and sources for their parameters.

Defect Name	**Diffusion constant** [cm^2/s]	**Activation energy** [eV]	**Source of Diffusion Constant**	**Attempt Frequency** [1/s]	**Binding energy** [eV]	**Source of Attempt Frequency**
B_i	4×10^{-4}	0.60	[5]	0	0	-
B_iH_2	0	0	-	2.8×10^{14}	1.28	assumed
B_iH_3	0	0	-	2.8×10^{14}	1.28	assumed
B_iB_S	0	0	-	0	0	-
$B_{i2}B_S$	0	0	-	0	0	-
B_iC_S	0	0	-	0	0	-
B_iH	0	0	-	2.8×10^{14}	1.28	[16]
B_iO	0	0	-	0	0	-
B_S	3.79	3.65	[5]	0	0	-
B_SH	0	0	-	2.8×10^{14}	1.28	[16]
C_i	4.40	0.88	[5]	0	0	-
C_iC_S	0	0	-	0	0	-
C_iH	0	0	-	1.0×10^{13}	1.32	assumed
C_iO	0	0	-	0	0	-
C_iOH	0	0	-	1.0×10^{13}	1.32	assumed
C_iP_S	0	0	-	0	0	-
C_S	6.11	3.29	[5]	0	0	-
C_SH	0	0	-	1.0×10^{13}	1.32	assumed
H	3.30	1.22	[17]	0	0	-
H_2	7.4×10^{-4}	0.81	[18]	1.0×10^{13}	1.32	assumed
I	1.1×10^{-6}	0.12	[19]	0	0	-
I_2	0.38	1.52	assumed	0	0	-
I_3	0	0	-	0	0	-
I_4	0	0	-	0	0	-
I_5	0	0	-	0	0	-
I_6	0	0	-	0	0	-
I_7	0	0	-	0	0	-
I_8	0	0	-	0	0	-
IO	0	0	-	0	0	-
O	$3.3 \times 10-1$	2.59	[18]	0	0	-
O2	$2.3 \times 10-2$	2.05	[18]	0	0	-
OH	0	0	-	1.0×10^{13}	1.32	assumed
OH_2	0	0	-	1.0×10^{13}	1.32	assumed
P_S	1.03	3.51	[5]	0	0	-
P_i	0	0	-	0	0	-
P_iI	0	0	-	0	0	-
$P_{S2}O$	0	0	-	0	0	-
P_SH	0	0	-	1.0×10^{13}	1.32	[20]

P_SO	0	0	-	0	0	-
V	3.9×10^{-4}	0.30	[21]	0	0	-
V_2	4.0×10^{-3}	1.30	[22]	0	0	-
V_2H	0	0	-	1.0×10^{13}	1.32	assumed
V_2H_2	0	0	-	1.0×10^{13}	1.32	assumed
V_2H_3	0	0	-	1.0×10^{13}	1.32	assumed
V_2H_4	0	0	-	1.0×10^{13}	1.32	assumed
V_2H_5	0	0	-	1.0×10^{13}	1.32	assumed
V_2H_6	0	0	-	1.0×10^{13}	1.32	assumed
V_2O	0	0	-	0	0	-
V_2O_2	0	0	-	0	0	-
V_2O_3	0	0	-	0	0	-
V_2OH	0	0	-	1.0×10^{13}	1.32	assumed
V_2P_S	0	0	-	0	0	-
V_3	0	0	-	0	0	-
V_3O	0	0	-	0	0	-
V_3O_2	0	0	-	0	0	-
V_3O_3	0	0	-	0	0	-
V_4	0	0	-	0	0	-
V_5	0	0	-	0	0	-
V_6	0	0	-	0	0	-
VH	0	0	-	1.0×10^{13}	1.32	assumed
VH_2	0	0	-	1.0×10^{13}	1.32	assumed
VH_3	0	0	-	1.0×10^{13}	1.32	assumed
VH_4	0	0	-	1.0×10^{13}	1.32	assumed
VH_5	0	0	-	1.0×10^{13}	1.32	assumed
VH_6	0	0	-	1.0×10^{13}	1.32	assumed
VO	0	0	-	0	0	-
VO_2	0	0	-	0	0	-
VO_3	0	0	-	0	0	-
VOH	0	0	-	1.0×10^{13}	1.32	assumed
VOH_2	0	0	-	1.0×10^{13}	1.32	assumed
VP_S	$9.70 \times 10-4$	0.93	[5]	0	0	-
B_SO	0	0	-	0	0	-
C_SO	0	0	-	0	0	-
VI	0	0	-	0	0	-
C_SOH	0	0	-	1.0×10^{13}	1.32	assumed
I_2O	0	0	-	2.15×10^{10}	1.00	assumed
C_iC_SI	0	0	-	0	0	-
$C_iC_SI_2$	0	0	-	0	0	-
C_iOI	0	0	-	0	0	-
C_iOI_2	0	0	-	0	0	-
IH	0	0	-	1.0×10^{13}	1.32	assumed

IH_2	0	0	-	1.0×10^{13}	1.32	assumed
IH_3	0	0	-	1.0×10^{13}	1.32	assumed
I_2H	0	0	-	1.0×10^{13}	1.32	assumed
I_2H_2	0	0	-	1.0×10^{13}	1.32	assumed
I_2H_3	0	0	-	1.0×10^{13}	1.32	assumed

6 Problems and Issues

After showing some of the results of this work, this section discusses drawbacks and problems this simulation still faces. Many of the input parameters of the included defects are not known and thus assumed or exhibit a great uncertainty. Simplifications of the real model had to be made, and some assumptions were necessary to make the simulation possible while trying to make it run at a reasonable speed. For this reason, the results shown in the previous section represent trends and tendencies, but unfortunately no exact results.

Errors and Uncertainties

As this is a simulation that features many different entities with different (most of the time unknown) properties and many simplifications introduced for various reasons, it is hard to exactly determine errors and uncertainties of the resulting data. The deviation of the result from the real case is composed of

- numerical errors, that stem from the discretization process and from the used numerical step size (also called "time step" in this work)
- simplifications and assumptions that were made because of the lack of physical knowledge (e.g. values for parameters/constants of defects)

A compromise has to be made between gathering results fast, but with less accuracy, or getting them only very slowly, with less errors and uncertainties. Normally and logically the second option is taken, because a more correct result will always be more important than a fast result. Nevertheless, faster results were favoured for this work, because the simulation speed (or rather: the numerical approach and setup) is not the only source of incorrect data. Mathematical simulations that use the same numerical methods as this simulation, usually use initial values and conditions that are certain. The only things that introduce errors are the numerical step size and the discretization itself. This is not the case for these simulations, because the initial values are defect parameters, which are often either not known or guessed based on similar defects. This means that the results show uncertainties from the get-go, and an exact solution to compare the obtained results to is not present. Only the numerical error (i.e. the impact of choosing a high time step and a discretization with a low amount of grid points) can be estimated, by running two identical simulations, one with a high time step and many grid points and one with a low time step and a low amount of grid points.

Simulation Speed

Even though the duration of the simulation was greatly decreased compared to its predecessor, it still takes some days to perform simulations with a moderate workload, due to the chosen parameters. The fastest simulation took roughly 31 hours to be completed (low process temperature, high implantation current, low oxygen concentration, low carbon concentration), while the longest one took approximately 283 hours (high process temperature, low implantation current, high oxygen concentration, low carbon concentration). In extreme cases, simulations with only a very low computational demand can be finished within a couple of hours of simulating. However, simulations with such parameters

are either only of minor interest, or only part of a chain of simulations, where other parts take significantly longer (i.e. a subsequent annealing step). Simulations with lower implantation currents, higher temperatures and higher implantation doses cause the duration to be longer, increasing the simulation time to several weeks for simulations with higher workload. This time frame extends even further when annealing processes are considered, which are performed at comparatively high temperatures. These aspects increase the simulation times of an implantation step combined with a subsequent annealing step up to several months, which is not viable for the semiconductor industry. As of now, the key to further improvement of the simulation speed lies in employing more computational power. Unfortunately, the problem at hand cannot be handled by parallel computing from the beginning to the end. Instead, only specific parts of the simulation code can be computed simultaneously, but they have to be merged again after that computation was done. While this constant separation and merging of information (which happens several million times within one simulation) is significantly slower than a complete parallelization of the problem, it is still substantially faster than not using this technique at all. Using more computational power for the simulation is a must, especially so, when the barriers from the lack of physical knowledge of the system vanish.

7 Summary and Outlook

The current simulation is capable of simulating the process of proton implantation into a silicon wafer. Depending on the manufacturing of the wafer (be it from silicon ingots prepared by the Czochralski Process, the Float Zone process or from epitaxial silicon wafers) different amount of impurities carbon and oxygen are initially present in the wafer. The addition of dopant atoms such as boron and/or phosphorus is also considered. If parameters are chosen, that exhibit a medium workload (e.g. implantation doses of up to 10^{19} m^{-2} and temperatures up to 423 K) the simulation of the proton implantation process takes a few days to several weeks to finish. Simulations whose parameters cause less workload take significantly less time, in extreme cases only a few hours. Simulations that need more computational resources (e.g. implantation processes with higher end doses or annealing steps) are currently still not viable, as the simulation time amounts to several months. A numerical compromise between accuracy of the results and duration of the simulation had to be made, that increases inaccuracy in the data and is the cause of erroneous or deviated results. As a result, the data gained from these simulations are unfortunately not exact and precise results, but rather show trends and tendencies.

Following trends and results have been observed in the simulations:

- The main effect of increasing the implantation current is a decrease in simulation time, because the implantation dose is reached faster. Doubling the implantation current will decrease the simulation time by a factor of 2, and so on. Changing the duration of the simulation by changing the implantation current also alters the concentration of competing defects. Bigger defects with more constituents need a longer time to be created. They show a higher concentration in longer simulations (where the implantation current is lower), at the expense of smaller defects with less concentration. One possible example are the concentrations of defects C_i/C_S and C_iOI/C_iOI_2. A higher current benefits the concentrations of both C_i and C_S. The opposite is true for low currents, because enough time could pass, that C_iOI and C_iOI_2 reach higher concentrations than C_i and C_S.

- A higher process temperature causes faster diffusion of diffusing defects. Reactions and dissociations are also affected, i.e. more defects react away and more defects dissociate. One common consequence is that peaks and dips in the concentration profiles are broadened. Moreover, the drop in concentration in regions after the implantation depth (from about 5.5 microns onwards) is not as steep anymore, as diffusing species occupy these regions faster.
  Higher process temperatures generally promote the formation of defects with multiple constituents at the expense of small and basic defects. Almost all of the basic point defects (hydrogen, vacancies, interstitials, carbon and boron) show decreases in concentration in various regions of the silicon wafer. Generally it can be said, that concentrations change, when their levels at low temperatures are either especially high or especially low and/or where many reaction partners are present. Hydrogen, for instance, decreases its concentration at the peak at the projected range, while carbon and boron reduce their concentration in the irradiated region, in which many vacancies, interstitials and hydrogen atoms are present.

- Raising the initial carbon concentration promotes simple carbon-related defects (e.g. C_i or C_SH) as well as compound carbon defects, such as C_iC_S. Carbon-related defects with other impurities,

such as oxygen or boron, also show an increase in concentration, but less than expected (i.e. raising the carbon concentration by a factor of two, increases the concentration of these defects by a factor of 1.6-1.8). These quantitative changes in C-related complexes are very sensitive other parameters. For example a more favourable formation of simple carbon defects can only be seen at higher process temperatures.

- A higher initial concentration of oxygen in the sample benefits vacancy- and interstitial-complexes. Their concentration levels yield ratio values of much higher than 4. This happens at the expense of oxygen-related defects that contain impurities (e.g. C_iO) as well as impurities (e.g. B_i). Since only the level of oxygen was increased, but not the one of the impurities, the impurity concentrations are a limiting factor for reactions with oxygen, resulting in either a lower or equal concentration than before. This can be observed for all combinations of parameters.

In the future, it is planned that the results of simulations with a heavier workload can be obtained faster. Moreover the aim is to be able to perform simulations of annealing steps at an elevated temperature within a feasible time frame. This is important, as samples usually do not only undergo the implantation step, but also a subsequent annealing step, that helps reaching a certain defect concentration profile faster due to the increased temperature. The following steps help to achieve this goal:

- Increasing the computational power by running the simulation on a computer cluster and computing the results with parallel computing. Even though the simulation code at hand used this feature, the parallelization process can be further optimized to guarantee higher speed.
- Changing the way available and used data are stored, retrieved and manipulated. This is closely tied to the previous point, where changing the format of the available data may lead to a new and faster way of treating and processing the data mathematically and numerically. Further, the simulation code can still be more simplifiedto get rid of unnecessary steps and calculations.

The following points have been heavily discussed to be added to the current simulation code, but they are either not available quite yet, or were too difficult to implement, which would have been beyond the scope of this thesis. These have the purpose of making the underlying model more accurate and thus yield more accurate results, at the cost of the simulation speed.

- Adding the dependencies of the chemical potential. The amount and distribution of defects (more precisely their charged states, see next point) change the chemical potential and vice versa. The chemical potential also influences the capture radii of the reactions, and could thus favour reactions of one specific defect complex or family. Currently, the chemical potential is assumed to be at a fixed position in the middle of the band gap of silicon.
- Implementing charged defects states for all the 86 present defects. Depending on the current chemical potential, defects may exist in either a neutral, (twice) negatively charged or (twice) positively charged state. The distribution of the charged state can be determined if the chemical potential is known. DFT-calculations can then be employed to get the exact distribution. Different charged states of defects might have different coefficients for the diffusion, which influences the reactions of defects just as well. For every step the chemical potential and the distribution of defects have to be determined again. This entails very much computational effort and is one of the main reasons why charged defect states have not been added yet.

References

[1] R C Newman. Defects in silicon. *Reports on Progress in Physicsl*, 45(10):1163 – 1210, 1982.

[2] A. Smakula and J. Kalnajs. Precision determination of lattice constants with a geiger-counter x-ray diffractometer. *Phys. Rev.*, 99:1737–1743, Sep 1955.

[3] C. Kittel. *Introduction to Solid State Physics*. Wiley, 2004.

[4] G. Gottstein. *Materialwissenschaft und Werkstofftechnik: Physikalische Grundlagen*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2013.

[5] P. Pichler. *Intrinsic Point Defects, Impurities, and Their Diffusion in Silicon*. Computational Microelectronics. Springer Vienna, 2004.

[6] G. Eranna. *Crystal Growth and Evaluation of Silicon for VLSI and ULSI*. Taylor & Francis, 2014.

[7] G.W.C. Kaye and T.H. Laby. *Physical and Chemical Constants And Some Mathematical Functions*. Longmans, Green and Co, 1911.

[8] A F B Braga, S P Moreira, P R Zampiere, J M G Bacchin, and P R Mei. New processes for the production of solar-grade polycrystalline silicon: A review. *Solar Energy Materials and Solar Cells*, 92(4):418 – 424, 2008.

[9] Zinaida A Salnick. Oxygen in czochralski silicon crystals grown under an axial magnetic field. *Journal of Crystal Growth*, 121(4):775 – 780, 1992.

[10] R W Series and K G Barraclough. Carbon contamination during growth of czochralski silicon. *Journal of Crystal Growth*, 60(2):212 – 218, 1982.

[11] J. Lutz. *Halbleiter-Leistungsbauelemente: Physik, Eigenschaften, Zuverlässigkeit*. Springer, 2006.

[12] J.G. Laven. *Protonendotierung von Silizium: Untersuchung und Modellierung protoneninduzierter Dotierungsprofile in Silizium*. Springer Fachmedien Wiesbaden, 2014.

[13] M Faccinelli, M Jelinek, T Wuebben, J G Laven, H J Schulze, and P Hadley. Simulation of the proton implantation process in silicon. *physica status solidi*, 13(10-12):750 – 755, 2016.

[14] J F Ziegler, M D Ziegler, and J P Biersack. Srim – the stopping and range of ions in matter (2010). *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 268(11-12):1818 – 1823, 2010.

[15] R. Smith. *Atomic and Ion Collisions in Solids and at Surfaces: Theory, Simulation and Applications*. Cambridge University Press, 2005.

[16] T. Zundel and J. Weber. Dissociation energies of shallow-acceptor-hydrogen pairs in silicon. *Phys. Rev. B*, 39:13549–13552, Jun 1989.

[17] M. Capizzi and A. Mittiga. Hydrogen in crystalline silicon: A deep donor? *Applied Physics Letters*, 50(14):918–920, 1987.

[18] Vasilii Gusakov. First principle study of the diffusion of oxygen and oxygen complexes in si, sige solid solutions and si nanocrystals. In *Gettering and Defect Engineering in Semiconductor Technology XV*, volume 205 of *Solid State Phenomena*, pages 171–180. Trans Tech Publications, 1 2014.

[19] VA Panteleev, SN Ershov, VV Chernyakhovskii, and SN Nagornykh. Determination of the migration energy of vacancies and of intrinsic interstitial atoms in silicon in the temperature interval 400-600° k. *JËTP Lett*, 23(12), 1976.

[20] K. Bergman, Michael Stavola, S. J. Pearton, and J. Lopata. Donor-hydrogen complexes in passivated silicon. *Phys. Rev. B*, 37:2770–2773, Feb 1988.

[21] Vladimir V Voronkov and Robert Falster. Fast and slow vacancies in silicon. In *Solid State Phenomena*, volume 205, pages 157–162. Trans Tech Publ, 2014.

[22] M. Mikelsen, E. V. Monakhov, G. Alfieri, B. S. Avset, and B. G. Svensson. Kinetics of divacancy annealing and divacancy-oxygen formation in oxygen-enriched high-purity silicon. *Phys. Rev. B*, 72:195207, Nov 2005.