University of Alberta

Comprehensive Simulation of Sputter Deposition

by

Francisco Javier Jimenez

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Microsystems and Nanodevices

Electrical and Computer Engineering Department

©Francisco Javier Jimenez
Fall 2012
Edmonton, Alberta

To my wife and parents

Abstract

The magnetron sputtering process is extensively used in industry to deposit thin films of a large number of materials for countless applications. However, to utilize the full potential offered by this process, it is essential to have a sound understanding of the complex dynamics occurring during a typical film deposition process. Computer simulations are a powerful, cost-effective tool to increase the understanding of the process which is otherwise difficult to obtain through theoretical models and experiments.

This thesis presents the integration and implementation of a comprehensive 3D computational framework for the simulation of magnetron sputtering discharges in arbitrary reactor geometries. The modular architecture of the proposed model allows the user to either concentrate on individual components of the deposition or to investigate the process as a whole. Every module within the framework is an independent, highly specialized unit, capable of simulating a specific component of the process. As key to the correct description of the process, a model to numerically solve for the magnetized plasma is presented in this dissertation. The proposed numerical model uses a robust algorithm which is able to handle inhomogeneous magnetic fields and a wide range of

process conditions using a global strategy. The interaction of the sputtered particles with the background gas, and the events at the target, are incorporated into the discharge model using an iterative scheme whereby modules comprising the computational framework are sequentially solved until convergence is attained. The resulting model is employed to study a variety of process parameters and effects such as gas rarefaction, plasma density and energy flux. Results, both from individual components and from the full iterative model, are compared against experiments to assess the validity of the model.

Acknowledgements

First, I wish to express my sincere gratitude to my thesis supervisor, Dr. Steven K. Dew, for his continuous support and guidance during the course of my research work. I would also like to thank my fellow labmates at Dr. Dew's Thin Film Laboratory. In particular, I would like to thank Dr. Samuel Ekpe and Mr. Jason Ng with whom I spent countless hours in the lab either doing some experiment, fixing something broken in the vacuum system or talking about random topics. I am also grateful to Glenn Elaschuk from the Nanofab who donated part of his time to help with troubleshooting the vacuum system in our lab. This thesis had not been possible without the financial support of the Natural Sciences and Engineering Research Council of Canada.

I want to thank Smith and Nephew (Alberta) Inc. (formerly NuCryst) for securing me a part time job in industry during the course of my PhD program. I am gratefully indebted to Dr. David Field for whom I only have words of gratitude. I particularly appreciate his support, words of encouragement and friendship. I also thank Dr. Martin Davis for being extremely flexible with my work schedule. Finally, I would like to thank the ingenious Mr. Terry Hunt whose vivacious personality makes working with him such a natural process.

Last but not the least, I wish to thank my wonderful wife, Karla Barron, for her continuous support, patience, and love. Had she not been this supportive, the writing of this PhD thesis, which was a long and winding road, would have been more difficult to endure.

Contents

Appendices

List of Tables

List of Figures

List of symbols

CHAPTER 1

Introduction

1.1 Thin Films

The world we know today would be a different place without thin films. They are the backbone to every electronic component and the cornerstone of a myriad of products of everyday use in the world. These films find applications in a wide range of disciplines given our ability to engineer their intrinsic properties to match those required in practice. Examples are numerous. The lifetime of equipment such as drill bits, gears and bearings can be extended by depositing a thin protective coating that can withstand more friction and wear without significant structural damage; nanocrystalline silver is used as an antimicrobial barrier in dressings for the treatment of burns and wounds that otherwise may readily infect with catastrophic consequences; SiO_2 and Al_2O_3 films are used as moisture and light barriers in the food packaging industry; and finally, the entire semiconductors industry is based on our ability as engineers to deposit a great variety of films with distinctive properties.

Over the last century, these engineered films evolved from a scientific curiosity to a well-developed technology. The term *thin film* eludes a formal definition; thin does not convey the scale to which it refers. A more formal definition is that a thin film is a coating of a material placed or deposited on top of another

(the substrate) and the thickness of this film is less than a few micrometres. A practical limit is usually taken at 1 μm but we may find that films are labeled as thin even when their thickness is about 2-5 μm. Once in thin film form, the properties of the material substantially differ from those that are palpable and measurable to our senses. In practice, real life applications of thin films require a good understanding not only about the fundamentals of thin film growth but also of how the material is brought in contact with the substrate. This knowledge is essential for the development of novel materials with tailored properties. With these thin films, we are fulfilling one of the dreams of ancient alchemists, that of changing the looks and properties of an unattractive and easy to find material to those that please our demands.

One of the methods to tailor the properties of the material is by carefully controlling the way it is deposited. In fact, the properties of these thin films are dictated by both the way they are deposited and by their physical dimensions. For instance, for films of few microns in thickness, the properties of the material are close to those of the bulk, but once the film is very thin (a few nanometres), the fundamental physical properties of the material changes. Additionally, by controlling the morphology of the film we can modify the intrinsic properties of the material according to our technological needs.

One of the most popular methods to deposit these thin films is known as physical vapor deposition (PVD). In PVD methods, the material to deposit is first vaporized in a controlled environment. Then, the vapor is put in contact with the substrate upon which it finally condenses. The condensation of these vapors is a natural phenomenon easily observed in daily life; we are all familiar with the condensation of water vapor on surfaces. The physical mechanisms that drive the formation of foggy windows in a cold winter day are essentially the same that are put forward during a typical PVD process.

1.2 Magnetron sputtering

Among PVD techniques, sputtering is one of the most widely used processes in both academia and industry due to its unique characteristics. There are three that stand out for their industrial impact: versatility, controllability and repeatability. The first one is all about size. In fact, sputtering is a malleable technology since it is simple to scale up or down according to the needs of the process. The second characteristic refers to control. In an industrial setup, full control over the process is essential for a successful implementation of the technology. Lastly, repeatability alludes to producing films with the same properties consistently. These distinct attributes of sputtering turn it into one of the preferred methods for industrial applications. Sputtering has permeated industrial environments, and it is nowadays widely used all over the world. According to BCC Research[1], the total area coated by sputtering techniques in 2011 is expected to be more than 1.2 billion m^2 and this number may reach 1.8 billion by 2016 representing almost 1/5th of the total global market for thin films worldwide[1].

A standard magnetron sputtering process consists of the bombardment of a surface (the target) with high energy particles. We can picture the surface of a material, a metallic surface for instance, as made of a myriad of atoms orderly arranged in structured arrays forming crystals. The impact of a fast particle with a highly ordered ensemble of atoms is similar to the breaking of the rack during a typical game of pool: the physics behind the cue ball dispersing the ordered colored balls during the opening shot are in essence the same that describes the fundamental principles of sputtering. The aftermath of these collisions is the eventual release of atomic particles from the surface. These particles then travel through a gaseous medium to finally reach the substrate where they condense.

Despite its popularity, there is no global mathematical model that helps to predict the behavior of a sputtering process in practice. Nevertheless, we may

---

[1]www.bccresearch.com

be able to use scaling laws heuristically derived. For instance, we may infer qualitatively the change in the rate at which particles arrive at the substrate, the deposition rate, as a function of one process variable, but this model would only apply to systems with a similar geometry. Another example would be that of the total heat flux bombarding the substrate. Depending on the process, we may be required to decrease the heat load at the substrate without significantly decreasing the deposition rate. A third example pertains to one of the main drawbacks of sputtering as a deposition process, that of the nonuniform erosion of the target. This nonuniformity restrains the continuous operation of the process since targets need to be replaced constantly. This necessary cycle greatly reduces the time the deposition system is up and running. It is clear that a deep understanding on the mechanisms that drive the modifications in the process would allow scientists and engineers to fine tune and optimize the process.

The use of analytical methods to derive scaling laws has also been explored over the years. However, these models are very complex to derive due to the great number of variables to consider in a sputtering process. As an alternative to traditional analytical and experimental models, computer based models offer a very distinctive advantage.

1.3 Computer Simulations

Computers are ubiquitous in our daily life. They slowly have broached into our mundane activities; it is difficult to imagine a day without interacting with computers in one way or another. The realm of science is in no way an exception. One of the revolutionary ways in which science has embraced computers is through a process known as computer simulations. Contrary to number crunching techniques in which models are extracted from experimental data, computer simulations are computational routines specifically designed to simulate a real system. Performing these simulations may be regarded as virtual experiments in the same sense that real experiments are conducted in

science. Worldwide, more scientists are turning to computational methods to assist them in their discoveries.

The process to build a successful computer simulation of a physical process involves several steps which can be encapsulated into four distinctive stages. In the first, we identify the fundamental physics that describes the phenomena of interest using well accepted theories. In the second stage, we build a model based on the theoretical framework laid down in the previous step. This model is an abstract description of the physical system most typically in the form of mathematical expressions. During the third stage, the built model is first translated into a language the computer can interpret to then proceed to form a set of rules and instructions indicating the roadmap for the computer to follow. These instructions are most commonly referred in computer science as algorithms. Then the algorithms are solved by letting the computer execute the rules we previously formulated. Finally, the last step involves the collection and visualization of the outcome of these simulations to derive our conclusions based on the new information inferred from the virtual experiment.

One of the appeals of computer models is their ability to provide solutions to complex phenomena otherwise intractable via traditional analytical methods. Magnetron sputtering processes are especially suitable to computational treatment given the nonlinear coupling between phenomena characteristic of sputtering.

1.4 Summary and dissertation outline

Magnetron sputtering is today one of the fundamental technologies to deposit a myriad of films in both industry and academia. Its inherent complexity has eluded analytical formalisms otherwise used to derive universal laws to describe the process. Computational methods provide us a great opportunity to build and solve models that can help us to elucidate the fundamental mechanisms that are in play during a typical sputter deposition.

This thesis is about the development, implementation and validation of a 3D comprehensive computational model of the sputtering process, a 'magnetronus electronica' [2]. In this thesis, I will introduce the main components of the simulator and describe in detail the development of one of the major modules of the project, the plasma module. I will also comment on the challenges faced for integrating a multiphysics project such as the present work. Examples of the feasibility of the simulator as a scientific tool will be provided. Finally, heat flux measurements using a scaled-down system representative of an industrial web coater are presented. I designed and built the heat flux measurement system which is based on a commercially available sensor as part of my research collaboration with Smith and Nephew (Alberta) Inc. (formerly Nucryst Pharmaceuticals).

The rest of this dissertation is organized as follows. Chapter 2 offers a short review of the basic principles behind the sputtering process. Complex concepts are introduced using only simplified models. The goal of this chapter is to provide the foundation upon which subsequent chapters are built upon and to have a better idea of the challenges this thesis intends to tackle. In Chapter 3, I introduce the SpudII project and its main components. I examine the basic computational framework from a high level perspective. In this chapter, my contribution to the project will be clearly stated. Chapter 4 starts with a review of two modules and provides experimental verification. Chapter 5 presents a new algorithm for modeling a DC glow discharge under the influence of a magnetic field. Chapter 6 describes the integration and implementation of the project focusing on the challenges faced during this stage. I briefly discuss the caveats of integrating a set of simulators into a single, multiphysics computational tool. Also, the full simulator is benchmarked against experimental data. Chapter 7 describes the development of a device to measure heat flux in industrial environments and presents a study of the effect of magnetic field on the total energy input at the substrate. Finally, Chapter 8 summarizes this thesis and makes some suggestions for further research.

CHAPTER 2

Background of the sputtering process

2.1 Introduction

Sputtering (also referred as to cathode disintegration, spluttering and impact evaporation) is a complex process. This chapter presents a brief introduction to the physics of sputtering and the fundamental principles behind this technology. I begin with a short historical perspective of thin films and a review of the history of sputtering to continue with a description of the physical phenomena occurring during a typical sputtering process. Given the multifaceted nature of the phenomena being described, the processes herein mentioned are portrayed using rather simple models; a thorough description of sputtering as a whole is beyond the scope of this thesis. The concepts that need further description are covered in subsequent chapters. The goal is not to be a comprehensive review of the field but to introduce the reader not familiar with the technology to the concepts I will address in the chapters to come. Interested readers searching for a more complete description of individual components are referred to books on magnetron sputtering [3–5], reviews and Chapters 4 and 5. Emphasis is made in introducing the fundamental computational methods relevant to the simulation of the sputtering process. This chapter wraps up with a detailed description of the challenges this thesis intends to tackle.

7

2.2   Origins of sputtering

The term sputtering was possibly first coined by K. H. Kingdon and I. Langmuir early in the 20th century [6] where they used the verb 'to sputter' to describe the disintegration of a cathode in an electric gas discharge. The first use of the word is still not clear though. For instance, Mattox [7] mentions that G. K. Weghner attributes the first use of the word to J. J. Thompson who used a similar term (spluttering) to apparently describe a similar phenomena in a book published before the article of Kingdon and Langmuir. To this day, the origin of the word to name what we now call sputtering is still open for debate.

As it often occurs in science, Kingdon, Langmuir and Thomson were not the first to observe sputtering. In fact, sputtering was first described in the 19th century. In 1852, Grove noticed that when a silver plate was placed in a 'good air pumped' device and acted as the cathode, the plate was free of oxide while the opposite occurred on the positive terminal [8]. The work of Grove, to this day, is considered the first scientific investigation of sputtering [7] and marked the beginning of an era. The Royal Society of London recently dedicated an issue in Groves honor to commemorate 150 years of sputtering [9].

At the time of its discovery, sputtering was a puzzle difficult to explain using the existent theoretical framework. For instance, Grove surmised [8]:

> The other theory which may be advanced is, that by dielectric induction the gases may be bodily separated, a layer, not molecular, but corporeal or voluminous, if I may be allowed these expressions, of oxygen being developed on the side next the anode, and one of hydrogen next the cathode, the gas intervening between the terminals being thus divided, as it were, into two halves.

During the years after Grove's discovery, others continued studying the phenomena but no theory was put forward to explain it [10, 11]. Besides, the

phenomena remained a laboratory curiosity with no practical use for decades. However, near the turn of the century, Tomas Alva Edison, in 1894, was granted a patent for a process to create phonographs that explicitly made use of a 'gaseous deposition' technique [12]. Edison's patent was challenged due to claims that the same invention had been already described before by Wright [11]. Edison argued, successfully, that his invention used a constant arc whereas that of Wright was a pulsed arc: what Wright was describing in his paper would be probably referred today as cathodic arc deposition (CAD) [13]. To this day, Edison is credited with the first industrial application of sputtering [7].

Despite Edison's success, sputtering remained industrially unexploited for years. Admittedly, one of the reasons for the minimal industrial impact of sputtering was that, as a technology, it was more expensive than other techniques widely used for depositing thin films such as evaporation. Another factor that contributed was that the applications that made use of thin films did not require strict control on the structure of the film and evaporation was entirely adecuated for this. Years passed before sputtering could match economically the evaporation industry. However, the industrial impact of sputtering drastically changed with the advent of the silicon age in the early 1960s. The process to build an integrated circuit (IC) entails the deposition of several layers of materials placed one on top of another. The wide range of materials needed for the fabrication of integrated circuit placed economic and technological constraints on contemporary evaporation techniques. The limited control on the microstructure of evaporated films turned out to be a big disadvantage when one needs to control variables such as conductivity and material compatibility between layers. Besides, the versatility of sputtering as a deposition process allowed engineers to design configurations to reduce costs, such as in-line systems for batch processing [14].

The simplest sputtering system is known as diode sputtering (Figure 2.1a). Its low efficiency [4] stimulated the research of techniques to increase the deposition rate for industrial applications. An example of such attempts was

the triode configuration where a filament was placed inside the chamber to increase the ionization efficiency of the process. The triode had little impact on industry since it normally occurred that the filament was constantly damaged. It took the invention of the planar magnetron [15] for sputtering to have a real impact in industrial environments (Figure 2.1b).
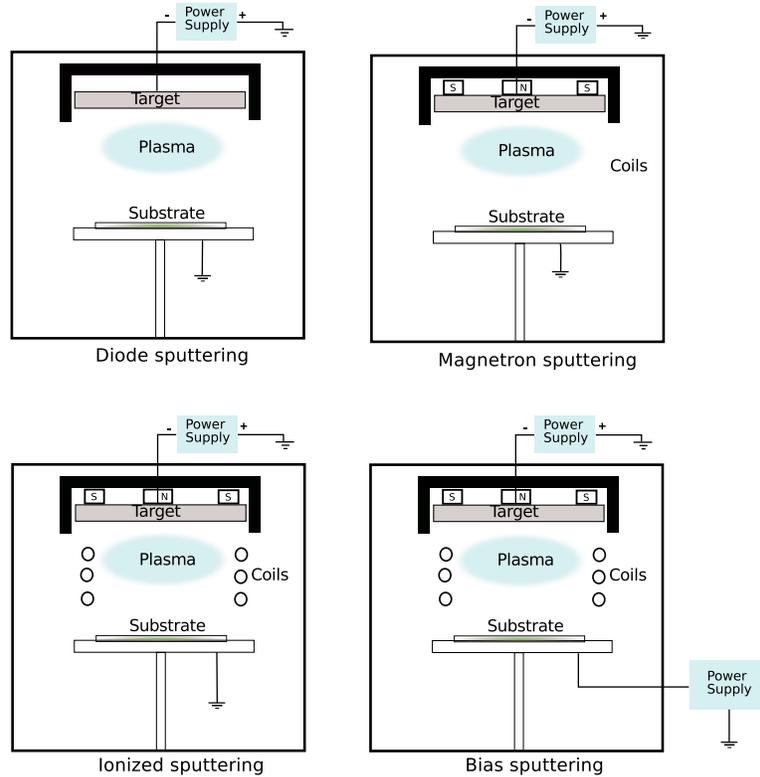
Figure 2.1: **Different configurations for a typical sputtering system: a) diode, b) magnetron, c) IPVD and d) biased.**

The magnetron was fully embraced by industry [16]. It soon became one of the most popular techniques to deposit industrially relevant films (see section 2.4). Further improvements (RF sputtering [17], rotary cylindrical magnetrons [18], unbalanced magnetron [19]) extended the range of materials to sputter and substrates upon which to deposit. The improvement of sputtering techniques was driven by the semiconductor industry which kept on reducing dimensions of its devices creating a new challenges. As an example, better coverage over

high aspect ratio vias as dimensions kept shrinking was demanded (see Figure 2.2). Collimated [20] and long throw sputtering were examples of some modification to the process that required little change to the process. However, these modification significantly reduced the deposition rate, reducing the appeal for batch production. The introduction of ionized magnetron sputtering [21] was the solution the semiconductor industry was looking for. This technique is still one of the work horses in the microelectronics industry for the deposition of challenging materials.
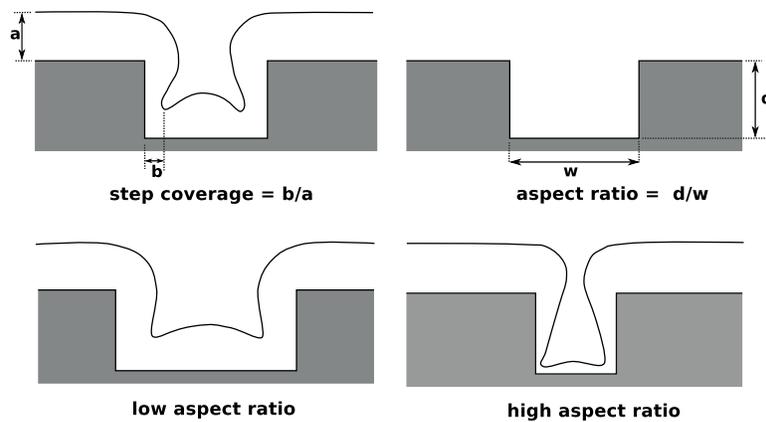
Figure 2.2: **Step coverage over high aspect ratio features for advanced metallization of semiconductor devices.**

In recent years, high power impulse magnetron sputtering (HiPIMS) [22] has been receiving a lot of interest in both academia and industry. The appeal of this technique is the full control on the energy of the depositing material [23]. Films with much better adhesion and microstructure are obtained using this state-of-the-art technology[24]. Furthermore, HiPIMS has been shown to be a very promising technology for the deposition of films otherwise very difficult to deposit using conventional sputtering techniques[25].

2.3   Physics of sputtering

Sputtering is the removal or erosion of atoms from a surface by particle bombardment. The outcome of this bombardment is the eventual ejection of atoms from the surface. Although the word sputtering only refers to the ion-surface interaction, it is a common practice to use this term to refer to the deposition method as a whole.

Figure 2.3 depicts a typical magnetron sputtering system. In this configuration, the main apparatus is a vacuum vessel where air is evacuated to a high vacuum ($< 1 \times 10^{-6}$ Torr). A process gas is then injected into the chamber and the flow is controlled to keep the pressure at the range of operation (typically between 1-40 mTorr). Argon is the most commonly used process gas due its inertness, availability and price. In a DC setup, the target material is connected to the negative rail of a power supply to create a self-sustained glow discharge. This glow discharge becomes the source of positive ions which are then pulled by the negative potential at the target. These ions are accelerated to high energies when they approach the surface and gain energy in the high electric field region next to the target. These ions then collide with the surface producing a collision cascade that removes atoms from the surface. The collision also extracts electrons from the surface which then are accelerated towards the discharge participating in the ionization of more atoms. The ionization efficiency is substantially enhanced by a static magnetic field which confines electrons to a region next to the target. Finally, sputtered atoms travel across the plasma to finally deposit onto the substrate and the walls of the apparatus.

Generally speaking, the sputtering process can we be divided into three stages [26]: generation, transport and deposition of sputtered particles. This is a convenient simplification for the purpose of describing the process as a whole. The theory behind each process is complex and extensive. In the descriptions that follow, I will cover only some of the key concepts shared by most systems. The reader interested in a detailed description of each topic is referred to the
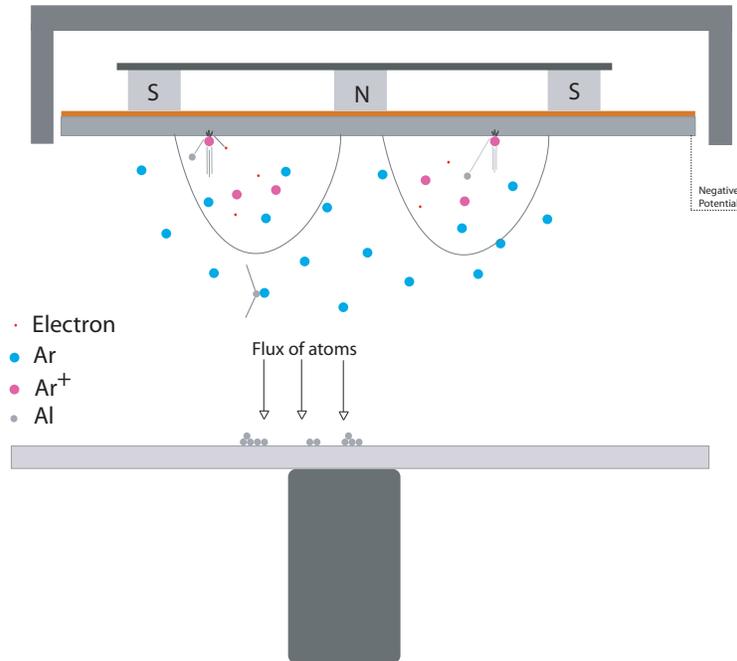
Figure 2.3: **Typical configuration of a magnetron sputtering deposition process. A glow discharge is generated between two electrodes in a low pressure environment ($\approx$ 1-100 mTorr). Argon ions from the discharge are attracted to the negatively biased electrode (the target), and as a consequence, some atoms from the cathode are vaporized. These atoms travel towards the substrate where they are deposited. The magnets behind the target are placed to confine electrons to increase the efficiency of the process.**

appropriate references mentioned in the Introduction and within the body of each section.

2.3.1 Target events

When an energetic ion hits (it may penetrate or not the surface), it undergoes a series of elastic and inelastic collisions that dislodge atoms from their equilibrium lattice positions. As a result, some atoms may be ejected from the surface if a recoil atom (that atom that has won energy by an elastic collision with the energetic ion) collides with an atom at the surface transferring it enough energy to overcome a certain energy barrier (see Figure 2.4). Another event derived from the collision is the generation of secondary electrons. These electrons, which then move towards the plasma, are key to sustain the discharge. The appropriate treatment of these secondary electrons is vital for the proper description of the discharge as we will see in Chapter 5.
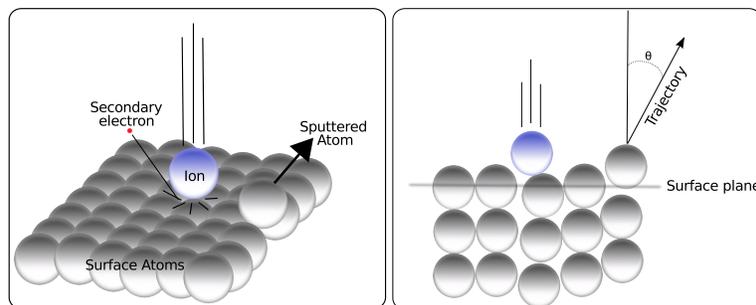
Figure 2.4: **Simple representation of an ion colliding with a surface. Conventionally, we only require an initial energy and trajectory to characterize the sputtered particle.**

In brief, the events that precede the collision at the first monoloayer of the surface are as follows. The energetic particle penetrates the lattice and continuously loses energy through collisions with the screen nuclei (elastic) and electronic clouds (inelastic) of the stationary lattice atoms [27]. The dominating effect depends on the elements involved in the collision and in the energy

of the incoming ion [5]. This slowing down process can be treated analytically using an effective collision cross section which is the sum of the nuclear and electron cross sections. However, these events differ in one aspect: elastic collisions modify the direction and energy of the incoming particle while electronic losses do not significantly change the original trajectory of the particle. Coulomb collisions are generally neglected for the range of energies associated with sputter deposition since the ion is neutralized within a few nanometres in the solid [28]. The original works of Lindhard [27], Sigmund [29] and the excellent review articles [28, 30–34] explain the sputtering process to a greater detail.

In general, the type of collisions occurring at the surface can be grouped into three regimes: the single knock-on, the collision cascade and the spike regions [29]. These three regions are depicted in Figure 2.5.
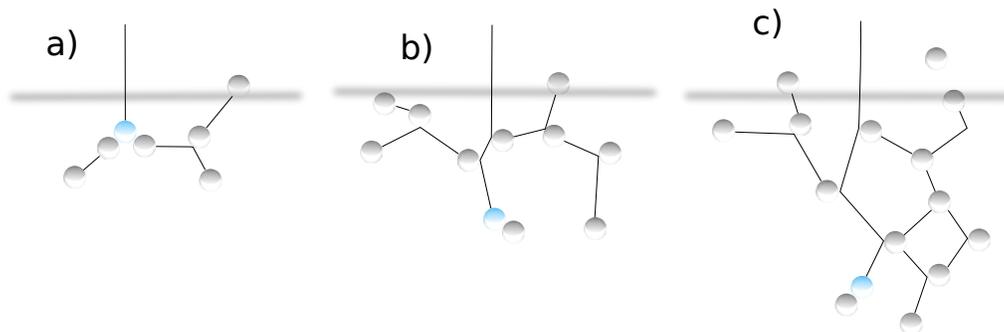
Figure 2.5: **Collision regimes relevant to sputtering. a) In the knock-on regime, only a few recoils are generated. The standard theoretical model for sputtering breaks down in this regime due to anisotropic scattering. b) In the collision cascade, the incoming particles have enough energy to create fully developed collision cascades. This region is well described by linear collision cascade theory [29]. c) In the spike regime, nonlinear effects dominates and the analytical treatment of the process is more complicated since collisions between moving particles are more frequent.**

In the knock-on regime, ions deposit little energy into the lattice producing

only a few recoils. These recoils have insufficient energy to generate a collision cascade and only a small number reach the surface with enough energy to escape. This regime is typical of low energy sputtering ($<$ 1 keV). In the linear collision region, recoil atoms do develop into full collision cascades within the solid. This is the region best understood and amenable to a theoretical treatment. Collisions between moving atoms are rare and a linear slowing down model is a good approximation to the process [29]. Finally, in the spike regime, the majority of atoms participating in collisions are moving which give rise to nonlinear effects that complicate the analysis. This region is of no practical interest for sputter deposition processes since it is generally observed at energies too high for any practical application.

Sputtering yield models

The number of atoms ejected per incoming ion is known as the sputtering yield Y, and it is one of the most important parameters in sputtering. Accurate analytical models to estimate Y are in general entangled and complex to derive from first principles, and they usually involve very sophisticated mathematical models. In 1969 [29], Peter Sigmund published one of the most influential works on sputtering to this day. His theoretical framework to describe the events at the target has been the underlying model for several works since its inception. The success of Sigmund's model stemmed from its accuracy to predict Y for a wide range of ion-target combinations [30]. The article also reinforced the momentum transfer theory demonstrated experimentally [35] to explain the fundamental mechanisms of sputtering as opposed to the evaporation theory some had suggested in the past [29].

The sputtering yield of a material depends on the ion energy and mass, the mass of the target atom and the incidence angle of impact. A universal response of the sputtering yield to changes in ion energy and angle of incidence is shown in figure 2.6. The figure is intended to be only an illustrative example of a general trend; thus, values should not be taken as guides but as estimates of qualitative variations. In general, the sputtering yield is zero for

energies below a certain value known as the threshold energy of sputtering, and its value depends on the target material [36]. As the ion energy increases, so does Y. The rate of change of Y with respect to energy is different for each ion-target combination but in general it follows the trend depicted in Figure 2.6. In practical sputtering deposition processes, unless we deal with very high voltage pulses, we are interested in the quasilinear region of the curve ($200\ eV < E < 600\ eV$).

Figure 2.6 also depicts the increase in Y as a function of angle of incidence. This in fact is a prediction of the momentum transfer theory [29]. In glow discharges, since ions are pulled from the plasma and reach the target with nearly normal incidence, the angle dependence does not play a significant role in practice.

A different mechanism to change Y is to vary the temperature of the surface [37, 38]. This effect is used in some techniques to improve the deposition rate [39]. We will discuss no further this effect since, in magnetron sputtering, we should keep the target below the Curie temperature (T_c) of the magnets (for a typical NdFeB $T_c \approx 300$) to avoid a significant reduction in the magnetic field which will increase the plasma impedance.

According to Sigmund's theory, Y can be estimated with the following equation [29]

$$Y = 0.042 \frac{\alpha(M_1, M_2)}{U_s} S_n(E_0) \tag{2.1}$$

where E_0 is the ion energy, M_1 and M_2 are the ion and target masses respectively, α is a function that depends on M_1 and M_2, U_s is the surface binding energy and S_n is the nuclear stopping cross section. S_n was defined by Lindhard [27] as

$$S_n(E_0) = 84.78 \frac{Z_1 Z_2}{\left(Z_1^{2/3} + Z_2^{2/3}\right)^{1/2}} \left(\frac{M_1}{M_1 + M_2}\right) s_n(\epsilon) \tag{2.2}$$
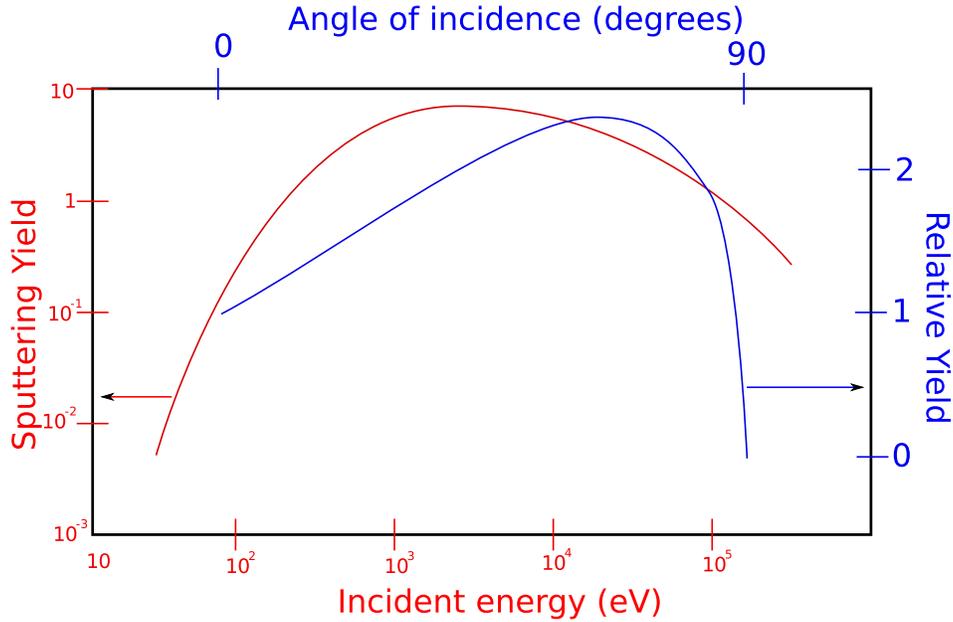
Figure 2.6: **Universal response of sputtering yield as function of ion energy and incidence angle. Below a specific threshold no sputtering occurs. This value is known as the sputtering threshold and it approximately four times the sublimation energy of the material. The color coded graph shows the yield response to varying ion energy (in red) and the relative yield variation with respect to angle of incidence (in blue).**

In the equation above, Z_1 and Z_2 are the atomic numbers of ion and target atoms respectively, s_n is the reduced elastic cross section and ϵ is the reduced energy which is defined differently depending on the screening length and the interatomic potential we choose to use. The expression for S_n also depends on the type of model we employ. Different expressions include those based on the Thomas-Fermi model [29], analytical approximations to Lindhard's data [40] or based on the KrC potential [41].

As mentioned before, Sigmund's formula proved to be a good approximation to experimental yields for energies greater than 1 keV but, at low energies, the values predicted by the theory were found to significantly deviate from

experiments [42]. The deviation was more pronounced the lower the energy of the ion. Several theories were put forward to elucidate the origin of the discrepancy but no theoretical model has proven to accurately describe the physics in the low energy region.

As an alternative to theoretical approaches, some authors proposed semi-empirical corrections to Sigmund's theory using fitting coefficients obtained from experimental data. The works of Yamamura [42, 43], Matsunami [40], Bohdansky [36] and Eckstein [44] are examples of such models. I briefly describe them below.

One of the first attempts was that of Yamamura et al. [42] who published an extensive article containing a long list of yields for several ion-target combinations based on the formula derived in the article. In general, reasonable agreement with experiments was obtained. One decade later, and with more experimental data to build a more reliable statistical model, the same research group [45] revised their original work and published an extensive and thorough monograph of sputtering yields that became one of the standard work for sputtering yields for many years [46].

The improved Yamamura formula reads:

$$Y(E) = 0.042 \frac{Q(Z_2)\alpha^*(M_2, M_1)}{U_s} \frac{S_n(E)}{1 + \Gamma \kappa_e \epsilon^{0.3}} \left[1 - \sqrt{\frac{E_{th}}{E}}\right]^s \qquad (2.3)$$

where Q, α^*, Γ, κ_e are fitting parameters that can be obtained from [45]. E_{th} is the species dependent threshold energy in the Yamamura model (not to be taken as the sputtering threshold described before) obtained via

$$E_{th} = U_0 \left[1.9 + 3.8 \left(\frac{M_2}{M_1} + 0.134 \left(\frac{M_2}{M_1}\right)^{1.24}\right)^{-1}\right] \qquad (2.4)$$

Another semi-empirical models is the work of Bohdanksy [36] later revised by Garcia-Rosales et al. [41]. The deficiencies of these models were addressed by Eckstein et al. who proposed a new formula [44]. This new equation reads [1]:

$$Y(E_0) = qS_n^{KrC}(\epsilon_L)\frac{\left(\frac{E_0}{E_{th}} - 1\right)^{\mu}}{\lambda_b/w(\epsilon_L) + \left(\frac{E_0}{E_{th}-1}\right)^{\mu}} \tag{2.5}$$

with $\epsilon_L = E_0\epsilon$ and E_{th}, q, μ, ϵ and λ_b tabulated in [5] for a wide range of ion-target combinations. S_n^{KrC} is the nuclear stopping power for the KrC interatomic potential

$$S_n^{KrC} = \frac{0.5ln(1 + 1.2288\epsilon_L)}{w(\epsilon_L)} \tag{2.6}$$

with

$$w(\epsilon_L) = \epsilon_L + 0.1728\sqrt{\epsilon_L} + 0.008\epsilon_L^{0.1504} \tag{2.7}$$

Within our research group, Stepanova and Dew [48] developed a model based on the concept of discrete-path transport theory [49]. The main modification to standard techniques is the inclusion of a few-collision approach to model the reduced number of recoils during low energy bombardment. The model has been benchmarked against the program TRansport of Ions in Matter (TRIM) and experimental data for a wide range of target materials. More details about this approach can be found in [48]. A comparison between the different models and experiments is shown in Figure 2.7.

Another approach to solve for sputtering yields is the use of computer simulations. In general, there are two methods to simulate events at the target:

---

[1]Eckstein points out in [47] that the term w(ϵ) is missing in both, the original publication [44] and in a recent book [5].
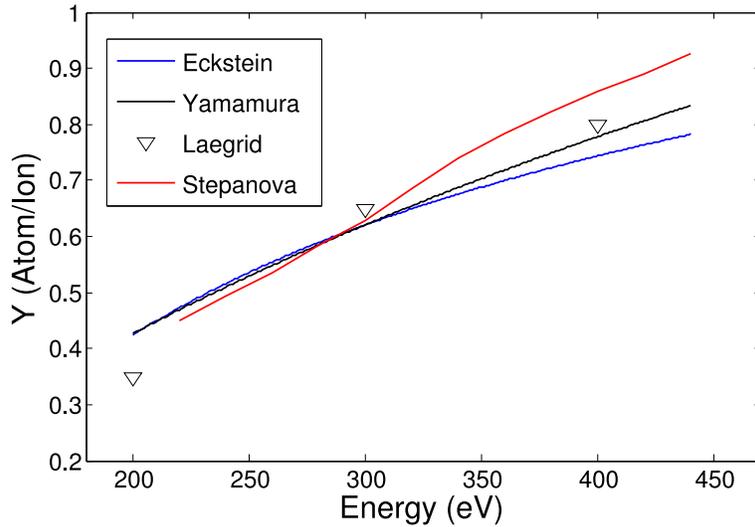
Figure 2.7:   **Results from Yamamura [45], Eckstein [47], Stepanova[48] compared against experimental yields from Laegreid [50]. Fitting coefficients taken from: Yamamura [45], Eckstein[5]. $U_s = 3.4$ in all models.**

those that employ the so called Binary Collision Approximation (BCA) [51] and those that explicitly solve every collision within the solid [52–55]. Examples of codes based on the BCA are TRIM [56], TRIM.SP [57], ACAT [43] and MARLOWE [51]. In general, predictions from these programs should be taken as guides since they are not valid in the low energy collision regime. For instance, it is well documented in literature that the program TRIM.SP [2] delivers results that differ from experiments in several cases [58]. For an accurate description of general trends, the software can be calibrated against experiments using the binding energy as a free parameter. Recent versions of the software may have improved their treatment in their low energy regime but this needs to be confirmed exhaustively as in [58]. For practical purposes, the estimates from equation 2.3.1 are in better agreement with published experiments [5].

----

[2]www.srim.org

Energy distribution of sputtered particles

In order to determine Y for an ion-target combination, the initial energy and trajectory of sputtered particles are key for an appropriate description of transport phenomena. In standard sputtering models, the Thompson model [59] is the most popular model to estimate the energy distribution of sputtered particles. The Thompson model predicts that the energy distribution of sputtered particles, J_T, follows the relation

$$J_T(E_s) \propto E_s(E_s + U)^{-3+2m} \tag{2.8}$$

where m is a cross section parameter [60], E_s is the energy of sputtered particles and U is the height of the surface energy barrier that the particle at the surface has to overcome, usually approximated with the surface binding energy [29]. But, in the sub-keV regime, the assumption of cascade isotropy breaks down and we are required to incorporate anisotropic events into the theory [48]. Experimental evidence of deviation from Thompson's model is well documented in the literature [61–64]. As an example, Figure 2.8 shows the comparison between a measured experimental distribution [63] with the Thompson formula for $U_{Al} = 3.4$. We can see in the figure how the experimental distributions falls-off faster than the expected E^{-2} relation predicted by Thompson's model.

Figure 2.8 bins the atom's energy irrespective of angle of emission. However, if we plot the energy distribution as a function of emission angle, the distribution broadens for large angles and becomes more Thompson-like [63]. One model that has been shown to accurately capture the anisotropy energy distribution is the equation proposed by Stepanova and Dew [65] which reads,

$$J(E, \theta) = J_T(E) \left(1 - \left(\frac{E + U}{E_{max} + U} \right)^n \right) \times$$
$$exp\left[-A \left(\frac{M_i(E \cos^k \theta + U)}{M_t E_i} \right)^l \right] \tag{2.9}$$
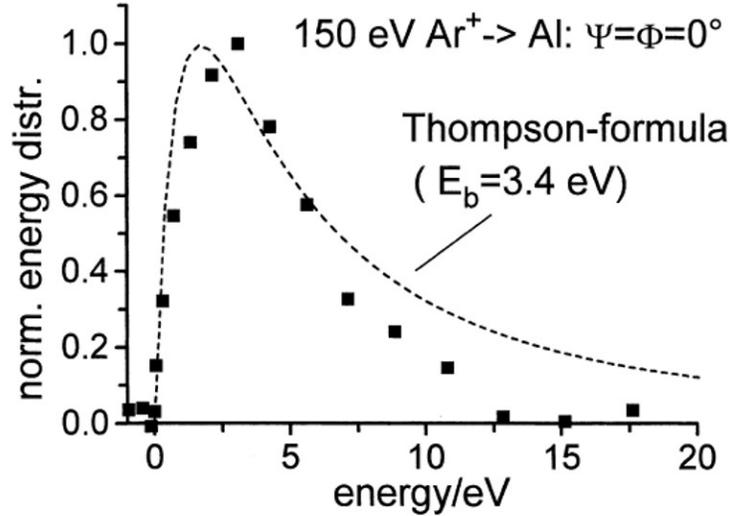
Figure 2.8: **Energy distribution of sputtered Al particles at 150 eV bombarding energy. The experimental distribution (shown by ■) deviates from the Thompson formula for energies greater than 5 eV. Reprinted from Ref. 63.**

where E_{max} is an angle dependent cut-off energy, M_i and M_t are the ion and target masses respectively, E_i is the ion energy and A,l,k are species-dependent fitting parameters [65]. Equation 2.9 has been shown to give good agreement with experiments given we appropriately choose the fitting parameters [65].

Angular distribution of sputtered particles

Angular distributions of sputtered particles based on the assumption of a fully developed collision cascades are cosine-like [29]. In other words, the polar angle θ is roughly cosine distributed while the azimuthal angle φ is uniformly distributed in the range $[0, 2\pi]$ (see Figure 2.9 for an explanation of the coordinate system). However, experiments with low energy ions on elemental [35, 66–70] and compound targets [67] have shown in some cases a significant deviation from the ideal cosine distribution. As an example, Figure 2.10 depicts the deviations from cosine distributions for 4 elements at two bombarding energies

calculated using a computational model described in [65]. The deviation from cosine-like distributions is more noticeable for low energy bombardment and is more pronounced for low mass target atoms. Figure 2.10 clearly shows that the distribution becomes more cosine-like when the ion energy is increased. However, this type of calculation should be taken only as an average since the surface roughness is expected to play a significant role in determining the true shape of the distribution [62, 68].
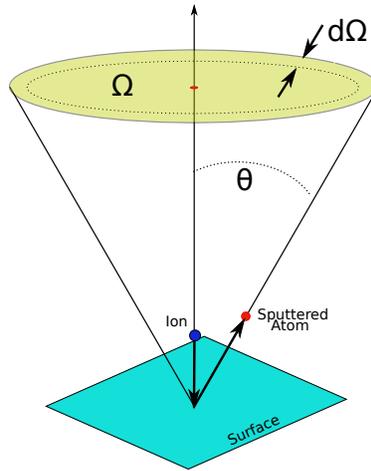
Figure 2.9: **Frame of reference for sputtered particles used in the rest of this dissertation.**

In the following, we will discuss some models that allow us to determine the initial trajectory of the sputtered particle for simulations. The simplest approach is to assume that the probability of emission is proportional to $cos^n\theta$ where the fitting coefficient n can be adjusted to match an experimental distribution. The value of n determines whether the distribution is cosine ($n = 1$), undercosine ($n < 1$) or overcosine ($n > 1$). It is also possible to modify the n coefficient to be a function of a shape parameter as proposed in Ref. 65. Here $n = 1 + ln(\Delta + 1)ln2$ and $\Delta = -4(U_s M_1 E_0 M_2)^{0.55}$ with U_s, M_1 and M_2 as previously defined. Nevertheless, a drawback of the cos^n relation is that it can not capture the heart-like shape distribution commonly observed in experiments [71]
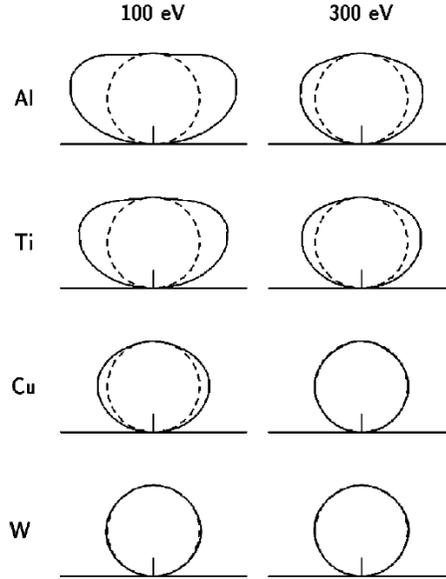
Figure 2.10: **Simulated angular distributions for four different materials. The angular distribution is broader (undercosine) at low energies for all elements except W. For higher energies, the distribution shifts towards a theoretical cosine-like form (shown as dash lines for reference). Reprinted from Ref. [65].**

To properly capture this type of distribution, we need to rely on more complex expressions. An example of an equation that captures the heart-like shape but also gives cosine-like distributions is the equation first proposed by Jones [72] which reads

$$\frac{dY}{d\omega} cos\theta \left(1 + a(1 - 2cos\theta)\right) \tag{2.10}$$

where a is a free parameter and θ the usual polar angle described before. Using this formula, the standard cosine distribution is obtained from equation 2.10 for $a = 0$. The elegance with which equation 2.10 captures a wide range of distributions with a single free parameter permits a parametric study of the effect of different distribution functions on the transport properties of sputtered particles.

The empirical heart-like shaped distribution can be also obtained from a combination of an analytical equation calibrated with experimental data. For instance, Horkel et al. [73] measured the angular distribution at the substrate and from their simulation they were able to estimate the angular distribution at the target. This technique looks appropriate in practice. Unfortunately, it requires accurate measurement of the angular distribution at the substrate for each element and geometry investigated. The equation used by Horkel et al. is

$$\frac{dY}{d\omega} \propto \sum_{n=1}^{5} a_n cos^n \tag{2.11}$$

The coefficients a_n for aluminum, titanium and copper for equation 2.3.1 are shown in table 2.1.

The difference between a cos^n relation and equation 2.3.1 is shown in Figure 2.11.

The effect of different models on the deposition rate at the substrate will be investigated in Chapter 4 using our computational model.

Table 2.1: Fitting coefficients for the analytical approximation $\sum a_n cos^n$ to the angular distribution of sputtered particles [74].

Material	a_1	a_2	a_3	a_4	a_5
Aluminum	1.573	-10.415	29.899	-34.696	13.152
Copper	2.55276	-10.8812	28.9098	-29.943	10.03786
Titanium	3.1798	-2.016665	-4.53936	6.18887	-2.42589
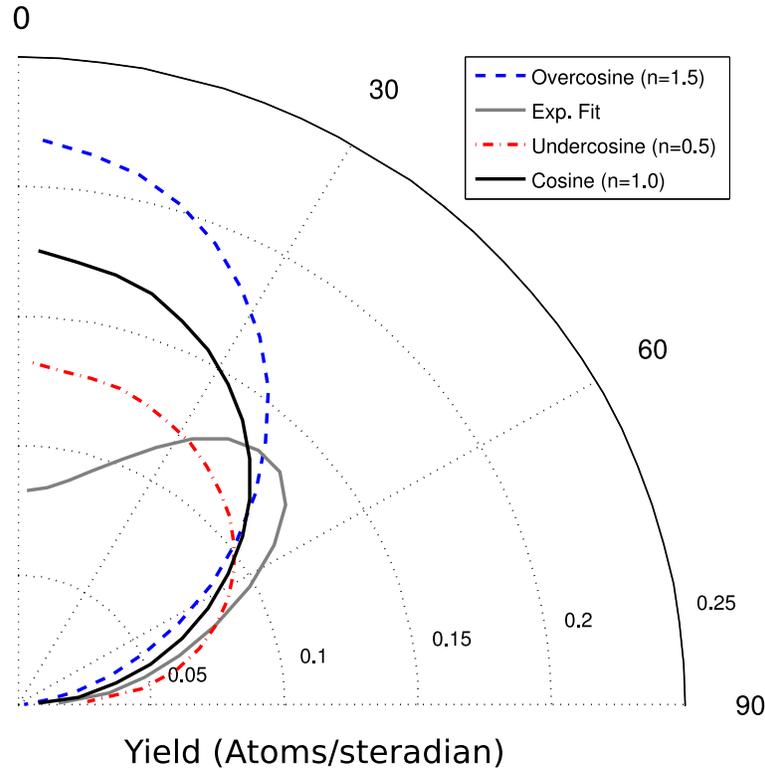
Figure 2.11: **Four different models for the angular distribution of sputtered particles. The figure shows the differential yield per solid angle in polar coordinates for three cos^n distributions (n=1.5, n=1, and n=0.5) and experimental distribution [74].**

2.3.2 Gas phase transport

Neutral particle transport

Once a sputtered atom leaves the surface, it moves towards the substrate and the walls of the chamber. The key to control the energy and angular distribution of sputtered particles at the substrate is during the gas phase transport. Each collision randomizes the angle and reduces the incident energy of depositing particles.

The energy at which these particles are emitted and the pressure regime commonly associated with sputtering depositions pose a problem for the modeling

of their transport properties. Sputtered atoms, having a wide range of energies (see Figure 2.8), are ejected from the target and may undergo none, few or many collisions in their way to the substrate or walls of the system. For instance, if the pressure is low (< 1 mTorr), the sputtered particles travel in straight lines suffering almost no collisions with the background gas. However, at higher pressures ($> 5 - 10$ mTorr), these particles experience many collisions with the background gas as a result of a reduced mean free path. A free path is the distance a particle travels between collisions. I must point out that an elastic collision between particles of comparable mass is a very efficient mechanism to lose energy, the efficiency of such mechanism depending on the mass difference. Chapter 4 explains this process in more detail.

The transfer of energy in each collision also affects the transport properties. Indeed, energetic particles undergoing several collisions deposit enough energy into the gas which increases its temperature. One of the consequences of this gas heating effect is a localized density reduction of the process gas. Both effects will have an impact on the free path of the sputtered particles. In addition, the sputtered particles also deposit momentum due to their directed velocity. We may expect that light sputtered atoms moving in a heavier gas will have a much different momentum transfer mechanism than a heavy sputtered particle traveling a lighter buffer gas. All these effects should be taken into consideration for an accurate description of the phenomena at hand.

An important concept that allows us to use methods that are more computationally efficient is that of thermalization. An energetic particle is thermalized when, as a result of continuous collisions with background gas atoms, it has lost most of its initial energy and is now comparable with the thermal energy of the buffer gas. A quantity that captures the degree of thermalization is the thermalization length. Figure 2.12 shows the effect of pressure on the thermalization of sputtered particles. In the figure, lines depicting particles trajectories end when the energy of the sputtered particle is close enough to the thermal energy of the gas. "Close enough" is a relative term, and will be defined in greater detail when the transport of neutrals particles is addressed

later in this dissertation. After thermalization, particles move with random trajectories: they have lost all memory of their original directed motion.

The key to simplify the computation is to divide the transport into two distinctive groups: energetic and thermal particles. We should then expect a very different transport model for each type of group given the great difference in free paths between these two group of particles. Recall that sputtered particles are emitted with energies well above the thermal energy of the buffer gas (for instance, the average energy is 5-10 eV while the thermal energy associated with a gas at 300 K is about 0.025 eV). An algorithm for the solution of the transport of sputtered particles is described in chapter 4. This chapter also describes in more detail the concepts behind the distinction between energetic and thermalized particles.

2.3.3 Charged particle transport

Plasmas

Plasmas are partially or fully ionized gases consisting of a mixture of electrons, ions and neutral particles. They are commonly referred as the 4th state of matter, however, a quick look at the evidence suggests that they should be considered the 1st state (it is estimated that most of the visible matter in the universe is in the plasma state [75]). Seen as fluids, no other state exhibits the collective behavior plasmas show. The way plasmas act was elegantly summarized by Chen[76]: 'plasmas behave as it they have a mind of their own'.

The history of human made plasmas dates back to the XIX century. Little was known about this type of phenomena; the technology limitations of the time limited generation of stable laboratory plasmas on which a thorough study could be performed. The works of Faraday (gas discharge in rarefied gas, 1838), Geissler (Geissler tube, 1857), Plucker (effect of magnetic fields on
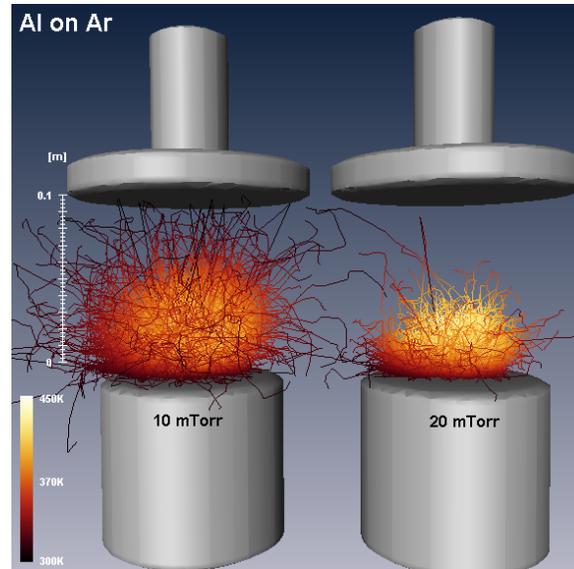
Figure 2.12: **Effect of pressure on the thermalization of sputtered particles. Each line represents a sputtered particles's trajectory after leaving the target. The end of each line is the location where the particle associated with that trajectory has been thermalized. At 10 mTorr, sputtered particles travel a longer distance before full thermalization occurs while at 20 mTorr the sputtered particles are thermalized closer to the target depositing all their energy in a smaller volume. The trajectories are color coded to show the increased temperature of the background gas as a consequence of the energy input into the gas from these collisions.**

gas discharges, 1858), Grove (sputtering, 1852) are examples of early scientific studies of low pressure laboratory plasmas[75]. The major breakthrough during this time came from the invention of a new technology to seal vacuum tubes which allowed scientists to work in sub Torr pressures [7].

In 1889 Friedrich Paschen, after an exhaustive study on glow discharges, published a formula that today is known as Paschen's law. This empirical law relates the minimum voltage for a discharge to appear between two parallel plates with the product pd where p is the pressure and d is the distance between the plates. Later, in 1929, Irving Langmuir coined the term plasma to describe the charge neutral region of the discharge [75]. It is interesting to note that two of the most used terms in the thin film industry, sputtering and plasma, were popularized by the same person. In today's world, besides being a scientist, Langmuir could have been seen as a brilliant publicist.

A wide range of phenomena in nature can be described as plasmas according to Figure 2.13. Any attempt to describe the whole spectrum would be a monumental task. The rest of this dissertation will focus on the region of the plot where sputtering plasmas happen to operate: the glow discharge.

The glow discharge

Glow discharges are characterized by the emission of visible light. This indeed can be used to our advantage. For instance, in sputtering, the glow from the plasma is used to control the process variables in industrial setups [77–80] especially in reactive environments where the control on the process parameters is essential to avoid poisoning the target [3, 81].

Figure 2.14 depicts the typical operation regions of a gas discharge between two parallel electrodes. Starting at the left region of figure 2.14a where the current is small, only a few electrons exist in the region between the plates; thus, few ionization events occur. As a consequence, the current measured in R is very low. This region is characterized by a nearly constant voltage: even a small increase in voltage results in a significant increase of ionization events
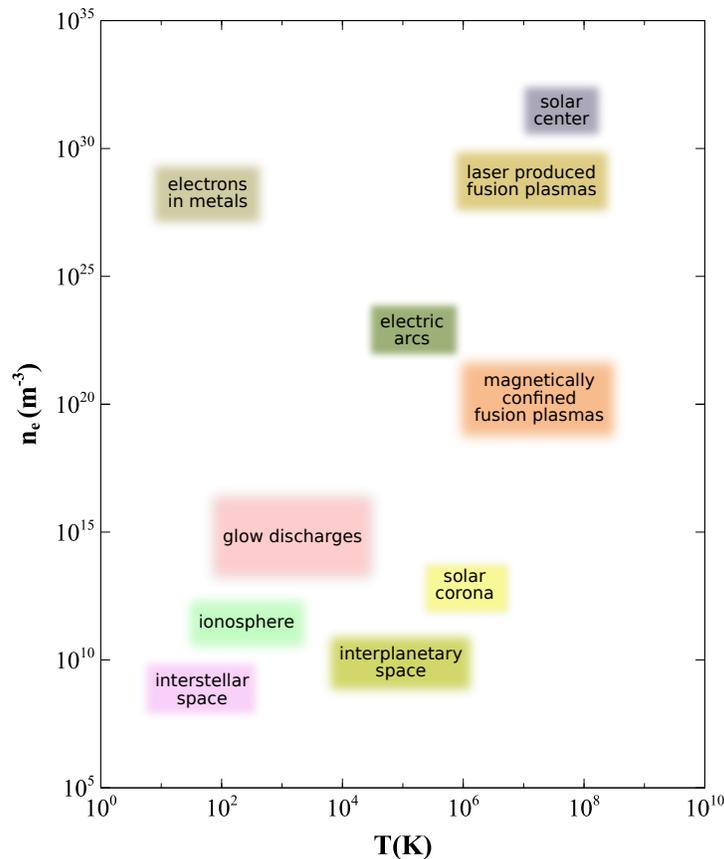
Figure 2.13: **Types of known plasmas characterized by their density and temperature.** n_e **is the electron density in** m^{-3} **and** T **is the electron temperature in Kelvin.Adapted from Ref. 75.**

that is reflected in a higher current passing through R. As current is boosted, we reach a point where the charge density in the bulk of the discharge starts to affect the potential distribution and a quasi-neutral region begins to form. This region is the subnormal glow discharge. Once most of the potential drop is localized near the negative potential, the discharge has reached the normal glow region. At this point a noticeable glow appears in a small region next to the cathode. If we continue increasing the current, the region covered by the glow expands until the whole electrode is fully covered. During this operation mode, the voltage stays constant. Once the area of the cathode is covered
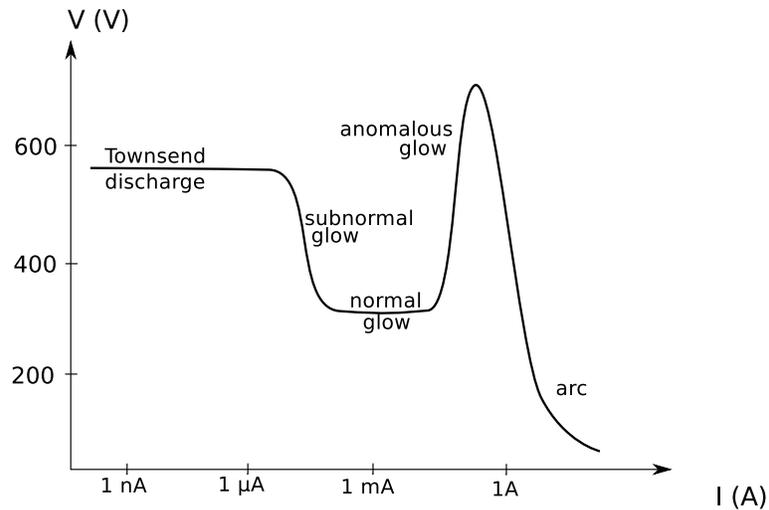
Figure 2.14: **Regions of the glow discharge as a function of current. The voltage in the discharge responds to changes in current in different ways depending on the region the discharge happens to operate. Figure adapted from Ref. 75.**

by the glow, any increase in current will result in higher voltage because the only way for the discharge to create more ionization events is by increasing the energy delivered to the discharge. We have reached then the abnormal glow discharge region typical of the operation of sputtering systems.

The basic fundamental principles of the magnetron glow discharge and the computational techniques to model this important phenomena are covered in Chapter 5. To avoid further confusion, whenever I mention the term plasma hereafter I implicitly refer to the abnormal glow discharge region described before.

The plasma criteria

So far, we have used a ambiguous definition of a plasma, that of a partially or fully ionized gas. There are three characteristics for a gas to behave as a plasma rather than a neutral gas [76]. First, the length scale over which electric field fluctuations are shielded should be much less than the characteristic

length of the system. This scale length is known as the Debye length, λ_D in honor of the Dutch physicist Pieter Debye [75]. Secondly, the time scale over which a perturbation in the plasma occurs should be smaller that the collision frequency with neutral particles ν . The frequency of typical plasma oscillations is known as the plasma frequency ω_{pe}. Finally, the number of charged particles within a Debye sphere (a sphere whose radius is λ_D) should be much greater than 1 so any electric field can be shielded efficiently. Mathematically

1. $\lambda_{D_e} << L$

2. $\omega_{pe} << \tau$

3. $N_{D_e} \ggg 1$

2.3.4   Substrate deposition

The events at the substrate are key to describe the conditions that lead to film growth. The fundamental mechanisms that follow the condensation of a vapor onto a surface to the eventual formation of a continuous film are governed by the energetics at the surface interface. A correct description is not only essential for capturing the details of the early stages of film formation but also plays a decisive role in understanding how the film evolves in time. Additionally, the substrate is constantly bombarded by several species with a wide range of energies which only complicates the analysis.

The growing process of sputtered films can be broadly grouped into three main stages: nucleation, coalescence and vertical growth. During nucleation, atoms (hereafter referred as adatoms) from the vapor adsorb on the surface. If they have enough energy to circumvent the diffusion barrier, adatoms move until they find a place more energetically favorable, otherwise they stay at the adsorbed location. The dynamics of nucleation are important since this stage plays a significant role in determining the grain structure of the film [82]. Adatoms at the substrate may also be re-emitted depending on their energy. A

group of adatoms that has clustered into a single unit is commonly referred as a nucleus. Processes such as thermally activated diffusion (determined by the temperature of the substrate), athermal diffusion (determined by the energy of the particle) and re-emission determine the nuclei evolution in space and time. In the coalescence stage, stable nuclei (islands) grow in size either by collecting more atoms or by 'eating up' smaller less stable nuclei (Ostwald ripening). At a critical size, nuclei merge with neighboring ones forming even larger islands. Eventually, if the conditions are appropriate, the whole surface is covered and the film starts growing vertically.

For sputtered films, some of the important parameters to consider are deposition rate, energy and angular distribution and temperature[83]. For instance, the microstructure of the film can greatly vary depending on process variables such as pressure or temperature. One of the early models is a structure zone model [83] depicted in Figure 2.15. The Thornton model is a modification of an early model first proposed for evaporated films [84]. Later modifications [85, 86] to the model have helped us elucidate the effect of process variables on the microstructure for thin films modifications and to find the process conditions that lead to novel films for new applications[87]. It is clear that a detailed description of all fluxes impinging on the substrate is essential if we attempt to understand growth at the very beginning stages.

Stress and adhesion of sputtered films are also key to control in real applications of thin films [88]. A typical plot depicting the type of stress in sputtered films as a function of deposition pressure is shown in Figure 2.16. In general, excessive compressive stress in films is associated with delamination and buckling while tensile stress produces cracking and decohesion [89]. The behavior depicted in Figure 2.16 can be explained using a very simplistic model for sputtered films where we assume a constant target-substrate distance (throw distance) in our description. At low pressures where films are found to be under highly compressive stress, particles arrive at the substrate with most of their initial energy since they have not lost it in collisions with the background gas. This energy activates mechanisms such as diffusion and grain migration.
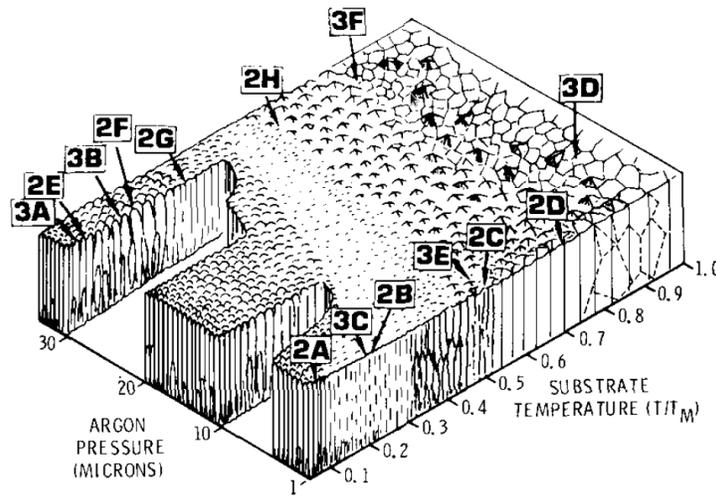
Figure 2.15: **Thornton's model describing the influence of substrate temperature and deposition pressure on the microstructure of sputtered films [83]. Low bombarding energy of condensing particles (low pressure) coupled with low deposition temperature forms films with a columnar structure. At a high deposition temperatures, the film grows denser irrespective of the incoming energy of depositing atoms.**

Adatoms can circumvent energy diffusion barriers and as a consequence they are more likely to move along the surface of the film which gives them the opportunity to find more energetically favorable places in the film. If the deposition rate is high, particles may not have enough time to find their stable position before another one buries it. The high energy of the arriving particles also causes atomic peening which in part is also responsible for the 'compression' of the atomic lattice. As the pressure is slightly increased, particles are now allowed to find their equilibrium positions and there is less strain. In this pressure regime, the film grows dense and smooth [90] but columnar and with randomly oriented grains unless the deposition temperature is sufficiently high and the deposition rate sufficiently slow to allow atoms to reaccommodate into the most favorable crystal structure. The lattice is under strain since

there are relaxing forces in the film once it grows thick [91]. As pressure is increased, sputtered particles lose energy on their way to the substrate which reduces their mobility in the surface. When the conditions are appropriate, the film can be grown with zero intrinsic stress. Increasing the pressure further, sputtered particles lose most of their initial energy and arrive at the surface fully thermalized. In this regime the film grows less dense and are rich in defects [92]. At the macroscopic scale, these voids produces a small force (Van der Waals) trying to pull atoms closer: this force is then associated with a tensile stress. I have intentionally left out the plasma bombardment in this discussion to simplify the model but, as one may have guess, it also plays a role in determining the total surface evolution of the film.

The descriptions depicted in the previous lines do not convey the complexity associated with a real deposition process. This complexity is exhibited in several layers, a complexity that has been the motivation to build elaborate mathematical models to describe the surface growth dynamics. As these models become more complex to solve analytically, scientists have turned to computational techniques as an alternative. Unfortunately, to capture the dynamics of all stages of film growth into a single model is a convoluted process, and we are forced to chose the appropriate technique according to the phenomena we are interested in studying. The main obstacle that the simulationist[3] faces in modeling thin film growth is related to the scale of the phenomena at hand [94]: nucleation takes place during the early stages of deposition and involves the interaction of few atoms during a very short time [95] while vertical growth occurs in minutes and it usually involves the interaction of thousands if not millions of atoms.

In modeling the mechanisms of film growth dynamics using computational methods, we identify two conceptually different approaches: continuum and atomistic descriptions. Their differences are briefly described below.

In continuum models, the growth is modeled by following the evolution of an analytic function associated with one or several parameters describing the

---

[3]Term I borrowed from [93]
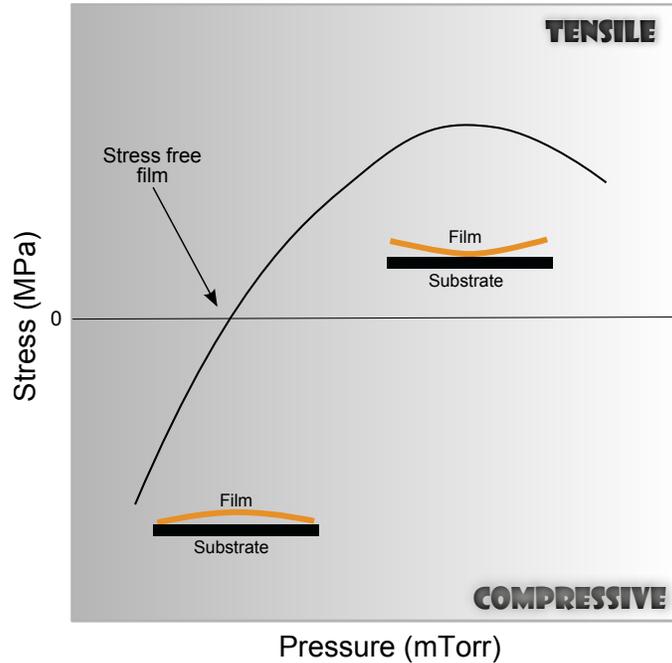
Figure 2.16: **Intrinsic stress of sputtered films as function of pressure. The figure depicts an universal trend of sputtered films. At low pressures, the sputtered particles arriving at the substrate are mobile enough to diffuse along the surface of the film. As pressure is increased, particles loss energy through collisions and arrive at the substrate quasi-thermalized (adapted from [92]).**

surface or by the solution of continuum kinetic models. By solving the equations derived from such models, the simulationist obtains detailed information about the process of interest and is able to derive scaling laws that relate to microscopic phenomena.

For instance, the height of the one of these continuum models follows the height function $h(x, t)$ via the solution of [96]:

$$\frac{\partial h(x, t)}{\partial t} = \phi(h, x, t) + \eta(xt) \tag{2.12}$$

where $h(x, t)$ is the surface height, ϕ is a function that represents the growth

effects to be modeled and $\eta(x, t)$ is a random noise function. The physics of the process is captured by the function ϕ which, in general, is very complex. Alternatively, in the kinetic model, general expressions that relates nuclei size and density with process variables can be obtained by solving rate equations that describe the kinetics of the process. Nuclei size can be characterized by a critical size [97] and the thermodynamics of stable nuclei can be summarized in rate constants obtained from experimental or *ab initio* calculations. Different derivations of these continuum models have been used to simulate several processes such as nucleation [97], diffusion [98], Ostwald ripening [99] and thermal stress [100].

The appeal of continuum models is their speed of computation and the fact that, in some cases, analytical solutions are possible which simplifies parametric analysis. However, continuum models do fail at capturing a wide range of phenomena at the atomic scale. To tackle this type of problems, atomistic models, dealing with the interactions at the atomic scale, were devised.

Starting at the smallest scale, that of the electron, density functional theory (DFT) looks at electron clouds to calculate the charge distribution within the atom. This is used to estimate the way atoms interact when placed in proximity with each other. Depending on the formulation, this method usually involves complicated mathematical expressions which are computationally demanding. DFT methods are usually limited to very small scale systems (a few atoms). However, DFT techniques are popular to obtain interatomic pairwise [101–103] or many body potentials [104] and rate constants for other methods [105].

At the atomic scale, molecular dynamics techniques describes the simultaneous interaction between two or more atoms in close proximity [106–109]. These interactions are solved using classical mechanics and involve the solution of several equations at each time step. In practice, only a few thousands atoms can be simulated using current computers.

In a different approach, Monte Carlo [90] methods are stochastic techniques to approximate the solution of a problem [110]. In short, pseudo-particles

are individually tracked within the domain of the simulation, and the type of interaction between these pseudo-particles is chosen at random from a set of processes each one with a different probability. The probability for each process can be estimated as a function of a certain activation energy which usually is an input to the simulation. Simulations of film growth using Monte Carlo techniques are well documented in literature [90, 94, 111]. One popular technique is solid-on-solid [112] which allow efficient but simple representations of 2D and 3D early crystal growth. The main drawback of solid-on-solid models is their inability to capture reentrancy, microstructure defects and voids [96].

Another popular particle based method is called ballistic aggregation [113, 114]. In this method, discs or spheres representing thousand of atoms are tracked and positioned to represent the surface evolution. This model allows the simulation of internal microstructure [96] otherwise difficult to capture by solid-on-solid implementations.

Combinations of models are of course possible and efforts to obtain a method that combines each other strengths are numerous. For instance, similar to the hybrid models of previous sections, there are models that combine continuum with discrete models models [94, 115–118].

A somewhat different concept to merge the realm of the continuum with that of the discrete into a single coherent algorithm is the Lattice-Boltzmann method. This model has recently gained popularity [119] due to its capability to model very complex phenomena. Lattice-Boltzmann methods follows the evolution of a ensemble of particles characterized by a single function. This method can be pictured as a discrete approach to of the Boltzmann equation. One of the strongest advantages of Lattice-Boltzmann methods is their accuracy for modeling two-phase processes or other otherwise intractable systems. However, the price to pay for such a model is a significant increase in complexity to build the model to solve and to program it.

Energy flux at the substrate

The heat flux also plays a decisive role in the growth dynamics of thin films [120]. The energy per atom plays an important role in determining. The importance of this parameter is not well reflected in the literature. Little is known about the relationship between process conditions and heat flux at the substrate in magnetron discharges. One of the reasons is the large number of sources that contribute to the total heat input in a sputtering system

The thermal flux at the substrate can be broadly divided into two main sources: particles and photons. Among the former group, sputtered atoms and electrons are responsible for most of the heat flux in unbiased DC magnetron sputtering. In other techniques, e.g. (High Power Impulse Magnetron Sputtering or HiP-IMS for short), the thermal flux is dominated by metallic ions which are about one or two orders of magnitude higher in energy than in DC or RF magnetron techniques. The energy distribution of these energetic ions can be tailored by selectively choosing a bias potential at the substrate. The second energy source, which includes photons from excited atoms and thermal radiation from the target, is more difficult to quantify. These inputs are commonly neglected but may play a significant role at high currents.

A third energy source in glow discharges is thermal conduction from the heated gas. Although the discharge is operated at a pressure where it is unclear whether classical fluid theory is applicable, back of the envelope calculations indicate that thermal conduction indeed may be a significant source of heat at the substrate. This topic will be addressed in chapter 6.

The main sources and sinks of energy at the substrate during a deposition process are summarized in Figure 2.17. Mathematically, The total heat flux can be dissected into several components [121] as

$$P_{tot} = \sum P_i + P_e + \sum P_n + \sum P_{rad} + \sum P_{film} \qquad (2.13)$$
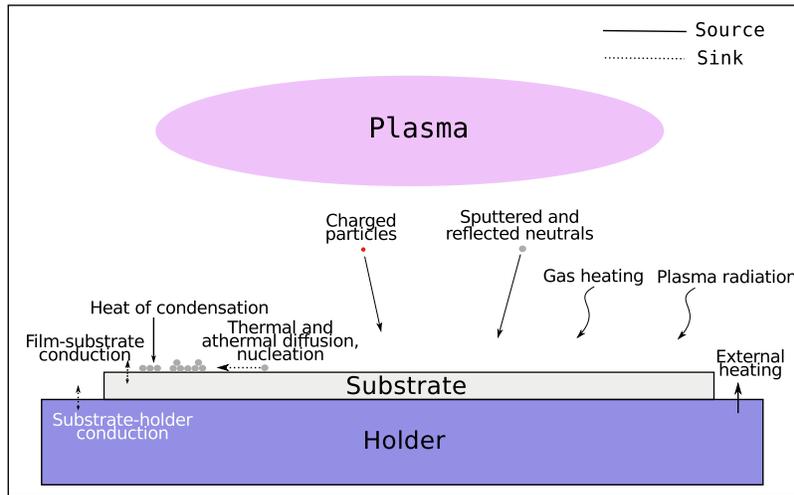
Figure 2.17: **Heat flux sources and sinks at the substrate during a typical sputter deposition process. The sources include particles fluxes from the metallic species and charged particles. Adapted from [121].**

where P_i and P_e are the heat input coming from ions and electrons respectively, P_n is the energy coming from neutral gas particles, P_{rad} the energy coming from the plasma and hot surfaces within the reactor and P_{film} the energy due to reactions and processes at the substrate. Table 2.2 summarizes the contributions to the total heat flux in a magnetron sputtering system

2.4   Industrial impact

Sputter deposition is one of the most used techniques to deposit thin films with industrial applications. As mentioned in Chapter 2.2, during the early days of thin films, the market was dominated by decorative and optical applications. During the last decades, we have seen an exponential increase of sputtered based coatings which are used for a wide range of applications. Some examples of industrially relevant films are:

- Decorative TiN, ZrN, ZrO_xN_y

Table 2.2: Range of particle's energy in sputtering

Species	Pressure	Energy
Energetic electrons	low	>10 eV
Bulk electrons	intermediate and high	$\approx 1 - 4eV$
Energetic ions	intermediate and high	>10 eV
Bulk ions	high	<0.1 eV
Plasma iradiation	high	n/a
Hot surfaces	high	n/a
Energetic sputtered particles	low	≈ 10 eV
Reflected neutrals	low	>10 eV
Thermalized neutrals	high	<0.1 eV

- Optics Ag, Al

- Glass (Low-e, self-cleaning glass, energy efficient glass, windshields, anti-reflective) TiO_2, SiO_2. Ag

- Photovoltaics: Mo, ZnO, ITO, Si

- Displays: ITO, Mo, Al

- Microelectronics: W, Ta, $AlTi$, $FeNi$, Co alloys, TaB, Cu, TaN, Au

- Medical films: Ag, TiO, Hydroxyapatite

- Tribological: TiN, CrN, $Si-C$, ZrO_2, ZrC, ZrN, DLC

- Food packaging: Al, Al_2O_3, SiO_2

With the wide range of applications, a computational tool that allows us to simulate the process would assist in both designing and optimizing of deposition conditions. However, from a computational perspective, the broad range of configurations this process may take poses a very complex problem to solve.

Even a small improvement in an industrial process, whether to improve the deposition rate or to increase the target utilization, in the long run, can represent a significant improvement in industrial environments where a constant search to optimize and reduce cost is ongoing.

2.5   Thesis challenges

The primary objective of this dissertation is the integration of a comprehensive computational tool for the simulation of the sputtering process. The integration involves the connection of several loosely coupled modules, each one having a very specific task. The main challenge is to find the appropriate structure that will allow us to integrate these specialized modules when the time scales associated with each process span several orders of magnitude. For instance, Table 2.3 describes the time scales normally associated with each phenomena in sputtering. The method to couple all modules in a self-consistent algorithm will be uncovered in chapter 6 when the integration of the plasma and neutral modules is addressed.

Table 2.3: Time scales for sputtering

Physical phenomena	timescale
Magnetic fields	static
Electron energy	< ns
Electron Transport	ns
Ions transport	μs
Excited species	μs
Energetic neutrals	μs
Gas heating	ms
Rarefaction	ms
Thermalized neutrals	ms
Deposition	min

The main objectives of this thesis are

- **Development and integration with the remaining components of a 3D magnetized plasma model for the simulation of a magnetron sputtering discharge**: This model deals with the numerical solution of a nonlinear system of partial differential equations to solve for the transport of charged particles in a cold plasma under the influence of a nonuniform magnetic field. This component is essential for the complete integration of the simulator.

- **Completion of missing, unfinished and untested components of the computational framework**: Some objects and modules of the framework lacked functionality and needed to be completed.

- **Integration of individual modules into a comprehensive suite of simulators**: The integration involves the correct communication between modules and the proper description of the physics. This stage needs a deep knowledge of the building blocks of the framework and sound understanding of the physical processes behind each module so the integration can be performed appropriately.

- **Experimental validation of key components of the framework**: Benchmark simulation results with experimental measurements on a magnetron sputtering system.

- **Characterization of the energy flux at the substrate in an industrial coater using a scaled-down magnetron as a model**: This objective deals an industrial component of this thesis and involves the development of a heat flux measurement system for industrial applications.

2.6   Summary

This chapter highlighted the basic principles of magnetron sputtering systems. While it did not include detailed description, I briefly described the

fundamental phenomena associated with the process to give the reader a basic introduction to the topic. First, I discussed the events that drives the release of sputtered particles from the target. I reviewed the most popular published models to estimate total and differential yields. Next, I examined the transport of neutral particles through a gaseous medium and discussed a few of the difficulties associated with this phenomena. I continued with a short description of a plasma in a sputtering discharge. The basic mechanisms that are in play in the range of operation of a gas discharge were addressed. Afterwards, I discussed the mechanisms of film condensation with a focus on computational techniques to model this interesting phenomena. I briefly dissected the contributions to the total heat flux at the substrate during a typical sputter deposition. Finally, I stated the specific objectives of this dissertation.

In the next chapter, the details of the computational framework are described. We will briefly cover the main concepts behind the software architecture and fundamental abstractions on which the software is built upon. Additionally, we will list the individual contribution to the project in an attempt to delimit the scope of this dissertation.

CHAPTER 3

SpudII: A framework for the simulation of physical vapor deposition processes

3.1 Introduction

This chapter introduces the computational architecture of the SpudII framework. The chapter's goal is to review the basic concepts of the project to lay a solid foundation for the incoming chapters. We will briefly describe the modular architecture and the interaction between modules using simple concepts to offer a brief, yet accurate, description of the overall structure of the project. For the sake of clarity and convenience, we will keep software engineering terms to a minimum and use simpler language when possible. Some of the concepts hereafter expressed are borrowed from Sean Leonard Master's thesis [122]. The reader looking for a more comprehensive description of the framework, one using a more appropriate software engineering terminology, is referred to this document. This chapter ends with a detailed description of my contribution to the project.

3.2   The SpudII project

The SpudII project was developed to increase the scope of a previous simulator [123] written by Dr. Steven Dew at the University of Alberta in the early 1990's. One of the original goals of this simulator, SIMulation of SPUtter Distributions (SIMSPUD), was the study of step coverage over high aspect ratio features for advanced IC metallization. This program evolved from a research environment to an industrial tool within a few years. However, the software soon reached its technical limits after several attempts to increase the applicability of the simulator to accommodate several industrial requirements [122]. It was clear that the software needed a major architectural change to keep with the demands of industry.

The choice of C++ as the development language for the SpudII project in favor of other alternatives was due to several factors. Some of them include speed to develop numerically intensive scientific applications, strong object-oriented support and the original estimate that the scientific computing community preferred C++ as the language for applications [122]. The price to pay was an overhead for programming in C++ for non-trained programmers that delayed the evolution of the program at the projected paced.

The structure of the code was a key concept during the early stages of development. Also, a key question was that of building a generic framework or a targeted application. An application, as opposed to a framework, is software that is designed to perform an specific task, to provide a specific solution for a very specific problem. In contrast, a framework is a set of components used to build a generic solution to a set of very similar problems. Leonard [122] argued that "one must consider if an application framework is more appropriate than just creating an application" (page 19) and expanded the reasoning behind the decision to build a simulation framework. From the description of the problem at hand depicted in the previous chapter, Leonard concluded that the development of a simulation framework was the appropriate choice given the wide range of phenomena involving a typical sputter deposition process.

The SpudII project addresses several aspects of a typical PVD process using a modular approach. This modular architecture where each module is a highly specialized component dealing with a very specific phenomena is required for the appropriate description of the sputtering process using computational methods. The modular architecture is better attained by following an object-oriented philosophy. Some of the concepts of object-oriented frameworks [124] on which this project is built are modularity, reusability and extensibility. Modularity is attained by encapsulating loosely coupled objects (modules) from the rest [122]. One of the keys of good modular architecture is the implementation of well-designed, robust interfaces [125]. Reusability, on the other hand, refers to building generic, reusable components to create new applications [124]. One of reusability's goals is to improve the quality of the software by reusing components already tested [125]. Finally, the concept of extensibility refers to the capability of the architecture to allow new components without a significant framework redesign [122]. By extending the current implementation of the framework, the developer can add features and increase the scope of the framework in a structured, systematic way [124].

The project in its actual revision consist of about 90,000 lines of code divided into several files. Figure 3.1 shows the output of running an open source program to count for lines of code in a project tree [1].

3.3 High level description

A typical simulation is controlled by an abstract entity which supervises the creation and destruction of objects during the lifetime of the simulation. This object controls all aspects of the simulation from initialization to finalization. From a very high level perspective, there are 3 basic types of objects other than the simulation object in our framework [122]:

1. Execution Modules (ExecutionModule): These are modules that execute

---

[1]cloc.sourceforge.net

```
    2372 text files.
    1585 unique files.
    1107 files ignored.

http://cloc.sourceforge.net v 1.09  T=4.0 s (229.8 files/s, 62494.0 lines/s)
-------------------------------------------------------------------------------
Language              files   blank   comment      code   scale  3rd gen. equiv
-------------------------------------------------------------------------------
C++                     190   15855     25499     58773 x  1.51 =       88747.23
C/C++ Header            625   17195     67442     44413 x  1.00 =       44413.00
XML                      34      86         0     12725 x  1.90 =       24177.50
make                     30     625       183      1940 x  2.50 =        4850.00
C                        16     601      1728      1905 x  0.77 =        1466.85
MATLAB                   16      45        86       404 x  4.00 =        1616.00
Teamcenter def            2      39        65        97 x  1.00 =          97.00
Bourne Shell              2      38         6        93 x  3.81 =         354.33
Bourne Again Shell        3      10        14        77 x  3.81 =         293.37
lex                       1      12         0        20 x  1.00 =          20.00
-------------------------------------------------------------------------------
SUM:                    919   34506     95023    120447 x  1.38 =      166035.28
-------------------------------------------------------------------------------
```

Figure 3.1: **Lines of codes as counted by the open source program cloc.**

a specific component of the process. An application is formed by inter-operability of Execution modules. These modules can only be owned by the Simulation Object.

2. Process Modules (ProcessModule): These modules can only perform an specific task if they are owned by a parent object.

3. Low Level Objects (LowLevelObjects): These are generic objects not associated in particular to any module. There are widely used during a typical simulation: vectors, grids, points, cells, etc.

The communication between modules is implemented using abstract interfaces to keep the architecture from depending on each other. Leonard gives a good description of the fundamental principles behind the need for interfaces [122]. Figure 3.2 is a simplified scheme showing the interaction between modules and of how these objects are connected during a simulation. In the figure, a proxy module is an implementation of an incomplete module to appear to the rest as a fully functional component. Proxy modules play an important role in the development of new modules in the framework.
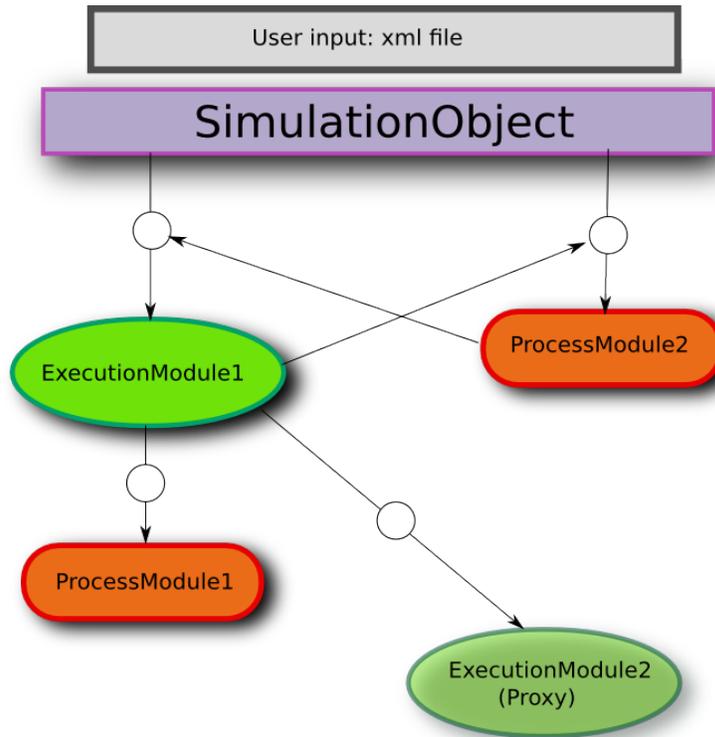
Figure 3.2: **High level description of the SpudII project. Circles refer to abstract interfaces which connect components of software. Arrows depict the flow of data between components for a specific application.**

3.4 Modules

With the basic architecture of the project already described, we are in a position now to relate the modules of the framework to the physical phenomena they model. As mentioned in the previous chapter, the sputtering process involves a myriad of intertwined processes characterized by very different time scales. Recalling Table 2.3 for instance, high energy electrons have a time scale in the range of ns, high energy neutrals in the order of ms and gas heating and diffusion of sputtered particles in seconds. Any attempt to build a comprehensive model that conglomerates all phenomena into a single global model would

be computationally inefficient, impractical and extremely complex to say the least. This great difference in time scales and different transport mechanisms is the motivation to construct a specialized scheme where phenomena are decoupled to facilitate the computation.

In general terms, the approach followed by the SpudII project is to decouple charged particle motion from neutral motion by encapsulating them in individual modules. This is done because of the intrinsic differences between these two processes. Each module deals with a specific component of the process. Moreover, within each module, a particle's transport is modeled using a different approach which depends on the energy of the particles. Thus, the model to solve for low energy or thermalized particles is structurally different to the model to solve for high energy particle's transport. The motivation, and the physics behind these approaches, will be detailed in chapters 4 and 5 respectively.

A simplified diagram of the main modules in the project is depicted in Figure 3.3. To simplify the description, each module or submodule is shown in boxes. The color coded boxes represent the type of physics they solve, whether they are purely kinetic (particle based) or thermodynamic (fluid approach) models. The communication and the type of information shared by the modules is also shown in the figure. For instance, the TargSpud module shares ions and secondary electrons with PS1 but not with any other component of the program. The same module shares information regarding sputtered and reflected neutrals with TS1. Then TargSpud can be labeled as the module creating energetic particles in the simulation region. The arrows going from module to module represent abstract interfaces to keep modules isolated from each other.

3.5 Running a simulation

There are five distinguishable stages during the course of a simulation:

Construction The simulation object parses the input xml file and constructs
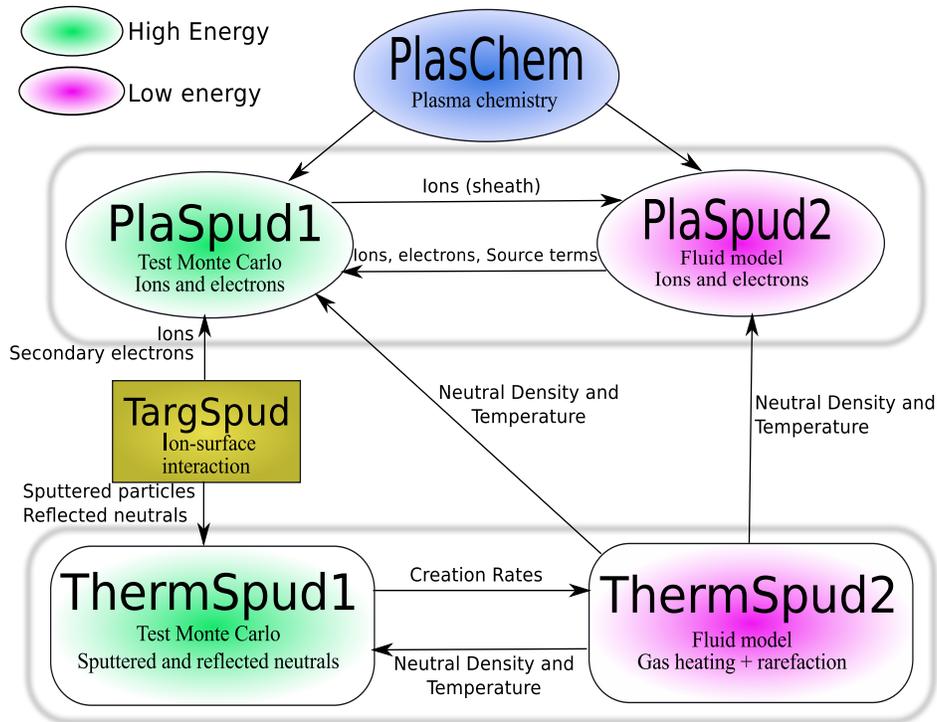
Figure 3.3: **The modular structure of the SpudII project. Depending on the application, we can instruct the simulation object to only build the modules that are relevant to the physics we attempt to simulate. Color coded boxes reflect the type of physics each module solves for.**

the simulation object. The proper construction of the objects involved in the simulation depends on the correct parsing of the xml file. Typical complications at this stage are corrupted xml files, missing components or badly formatted input files.

Initialization This stage is characterized by the initialization of all modules created in the simulation. Grids are created and the basic relationship between components is built.

Execution The core of the computations takes place during this stage. It is at

this stage that all ProcessModules scheduled to execute are sequentially run according to a virtual queue controlled by the SimulationObject object.

Output Once the simulation is deemed finished, data collected during the run is dumped to output files. Examples are grid files, surface statistics, particle tracking, etc.

Destruction Finally, all objects in the simulation space should be properly destroyed to liberate memory in a safe way.
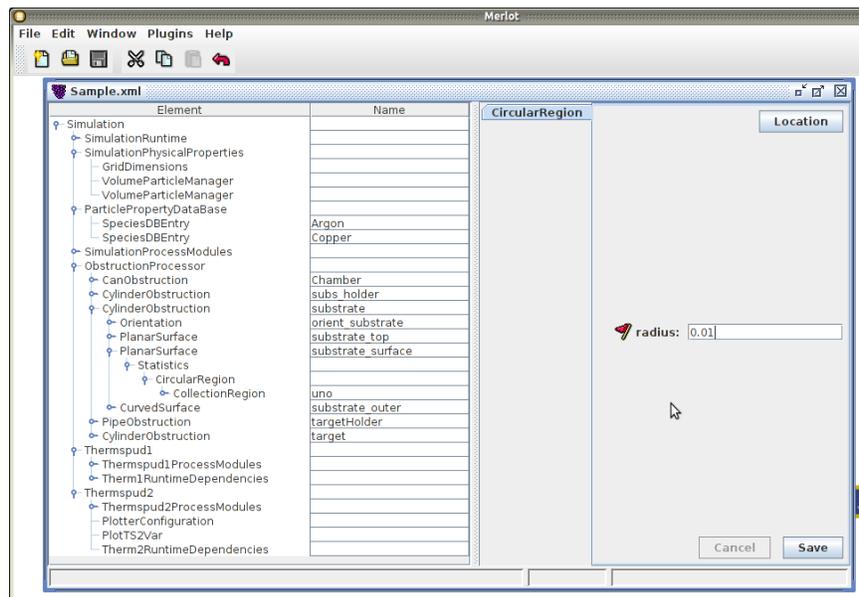
Figure 3.4: **Example of an xml input file for the simulation of the sputtering of copper using argon as the process gas. The input file shows a typical chamber configuration consisting of a cylindrical chamber, a substrate, a substrate holder, a target and a target holder. To simplify the editing of the input file, the xml file was opened using an xml editor with a graphical user interface.**

As an example, Figure 3.4 shows an example of an input file used for the simulation of a typical copper sputtering process. Each line of the input file

is parsed by a third party parser and the input information is sent to the simulation object which then creates the virtual simulator.

The SpudII program has been tested on a great variety of PC configurations, ranging from a typical desktop computer running any distribution of the Linux operating system or high-performance Linux cluster. The amount of RAM required to run an instance of the program strongly depends on the size of the grids employed to discretize the transport equations.

3.6 Individual contribution to the project

For a complex multiphysics project to be successful, it requires the participation of several people ranging from computer scientists, plasma and surface physicist, engineers and scientists. This project is no exception. Several people have contributed along the years since the original project was first described. Figure 3.5 gives a summary of the people that have been involved in the project along the years.
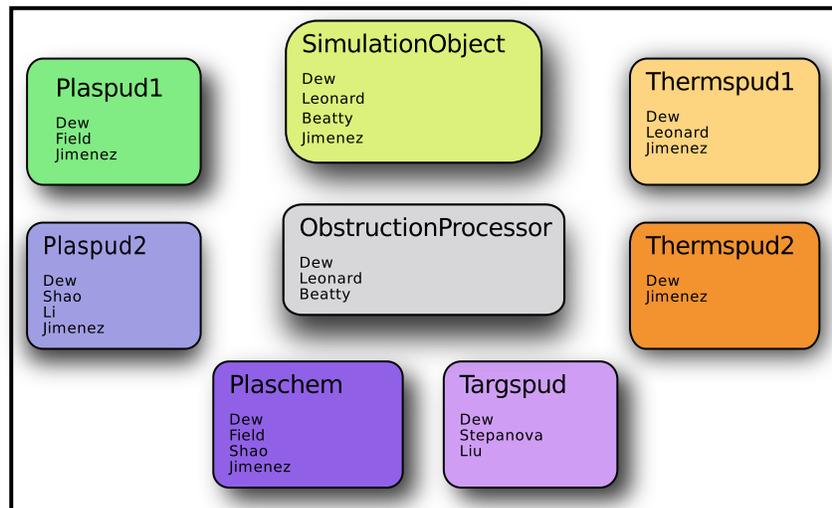
Figure 3.5: **Historic collaboration on the spudII prject.**

It's difficult to delimit the contribution of each scientist to the overall structure of the whole in simple terms. However, we must explicitly describe the components of the project on which I have been involved so my contribution could be gauged appropriately. The following is a representative but not comprehensive list of my contribution to the development of the project. I have avoided small details that, although minor to comment about, demanded a substantial amount of my time.

SpudII Architecture: Iteration order, implementation of integrating factors of the multiscale model of sputtering depicted in Figure 3.3, refinements of input parameters. Debugging all modules.

Plaschem: Plasma chemistry involving metal gas atoms and metastables, ion mobility model.

ObstructionProcessor: Collision model at surfaces, implementation of surface statistics at objects.

Plaspud1: Minor modification to collision model, handling of metal vapor densities and argon metastables.

Thermpsud2: Optimization of numerical method for thermalized neutrals [126], experimental validation of the thermalized neutrals transport and extension to include solution for more neutrals (Chapter 4).

Plaspud2: Redesign, implementation and validation of 3D numerical fluid model for a magnetized discharge (Chapter 5).

Thermspud1: Minor change to collision model, experimental tuning and validation of high energy neutrals transport (Chapter 4).

LowLevelObjects: Integration of linear algebra libraries for sparse matrix computations (Chapter 4).

Integration: Integration of modules (Chapter 6).

Validation: Experimental results in this thesis

Code management: Maintenance of software repository

3.7 Summary

In this chapter, I described the architecture of the SpudII framework and delineate the basic computational concepts upon which the project is built. I provided a high level description and introduced the modular architecture of the project which allowed us to relate each module with the physical phenomena each module is designed to solve. Finally, I delimited the individual contribution to the project so the reader is in a position to weigh my participation in the development of the computational framework presented in this dissertation.

In the next chapter, I present an algorithm for the solution of mass, heat and momentum transport in a magnetron sputtering system. I will start with the basic premises of our method of solution and end with an experimental verification of the model.

CHAPTER 4

ThermSpud: A module for heat, mass and momentum transport [1]

4.1 Introduction

In this chapter, I will review the development of a computational module that solves for the transport of neutral particles in a magnetron discharge. First, I briefly review the literature on computational methods for the simulation of this phenomena. Next, I continue with the description of the ThermSpud module. The motivation for the computational treatment of the problem is addressed within this part of the chapter. I first describe the energetic neutral module previously introduced in Ref. 122. Afterwards, the focus is on the modules that solve for the thermalized neutral's transport. A full description of the model is available elsewhere [126]; I summarize it for completeness. The key components of the module that were added as part of this dissertation will be addressed in this section.

The last half of the chapter is devoted to a thorough benchmark of this module. For this purpose, I will compare simulations with experimental results available

---

[1] Jimenez et al. J. Vac. Sci. Technol. A. 24, 1530 (2006)
Jimenez and Dew. J. Vac. Sci. Technol. A 30, 041302 (2012)

in the literature and from experiments performed in our test chamber. These experiments are representative of a wide range of pressures and powers. This chapter wraps up with a brief discussion of the importance of incorporating a neutral gas simulation for the modeling of magnetron discharges.

4.2 Transport of neutral particles in magnetron sputtering

The main obstacle when modeling the sputtered particles' transport phenomena stems from the pressure regime wherein typical magnetron discharges operate. Sputtering, as a deposition process, takes place in the transition or Knudsen regime flow right between free molecular and viscous flow [123]. The parameter that characterizes the type of flow is the distance between collisions or free path, λ. When λ is greater than the characteristic length of the system, we deal with a free molecular flow. In this type of flow, collisions between alike particles are rare and the flow is dominated by the interaction between particles and surfaces. On the other hand, when λ is much shorter than the characteristic length of the system, the flow is viscous and amenable to analytic treatment. The transition or Knudsen regime lies between these two types of flows [127]. In this pressure range, it is hard to establish a priori whether a molecular or a viscous treatment is more appropriate to solve for the transport of sputtered particles.

Despite the pressure regime challenge, there have been several efforts to model this stage of the sputtering process. In one of the first works on the topic, Ecker and Emeleus [128] identified the need to divide the flux into two components. They labeled these components as persistent and diffusive. A persistent flux, as they defined it, is the type of flow associated with particles colliding anisotropically whereas a diffusive flux is one where particles are randomly scattered. Using this approach, they were able to predict the sputtered material density between the electrodes of a diode sputtering system. Westwood [129] treated

the problem in more detail and developed a deposition rate model for diode sputtering. The model used two parameters to approximate the deposition rate: the number of collisions a particle undertakes before thermalization and the product pd_0 where p is the pressure and d_0 is the target-substrate separation or throw distance. Keller and Simmons [130] used an analytical approach based on a phenomenological model derived from the concepts laid down by Ecker and Emeleus [128] whereby particles are divided into two populations, one streaming and the other diffusing. In their model, the source of these streaming particles is the target while the source of the diffusing particles is formed by all streaming particles that have been slowed down by collisions with the background gas particles. The equation they obtained from this simple model is known as the Keller-Simmons equation, and it is still used today to predict the particle flow at the substrate [131–137]. One key argument against this model is that it does not include other phenomena commonly associated with sputtering discharges such as the nonuniform target erosion, gas heating, rarefaction and the plasma interaction. To address some of these problems, Palmero et al. [138] modified the Keller-Simmons equation to include other processes such as gas heating and the interaction with the plasma, whereas Ekpe et al. [139] included the target events (nonuniform erosion, undercosine emission) and extended the model to three dimensions. Finally, a similar model is presented in [140, 141] whereby, to approximate the source term of a diffusion equation, the authors use a thermalization length as a free parameter of the simulation. All these analytical models are able to predict the particle flux at the substrate but only for a specific combination of gases and geometries. Furthermore, the analysis is constrained in most cases to only one dimension, and they strongly depend on the correct choice of the underlying free parameters. Overall, these simplified analytical models fail at describing the dynamics of the process as a whole and can only provide general trends when they are applied to real deposition systems.

It was soon discovered that, in addition to deposition fluxes, the energy deposition at the substrate was also key to describe the microstructure of the film. Most of the models described above did not include the energy com-

ponent. Cadieu and Chencinski [142] developed a model based on a slowing down approximation where a particle loses energy as a function of the traveled distance. With their model, the authors were able to predict the average energy of sputtered particles at the substrate as a function of the pd_0 product. With this information, they correlated their findings with the properties of the superconducting films they deposited. For instance, they observed that T_c, the superconducting transition temperature of their films, was a function of the average energy at the substrate and concluded that a certain degree of thermalization promoted the type of growth needed for their films to exhibit a high T_c. Later, based on a similar concept, Gras-Marti et al. [143–145] developed a model to approximate the average energy using a similar slowing-down approach. However, these analytical approaches share the same disadvantages with the previous models: they may produce accurate predictions for a specific set of conditions, but they do not capture the physics of the process. Furthermore, they are of little help to engineers working in complex geometries typical of industrial environments.

To circumvent the shortcomings of analytical models, several authors searched for other alternatives for simulating the sputtering process. One of such alternatives is based on a kinetic treatment of the phenomena. The epitome of kinetic theory is the Boltzmann equation which describes the state of a gas using an statistical description of the medium (see 5). However, the solution of the Boltzmann equations is extremely complicated even for simple geometries. The complexity stems from the nonlinear nature of the collision term and the great number of variables in the equation. We therefore should rely on methods that approximate the solution of the Boltzmann equation.

The Monte Carlo method is an example of a technique aimed at approximating the Boltzmann equation. This technique has been widely used to describe the evolution of sputtered particles in several geometries. The most popular method in sputtering simulations is the particle test Monte Carlo (ptMC) [146]. In this method, particles are individually tracked during their life time by solving equations describing the forces acting upon them. The application of the

ptMC method to sputtered particle transport in magnetrons was pioneered by Motohiro et al. in the early 1980's [147]. In their model, hard spheres of constant radius were individually followed from the target until they were lost to the walls of the simulation. Particle scattering was assumed isotropic in the center of mass frame as predicted by classical mechanics theory. Somekh [148] extended the model to include soft spheres whose effective radius was energy dependent. The model to approximate the effective radius was based on the 6-12 Lennard-Jones potential and some empirical relations extracted from previous experiments. Turner [149, 150] and Serikov et al. [151] further improved the model to include a nonuniform temperature distribution in the chamber. The temperature distribution was obtained by coupling an energy equation with their tpMC model. Myers et al. [152, 153] identified the need for a better scattering treatment and used a universal potential to calculate the scattering angles in the collision. Their model reveled subtle differences when different interatomic potentials were used. Yamamura et al. [45] used their own program (ACAT) to obtain better estimates of the sputtering yield and the angular and energy distributions of sputtered particles. They also included the effect of reflected neutrals in their modeling. Dew [123] used an energy dependent cross section to speed up the computation model and coupled his algorithm with a feature scale model for the deposition over high aspect ratio features. Bogaerts [154] used a network of simulators where most aspects of the simulation were accounted for and found a significant heating of the process gas as a function of the sputtered material. Vyas et al. [155] also developed a comprehensive model of sputtering and used it for the improvement of an ionized sputtering deposition reactor. Mahiue et al. [87] improved the collision model by considering the effect of a non-stationary gas atom. Their modeling is suitable for a large set of geometries and has been extensively tested against experimental data at low pressures [73, 74]. Kolev et al. [156, 157] used the model of Ref. 154 to include the plasma events and the nonuniform erosion rate. Other approaches that use the ptMC for the simulation of sputtered particles include [158–163]. While this method is best suited for the low pressure regime, it is extremely inefficient at high pressures where collisions are

more frequent and particles lose energy relatively quickly. Computationally, simulations using the ptMC in the high pressure regime ($\gtrsim 15 mTorr$) are lengthy and inefficient. Finally, a limitation of ptMC models is their inability to simulate the process self-consistently.

Another approach to simulate the sputtering flux is the direct simulation Monte Carlo (dsMC) first proposed by Bird [164] and later modified by Nanbu [165] for sputtering. Briefly, the domain to simulate is partitioned in cells each one containing a determined number of particles. To start the simulation, a set of velocities is sampled from a velocity distribution and all particles are moved one time step δ_t according to their sampled velocity. The choice of δ_t is critical: one normally is limited to $\delta_t < \tau$ where τ is the time between collisions. This assures that we are decoupling the collision process from the collisionless molecular motion [165]. After a δ_t has elapsed, the collisions between particles are solved for. The requirements for a statistical representation of the process puts a constraint on the minimum number of particles per cell that need to be simulated. Nanbu showed that one can recover the Boltzmann equation using this method [165]. While this is a robust method, the constrains of the dsMC imposed by the small time step restriction limits the range of application of the model. The dsMC method is usually combined with some other techniques to avoid excessive computations.

The solution of the Boltzmann equation can also be approximated by solving for the averages or moments of the distribution function. The set of equations thus obtained is known as the Navier-Stokes equations and it is the backbone of many computational fluid dynamics simulations. The mathematical derivation of such equations is beyond the scope of this thesis and the reader is referred to a more appropriate literature on the topic [166–168]. In this approach, a set of differential equations describing the transport of mass, energy and momentum is solved. The underlying assumptions of the model only allow us to apply it if the medium can be approximated as a fluid. This assumption means that collisions between alike particles should be more frequent than with the walls of the enclosure. Therefore, this scheme is not appropriate for the energetic

component of the sputtered flux which has a long λ_m.

An interesting approach is to combine a particle method with a continuum approach to increase the range of validity of the model. This scheme is known as hybrid modeling and it has been used with success for the simulation of sputtering transport [126, 169, 170]. The premise of hybrid schemes is to combine the advantages of two techniques in such a way that the weakness of one method is covered by the second and vice-versa. For instance, we can model the highly nonequilibrium component of the flow by a particle model and switch to a faster, continuum model once a certain set of conditions are fulfilled.

To the extent of the author's knowledge, only a few models have incorporated a nonuniform gas density and temperature into their calculations iteratively [126, 154–156]. This is an important limitation of earlier models since the free path of sputtered particles is a function of the gas density and temperature. In an earlier work [126], I showed that the inclusion of inhomogeneous gas heating and rarefaction effects in the ptMC algorithm has a substantial effect on the transport properties of sputtered particles at pressures above 10 mTorr. There are fewer global models that couple the plasma, the interactions at the target and the gas phase transport in a self-consistent simulation of a magnetron sputtering discharge [155, 156, 171, 172]. Self-consistency is the key to properly couple the interactions at the target with the gas phase phenomenon found in a typical sputter deposition process. This chapter only deals with the neutrals module; the interaction with the plasma is left for a subsequent chapter.

4.3 ThermSpud module

The ThermSpud (TS) module is a hybrid approach that combines a ptMC algorithm with a fluid model. Both sub-modules are designed such that the transport of high energetic particles is decoupled from the thermalized transport. The coupling between the modules is solved iteratively. The internal

structure of the TS module is shown in Figure 4.1. The TS module divides the transport into two components [126]: one energetic and another in quasi-thermal equilibrium similar to the Ecker and Emelus model. The novelty of the model is the inclusion of the momentum equation to solve for momentum transport in the gas [126].

The algorithm of the TS module is shown in Figure 4.2. The simulation starts by parsing an xml input file to extract all the required information to build the virtual environment. The algorithm first solves for the energetic particle transport in TS1. Using the algorithm shown in Figure 4.2, TS1 does not need to create a grid to operate. In fact, when energetic particles are tracked, they only need information regarding the local temperature and density of the cell on which they happen to collide. The required information is extracted from the TS2 module using an abstract interface as explained in Chapter 3. In the same manner, all thermalized particles are dumped into the TS2 grid as source terms to the continuum equations. Once TS1 has sequentially simulated enough particles, the algorithm jumps to the TS2 module for its execution. In the TS2 module, the density and temperature grid is updated by solving the Navier-Stokes equations, and the updated information is fed back to the TS1 module. With the updated temperature and density grid, the original trajectories calculated in the previous iteration of the TS1 module need to be recalculated resulting in a new TS1 run and a new distribution of sources for TS2. The process is repeated until the difference between two successive temperature and density grids is small. This process normally takes about 3 or 4 global iterations depending on how inhomogeneous the gas density is.

4.3.1   TS1: High energy neutrals transport module

TS1 is an improved version of SIMSPUD [123] allowing for transport in an inhomogeneous gas. It uses the test particle Monte Carlo method (tpMC) to solve for high energy particles transport [146]. In the tpMS, test particles are created and individually followed during their life time in the reactor.
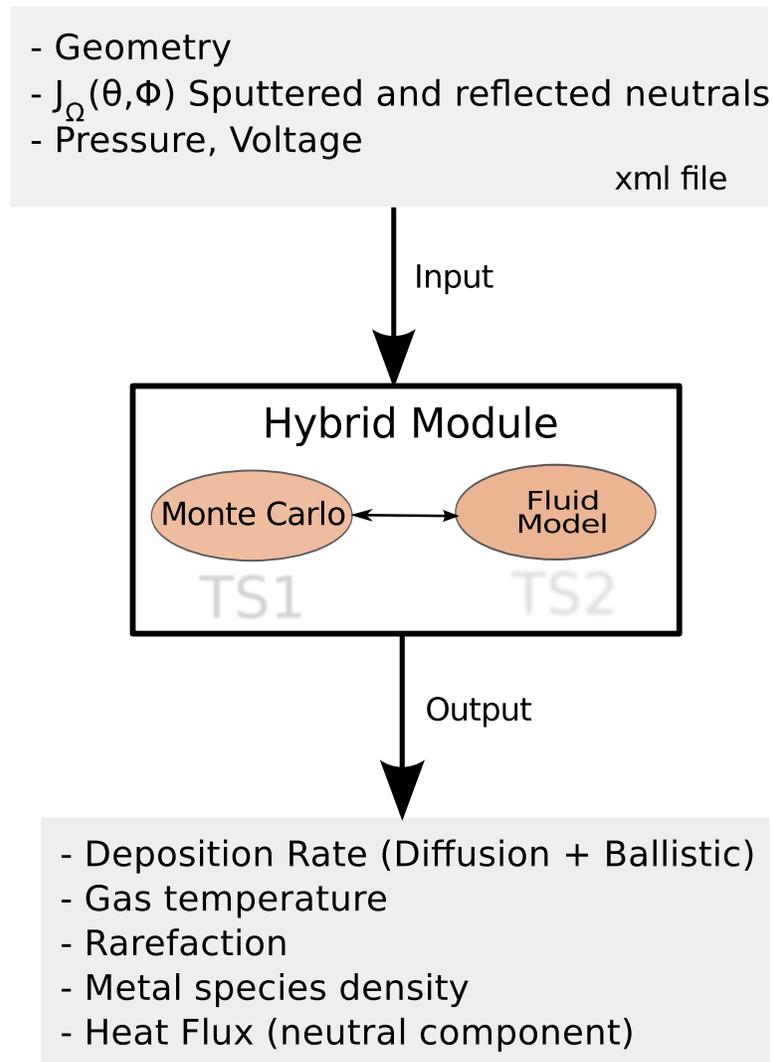
Figure 4.1: **Schematic of the ThermSpud module. The module takes several input parameters to construct the appropriate virtual environment. The module updates several variables associated with the simulation region and outputs information such as deposition rates and heat fluxes at the substrate.**
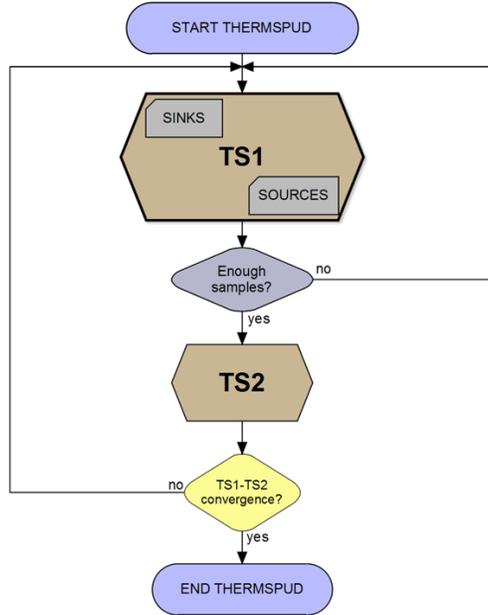
Figure 4.2: **Algorithm of the ThermSpud module. The module starts by running TS1 until sufficient test particles have been simulated. The criterion is based on the number of thermalization events. TS2 updates the density and temperature grids in the reactor using source information fed from the TS1 algorithm. The new temperature and density grid is compared against the old one and if differences are found, the algorithm repeats the cycle.**

These particles are assigned with an initial position and velocity based on sputter emission distributions. The initial position is usually a point on the target surface and the initial velocity is randomly sampled from a theoretical or experimental distribution using an inverse transform sampling technique. Once the initial position and velocity are known, particles are moved within the simulation region until they collide with a gas particle. A key step during this stage of the simulation is to determine λ. In the ptMC method, λ is assumed to be Poisson distributed according to [127]

$$f(\lambda)d\lambda = \frac{d\lambda}{\lambda_m}e^{-\lambda/\lambda_m} \tag{4.1}$$

where λ_m is the mean free path of the particle. In a mixture of gases, λ_m for a particle of species i is approximated via:

$$\lambda = \sum_j^N \frac{1}{n_j \sigma_{ij}} \tag{4.2}$$

where n_j is the number density of species i, σ_{ij} is the collision cross section between species i and j and N is the total number of neutral species in the cell.

If Equation 4.1 is the probability density function $p(\lambda)$ describing the distribution of free paths, we can randomly sample free paths using the cumulative distribution function (cdf) $P(\lambda)$ defined as

$$P(\lambda) = \int_0^\lambda p(\lambda) d\lambda \tag{4.3}$$

The inverse transform sampling technique is based on the property that if a continuous random variable, in this case λ, then $X_1 = P(\lambda)$ has a uniform distribution on $[0, 1)$ where X_1 is another random variable [173, pp 26, theorem 2.1]. All random generated numbers X for the Monte Carlo algorithms in this thesis are uniformly distributed on $[0, 1)$. Thus, to randomly sample free paths from a Poisson distribution, we generate a random number X_1 and obtain λ from

$$\lambda = P^{-1}(X_1) = -\lambda_m ln(1 - X_1) \tag{4.4}$$

After the particle has traveled a free path, a two-body collision event is simulated. During the collision, particles exchange energy and momentum with the background gas, and as a consequence they get new post-collision trajectories. The classical approach for the energy transferred (E_{trans}) in a two-body collision is obtained via (in the center of mass frame) [174]

$$E_{trans} = \frac{4E_c M_c}{M_2} sin^2 \frac{\Theta}{2} \tag{4.5}$$

where E_c is the energy in the center of mass frame, Θ is the projectile's scattering angle in the center of mass frame which is related to the lab frame scattering angle ϑ by:

$$\vartheta = tan^{-1}\left(\frac{sin\Theta}{cos\Theta + \frac{M_1}{M_2}}\right) \tag{4.6}$$

To calculate the collision cross section and the scattering angle for potentials other than coulombic, we are required to use an energy-dependent collision cross section to include short-range interactions. The first models that included this effect were developed by Robinson [175] and later modified by Somekh [148]. Both models simulate short-range interactions (repulsive potential) using an empirical cross-section parameter. Other authors have developed more realistic schemes using interatomic potentials to calculate the effective distance of closest approach, i.e. the Lennard-Jones potential [150], screened coulomb potential [74, 176], Thomas-Fermi potential [45] or Born-Mayer type potential [159, 162, 177]. In this work, we used the Ziegler-Biersack-Littmark (ZBL) interatomic potential that also includes this short-range interaction [174]. Similar schemes have been used previously in sputtering simulations [152, 178], and Kuwata et al. [101] compared the ZBL and Born-Mayer interatomic potential against a quantum chemical potential. Their results show that the ZBL potential is a better approximation to the quantum chemical potential than the Born-Mayer and in general to all those models that use the Thomas-Fermi-Dirac approximation.

With the choice of the interatomic potential, the scattering angle is obtained by solving the scattering integral [174]:

$$\Theta = \pi - 2 \int_{r_{min}}^{\infty} \frac{bdr}{r^2 \left[1 - \frac{V(r)}{E_c} - \frac{p^2}{r^2}\right]^{\frac{1}{2}}} \tag{4.7}$$

where $V(r)$ is the interaction potential, b the impact parameter and r is the distance of closest approach

The scattering cross section σ is defined as:

$$\sigma = \pi(b_{max})^2 \tag{4.8}$$

where b_{max} is the maximum impact parameter for which a significant deflection is observed. b is found by solving the denominator of Eq. 4.7:

$$b^2 = r_0^2 \left(1 - \frac{V(r_0)}{E_c}\right) \tag{4.9}$$

In practice, solving Eqs. 4.7 and 4.9 for potentials other than coulombic for every collision is computationally intensive. TS1 uses the ZBL model to approximate the potential and the Biersack-Haggmark *magic formula* [174] to simplify the solution of equation 4.7. The *magic formula* and the ZBL universal potential are the cornerstone of the scattering model used in TS1. A 1D look-up table for cross sections and a 2D look-up table for scattering angles as a function of energy are calculated at the beginning of the simulation for efficiency. The impact parameter for each collision is randomly chosen as $b = \sqrt{X_2}b_{max}$ where X_2 is another uniform random number on $[0, 1)$.

With θ already determined, the calculation of the post-collision velocities is as follows. Given $\boldsymbol{x_s}$ and $\boldsymbol{x_g}$, the pre-collision normalized trajectories of sputtered and gas atom in the center of mass frame respectively, and assuming an elastic collision, the post-collision normalized trajectories $\boldsymbol{x_s}'$ and $\boldsymbol{x_g}'$ are simply

$$\boldsymbol{x_s}' = \boldsymbol{x_s}\overline{\boldsymbol{S}}$$
$$\boldsymbol{x_g}' = \boldsymbol{x_g}\overline{\boldsymbol{S}} \tag{4.10}$$

where $\overline{\boldsymbol{S}}$ is a transformation matrix defined as:

$$\boldsymbol{S} = \overline{\boldsymbol{T}}\boldsymbol{x} \tag{4.11}$$

where $\boldsymbol{x} = (\cos(\theta)\sin(\varphi), \sin(\theta)\sin(\varphi), sin(\varphi))$. $\overline{\boldsymbol{T}}$ is a matrix defined as:

$$\overline{\boldsymbol{T}} = \begin{bmatrix} c_x & c_y & c_z \\ b_x & b_y & b_z \\ a_x & a_y & a_z \end{bmatrix} \tag{4.12}$$

with $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ forming an orthonormal basis in $\Re^3$. To calculate this orthonormal set, we first make $\boldsymbol{a} = \boldsymbol{x_s}$, then generate a random vector $\boldsymbol{v_r}$ $|\boldsymbol{a} \cdot \boldsymbol{r} \neq 0$ and make $\boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{x}$. With $\boldsymbol{a}$ and $\boldsymbol{b}$ defined, we obtain $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$. It is implicitly assumed that $\boldsymbol{x_s}$, $\boldsymbol{x_g}$, $\boldsymbol{x'_s}$, $\boldsymbol{x'_g}$, $\boldsymbol{x}$, $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are unitary vectors.

After the collision, the post-collision trajectories and velocities are converted back to the lab frame, and the test particle is tracked again according to its new velocity. The process described in the previous lines is repeated until the energetic particle is lost to the simulation. To be as considered lost, particles need to either hit a wall or lose most of their initial energy. The latter process is known as thermalization. Once the particle is thermalized, it is sent to TS2. The number of particles N_t to stop the simulation depends on several factors such as pressure, current and geometry. We usually find this value heuristically by monitoring statistical fluctuations between iterations. Finally, results are converted from simulation particles to real particles using a scale factor SF_{TS} which is obtained via $SF_{TS} = N_{target}/N_t$ where N_{target} is the particle flux at the target approximated by $N_{target} = YI/(1+\gamma_e)$ with Y being the sputtering yield, I the target current and γ_e the secondary electron emission coefficient.

TS1 handles inhomogeneities using a pseudo-inhomogeneous gas density model and an octree grid. Uniform gas density and temperature are assumed within each cell but they are allowed to vary between cells to incorporate gas rarefaction and gas heating in the simulation. Since adjacent cells may have a different density and temperature, every time a particle moves from one cell into another, a new λ_m is calculated and weighted back to determine the distance it has to travel according to the new local density. If there are more than 2 gas species, a weighted average neutral density data structure extracted

from the current cell is used to determine randomly the appropriate species to collide with [122].

TS1 only simulates elastic collisions. Charge exchange and other inelastic processes involving high energy neutrals are assumed negligible. If a background gas atom becomes energetic after a collision event, it is pushed into a high-energy particle container called ParticleBucket to be followed afterwards. The ParticleBucket is a queue on which energetic particles are temporally stored to be simulated later in the algorithm.

The boundary conditions for TS1 are yet to be extended, but for simplicity, we make $S = 1$ for all sputtered target particles hitting any obstruction on the domain where S is the sticking coefficient representing the probability of a particle to be deposited on a surface. On the other hand, all gas particles are initialized with $S = 0$ and use specular reflection. These coefficients can be configured at run time in case they need to be changed. The particle-wall collision model also includes resputtering, etching and re-emission from surfaces but it is omitted for simplicity in this description. The flow chart of the TS1 algorithm is summarized in Figure 4.3.

4.3.2 ThermSpud2(TS2): Transport of thermalized neutrals

ThermSpud2 (TS2) is a computational fluid dynamics (CFD) approach to solve for mass, heat and momentum transport of thermalized neutrals [126]. The heart of the TS2 module is the modified set of Navier-Stokes equations [167]:

$$\nabla \cdot \Gamma_n = S_t \tag{4.13a}$$

$$\nabla \cdot \Gamma_h = S_h \tag{4.13b}$$

$$\nabla \cdot (\rho v_x \boldsymbol{v}) = -\nabla_x P + \nabla \cdot (\mu \nabla v_x) + S_x + A_x \tag{4.13c}$$

$$\nabla \cdot (\rho v_y \boldsymbol{v}) = -\nabla_y P + \nabla \cdot (\mu \nabla v_y) + S_y + A_y \tag{4.13d}$$

$$\nabla \cdot (\rho v_z \boldsymbol{v}) = -\nabla_z P + \nabla \cdot (\mu \nabla v_z) + S_z + A_z \tag{4.13e}$$
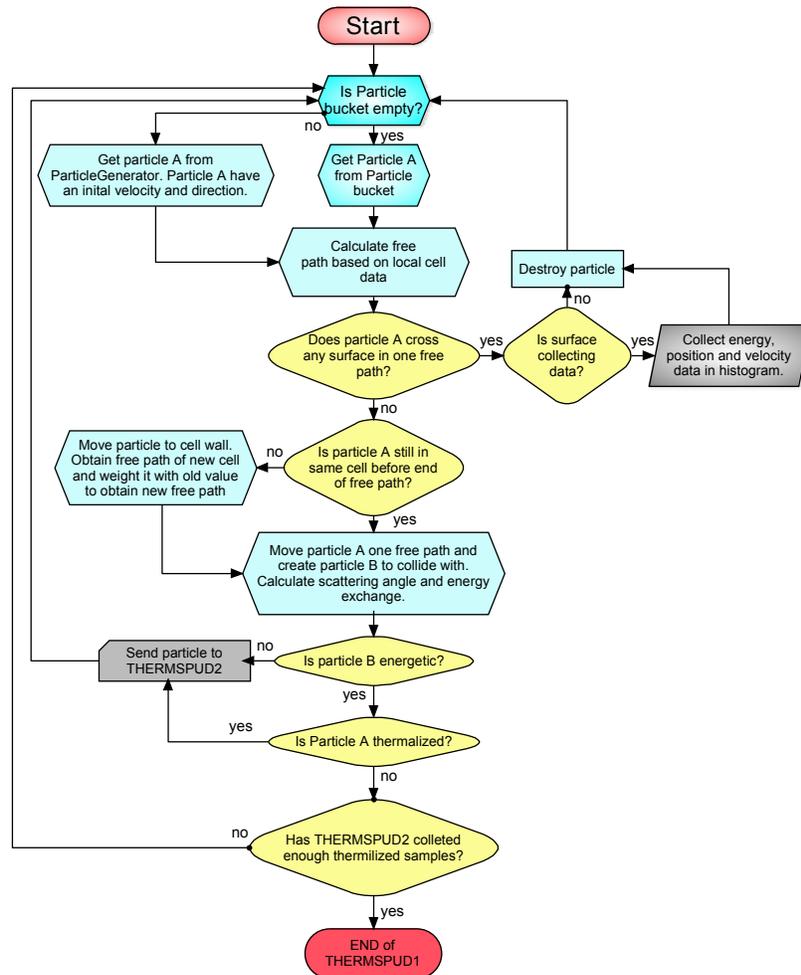
Figure 4.3: **Algorithm of the energetic particle transport algorithm used in the ThermSpud1 module.**

with

$$\Gamma_h = -\kappa \nabla T \tag{4.14a}$$

$$\Gamma_n = -D\nabla n + \boldsymbol{v}n \tag{4.14b}$$

where n is the neutral's density, Γ_n and Γ_h are the particle and energy fluxes, ρ is the density, P the pressure, $\boldsymbol{v}$ the gas velocity, μ the viscosity, $\boldsymbol{A}$ captures the extra viscous terms not included in the equation [166], S_t the source of thermalized particles, S_h the energy and $\boldsymbol{S}$ the momentum source terms coming from the MC code. Particles fluxes are assumed to be described by the diffusion-convection approximation. The R_m term also collects the remaining terms of the stress tensor in the momentum equation. The particle fluxes are approximated with a diffusion and convection model. Transport coefficients D and κ are calculated using the hard-sphere model [179]. This model has been found to be a good approximation for thermal cross-sections [180]. Equations 4.13 are discretized using the finite volume method [166] in a collocated octree grid and the obtained equations are solved using an iterative scheme.

In the original implementation of the solver [126], equations 4.13 were first discretized in an octree grid using the finite volume method and each equation solved separately. This type of solution is known as a segregated algorithm [181]. The linear system of equations formed from each discretized equation was solved using an overrelaxation algorithm while the coupling between the momentum and continuity equation was solved using underrelaxation [168]. This method, while robust, was slow and usually depended on the correct choice of the relaxation parameters. However, the need for a faster algorithm to reduce global execution times was determined as one of the key steps for the integration of the project. To speed the computation of TS2, I had to rewrite the core numerical scheme of TS2 and use a different method of solution.

Before I proceed to describe the new solution method, it is appropriate to briefly review some of the fundamental concepts behind the solution of a sys-

tem of linear equations. The discretization of equations 4.13 using the finite volume method on the kind of grid used in TS2 leads to a system of equations that can be cast into the general expression

$$\overline{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b} \tag{4.15}$$

where $\overline{\boldsymbol{A}}$ is an $N \times N$ nonsingular matrix normally known as the coefficients matrix, $\boldsymbol{x} \in \Re^N$ is an unknown vector and $\boldsymbol{b} \in \Re^N$ is a right-hand vector normally associated with boundary conditions, source terms and external forces. N is the rank of the matrix. In this particular case, the matrix $\overline{\boldsymbol{A}}$ depends on the grid structure, the discretization scheme and the transport coefficients.

There are, in general, two approaches to solve equation 4.15, namely direct and iterative methods. In direct methods, the system of equations is solved by computing $\overline{\boldsymbol{A}}^{-1}$ and then $x = \overline{\boldsymbol{A}}^{-1}\boldsymbol{b}$. However, the memory requirements to invert $\overline{\boldsymbol{A}}$ when N is large limits the application of direct methods to relatively small systems[182]. In our case, we deal with systems when N ranges from 10^5 to 10^7 for which direct methods are still impractical to implement.

The premise of iterative methods is to arrive to the solution $\boldsymbol{x}^*$ by iterating

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + F(x^{k+1}, x^k) \tag{4.16}$$

until $\boldsymbol{x}^{k+1} \approx \boldsymbol{x}^k$. In the equation above, F is a linear operator and $\boldsymbol{x}^k$ is the k iterate of $\boldsymbol{x}$. The iteration is started with an initial guess $\boldsymbol{x}_0$. A good initial guess dramatically speeds up the convergence rate of the system. If F is a function if x^k alone, the iteration is explicit. If, on the other hand, F is a function only of x^{k+1} the iteration is implicit. A measure of convergence of the iteration is the residual $\boldsymbol{r} = \overline{\boldsymbol{A}}\boldsymbol{x} - \boldsymbol{b}$. Equation 4.16 has converged when $||\hat{\boldsymbol{r}}|| < \delta$ where $\hat{\boldsymbol{r}}$ is the normalized residual, δ is an error bound and $||.||$ is a vector norm. When this occurs, it is said that the iterative scheme solved the linear system $\overline{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{b}$ in k iterations. The normalized residual is obtained via

$$r = \frac{\overline{A}x - b}{\overline{A}x} \tag{4.17}$$

The theory behind iterative techniques is extensive[183, 184]. Generally speaking, iterative methods can be divided into point stationary and projected. We briefly summarize the key characteristics of each one in the following lines.

In stationary methods, equation 4.16 is solved sequentially element by element. This approach is equivalent to *explicitly* solving the discretized equations by sweeping the grid point by point. The simplest point stationary iterative method is commonly known as the Jacobi iteration. In this scheme, the coefficient matrix elements $A_{i,j}$ and the vector elements x_i are constant throughout the k iteration. Thus, each iteration k in a point stationary iterative technique involves N operations where x_i^{k+1} is obtained using an explicit expression. To improve the rate of convergence, we can calculate x_i^{k+1} using x_{i-1}^{k+1} available from the previous step. This iteration scheme is known as Gauss-Seidel iteration. The appeal of stationary methods is that they are easy to understand, and their implementation in a computer algorithm is straightforward. Additionally, the memory requirements of this approach are low. However, the low rate of convergence severely limits their use for real applications.

On the contrary, in projection methods, the iteration is performed in such a way that the next iterate is an *implicit* function of the next iterate. In some cases, the next iterate is approximated with a linear combination of the previous iterates. Examples of popular projection methods are the conjugate gradient method (CG) [185] and generalized minimum residual (GMRES) [186]. The disadvantage of projection methods is their higher memory requirements when compared to stationary methods, a greater computational complexity and the need to perform operations involving matrices and vectors. In my experience, the benefits of projection methods outweighs their disadvantages by a large margin when compared to point iterative techniques, especially when execution times are compared.

The modified numerical method to solve equations 4.13 is the following. Indi-

vidual equations are solved using a preconditioned Choleski CG method [185] for the symmetric system (equation 4.13b) and an ILU preconditioned GM-RES [186] for the unsymmetric systems (equations 4.13a, 4.13c, 4.13d and 4.13e). The coupling between the momentum and the continuity equation is solved using underelaxation. The modification of the numerical scheme from the old stationary to a projection method reduced the execution time of a full TS1-TS2 iteration from 15-20 hours to only 4-5 hours. To perform all of our linear algebra operations, I have integrated in our framework the MV++ set of vectors and sparse matrix classes [2] and the templated library IML++ [3]. The previous libraries are seemingly integrated in the class library for numerical computing SparseLib++ [187]. The steps to integrate are then simple. The SparseLib++ library is compiled using the same c++ compiler we use to compile our code. Then, the static libraries and header files are added to our framework. The structure of the MV++ classes is well documented in the Sparselib++ user's guide available at the project's homepage [4]. We have used the SparseLib++ v 1.7 in all of our calculations.

Another significant change to the algorithm was to modify the data structure of the TS2 grid to allow the solver to handle more than 2 neutral species. Originally, the algorithm was hard-coded to only allow the solution of 2 neutral particles, the background gas and the metal species. However, the integration with the plasma solver required the solution of the metastable population. Additionally, the computational structure was modified to allow a better integration with the plasma module which required special handling of creation rates. Details of these changes are described in chapter 6.

A key assumption of the TS2 module is to treat thermalized neutrals with a continuum approach. The Knudsen number, Kn, which the ratio of the mean free path λ to the characteristic dimension L of the system [127] is a good metric for the validity of the continuum approach. For it to be valid, Kn should be lower than 1; a higher value breaks the model, and the Navier-

---

[2]math.nist.gov/mv++
[3]math.nist.gov/iml++
[4]math.nist.gov/sparselib++

Stokes equations are no longer valid or do a poor job at describing the transport properties of the fluid. Kn can be approximated via [179]:

$$Kn \approx \frac{k_B T}{\sqrt{2}\sigma p} \tag{4.18}$$

where k_B is the Boltzmann constant, T the temperature, σ is the thermal collision cross-section and p the pressure. For Ar, $\sigma \approx 40\text{Å}^2$ and for a pressure and temperature of 10 mTorr and $300K$, $\lambda \approx 5.3$ mm (below 10 mTorr there is no significant heating and the flux is mainly ballistic). All simulations in this section are for $L = 10$ cm which gives $Kn = 0.0563$ for 10 mTorr. This suggests that even though we are in the transition regime (according to the practical division proposed by Dushman and cited in Ref. 127), we are still in the region where a continuum approach is a valid approximation for mass, heat and momentum transport.

The boundary conditions are of Dirichlet type for heat transfer ($T_{wall} = T_w$) and metallic vapor transport ($n_{wall} = 0$), non-slip boundary conditions for the momentum equation ($v_\perp=0$) and Neumann boundary conditions for background gas atoms ($\nabla n=0$).

A key component of the model is the concept of thermalization [126, 143, 148, 149] which is defined as the process in which an energetic particle loses most of its initial energy through collisions with background thermal atoms. After thermalization, particle transport is assumed to be diffusive [129]. Following particles that are almost thermalized in the MC algorithm is very inefficient since they do not deposit much momentum and energy into the gas, slowing down the simulation. We have found that the best agreement between simulation and experiments is attained for $E_{thr} = 5k_b T$ where k_b is the Boltzmann constant and T is the local temperature.

4.4 Results

4.4.1 Simulations

Our standard simulation region consists of a 20 cm length × 15 cm radius cylindrical chamber divided into hexahedral cells with dimensions d_x, d_y and d_z where $d_i = L_i/N_i$. For the data shown in this section we have used $N_i = 85$, $N_y = 85$ and $N_z = 120$. Pressures representative of ballistic, transition and purely diffusive transport were chosen (5 to 40 mTorr). We used an experimental erosion profile to randomly launch sputtered and reflected neutrals from the target using the model described in Ref. 48 as described in section 4.4.3. The model of Ref. 48 predicts sputtering yields somehow higher than the values obtained from other models (SRIM[5], Yamamura et al. [45], Ekstein et al. [188]). However, it is well understood that the computation of these yields from first principles is highly sensitive to the choice of the surface binding energy and factors such as surface erosion and crystal structure of the target which makes this process rather uncertain. In an attempt to be consistent, we have modified the value of E_b in the Stepanova and Dew equation to obtain sputtering yields in accordance to the equation of Ref. 188 which has been extensively studied and compared against experiments. Figure 4.5 shows the sputtering yields as functions of energy that were used in this work. The energy distribution predicted by the Stepanova and Dew formula was found to appropriately describe the measured distribution when sputtering aluminum [189].

All simulations were run on a linux cluster consisting of several dual core AMD 280 Opteron processors. We have compiled our code using gcc version 4.1.2. Typical simulation times range from 48 to 60 hours per run. length of simulations

A key assumption of the simulation is that of a monoenergetic flux of ions at the target [150]. Experimental and simulation efforts have been devoted to this

---

[5]www.srim.org
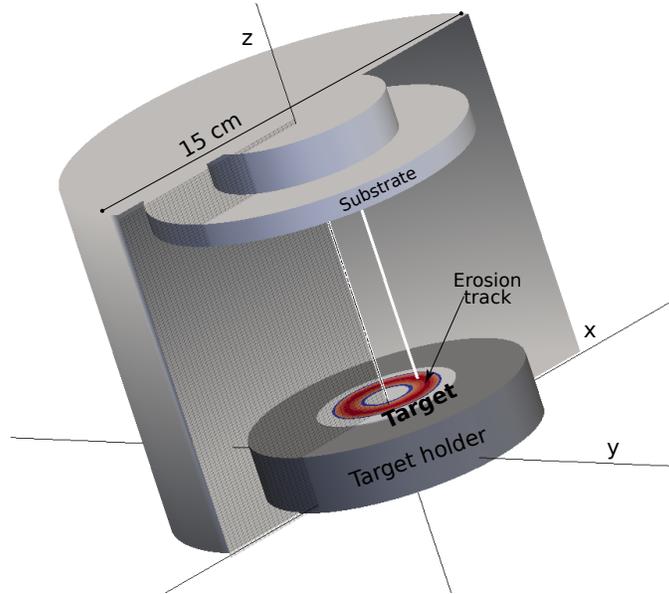
Figure 4.4: **Computational domain for the simulation region used to obtain the results reported in this report. The shaded region on the target is a gray scale representation of the measured erosion profile we input to the TargSpud module.**

problem [190–194] suggesting that the flux is far from being monoenergetic, although it is difficult to extrapolate these results to magnetron discharges since the magnetic field alters the dynamics of the plasma. Although we can safely assume that ions are not magnetized, the electron confinement pushes the sheath edge towards the cathode reducing the dark space. The sheath is shrunk and ions are far less likely to collide with gas atoms which is one of the mechanisms attributed to observed ion energy distributions (IED) at the target. Additionally, the smaller sheath reduces the amount of ionization at the sheath which lowers the population of low energy ions at the target surface. For glow discharges, results have shown [190, 192, 193] that the IED at the cathode strongly depends on the ratio δ_s/λ_n where δ_s is the sheath thickness and λ_n is the ion mean free path. Shon et al. [195] measured the average ion energy at the target of a magnetron discharge and found that the average energy is lower than the cathode potential. Goeckner et al. [194]
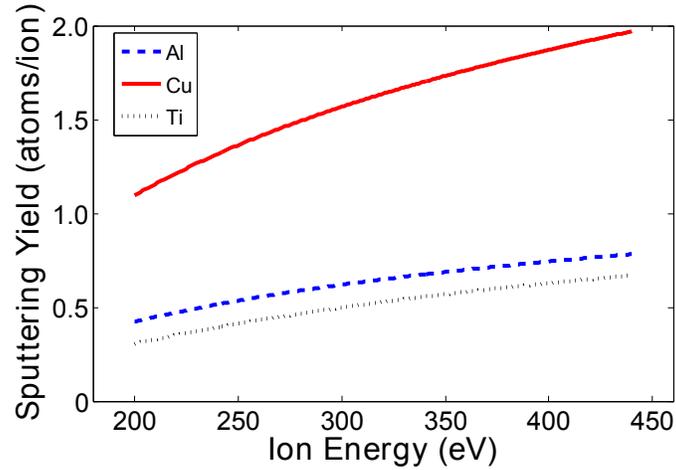
Figure 4.5: **Sputtering yields as function of ion energy for three different materials. Values were taken from the the generalized sputtering yield formula presented in Ref. 188.**

estimated that the average energy is about 73% of the cathode potential even considering the collapse of the dark space region. However, Turner [149] argues that the reduced thickness of the sheath significantly reduces the probability of collisions in the sheath and he suggests that ions entering the sheath hardly collide. In his simulations, he assumes that the ions striking the cathode have energies equivalent to the cathode potential which means that charge exchange, ionization and elastic collisions in the sheath are negligible. The simulations presented in this chapter have used a value that lies between these two models. We therefore have assumed that ions strike the target with an energy of 0.85 times the target voltage in our simulations. With respect to the total current at the target, we have assumed that the total current at the cathode results only from incident ions and emitted secondary electrons. Thus, the total current at the target has been calculated as $J_t = J_i + \gamma J_i$ where γ is the secondary electron coefficient adapted from the simple model of Ref. 196. Finally, the input values for the simulator are obtained by multiplying J_t by the sputtering yield or the neutral reflection coefficient. The calculated numbers are then entered in the configuration file where they are automatically scaled

according to the input erosion profile and the target area. Simulations results are shown for three materials: Al, Ti and Cu.

The number of particles launched at the target has an impact on the accuracy of the simulation given the statistical nature of the Monte Carlo method. The number of thermalized particles per cell is a key parameter since it is this value that governs the accuracy of the fluid model. We have found empirically that in the computational domain of Figure 4.4, a value of 1×10^8 created particles leads to enough thermalization events to minimize statistical fluctuations between iterations.

4.4.2   Experiments

We deposited aluminum films in a DC magnetron sputter system using a 75 mm cylindrical planar target (99.99%) in a pure argon environment. The sputtering system used was a Kurt J. Lesker magnetron system. The system was pumped down using a combination of a Leybold Trivac D V rotary pump and an APD Marathon 8 cryopump to a base pressure below 3×10^{-7} Torr. The process gas pressure was kept in the range of 5-40 mTorr using a MKS baratron and MKS 250 flow controller and by throttling the high vacuum valve. Films were deposited onto glass slides located at 10 cm in front of the target without external heating; the magnetron was operated in the constant power mode. Deposition rates were calculated using two methods: using an Alphastep 250 Profilometer and measured in-situ using a TM-400 Inficon thickness measurement system.

The IV characteristic of our experimental setup is shown in Figure 4.6. The IV curve, which is determined by the magnetic field shape, the target material and the design of the target shield, is only relevant when we compare different experiments using the same sputter gun. In our experiments, we used the same gun for all three targets; however, the Ti target was half the thickness of the other two. This, and the higher value of the secondary electron coefficient

of Ti [197], explain the higher currents for Ti sputtering observed in Figure
4.6.
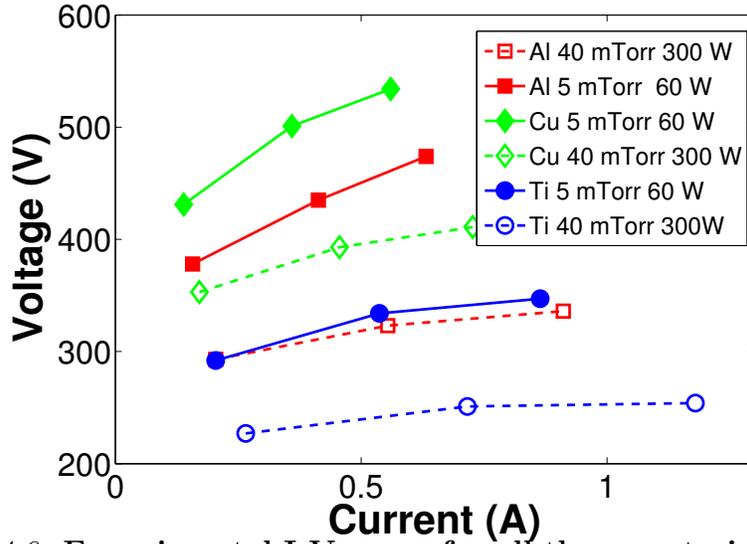
Figure 4.6: **Experimental I-V curve for all three materials used
in this chapter. Only data for pressures of 40 and 5 mTorr
and powers of 60 and 300 W is plotted. Lines are to guide the
eyes.**

4.4.3   TargSpud

TargSpud is the module responsible for the creation of sputtered and reflected
neutral particles at the target. This module takes as input arguments an
erosion profile, an angular and energy distribution of sputtered particles and
a IV curve for the process to simulate. The process to pick particles from the
target surface follows a simple approach. We first specify the total particle
current N_t calculated as described before and a file representing the erosion
profile at the target surface. This file contains a list of values in the format
$x, y, p(x, y)$ where (x, y) are the coordinates of a point in the surface of a plane
and $p(x, y)$ is the probability of a particle of being ejected at the point (x, y).
If we take $p(x, y)$ as a pdf, we can construct a cdf using the same technique

described before. Finally, using the inverse random sampling technique, we can pick points (x,y) at random from the cdf by generating a random number X_3.

4.4.4 Angular and energy distributions

To simulate the transport of sputtered particles, the ejection velocity is of critical importance at low pressures where particles undergo few collisions and their motion is fully determined by their initial trajectory. However, it is not always clear what distribution one needs to use in simulating the process using particle based methods. For instance, there are simulation studies showing that results obtained using a simple $cos(\theta)$ [153, 160, 162, 172, 178, 198–200] or $cot(\theta)$ [161] initial angular distribution, where θ is the angle between the initial particle trajectory and the surface normal(see Figure 2.9), agrees reasonable well with experiments. Similarly, some authors have fitted a $cos^n(\theta)$ [177, 201, 202] or an empirical relation [203] to obtain an initial distribution to match experimental results. For instance, Eisenmenger-Sittner et al. [201] used a cos^n distribution with n as a fitting parameter. Interestingly, they needed to modify the value of n as function of pressure but found no theoretical justification for this modification. The values of n ranged from 1-2 in a pressure range of 0.1-1 Pa. In contrast to purely analytical models, some authors used the popular code SRIM[6] to obtain the angular distribution [87, 153] even though there is evidence that SRIM does not properly capture the events at the target in the energy range of typical of magnetron discharges [58, 65]. Good agreement between simulations and experiments was also found using this method. Lastly, for distributions that can not be approximated by a cos^n relation, some authors have taken a different approach. They used their simulation results and their experimental results to calibrate their initial distributions [73, 74] in an iterative manner. For instance, van Aeken et al. [74] used a fitting function in the form of $\sum a_n cos_n(\theta)$ to generate a distribution and proceeded to simulate

---

[6]The Stopping and Range of Ions in Matter www.srim.org

their process. Afterwards, their simulation results were compared against their pin-hole experiment adjusting the fitting parameters a_n until the difference between simulation and experiments was small. The resulting distribution was significantly different from the one obtained using SRIM (see fig. 3 in [74]) increasing the uncertainty about what is the proper distribution to use for realistic simulations. The model to obtain the angle and energy distribution in TS is taken from a formula proposed by Stepanova and Dew [48] as an empirical modification to the Thompson distribution commonly used in sputtering simulations[59].

4.4.5 Gas heating

In this section, gas heating is only considered from sputtered particles and reflected neutrals; other heat sources are not included in the simulation results. The temperature of the background gas for 2 powers and all three materials investigated in this section is shown in Figure 4.7. I only show data for 40 mTorr since it is at this pressure that we observed the maximum temperature variation between low and high power.

In Figure 4.8, I show the maximum temperature along the axis of the chamber as a function of pressure and power for three elements. The maximum temperature is lower for Al than for Ti and Cu. A couple of effects explain this results. First, Al is less efficient at depositing energy to the background gas given the mass mismatch between Al and Ar. Second, Al has the lowest neutral reflection coefficient of the three elements investigated [65] for similar reasons. High energy reflected neutrals, mostly argon atoms in our case, are very efficient in depositing energy into the background gas.

The location of the maximum temperature shifts towards the target as pressure is increased since particles are thermalized faster and deposit most of their energy closer to the target. The implication is that at high pressures the target acts as an energy sink needing better cooling. Depending on the cooling efficiency and the target material, we may observe that the rate at which
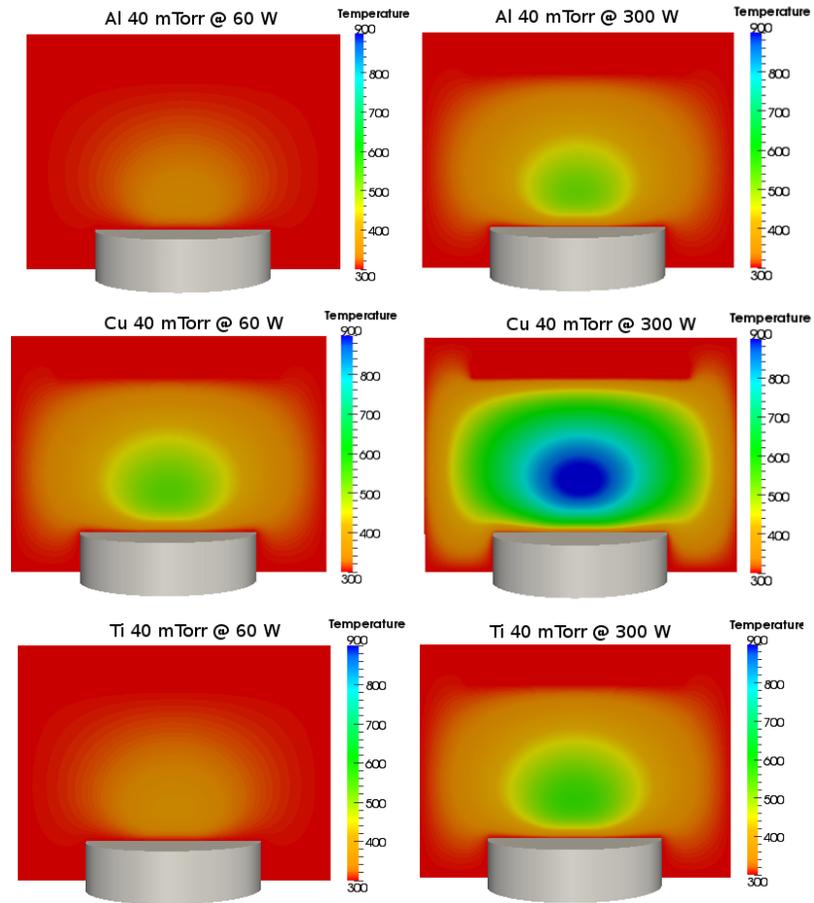
Figure 4.7: **Simulated temperatures of the background gas for 2 powers and three elements. Data is for 40 mTorr.**
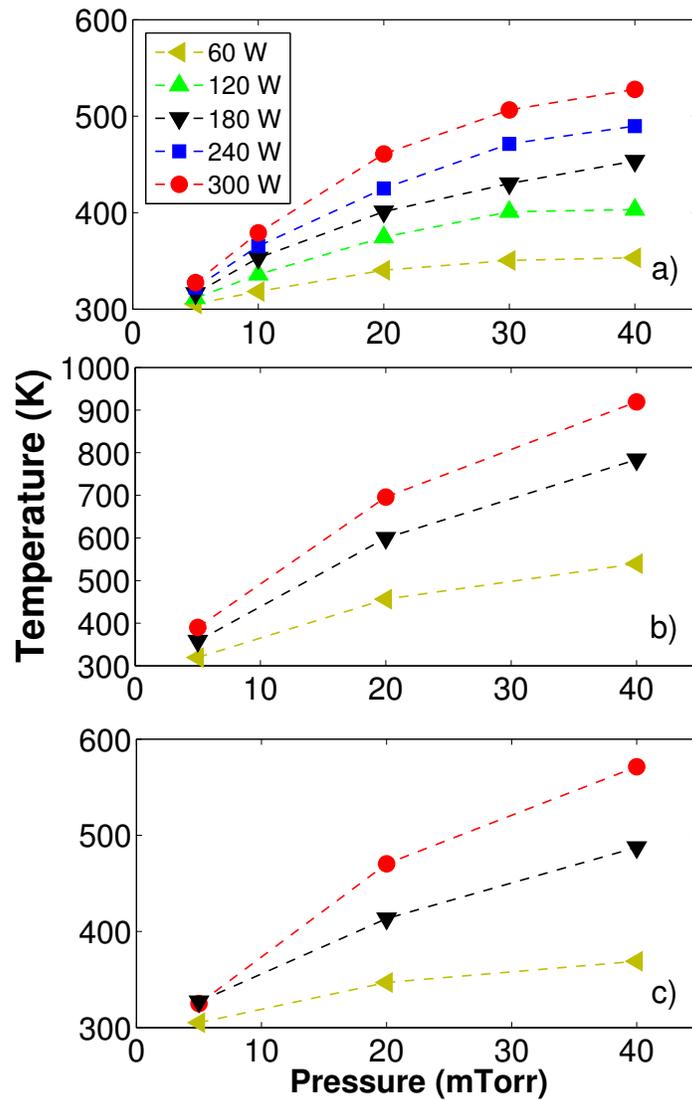
Figure 4.8: **Maximum temperature along the center of the chamber for 3 different materials as function of power and pressure: a) Al, b) Cu, and c) Ti.**

the gas temperature increases may vary. This suggests that we may need to consider the temperature of the target as an important parameter in the overall calculation of gas heating and plasma impedance as previous studies have suggested [154, 204]. Figure 4.8 also implies that further increases in pressure do not heat the gas at the same rate: a significant part of the generated heat is absorbed by the target, and geometric factors may be more important. We should note that the simulations were performed under the assumption of a constant thermal accommodation coefficient [150] and perfect cooling of the target (the temperature of the target is assumed constant at 300 K).

Of the three elements investigated, Cu shows the greatest temperature increase at comparable power levels. This is due to the higher sputtering yield of Cu when compared to that of Al and Ti and the greater neutral reflection coefficient of Cu. Although Cu is less efficient in transferring energy to Ar by collisions than Ti, the great difference in sputtering yields plays a decisive role in determining the steady state temperature profile. Recall that the energy transfer coefficient in a binary collision is expressed by Equation 4.5, then for all elements investigated in this report, Al, Cu and Ti, the corresponding transfer energy coefficients are (assuming a stationary argon buffer gas and a head-on collision) 0.04, 0.01 and 0.05 respectively. From these values, it is evident that Ti would transfer more energy to the background gas per collision than Al and Cu, but, as shown in Figure 4.5, the sputtering yield of Cu is at least 3 times larger than that of Ti. Thus, at comparative currents, Cu will deposit more energy since there are more Cu particles colliding per unit time than in the case of Ti and Al sputtering. On the other hand, Ti sputtering shows a higher maximum than that of Al even though the sputtering yield of Ti is slightly lower than that of Al. A couple of factors are in play to explain this: the higher energy transfer coefficient of Ti when compared to that of Al and there the are more reflected neutrals in Ti sputtering. The neutral reflection coefficient of titanium is significantly higher than that of aluminum according to our own calculation (by an order of magnitude). According to the model of Ref. 48, the neutral reflection coefficient of Al at 300 eV is 0.01 while that of Ti at the same energy is about 0.2. A greater coefficient means

that more ions are reflected as high energy neutrals after colliding with the target, implying a higher flux of these high energy neutrals emanating from the target. These reflected neutrals are very efficient in transferring energy to the background gas since they have essentially the same mass as the background gas atoms and the energy transfer coefficient is 1.

Direct comparison against experimental data may be misleading in some cases given the great number of variables that we need to consider. Factors such as chamber geometry, magnetic field configuration and target size only allow us to use empirical models for qualitative comparisons of similar configurations. Furthermore, the intricacies of the process makes the measurement of temperature and density in the conditions typically found in sputtering systems a complex task. There are several experimental techniques to characterize the sputtering discharge, most of them based on optical techniques [205–210]. From the list, the work of Wolter et al. [207] better resembles our setup, and it is therefore used to benchmark our simulations. In their setup, Wolter et al. used optical absorption spectroscopy (LAS) to measure both the gas temperature and the aluminum density in a sputtering reactor by measuring the signal losses due the absorption of a blue laser. The frequency of the laser was chosen to excite a ground state population to a known metastable state. The gas temperature is obtained from the effective full width at half maximum (FWHM) of the obtained signal, while the atomic density was proportional to the integrated absorption profile. The accuracy of the experimental results is about $\pm 30\%$ [207], and this is represented with error bars in the plots when applicable. The simulation data to compare against the experimental values was collected considering the geometry of the experimental setup [207]. To collect the effective density of the experimental measurements from our simulations, results were post-processed using vtk and python scripts. First, the full 3D grid was loaded, and then only the region where the laser passed through the discharge (a line 5.6 cm away and parallel to the target surface) was extracted from the 3D file to simulate the experimental absorption length. Then, the values thus obtained were integrated and averaged to obtained a representative density from the region swept by the laser. See Figure 4.9.
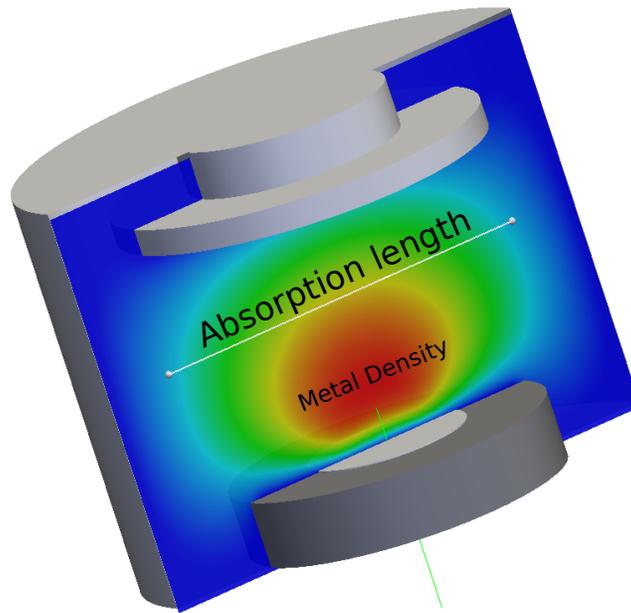
Figure 4.9: **Absorption effective working length of the optical absorption spectroscopy experiment. The line in the simulation region is where data was integrated in our simulation results.**

The measured and simulated temperature profiles are shown in Figure 4.10. In the figure, I compare simulated gas temperatures with the experimental values of temperature as measured by Wolter et al. In the plot, I have included both the integrated and maximum temperature along the optical length of the experiment. Overall, we observed good qualitative and quantitative agreement between the simulated and the maximum values. The maximum values should be the ones associated with the measured temperature since the method to measure it (Doppler broadening) would be only sensitive to the maximum temperature along the optical path of the laser.

From Figure 4.10, we can observe that the difference is larger at low pressures but this is indeed an expected result since in this pressure regime aluminum atoms are not fully thermalized. The TS2 model does not include the energetic population which is expected to be a significant proportion of the thermalized
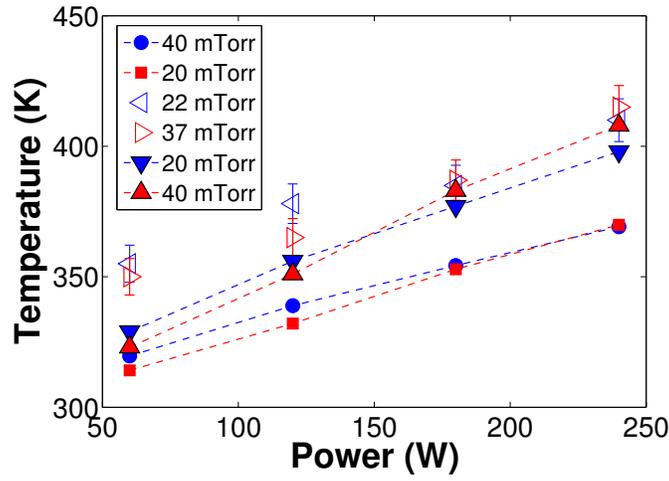
Figure 4.10: **Measured and simulated gas temperature during an aluminum deposition taken at 5.6 cm from the target.**

density in this low pressure regime. By including this population, we may increase the effective temperature of the region since the effective temperature of the energetic population is greater than that of the thermalized particles. By effective temperature, I of course mean the average energy of the energetic population since the energy distribution is evidently not maxwellian. We can also observe a systematic deviation of about 5 to 10 degrees between our results and the experiments. In an effort to find an explanation for this discrepancy, we repeated our simulations with a cathode temperature of 350 K since we originally attributed the difference to our boundary conditions at the target ($T_{target} = 300K$). Results from this change are shown in Figure 4.11 for 300 W. Contrary to what we would have expected, the maximum temperature across the optical length is lower for the 350 K simulations when compared to the one at 300 K and, as it is evident in the plot, the difference is greater at 40 mTorr. This behavior may be explained by plotting the temperature along the z axis at the center of the cylindrical chamber as shown in Figure 4.12. In the plot, we observe that the effect of using a hotter boundary condition is to move the location of the maximum temperature closer to the target, meaning that heat is dissipated in the region closer to the cathode. As we

have increased the target temperature without changing the energy input to the gas from the sputtered particles (all other parameters are kept the same), energy is dissipated in a volume closer to the target, thus attributing the difference to this volumetric effect. These type of information is not captured by models that do not incorporate volume effects into their calculations. No significant difference between both curves at 20 mTorr are observed. This may lead us to conclude that when the energy input to the gas is relatively small, the sensitivity of the simulation results with respect to the target temperature is small. However, when the energy input increases, equivalent to increasing the pressure, we expect to have a greater effect on the simulated temperature profiles with respect to the target temperature.
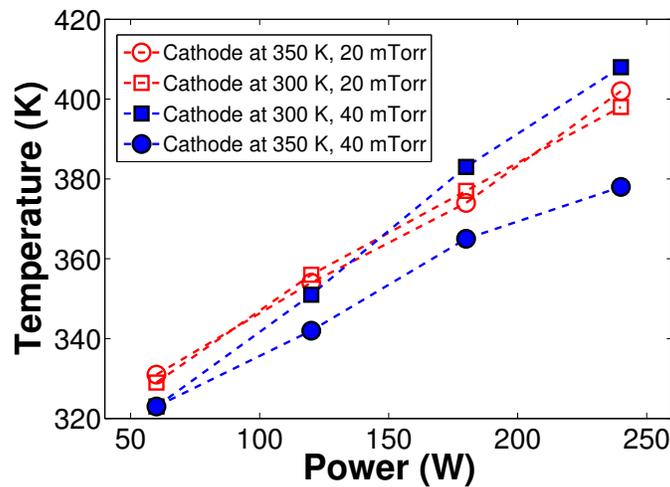
Figure 4.11: **Comparison between two temperature boundary conditions and the effect on the simulated temperature.**

4.4.6 Sputtering wind

The driving force for the observed gas reduction in sputtering systems is due to gas heating induced rarefaction and momentum transfer from sputtered particles. The latter was first described by Hofmann as the sputtering wind [211]. To find the driving force, we performed a simple computer simulation. We
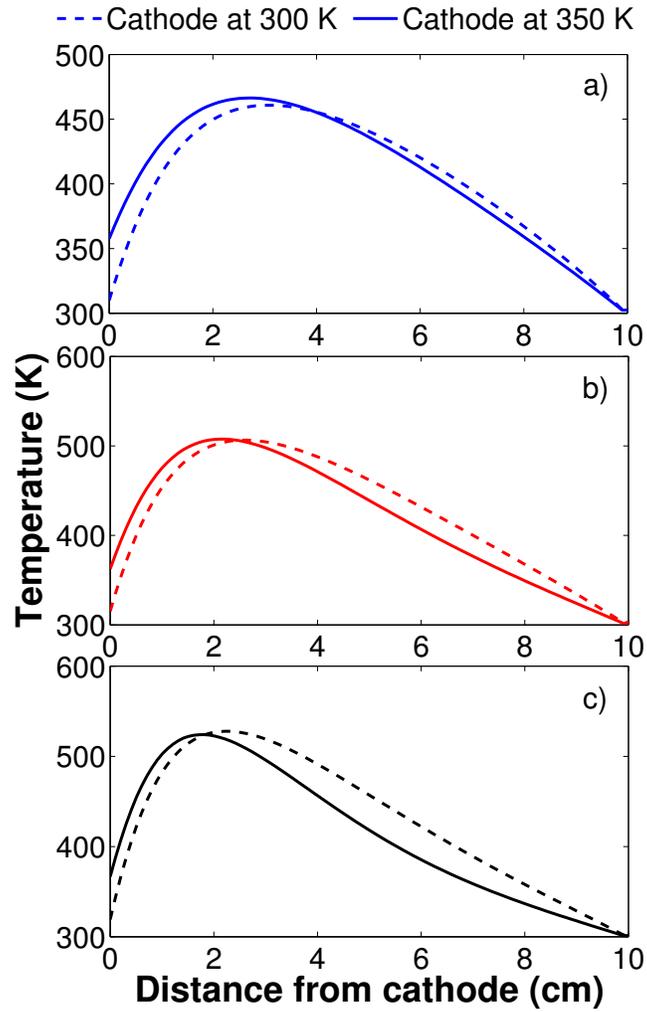
Figure 4.12: **Effect of target temperature on the simulated temperature profile along the z axis for three simulated pressures: a) Al, b) Cu, and c) Ti. Data is for 300 W.**

used the computational domain described above and simulate the transport of 3 different materials on 3 different processing gases. We chose the materials as representative of light, medium and heavy atoms for a comparative analysis. Likewise, the process gases were chosen as representative of a light, medium and heavy atoms. The results of the simulations are summarized in Figure 4.13.
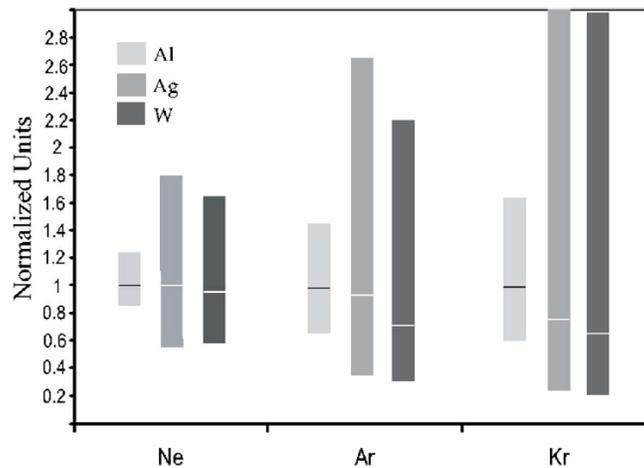
Figure 4.13: **Contribution of gas heating and sputtering wind to the rarefaction of the process gas. The top of the bars represents the normalized temperature of the gas while the bottom refers to the normalized gas concentration. Data were obtained at regions of maximum temperature along the central axis.**

Figure 4.13 could better explain the effects of the sputtering wind and gas heating on rarefaction levels. There is plentiful information encoded in this simple figure. First, each bar represents a pair target-gas combination. The position of the bar on the plot connotes the process gas and the color defines the target species. The length of the bars encodes the extent of gas heating and rarefaction with respect to normalized values along the z axis of the chamber. The end values of the bars refer to normalized values of temperature and density. The top ends show the simulated temperature normalized with respect

to the ambient temperature $\bar{T} = T_z/T_0$ for that specific target-gas combination where T_z and T_0 are the maximum and room temperature respectively. The bottom end depicts the simulated normalized density $\bar{n} = n_z/n_0$ where n_z and n_0 is the minimum and nominal gas density at the same point where the maximum temperature was obtained. The line within the bars is the product $\bar{T}\bar{n}$. For instance, the first bar on the left tells us that the maximum temperature observed for the simulation of Al on Ne is about 1.2 times the room temperatures and at the same time that the minimum density was about 0.85 times the density associated with the process pressure.

According to the figure, if gas heating were the principal mechanism for causing rarefaction, assuming an ideal gas, we would expect that the product of the normalized values of density and temperature would be close to 1. When this product is less than 1, there is another effect responsible for the observed lower rarefaction. A possible explanation is a non-constant pressure. This effect is the sputtering wind where the flux of particles coming off the target deposit enough momentum to produce a pumping effect. The external gas flow has been neglected in our calculations. Thus, in those cases where the line lies below 1, the sputtering wind effects kicks in, effectively 'pumping' the gas away from the target. This pumping has been incorporated in the model solving a momentum equation for the gas species.

For Al with all gases considered, the horizontal bar is located about 1, suggesting that gas rarefaction is caused mainly by gas heating. With Ag and W, the sputtering wind effect becomes significant with higher gas masses, due to a more efficient momentum transfer between sputtered and reflected neutrals and background gas atoms. The difference in relative rarefaction between W and Ag in Ar and Kr may be attributed, based on the numerical results, to the difference in their respective sputtering yields. The results generally show that rarefaction might be caused by both gas heating and pumping effects and the extent of the contribution of each individual factor depends on the combination of gas-target species and on the process conditions.

4.4.7 Gas rarefaction

There is experimental evidence that the pressure in the vessel is not constant [212], especially in the region next to the target. The competing mechanisms that drive the reduction of the background gas density are dependent on factors such as pressure, buffer gas and target material. For instance, we saw in the previous section that for light atoms in a heavier buffer gas, the dominating mechanism is gas heating [213]. As the mass of the sputter atoms increases, the sputtering wind (deposition of momentum into the buffer gas) starts contributing to the rarefaction effect. This indeed may be an important factor in pulsed and High Power Impulse Magnetron Sputtering techniques where a significant amount of momentum is deposited into the gas due to the high flux of sputtered particles and reflected neutrals emanating from the target [214, 215]. As the mass mismatch increases, momentum deposition dominates since the energy deposition becomes less efficient [213].

The relative background gas density for 2 powers is shown in Figure 4.14 for 40 mTorr. The figure shows how the rarefaction of the background gas is stronger for Cu (30% at 300 W) presumably due to gas heating. The sputtering wind would only kick-in for this combination of target-buffer species when the discharge reaches the saturation temperature for this geometry (see section 4.4.5).

Figure 4.15 shows the rarefaction of the buffer gas as a function of power and pressure. Similar to the observed trends in the previous section, Cu sputtering shows a higher rarefaction when compared to Al and Ti at the same power level. The observed rarefaction for copper is not only due to gas heating but also to a more efficient momentum transfer from heavier Cu to Ar. Both mechanisms contribute to the rarefaction levels depicted in Figure 4.15.

Another way in which rarefaction affects the overall deposition process is by changing the dynamics of the plasma. This is an important parameter in the plasma module since the plasma transport coefficients depend on the local density of the background gas [216]. Furthermore, the probability of collision
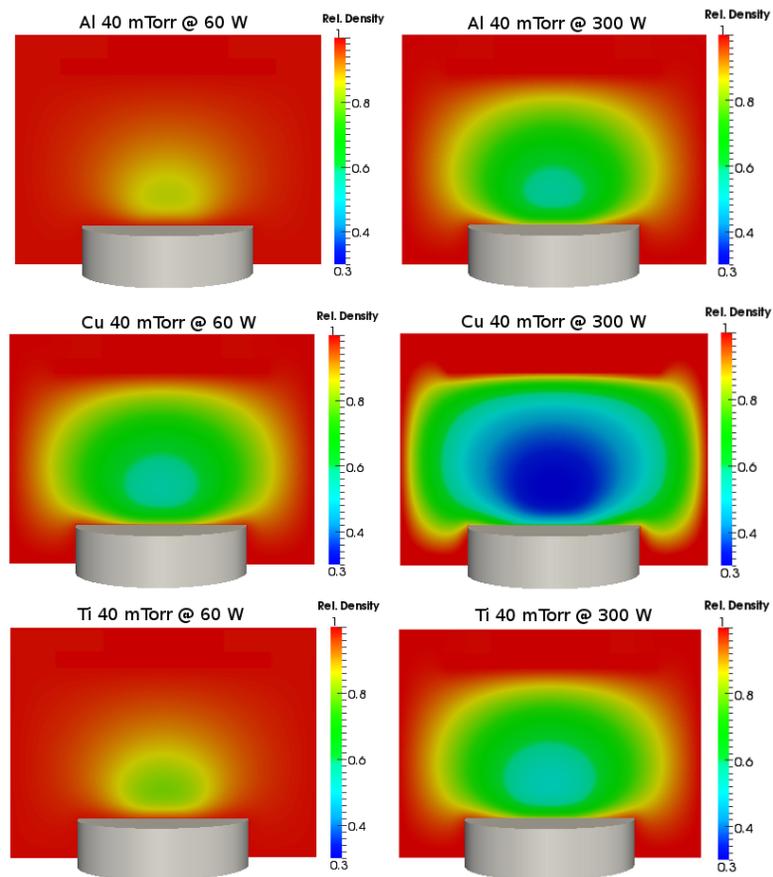
Figure 4.14: **Background gas density reduction for 2 powers and three elements. Data is for 40 mTorr.**
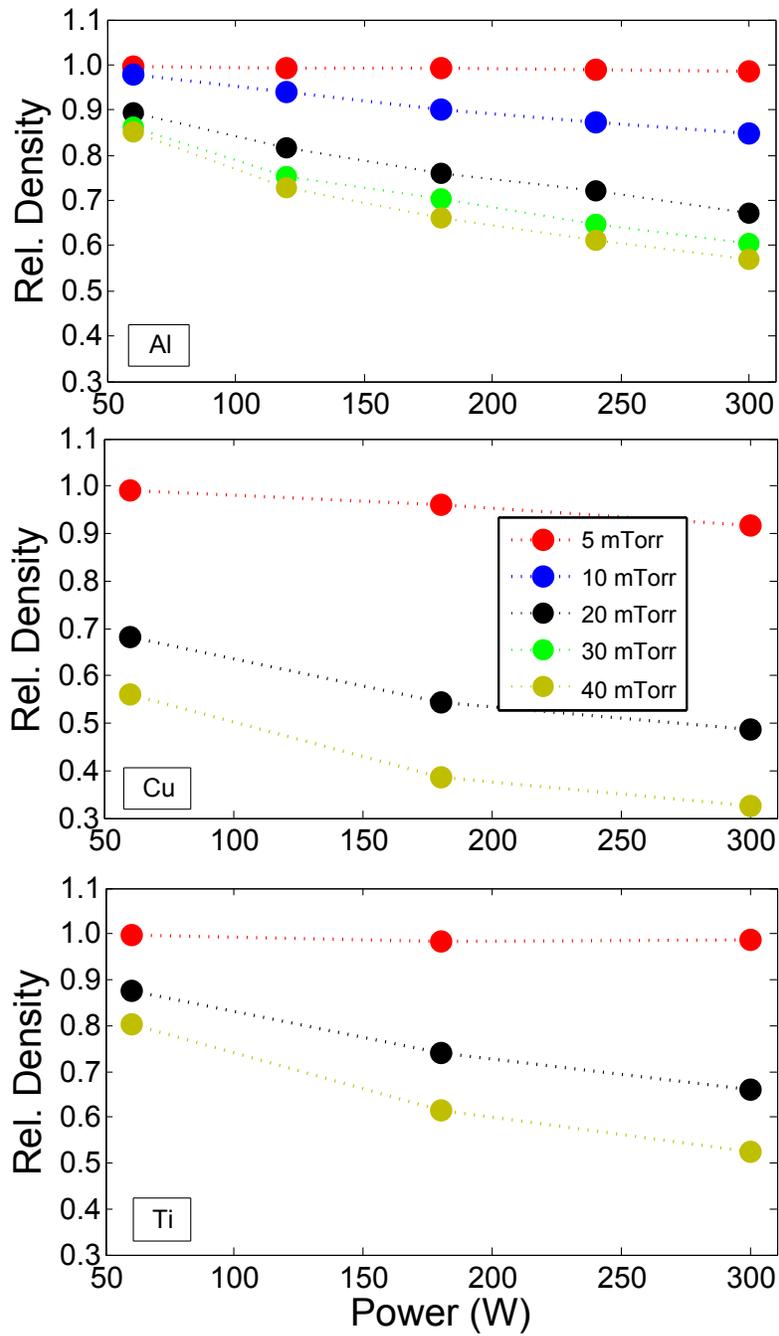
Figure 4.15: **Gas rarefaction as a function of power and pressure for three different materials.**

of energetic charged particles (secondary electrons ejected from the target and accelerated to high energies, ions in the sheath) also depends on the density of the background gas.

4.4.8 Deposition Rate

The deposition rate as function of pressure for aluminum is showed in Figure 4.16. The deposition rate was measured at the center of the substrate, and the reported values include the contribution of ballistic and diffusive fluxes. The agreement is good considering the range of pressure and power covered in the plot. In Figure 4.16, experimental results were changed from nms^{-1} to $m^{-2}s^{-1}$ using the bulk density of the material. In the case of Al, this is not a bad approximation since the expected density of these films under the conditions studied in this report should be close to the bulk density [83].
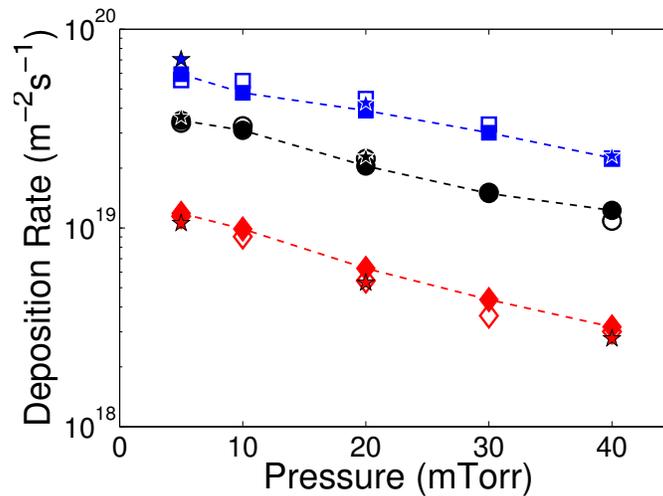
Figure 4.16: **Aluminum deposition rate as function of pressure for 3 different powers. The deposition rate was measured at the center of the substrate. Simulations include both diffusive and ballistic contributions. Experiments: $\diamond$ Depth profiles measurements, ★ Microbalance measurements. Simulations: ■ 300 W, ● 120 W and ◆ 60 W. Lines are just to guide eyes.**
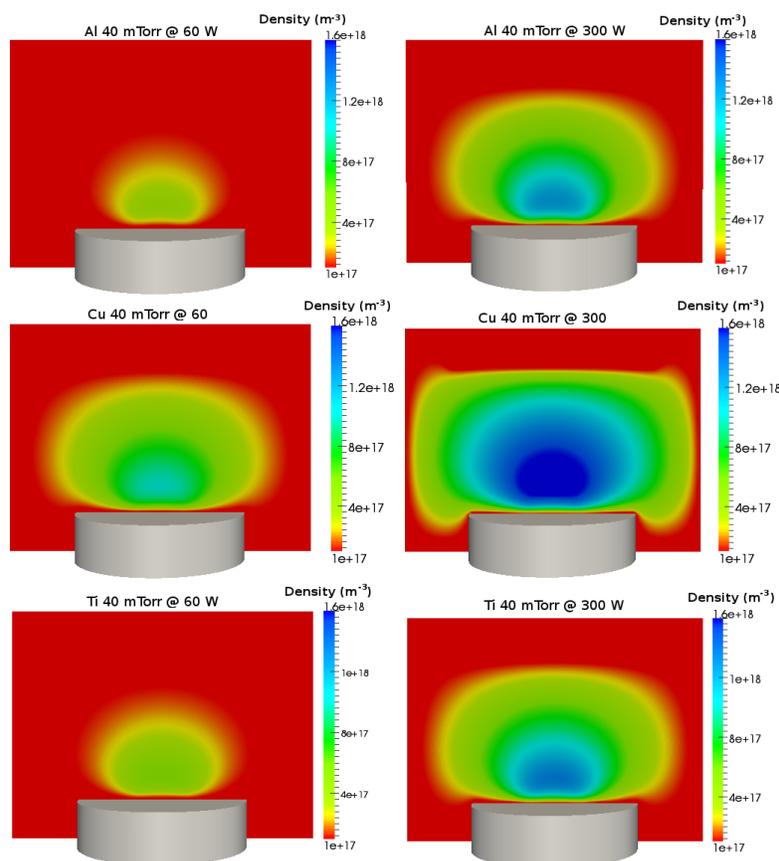
Figure 4.17: **Simulated metal vapor density for 2 powers and three elements. Data is for 40 mTorr.**

4.4.9   Metal density

The simulated metal density in the reactor is shown in Figure 4.17 for 3 materials and 2 powers. As expected, Cu density is higher at comparable power levels even when the current associated with this power is lower than that of Al and Ti (see Figure 4.6). The higher Cu density is due to the large value of the sputtering yield of Cu (almost 4 times as higher than Al and Ti).

I calculated the metal density in the chamber and benchmarked the results to those of Ref. 207 in Figures 4.18 and 4.19. In Figures 4.18 and 4.19, we observe a difference from the experimental data of about a factor of 2. However, the

overall trend is captured by the simulation. This difference may be attributed to several factors such the ones exposed in section 4.4.5. We modified our computational model, Figure 4.4, to better match the experimental chamber of Ref. 207 (we increased the volume of the chamber) in an attempt to find a geometrical factor that may contribute to this difference. Using this new set-up, we repeated our simulations but found no significant difference with respect to results of Figure 4.4. This indicated that the simulation volume did not affect the results and the difference may lie in another factor.
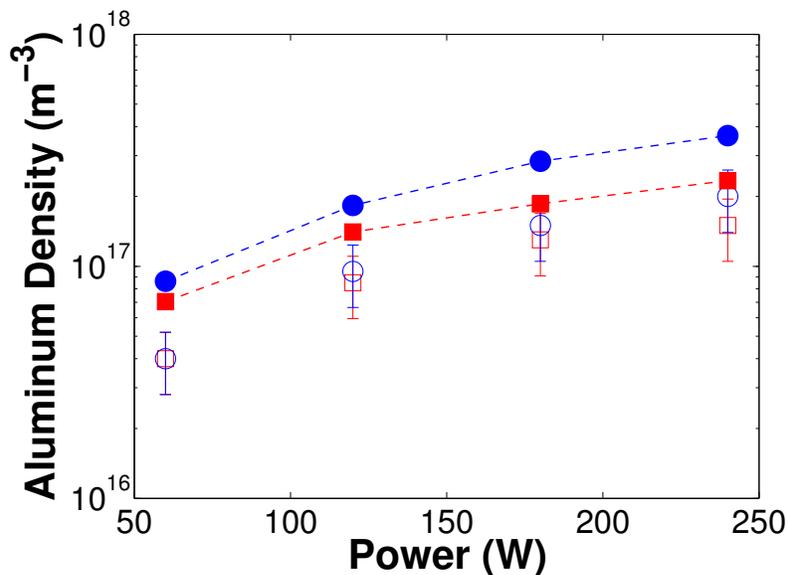
Figure 4.18: **Aluminum density taken at a distance of 5.6 cm from the target. The experimental data (open symbols) was taken from Ref. 207. The simulation values (filled symbols) were obtained integrating along a 15 cm line passing through the center of the discharge and parallel to the target. ■ 40 mTorr, □ 38 mTorr, • 20 mTorr, ○ 22 mTorr.**

We have also benchmarked our simulation with other experiments [210] where the absolute Ti density was measured using optical spectroscopy. The experimental setup used in this work is different to ours; we had to use the description of their system to modify our configuration file to resemble their geometry. Results of our modeling and their experiments are plotted in Figure 4.20 for Ti.
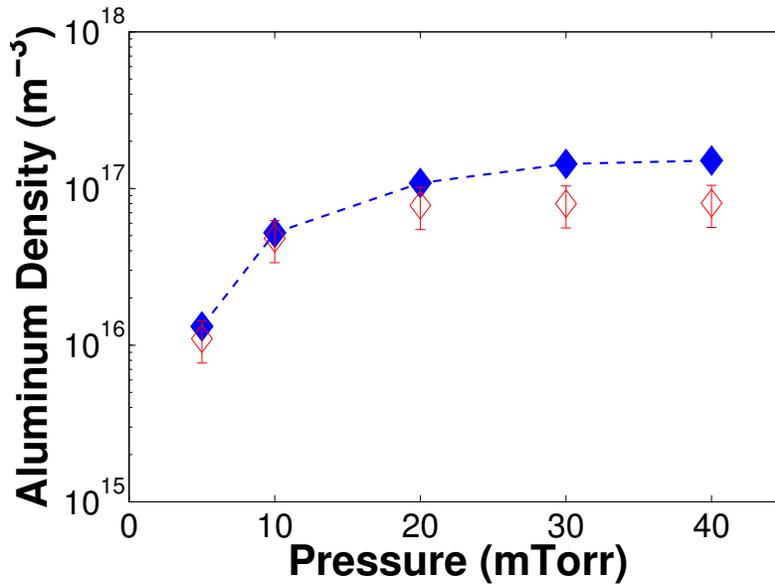
Figure 4.19: **The aluminum density as a function of pressure. The simulation (filled symbols) correctly captures the aluminum density saturation as pressure is increased observed in experiments (open symbols). Data is for 180 W.**

The uncertainty on the measurements is also about 30 %. We followed the same procedure described before to simulate the absorption length in the measurement technique. It is evident from Figure 4.20 that our results differ from experiments by a factor of ≈ 4. However, it is stated in the document that their experimental results are somehow higher than other reported values from similar experiments. To explain these higher values, several hypothesis were put forward. One of them is due to the calculation of the absorption length. Gaillard et al. estimated the absorption length of their experiment assuming a conic distribution of sputtered particles next to the target. In this way, they estimated an absorption length which was a function of the distance to the target. However, at this pressure (30 mTorr), transport of sputtered particles is dominated by diffusion, and the distribution of sputtered particles is not conic [126]. Based on this, the estimated absorption length in the experiments may have been underestimated which could bring their densities closer to our

simulated values. Despite the quantitative discrepancy, our modeling captures the overall relative trend of the experimental values quite well.
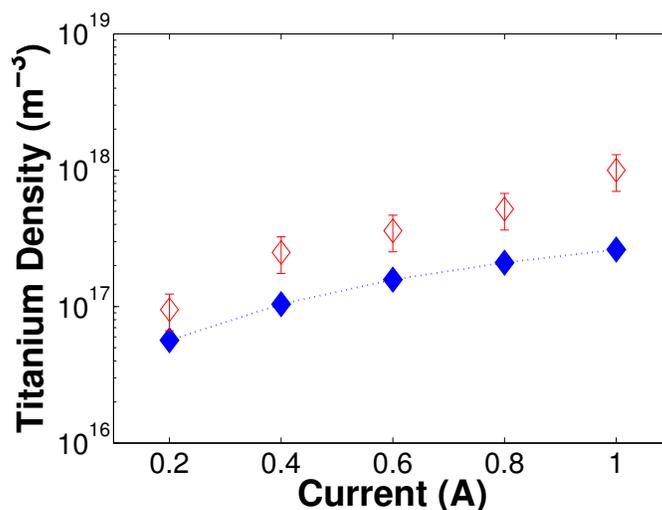
Figure 4.20: **Experimental (open symbols) and simulated (filled symbols) Ti densities corresponding to the experimental setup of Ref. 210. Data is taken at 8 cm from the target at 180 W.**

4.5   Rarefaction as a mechanism to increase ionization of metal species

Rarefaction is also a mechanism to increase the ionization of the sputtered particles in the discharge since the collision probability between an electron and a metal vapor atom depends on both the gas and the metal vapor density. This probability may be indeed too small to play a significant role at low pressures where the thermalization length is long and the majority of the sputtered flux is far from being in thermal equilibrium with the background gas. As pressure increases, however, particles are thermalized faster, and the density of sputtered particles close to the target grows. If the collision frequency of electrons and metal atoms increases, so does the ionization of metal atoms. The

ionization potential of metal atoms is significantly lower than that of argon, and both energetic and thermalized electrons are prone to ionize sputtered particles if the collision frequency increases. However, this is not commonly observed in normal sputtering discharges because the metal vapor density is significantly lower than that of the background gas. However, by reducing the density of the gas, the discharge itself promotes the growth of ionized metal particles. In Figure 4.21, the ratio $\frac{n_s}{n_n+n_s}$ is plotted where n_s is the sputtered atoms density and n_n is the background atom density. From the figure, it is clear that by increasing the sputtering power, the relative density of sputtered particles with respect to that of the buffer gas also increases. The effect is more pronounced for Cu; a greater rarefaction than for Al and Ti is observed for this element.
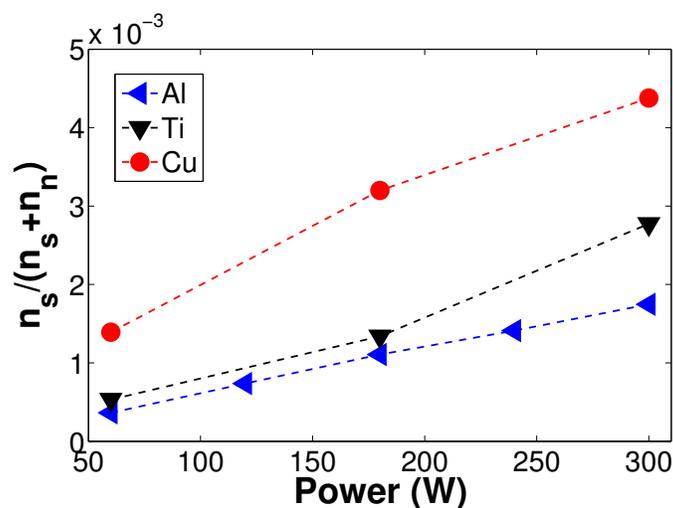
Figure 4.21: **The ratio $\frac{n_s}{n_s+n_n}$ for 3 different powers indicating the increased probability for electron impact ionization of the metal vapour with power. The pressure is 40 mTorr.**

Figure 4.22 shows the same ratio as in Figure 4.21 but now as function of pressure and at constant power. The observed effect is indeed driven by a combination of rarefaction and thermalization lengths. At low pressures, the mean free path of sputtered atoms is long and the probability of collision low. This is why methods to enhance the ionization of sputtered particles operate

at relatively high pressures. Typical pressures for ionized metal deposition and HiPIMS range from 10 to 30 mTorr. At these pressures, the mean free path of sputtered particles is just the appropriate value to balance the ionization of metal atoms with losses via diffusive transport. There is, however, with every process a point where all variables collude to increase the metal ionization but this optimum pressure is not easily predictable.
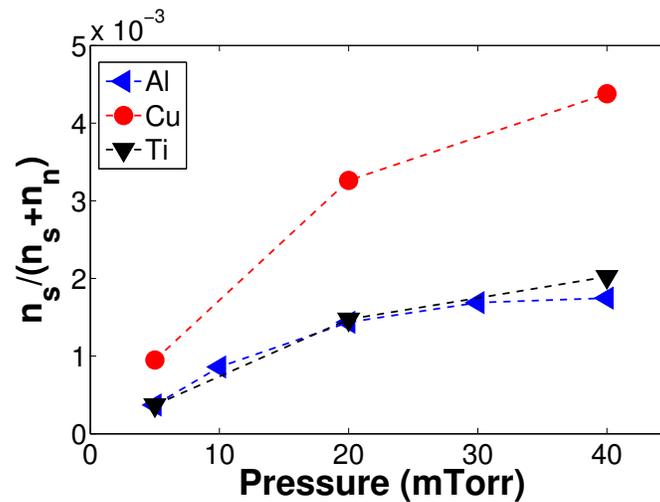
Figure 4.22: **The ratio** $\frac{n_s}{n_s+n_n}$ **for 3 different pressures for the materials investigated in this report.**

Rarefaction is important for the accurate simulation of the plasma in sputtering discharges. During a Monte Carlo integration of electron trajectories in an argon discharge, electrons are tracked in the magnetic field, and their mean free path is inversely proportional to the total gas density. An energetic electron can collide elastically or inelastically with any neutral atom but they collide more frequently with Ar atoms since they are the dominant species in the discharge. As long as the electron has enough energy to collide inelastically with Ar atoms (the lowest excitation energy of argon is about 11.5 eV [217]), they will continue slowing down until they eventually are absorbed by the walls of the system. However, if the ratio between argon and copper density gets smaller, electrons have other mechanisms to cool off since now the collision frequency with copper atoms increases. Moreover, electrons that have

lost most of their initial energy but are unable to collide inelastically with Ar have enough energy to ionize metal atoms. These electrons may be the ones responsible for the increased ionization observed in normal discharges. The same reasoning applies to electrons in the tail of the energy distribution function of the cold electron group. There is experimental evidence that this indeed is the case for cold electrons in a normal magnetron discharge where a reduction of the cold electron temperature is observed as power is increased [218]. This same effect was also observed by Field in his plasma survey using silver [2]. In this discussion, we have neglected three-body collisions and electron recombination for simplicity: these events are rare in the type of plasmas normally found in sputtering systems [127].

However, according to Bohlmark [219] and Christou [220], Penning ionization in the dominant source of metal ions, although Hopwood [221] indicates that the main ionization process for metallic species is electron impact ionization, especially at high pressures. It is evident that to determine the contribution of each phenomena by computational methods towards the increase is a topic of further research. A systematic analysis of Penning versus direct impact ionization is needed to find the relative contribution of each process for a specific material.

4.6 Summary

I have presented in this chapter the module within our computational framework that solves for the transport of neutral particles. The model is suitable for a great variety of chamber configurations and process conditions. I presented a study of rarefaction and gas heating as a result of energy and momentum deposition into the gas by energetic sputtered and reflected neutrals for three elements. The different mechanism driving the reduction of the process gas were addressed. A model for emission of sputtered and reflected particles that accounts for the noncosine distribution has been used in the simulation. The maximum temperature and rarefaction of the process gas has been found to

be dependent on parameters such as the target material, pressure and geometry of the chamber. Simulation results have been compared to experimental measurements of gas density and temperature showing good agreement in the whole range of pressures and powers investigated in this report. The simulated deposition rates were found to be in excellent agreement with experimental measurements. The ratio of metal to background species density increases for increasing power suggesting that direct impact ionization may play an important role in explaining the observed increase in metallic ion species in a conventional magnetron sputtering deposition. The relative contribution of direct impact ionization can only be calculated once Penning and asymmetric charge transfer processes are incorporated into the discharge model.

CHAPTER 5

PlaSpud: Numerical solution of glow discharges

5.1 Introduction

This chapter is about the numerical solution of the type of plasmas found
in sputtering discharges. The chapter starts with a brief introduction to the
mathematical models commonly used to described a glow discharge includ-
ing a simple derivation of the plasma fluid model. Next, the computational
model used to describe the magnetized discharge and the development of the
hybrid algorithm proposed in this chapter is described in detail. This consists
of the energetic particle's module which is fully described elsewhere [2]. Then,
the part of the algorithm that solves for the transport of low energy charged
particles is presented. In this section, I present a fluid plasma model of the
magnetron discharge where the effect of the magnetic field on the transport
properties of the plasma has been included, resulting in modified set of equa-
tions where the plasma transport coefficients are no longer considered isotropic.
The numerical model used for the solution of the magnetized fluid model is
described in detail in the following section. This model is a general scheme
that can be used to solve for any planar magnetron configuration (cylindrical,
rectangular, etc). Finally, a simple one dimensional model is used to test the
numerical algorithm proposed in this chapter. Results from this one dimen-

sional model are compared against published simulations to benchmark the plasma model.

5.2 Mathematical description of a glow discharge

The mathematical model that uses the least assumptions and simplifications to describe the transport characteristics of a plasma is the Boltzmann equation. As mentioned in the previous chapter, the Boltzmann equation is the quintessence of kinetic theory. This equations can be described as follows. Let $\boldsymbol{r}$ be a vector in $\Re^3$ and $\boldsymbol{v}$ a vector in velocity space. Then, the goal of kinetic theory is to find a general expression that approximates the statistical particle distribution, $\mathcal{F}(\boldsymbol{v}, \boldsymbol{r}, t)$. This distribution function is a mathematical representation of the number of particles that reside inside a volume in phase space at a given time. In other words, $\mathcal{F}$ represents the number of particles inside the volume $x + dx, y + dy, z + dz, v_x + dv_x, v_y + dv_y + v_z + dv_z$ at time t. One can find the relation between $\mathcal{F}$ and the total number density n by evaluating the expression [76]:

$$n(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} \mathcal{F}(\boldsymbol{r}, \boldsymbol{v}, t) \, d\boldsymbol{v} \tag{5.1}$$

$\mathcal{F}$ is commonly obtained solving the so-called Boltzmann equation (BE) which reads[75]:

$$\frac{\partial \mathcal{F}}{\partial t} + \boldsymbol{v} \cdot \nabla \mathcal{F} + \frac{\boldsymbol{F}}{m_\alpha} \cdot \frac{\partial \mathcal{F}}{\partial \boldsymbol{v}} = \left(\frac{\partial \mathcal{F}}{\partial t} \right)_c \tag{5.2}$$

where $\boldsymbol{F}$ is a force acting on the particles and $\left(\frac{\partial \mathcal{F}}{\partial t} \right)_c$ is the time change of $\mathcal{F}$ due to collisions. The right-hand side term is commonly referred as the collision integral and represents the number of particles that leave phase space

by collisions [75]. In magnetron discharges, $\boldsymbol{F}$ is a function of the electric and magnetic field, $\boldsymbol{E}$ and $\boldsymbol{B}$, respectively.

Numerical solutions of the Boltzmann equation are only possible for limited cases due to its extreme complexity. To solve this equation in realistic geometries, the simulationist faces insurmountable problems when building a model that can treat the discharge within complex geometries in more than one dimension. Finding methods for the solution of the Boltzmann equations is a topic of extensive research, and some authors suggest that there could be an efficient computational method in the foreseeable future [222].

A very popular way to approximate the Boltzmann equation is to use a simpler, more tractable model. The plasma fluid model is such model. The fluid model is a simple but powerful mathematical abstraction to describe the glow discharge. In the fluid model, the transport of charged particles is approximated by two or more impenetrable fluids [76]. The basic equations are obtained by integrating Equation 5.2 over velocity space [76]. This derivation of the plasma fluid equations is beyond the scope of this thesis; the reader is referred to the books of Chen [76], Lieberman [216] and Piel [75]. In the following, we derive the fluid equations using a simpler analysis [75, 76, 216].

The starting point is the continuity equation for species a which reads [76]:

$$\frac{\partial n_a}{\partial t} + \nabla \cdot (n_a \boldsymbol{u}_a) = R_a \tag{5.3}$$

where $\boldsymbol{u}_\alpha$ is the plasma fluid velocity and R_a is the source term of species a. The fluid velocity can be approximated by the momentum transport equation [216]

$$n_a m_a \left[\frac{\partial \boldsymbol{u}_a}{\partial t} + \boldsymbol{u}_a \cdot \nabla \boldsymbol{u}_a \right] = q n_a \boldsymbol{E} - \nabla p_a \tag{5.4}$$

where m_a is the mass of particle a, q is the electric charge, p_a is the partial pressure due to species a and ν_{na} is the neutral collision frequency between

particle a and a neutral gas. To avoid further confusion, we will drop the subscript n from ν; it is implicitly assumed that all collision frequencies involving particles a are with neutral species. Also, for simplicity, the force acting on the fluid is only due to an electric field $\boldsymbol{E}$.

If we neglect inertia $(\boldsymbol{u}_a \cdot \nabla \boldsymbol{u}_a)$ and use the approximation $nm\partial u/\partial t \approx \nu nm$ [223] in equation 5.4, we obtain [216]:

$$qn\boldsymbol{E} - \nabla p_a - mn\nu_m \boldsymbol{u}_a = 0 \tag{5.5}$$

Solving for $\boldsymbol{u}_a$ yields

$$\boldsymbol{u}_a = \frac{q\boldsymbol{E}}{m_a \nu_a} - \frac{kT}{m\nu_a}\frac{\nabla n_a}{n_a} \tag{5.6}$$

Finally, substituting Equation 5.6 into 5.3 we obtain

$$\frac{\partial n_a}{\partial t} + \nabla \cdot \Gamma_a = 0 \tag{5.7}$$

with

$$\Gamma_a = Z_a \mu_a \boldsymbol{E} n_a - D_a \nabla n_a \tag{5.8}$$

where Z_a is the particle's charge, μ_a and D_a are the particle's mobility and diffusivity, respectively, defined by

$$\mu_a = \frac{T_a}{m_a \nu_a} \tag{5.9}$$

$$D_a = \frac{q}{m_a \nu_a} \tag{5.10}$$

To close the system of equations, we need an expression for $\boldsymbol{E}$, which can be approximated by the solution of the Poisson equation

$$\nabla^2 \phi = -\frac{\rho}{\epsilon} \tag{5.11}$$

where ϕ is the plasma potential ($\boldsymbol{E} = -\nabla\phi$) and ρ is the charge density defined as

$$\rho = q(\sum Z_a n_a) \tag{5.12}$$

Equations 5.7 and 5.11 are the core of the plasma fluid model. For an electropositive discharge made of only electrons, positive ions and neutral particles, the plasma fluid model reads (in steady state):

$$\nabla \cdot (-\mu_e \boldsymbol{E} n - D_e \nabla n) = S_n \tag{5.13a}$$

$$\nabla \cdot (\mu_i \boldsymbol{E} p - D_i \nabla p) = S_i \tag{5.13b}$$

$$\nabla^2 \phi = -\frac{q}{\epsilon_0}(p - n) \tag{5.13c}$$

where n is the electron density and p is the ion density (assuming singly charged ions). The new terms S_e and S_i are the electron and ion creation rates respectively. I have opted for this notation instead of the traditional one used in plasma physics (n_e for the electron energy and n_i for the ion density) to simplify the description of the numerical algorithm later in the chapter. Subscripts e and i refer to electrons and ions respectively.

The treatment of magnetron discharges using the fluid model, or a simplified version of it, is documented in several articles in the literature. Bradley and Lister used a simple one-dimensional analysis of a magnetron discharge [224–226] to perform a parametric study of the cathode fall region [225]. Cramer [227] also performed a one dimensional analysis and included the effect of the magnetic field in the cathode region. Creation rates were approximated using a modified scheme using the Townsend coefficient approximation. Palmero et al. [228] used a similar one dimensional approach to the glow discharge to describe a radio-frequency sputtering plasma. All these studies dealt with

the discharge using a one dimensional approach. Theses models, while fast and relatively simple to implement, are not able to capture the complex electron dynamics of typical magnetron configurations [229]. A more appropriate treatment of the magnetized discharge should include a two (2D) or three (3D) dimensional description of the magnetic and electric fields in the reactor. Such models, however, are very difficult to solve numerically. The main complication is the strong coupling between the electron density and the electric field. Surzhikov and Shang [230], for example, developed a theoretical framework where the momentum equation for electrons (Equation 5.4) implicitly includes the effect of the magnetic field. The problem of numerically instability in dealing with the magnetized system of equations was tackled by first solving the system without a magnetic field and then using this solution an initial condition for the magnetized system. Costin [231, 232] developed a 2D model (cylindrical symmetry) that includes a more realistic description of the complex electron dynamics of a magnetron discharges since the treatment of the discharge in a cylindrical geometry allows one to include the $\boldsymbol{E} \times \boldsymbol{B}$ effect in magnetrons without the complications of a full 3D model. The solution method involves splitting the electron flux into two components, one classical and one magnetized. The classical model does not include the effect of the magnetic field. The key for solving this model numerically was to treat the classical flux implicitly in time while the magnetized flux was treated explicitly. Hagelaar [233] also solved the 2D fluid equations for a magnetized discharge assuming a constant, homogeneous magnetic field in the third dimension. His model was more in the line of examining the numerical aspect of the algorithm. Leray et al. [234] extended the fluid equations to include negative ion species and neutral particles. The numerical solution is simplified by using the plasma quasineutrality approximation ($p^+ = p^- + n$) where p^+ and p^- are the positive and negative ion densities respectively. By doing this, the strong coupling between the electron continuity and Poisson equation was avoided since the premise of the plasma approximation is to avoid the solution of the Poisson equation. Other applications of fluid models in plasma simulations include modeling direct-current (dc) discharges in argon [235–239] or helium [240–242],

capacitively [243–249] or inductively [250] coupled radio-frequency discharges, plasma enhanced chemical vapor deposition [251, 252] and high density plasma processing reactors [253, 254]. One disadvantage of fluid models when applied to magnetron discharges is the treatment of the cathode sheath whenever secondary electrons are created. This problem has been solved by modifying the boundary conditions at the target [231]. However, it is not always clear if this model can be generalized to different magnetron configurations or if the model is pressure dependent. Surzhikov and Sham [230] addressed the cathode sheath problem using a different approach. In their model, the plasma fluid equation was modified under the assumption of a highly directional electric field. The equations are then incorporated in their global model together with the modified boundary conditions. Another disadvantage of schemes based solely on the fluid approach is the calculation of creation rates for all processes in the plasma.

Another alternative to approximate the Boltzmann equation is a particle kinetic approach. Similar to the treatment of neutrals' transport in Chapter 4, the premise of particle-based algorithms is to track a great number of particles during their lifetime in the simulation. By indivdually solving for a large number of particles, we can approximate the solution of the Boltzmann equation without the need to solve Equation 5.2. Sheridan et al. [255] developed a Monte Carlo model to describe in detail the interaction between electrons and the magnetic filed. In a parallel work, Miranda et al. [256] published a study using the same algorithm to plot two-dimensional maps of ionization events and the effect of different magnetic fields on the ionization distribution. Ido [257] also used a Monte Carlo approach to investigate the effect of electron confinement for different configurations of the magnetic field. Holik et al.[258] used a Monte Carlo model to investigate the cross-corner effect of rectangular magnetrons. While particle approaches are powerful and relatively easy to implement, they are not self-consistent. Self-consistency in plasma simulations refers to the algorithm's ability to capture the change in the electromagnetic field resulting from the charged particle motion since these modified fields will disturb the original motion of particles in a nonlinear manner. Nevertheless,

particle based codes have been extensively used in magnetron discharges to helps us understand the complex relation between process variables otherwise difficult to observe using other methods. Other applications of Monte Carlo methods in glow discharges include [259–263].

To address the absence of self-consistency of pure Monte Carlo models, Birdsall [264] developed the Particle-in-Cell Monte Carlo technique (PIC-MC) where a solver for the electromagnetic field is included in the solution strategy. This method has been used to describe a magnetron discharge, and it is one of the prefered methods to describe it [156, 157, 171, 171, 265–277]. The method is similar to the dsMC technique introduced in the previous chapter in the way that all particles are followed during a time step. At the end of the time step, the electric and magnetic field are calculated based on the current densities. In general, PIC-MC models provide enough information about the discharge and are arguably one of the best algorithms to treat the magnetron discharge. However, from an engineering perspective, the PIC-MC algorithm does suffer from severe limitations both in physical dimensions and execution time. A comprehensive model of magnetron discharges such as the model presented in this dissertation would be severely limited to a small number of geometries where the model could be used. High execution times of up to months are observed even in state-of-the-art algorithms where the solution region is restricted to small geometries [278].

Arguably, the best time-accuracy compromise for magnetron discharges is provided by hybrid algorithms. The basic premises of hybrid models were addressed in Chapter 5. Hybrid models have been successful in describing the type of plasmas normally observed in magnetron discharges, but only if the transport coefficients are carefully chosen [222, 279]. Examples of hybrid modeling in magnetically confined plasmas are numerous. For instance, Porteous and Graves [280] used a constant magnetic field in the z direction to simulate a dc discharge. Grapperhaus et at. [281] developed a modeling platform they named Hybrid Plasma Equipment Model (HPEM) that has been extensively used to simulate several type of plasmas [282–284]. This platform has been

recently modified to simulate an IPVD reactor [155]. Shidoji [285] developed a two-dimensional model of the magnetron discharge that included the magnetic field. The model has been applied to study the erosion dynamics of the target and the effect of the magnetic field on the discharge characteristics[286–288]. Rondanini et al. [289] developed a comprehensive model that includes the neutral gas flow to predict electron and ion densities in magnetically confined plasmas. They found good agreement between simulations and experiments even considering the inherent assumptions of the fluid model. Kolev [290] expanded the modeling network of Bogaerts and co-workers[291] to simulate the magnetron discharge. He argued that hybrid models were numerically unstable in describing the discharge for magnetic fields typical of magnetron apparatus. I will describe later in the chapter a numerical model aimed at stabilizing the solution of the magnetized fluid model. Hybrid modeling has also been used to described the operation of several types of plasmas including dc [246, 292–295], nuclear reactors [296], plasmas for optical emission spectroscopy [297] and inductively coupled plasmas [298].

The model presented in the remainder of this chapter follows the hybrid scheme. Hybrid models, when appropriately defined, offer several practical advantages over PIC-MC algorithms such as shorter execution times, simulation of large systems and ease of implementation.

5.3 PlaSpud: A plasma model for magnetron sputtering discharges

PlaSpud is a hybrid Particle-Fluid algorithm for the solution of the transport of charged particles in a magnetron discharge. The concept behind the model is similar to the one described in Chapter 4. PS combines a MC algorithm with a fluid model in an attempt to reduce the computational load on the numerical method without compromising its accuracy. One of the goals of this module is to build a robust numerical model that can be applied to general geometries

and magnetic fields configurations.

Similar to the structure of the TS module, particles are divided into two groups: energetic and thermalized. The group in thermal equilibrium is referred hereafter as the low-energy group. This group is characterized by a single temperature T which is a measure of the average energy of the real electron and ion energy distribution function. The high energy group, hereafter referred as the *hot* group, is far from being in thermal equilibrium and cannot be characterized by a single temperature.

The general scheme of the PlaSpud model is illustrated in Figure 5.1. In the figure, PlasChem is the submodule in charge of providing the plasma chemistry to the ptMC algorithm and transport coefficients to the fluid model. The role of the PlasChem submodule in the overall solution strategy will be clear as we describe both the PS1 and PS2 algorithms. As in TS1, the PS1 module does not own a grid requiring interfaces to other modules where the grids are built. In the PS1 module, the algorithm to track energetic particles requires, apart from the local electric and magnetic field, information regarding the local gas temperature and densities at the position of the particle. In the computational framework presented in this dissertation, the background (neutral) gas density and temperature grid is owned by TS2 as presented in Chapter 4, the electric field is owned by the PS2 module and the magnetic field is an independent entity. As with all communications between modules in SpudII, the exchange of information between modules is mediated through abstract interfaces which are defined at the initialization stage of every instance of the SpudII program.

5.4 PlaSpud1 (PS1)

This module was written by Dr. David Field as part of his PhD dissertation [2]. Although I have modified some aspects of the module for the final integration of the framework, the key concepts used in this module were implemented by Dr. Field. The interested reader can find a more detailed description of this
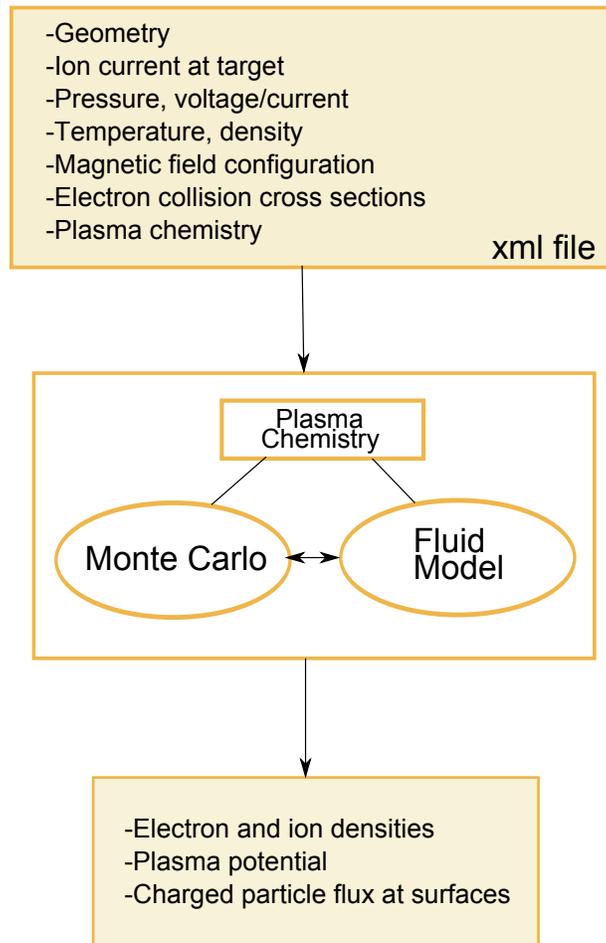
Figure 5.1: **PlaSpud scheme**

module in Ref. 2. In the following, I summarize the module for the sake of completeness.

The core of the PS1 algorithm is the Lorentz force equation. This equation governs the trajectory of a charged particle when it is under the influence of an electric and magnetic field. The Lorentz equations reads [76]

$$m_a \frac{d\boldsymbol{V}_a}{dt} = q\left(\boldsymbol{E} + \boldsymbol{V}_a \times \boldsymbol{b}\right) \tag{5.14}$$

where m_a is the mass of charged particle a, and $\boldsymbol{V}_a$ the velocity of particle a.

Equation 5.14 is simultaneously solved using a 4th order Runge-Kutta integrator [299] with an adaptive time step Δt. Hot particles are followed until they are thermalized or lost to the walls of the chamber. The current implementation of PS1 only tracks energetic ions inside the plasma sheath.

A particle is moved one time step after which it is determined if it undergoes a collision. The probability P of collision after traveling a distance δ_d is [300]:

$$P = 1 - \exp\left(-\delta_d \sum_{i=1}^{n} n_i \overline{\sigma}_i(E)\right) \tag{5.15}$$

where $\overline{\sigma}_i$ is the total collision cross section of processes between electrons and neutral particle i, n_i is the neutral number density of species i, and n is the total number of neutral particles with which an electron can collide. The total collision cross section for particle i is obtained via

$$\overline{\sigma}_i = \sum_{j=1}^{m} \sigma_{i,j} \tag{5.16}$$

where m is the total number of processes associated with the charged particle and the neutral i. An example of the energy dependent cross-sections for various collision processes between an electron and an argon atom is shown in Figure 5.2.

The procedure to move a charged particle and to determine whether a collision has occurred or not is as follows. First, a particle is moved a distance δ_d by solving equation 5.14 for an ascribed Δt. The method to chose an appropriate value of Δt is described in detail in [2]. The choice of Δt is extremely important for the stability of the simulation. Once the particle's trajectory has been determined, a random number X_1 is generated and compared against P. If $X_1 < P$, the particle continues its trajectory without a collision, a new integration of Equation 5.14 is performed and a new random number is generated for
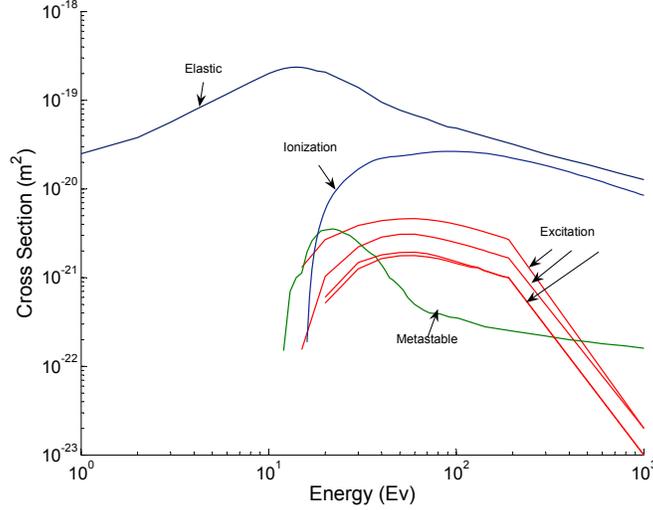
Figure 5.2: **Electron collision cross-sections as a function of energy for a collision between an electron and an argon atom [301].**

comparing against the new P. This process is repeated until $X_1 > P$. At this point, we simulate a collision process using a model proposed by Nanbu [300]. First, we chose the type of particle to collide with by building the fractional probability list

$$P_{i+1} = \frac{Pi + n_{i+1}\bar{\sigma}_{i+1}}{\sum_{i=1}^{n} n_i \bar{\sigma}_i} \tag{5.17}$$

with $P_0 = 0$, $P_1 = n_1\bar{\sigma}_1$ and $P_n = 1$. Secondly, we generate a new random number X_2 and choose the neutral particle i to collide with by finding the P_i from the list above for which $X_2 < P_i$ and $X_2 > P_{i-1}$.

With the collision partners chosen, e and n_i, the corresponding process is determined. To accomplish this, we build a new fractional probability list using

$$P_{j+1} = \frac{P_j + \sigma(E)_{i,j+1}}{\sum_{j=1}^{m} \sigma_{i,j}} \tag{5.18}$$

with $P_0 = 0$, $P_1 = \sigma_{i,1}$ and $P_{i,m} = 1$. We generate a third random number X_3 and find the process j by finding the P_j for which $X_3 < P_j$ and $X_3 > P_{j-1}$.

With neutral particle i and plasma process j selected by the method above, we compute the collision event according to the rules associated with each plasma process included in the simulation. New trajectories are found using procedures similar to those described in Chapter 4 with the exception that now inelastic processes are considered so that kinetic energy is not conserved.

5.4.1 Plasma processes

In a non-reactive plasma process, like the one considered in this dissertation, the most relevant processes occurring between an energetic electron and a neutral atoms in the bulk of the plasma are: elastic collisions, ionization and excitation. These events are characterized by a threshold energy for the process to occur. Only for inelastic processes (ionization and excitation), this threshold energy, hereafter referred as the inelastic energy of the process (E_{ine}), is non zero. Some of the processes included during a typical run of PS1 are listed in Table 5.1. To add/remove processes from the simulation we just require to edit the input xml file at the beginning of the simulation.

Each process is briefly described in the following lines. To simplify the description, the hot electron being tracked in the Monte Carlo code will be referred as the primary electron while the neutral gas atom the electron collides with will be referred as the gas atom. Further details on the algorithm can be found in Ref. 2.

5.4.2 Elastic scattering

The model used in PS1 to treat elastic collisions is based on the work of Surendra et al. [293] who proposed an analytical expression to calculate the scattering angle probability as a function of the primary electron energy. This

Table 5.1: Plasma processes between electrons and neutral particles in a dc magnetron discharge. The list includes elastic collisions, excitation events and ionization of both background gas and metal species [2, 300, 301].

Process	**Type**	E_{ine} **(eV)**
$e + Ar \rightarrow e + Ar$	Elastic collision	0
$e + Ar \rightarrow 2e + Ar^+$	Argon ionization	15.76
$e + Ar \rightarrow e + Ar^*$	Argon excitation 1	11.79
$e + Ar \rightarrow e + Ar^*$	Argon excitation 2	14.25
$e + Ar \rightarrow e + Ar^*$	Argon excitation 3	14.79
$e + Ar \rightarrow e + Ar^*$	Argon excitation 4	15.48
$e + Ar \rightarrow e + Ar^m$	Argon metastable	11.60
$e + Ar^m \rightarrow 2e + Ar^+$	Argon step ionization	4.14
$e + M \rightarrow e + M$	Elastic collision with metal atom	0
$e + M \rightarrow e + M^+$	Metal atom ionization	Depends on atom
$e + M \rightarrow e + M^*$	Metal atom excitation	Depends on atom

model is characterized by a higher probability of forward scattering at high energies while, at low energies, the scattering is isotropic. The process to simulate an elastic collision is simple. First, a dummy gas atom of species i is created from the TS2 grid. We use a dummy particle since the elastic event will hardly modify the energy and direction of the gas particle given the great mass difference between the gas and the electron. Next, the electron trajectory is modified by randomly selecting a scattering angle according to the Surendra model. To randomly choose the scattering angle ψ of each collision, Field followed the following procedure [2]. First, a random number X_4 is generated and the equation

$$X_4 = 2\pi \int_0^\psi \sigma(E, \psi') \sin \psi' d\psi' \tag{5.19}$$

is solved until a value of ψ is found for which the expression above is true. The expression to calculate $\sigma(E_0, \psi)$ is given also by Surendra et al. [293]

$$\sigma(E, \psi') = \frac{E}{4\pi \left[1 + E \sin^2(\psi'/2)\right] ln(1 + E)} \tag{5.20}$$

Similar to TS1, $\varphi = 2\pi X_5$ where X_5 is another random number. To calculate the post-collision velocities, a procedure similar to the one described in section 4.3.1 is followed. All calculations during this step are performed in the lab frame of reference since, given the negligible electron mass with respect to that of gas atoms, both reference frames, lab and center of mass, are practically the same. After the collision, the energy transferred in the collision E_{trans} is obtained by [174]

$$E_{trans} = E \left(1 - 2\frac{m_e}{m_n}(1 - \cos \psi)\right) \tag{5.21}$$

where m_e is the electron mass and m_n is the mass of the neutral particle. After the collision, the primary electron continues its trajectory with an energy $E - E_{trans}$.

The choice of accurate cross-sections is an important step in Monte Carlo algorithms [302]. Several cross-sections for elastic electron scattering on argon are available in literature [268, 303–306]. PS1 originally used the cross-sections used in Ref. 268 which are based on the work of Hayashi et al.[307]. Phelps et al. [301] revised these cross-sections and compiled a list they recommend for Monte Carlo codes [301]. These cross-sections are available online [301] and are the ones I use in the current implementation of PS1.

Data on elastic cross-sections between electrons and metal atoms is, on the other hand, more difficult to obtain. Unfortunately, experimental data is only available in a limited energy range with no practical use for typical magnetron discharge models [308, 309]. Thus, I have approximated these cross-sections with a screened Rutherford cross section using the equation proposed in Ref. 310. Elastic collisions between metal vapors and electrons were not included in the original implementation of PS1.

5.4.3 Excitation

In general, two types of excitation events occur in the plasma bulk. The first type includes all short-lived excitation events. These excited states are commonly de-excited by electric dipole radiation rather than by collisions [216]. This fast transition allow us to assume that excited atoms return to the ground state fast enough so they do not collide with another particle. Field [2] used the cross sections from Nanbu et al. [268] in which all argon exited states are grouped in 4 processes (see figure 5.2). These cross-sections are based on the work of Hayashi et al.[307]. To speed up the calculations, I have combined all of Hayashi's cross-sections into a single effective excitation cross section in PS1.

The second group of excited particles consist of those that do not immediately emit radiation [311]. The lifetime of these atoms is long compared to the diffusion length; thus, they remain in the chamber for longer periods of time. Similar to the previous group, the two metastable states of argon, namely the $^3P_2$ and $^3P_0$ states, are combined into an effective metastable state (inelastic energy of 11.55 eV). We followed this approach since our interest is on the total metastable population [301, 312, 313]. The cross-sections for metastable excitations were obtained from the Phelps's dataset available elsewhere [301].

It is not clear what direction the primary electron should follows after an excitation event. However, Field considered that the Surendra model is a good approximation for the scattering process of an excitation event [2]. Thus, the same model used for elastic collisions is used to calculate the post-collision velocity of the primary electron. After the collision, the electron continues its trajectory with an energy equal to E-E_{trans}-E_{ine} where E_{ine} depends on the inelastic energy of the excitation event.

5.4.4 Ionization

During an ionization event, the primary electron collides with the gas atom, stripping one or more electrons from the outer core. To simulate an ionization event in PS1, we first pull a gas atom from the density grid (TS2) and proceed to create a new electron and ion. The new electron is given an energy E_{new} and velocity V_{new} according to a prescribed model. Field suggested a simple scheme based on the results of Carman [314] whereby the energy transferred to the extra electron is uniformly distributed in the range $(0, 4)$ eV and the trajectory is uniformly distributed on the surface of the unit sphere. The created ion maintains the same energy and direction of the original gas atom. Both, the new electron and the new ion, are sent to the ParticleBucket to be simulated afterwards. The ParticleBucket object in PS1 is similar to the container used in TS1 that temporally holds energetic particles for later processing. These new particles may need to be followed if they were created in the sheath where they can gain/lose energy very rapidly. If the particles in the ParticleBucket are not in the sheath and their energy is small to be considered as thermalized, they are sent to the PS2 grid to be treated as particle creation rates for the conservation equations.

Cross-sections of argon impact ionization are also obtained from the Phelps's dataset [301]. Unfortunately, cross-sections for direct impact ionization of metal vapors is scarce. In this dissertation, data for direct impact ionization of aluminum was extracted from Ref. 315 and data for copper is from Ref. 316. In cases where no experimental data is available, the formula proposed by Lotz [317] is used.

The choice of an appropriate model for the post-collision velocities is also uncertain. Several alternatives are discussed in Ref. 318, however, Field considered that Surendra's model was appropriate for the conditions normally found in magnetron discharges. Thus, after an ionization event, the primary electron continues its trip in the discharge with an energy equal to E-$E_{ioni} - E_{new}$ where E_{ioni} is the ionization potential of the gas atom with which the primary

electron is colliding.

5.4.5 Thermalization

Once the outcome of the collision event has been determined, the algorithm computes the energy and position of the particle to decide if it should keep on tracking the particle. If the particle is located inside the cathode sheath, the algorithm keeps on following it regardless of its energy E. If, on the contrary, the particle is not inside the sheath, E is compared against an arbitrary threshold E_{Th}. If $E < E_{Th}$, the particle is deemed as thermalized and is sent to the PS2 grid to be processed in the fluid model. The numerical value of E_{Th} is chosen as the smallest among all inelastic processes defined in the simulation. For instance, based on Table 5.1, and with no metal atom ionization considered, $E_{Th} = 11.5eV$, the inelastic energy for the excitation energy event 1 in in Table 5.1. However, if we also include ionization and excitation events involving metal species, E_{Th} should be equal to the lowest inelastic energy of all processes defined in the xml input file.

5.4.6 Secondary electron emission coefficient and reflection coefficient

The treatment of the secondary electron emission coefficient γ_e is more complicated. Field [2] explains in detail the problem of adopting a certain model and the need for treating this value as a fitting parameter in the simulation. In general terms, γ_e can be approximated using equations

$$\gamma_e = 0.032(0.78E_{ioni} - 2\phi_m) \tag{5.22}$$

$$\gamma_e = 0.2(0.8E_{ioni} - 2\phi_m)/E_F \tag{5.23}$$

where E_{ioni} is the ionization potential of the parent gas of the ion hitting the surface, ϕ_m is the work function of the surface and E_F is the Fermi energy of the

surface material. Equations 5.22 [196] and 5.23 [319] are only approximations to the actual value, and they have been found to be valid only for ion energies below 500 eV [320]. However, the exact value is difficult to obtain since it depends on numerous variables. For instance, Phelps found that γ_e strongly depends on the cleanliness of the surface [320], and may significantly vary depending on the treatment of the surface. In another study, Depla et al. [321] measured the discharge characteristics for several materials and correlated the SEEC of the material with the inverse of the discharge voltage finding a strong correlation between the two parameters for pure target materials. The authors also observed a strong SEEC dependence on the surface conditions of the target.

Another surface parameter is the so called electron reflection coefficient (ERC) [156, 322]. This parameter reflects the likelihood of electrons to be reflected when they hit a surface. In the case of secondary electrons in magnetron sputtering configuration, a finite ERC means that, if they happen to return to the target suffering no innelastic collision in their trajectory, there is a possibility for the electron to return to the discharge region without being absorbed by the target. Field observed a small dependence on creation rates as function of the ERC [2] when the value changed from 0.25 to 0.35 at 5 mTorr. The effect of the ERC is greater at low pressures when the free path of electrons is large and electrons have a greater probability of being recaptured at the cathode.

These coefficients are not well characterized yet, and their numerical values span a wide range depending on several variables. See Ref. 320 for a discussion on the SEEC coefficient. In the results presented in next chapter, I have opted for average values of these coefficients. Thus, $\gamma_e^{Al} = 0.095$ [321] and $ERC_{Al} = 0.3$ [2]. The effect of these two parameters on the results predicted by the algorithm described in this section is explained in detail in Ref. 2.

5.4.7 Scaling factors

To scale simulation time to real time in the plasma module, the scaling rate factor SF_{PS} is calculated in a similar manner to SF_{TS}. An input parameter of the simulator is the number of ions to simulate N_i. Ions are pulled from the sheath edge and traced until they reach the target surface, and PS1 will stop once N_i sheath ions have been simulated. These ions, and the ones born within the sheath N_s, comprise the total ion flux at the cathode surface. Thus, to weight this ion flux with the real current

$$SF_{PS} = N_t/I_t \qquad (5.24)$$

where $N_t = N_i + N_s$.

5.4.8 n^{hot}

The method to calculate the hot electron density in the reactor consists of measuring the time hot electrons spends in each cell in the grid. Every cell in the grid has a timer that is increased an amount δt_{hot} every time a hot particle spends time on the cell. Once PS1 has finished simulating ions, this timer is first averaged over the number of hot electrons and then scaled by SF_{PS} and divided by the cell volume. This value represents the hot electron density in units of m^{-3}.

5.4.9 Algorithm

The process described in the previous subsection is summarized in Figure 5.3. To start the simulation, the algorithm extracts an ion from the plasma sheath edge where ions arrive by diffusion from the bulk of the plasma. Once in the plasma sheath, ions are accelerated towards the cathode by the high electric field in the region next to the cathode. The resulting impact, as explained

in Chapter 2, may result in several particles being ejected from the target. When a secondary electron is ejected, the algorithm then follows the electron and all byproduct particles (temporarily stored in a particle bucket) created as a result of the interaction with the process gas. Once all particles in the ParticleBucket have been simulated, the algorithm pulls another ion from the sheath, and the process just described repeats itself until a sufficient number of primary ions have been simulated.

5.5 PlaSpud2 (PS2)

The goal of the PlaSpud2 module is to solve for the transport of low-energy charged particles in a magnetized plasma. The mathematical derivation presented in this module is based on the work of Dr. Yun-Feng Shao and Dr. Ming Lee.

In the absence of a magnetic field, the plasma fluid model of a dc discharge is approximated by the solution of Equations 5.13. To solve this nonlinear system of equations, first we need to specify the source terms and transport coefficients.

In PS2, S_n and S_i are obtained from the PS1 module. When particles are run in the PS1 algorithm and deemed thermalized, they are sent to the PS2 grid where they are collected and added to the total creation rates per cell. When PS1 finishes, it needs to scale all the rates resulting from the interaction of hot electrons with background atoms. To scale these rates, PS1 first computes Equation 5.24 to obtain SF_{PS} and then it scales all processes using

$$S_x = \frac{S_x^{PS1} SF_{PS}}{V_{cell}} \qquad (5.25)$$

where S_x^{PS1} are number of events collected in the PS2 grid from process x, V_{cell} is the grid cell volume and S_x are all creation rates produced in the plasma
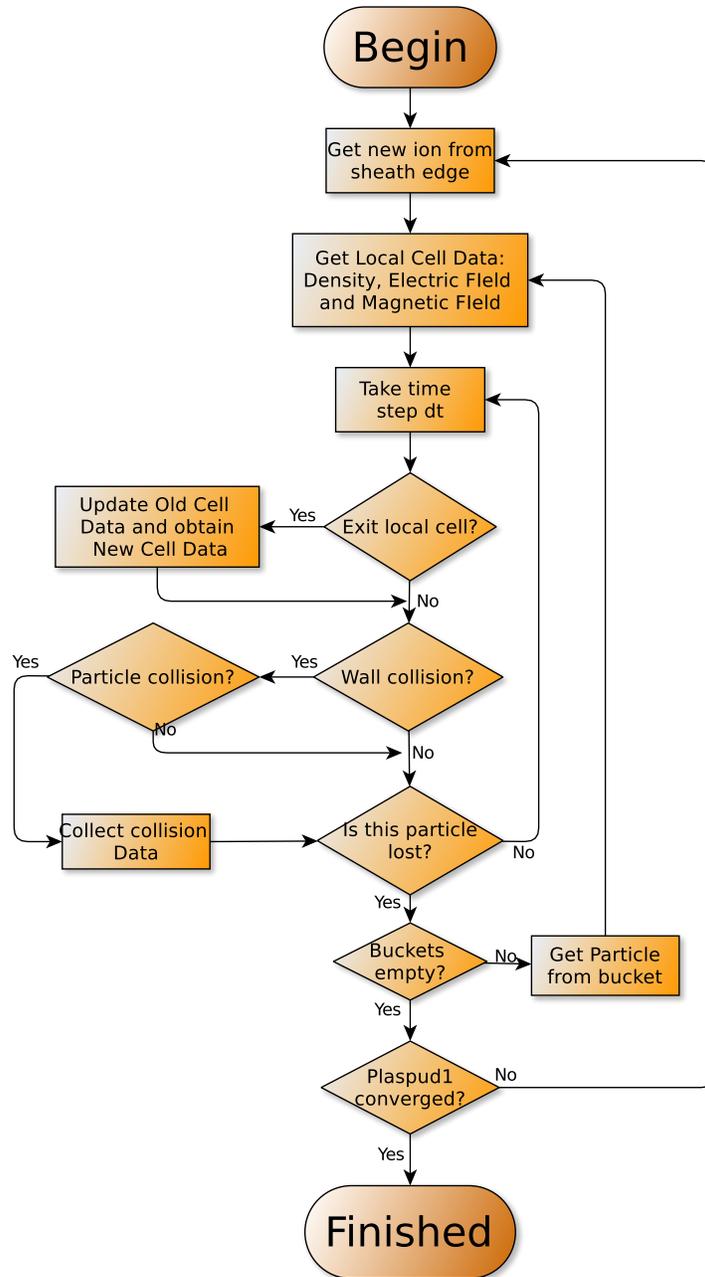
Figure 5.3: **PlaSpud1 algorithm for the simulation of high energy particles. Figure taken from Ref. 2.**

module (thermalized electrons, process and metal gas ionization, metastables and excitation events.)

The transport equations μ_e, μ_i, D_e and D_i can be approximated using the relations derived in Section 5.2, namely

$$\mu_a = \frac{q}{m_{an}\nu_a} \tag{5.26}$$

$$D_a = \frac{k_B T_a}{m_{an}\nu_a} \tag{5.27}$$

where T_a is a temperature associated with particles a, ν is the collision frequency between particle a and neutral atoms, m_{an} is the effective mass and n_n is total neutral particle density $\sum n_i$. The the transport coefficients follow the Einstein relation

$$D = \mu k_B T \tag{5.28}$$

Instead of determining the collision frequencies for both electrons and ions, we have used a simpler approach to obtain the transport coefficients based on experimental data. For ions, we have used the expression proposed by Phelps [320] to estimate the ion mobility. The equation was obtained from a fitting to experimental results. The Phelps equations reads

$$W = 4(E/n)/\{1 + [0.007(E/n)]^{1.5}\}^{0.33} \tag{5.29}$$

where W is the argon drift velocity ($\mu_i|\boldsymbol{E}|$) and E/n is the reduced field in units of Townsend (1 Td $= 10^{-21}\mathrm{Vm}^2$). Figure 5.4 compares the Phelps formula against experimental values and other models to calculate μ_i. D_i is obtained from Equation 5.28 assuming ions are in thermal equilibrium with the background gas atoms.

The electron coefficients are taken from a numerical algorithm of the Boltzmann equation [326]. The value extracted from this reference is in reasonable
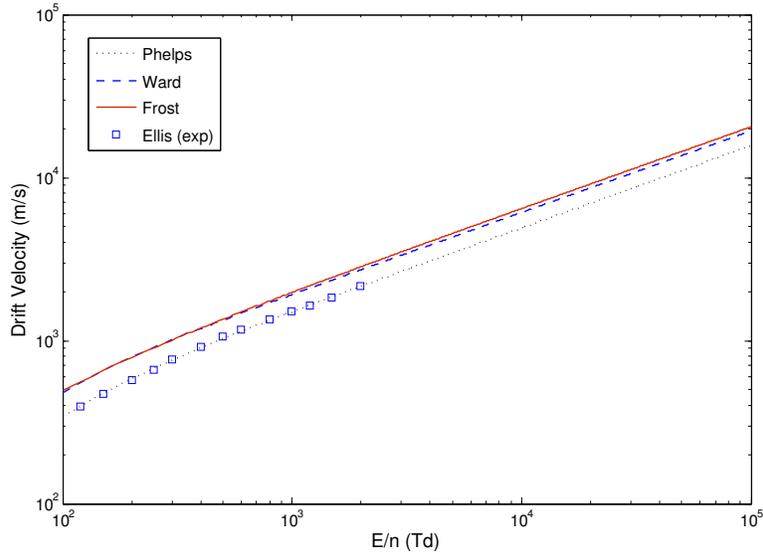
Figure 5.4: **The ion mobility models of Phelps and Petrovich [320], Frost [323] and Ward [324]. Experimental data from Ellis et al. [325] is included for comparison**

agreement with reported values of μ_e used in literature [235, 237, 295, 303]. D_e is assumed to follow the Einstein relation. To approximate the electron temperature, we have used a constant value approximation [295]. It is argued in [230] that the use of a constant electron temperature to solve the magnetized plasma equations is a good approximation. Treating the mean electron temperature as another variable would significanltly increase the complexity of the numerical solver since the equation governing the energy transport in plasmas is a function of the electron flux, and, as it will be seen in the next section, this quantity is difficult to obtain even using simplified models.

5.5.1 The Magnetized Plasma

Equations 5.13 were derived under the assumption of zero magnetic field. However, when $\boldsymbol{B} \neq 0$, the transport properties of the plasma greatly changes and Equations 5.13 should be modified. The main implication is that now

the transport coefficients are no longer isotropic but have different values depending on the direction of the magnetic field. The magnetic field alters the transport properties of electrons in such a way that they are less prone to move perpendicular to the magnetic field.

To obtain the fluid model for a finite $\boldsymbol{B}$, the force exerted by the magnetic field on the charged particles should be included in Equation 5.4. For magnetic fields typical of sputtering systems, we can safely assume that only the electrons are tightly bound to it while ions are non magnetized [230, 231, 290, 326]. Therefore, to obtain the electron transport equation with a magnetic field, including $\boldsymbol{B}$ in Equation 5.4 and dropping the subscript e for convenience yields

$$nm \left[\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} \right] = q(n\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B}) - \nabla p \tag{5.30}$$

Following a similar analysis to that of section 5.2 and after some lengthly algebra, the total electron flux $\boldsymbol{J}$ in a magnetic and electric field is [76, 233]:

$$\mathcal{J}_x = \frac{1 + \Omega_x^2}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial x} - D \frac{\partial n}{\partial x} \right] + \frac{\Omega_x \Omega_y}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial y} - D \frac{\partial n}{\partial y} \right]$$
$$\frac{\Omega_x \Omega_z}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial z} - D \frac{\partial n}{\partial z} \right] \tag{5.31a}$$

$$\mathcal{J}_y = \frac{1 + \Omega_y^2}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial y} - D \frac{\partial n}{\partial y} \right] + \frac{\Omega_y \Omega_x}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial x} - D \frac{\partial x}{\partial x} \right]$$
$$\frac{\Omega_y \Omega_z}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial z} - D \frac{\partial n}{\partial z} \right] \tag{5.31b}$$

$$\mathcal{J}_z = \frac{1 + \Omega_z^2}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial z} - D \frac{\partial n}{\partial z} \right] + \frac{\Omega_z \Omega_x}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial x} - D \frac{\partial n}{\partial x} \right]$$
$$\frac{\Omega_z \Omega_y}{1 + \Omega^2} \left[\mu n \frac{\partial \phi}{\partial y} - D \frac{\partial n}{\partial y} \right] \tag{5.31c}$$

where $\boldsymbol{\Omega} = \Omega \boldsymbol{B}/B$, and $\Omega =$ is the hall parameter obtained from

$$\Omega = \frac{|q\boldsymbol{B}|}{m\nu} \tag{5.32}$$

The modified system of equations to solve is then reduced to:

$$\nabla^2 \phi = -\frac{q}{\epsilon_0}(n_i - n_e - n_{hot}) \tag{5.33a}$$

$$\nabla \cdot \tilde{\boldsymbol{J}}_{\boldsymbol{e}} = S_e - \nabla \cdot \hat{\boldsymbol{J}}_e \tag{5.33b}$$

$$\nabla \cdot \boldsymbol{J}_{\boldsymbol{i}} = S_i \tag{5.33c}$$

$$\tilde{\boldsymbol{J}}_e = -\overline{\boldsymbol{\mu}}_e \boldsymbol{E} n_e - \overline{\boldsymbol{D}}_e \nabla n_e \tag{5.34a}$$

$$\boldsymbol{J}_i = \mu_i n_i \boldsymbol{E} - D_i \nabla n_i \tag{5.34b}$$

$$\hat{\boldsymbol{J}}_e = -\hat{\overline{\boldsymbol{\mu}}}_e n_e \boldsymbol{E} - \hat{\overline{\boldsymbol{D}}}_e \nabla n_e \tag{5.34c}$$

where n_{hot} is the hot electron density, $\overline{\boldsymbol{\mu}}_e = \mu_e \overline{\boldsymbol{H}}$, $\overline{\boldsymbol{D}}_e = D_e \overline{\boldsymbol{H}}$, $\hat{\overline{\boldsymbol{\mu}}} = \mu_e \overline{\boldsymbol{G}}$, $\hat{\overline{\boldsymbol{D}}}_e = D_e \overline{\boldsymbol{G}}$. $\overline{\boldsymbol{H}}$ and $\overline{\boldsymbol{G}}$ are obtained from the matrix $\overline{\boldsymbol{L}}$ as $\overline{\boldsymbol{H}} = diag(\overline{\boldsymbol{L}})$ and $\overline{\boldsymbol{G}} = \overline{\boldsymbol{L}} - diag(\overline{\boldsymbol{L}})$. The matrix $\overline{\boldsymbol{L}}$ is defined as

$$\overline{\boldsymbol{L}} = \begin{bmatrix} \dfrac{1+\Omega_x^2}{1+\Omega^2} & \dfrac{\Omega_x\Omega_y}{1+\Omega^2} & \dfrac{\Omega_x\Omega_z}{1+\Omega^2} \\[2ex] \dfrac{\Omega_y\Omega_x}{1+\Omega^2} & \dfrac{1+\Omega_y^2}{1+\Omega^2} & \dfrac{\Omega_y\Omega_z}{1+\Omega^2} \\[2ex] \dfrac{\Omega_z\Omega_x}{1+\Omega^2} & \dfrac{\Omega_z\Omega_y}{1+\Omega^2} & \dfrac{1+\Omega_z^2}{1+\Omega^2} \end{bmatrix} \tag{5.35}$$

Boundary conditions

Since the electron flux next to the target due to hot particles is handled in PS1, the boundary conditions for the fluid model are greatly simplified. In this model, the choice of simple boundary conditions has been driven by the the need to avoid numerical instabilities in the algorithm. The boundary conditions in our model are:

At the target:

$$\nabla p = 0$$
$$n = 0$$
$$\phi = V_{target}$$

any other surface:

$$p = 0$$
$$n = 0$$
$$\phi = 0$$

These simple boundary conditions are commonly used in plasma fluid simulation [237, 245]. An alternative model is the one proposed by Hagelaar [327] where particle fluxes are specified at each surface. However, given the high electric field at the target, we expect that the particle fluxes are dominated by the convective component. Therefore, the use of the simplified boundary conditions is justified. I tried the Hagelaar boundary conditions in the simple one dimensional model described later in the chapter and observed no significant difference at the target between the simple boundary conditions described before and those of Hagelaar. Furthermore, the use of the Hagelaar set reduced the convergence and stability of our solver. For this reason, only the simplified set of boundary conditions is used in the simulations presented in the next chapter.

5.6   Numerical method: one dimensional test model

In this section, I describe the numerical method used in PS2 to solve equations 5.33. For simplicity, the model used to derive the numerical method is only one dimensional with $\boldsymbol{B} = 0$. The method, while inefficient for the one dimensional

model described in this section, is designed for the solution of the plasma equations when the number of cells is large.

The plasma fluid model is representative of what is known as a stiff problem [299]. The strong nonlinear coupling between the electric field and the electron transport is the origin of this stiffness. The high nonlinearity of the plasma equations is more evident at the initial stages of the discharge, and care should be taken to avoid numerical instabilities. The direct application of general solvers for nonlinear equations is not justified when the initial guess is not close to the solution [183, 328]. In such cases, the solution may fatally diverge even if special measures are taken to stabilize the method. To address this problem, I have incorporated a fictitious time dependence into the plasma equations. These types of numerical schemes are known as continuation methods [329–331]. The new variable t (do not confuse with the real time) is just a numerical artifact to alleviate numerical instabilities due to large changes at the beginning of the simulation [331]. The time marching solution thus obtained is not representative of the time evolution of the discharge. Thus, the equations to solve in the model are (in scaled units):

$$\nabla^2 \hat{\phi} = -\frac{\hat{q}}{\epsilon_0} \left(\hat{n} - \hat{p} \right) \tag{5.36a}$$

$$\frac{d\hat{n}}{dt} + \nabla \cdot \boldsymbol{J}_{\hat{n}} = S_{\hat{n}} \tag{5.36b}$$

$$\frac{d\hat{p}}{dt} + \nabla \cdot \boldsymbol{J}_{\hat{p}} = S_{\hat{p}} \tag{5.36c}$$

$$\tag{5.36d}$$

where $\hat{\phi} = \dfrac{\phi}{\phi_0}$, $\hat{n} = \dfrac{n}{n_0}$, $\hat{p} = \dfrac{p}{p_0}$, $\hat{S}_n = \dfrac{S_n}{n_0}$ and $\hat{S}_p = \dfrac{S_p}{p_0}$ and the 0 subscript denoting the maximum value of the variable. In the remainder of this document, the hat has been dropped for ease of representation.

The test model consists of two infinitely large electrodes separated by a length L. An argon gas discharge is formed between these electrodes by applying a potential between the electrodes. This simple one dimensional model is

summarized in Figure 5.5. The set of partial differential equations describing the plasma was discretized using the finite volume method (FVM) [166, 168]. The domain is divided into N cells of equal volume and size. Let us call x the axis perpendicular to these electrodes and i the index of each cell. Therefore, there are N cells in our model and $3N$ variables defined at the center of each cell on the grid (n, p, ϕ). At the electrodes, $\phi_A = -V$ and $\phi_B = 0$. The creation rates used in the model have been chosen for comparison purposes only.
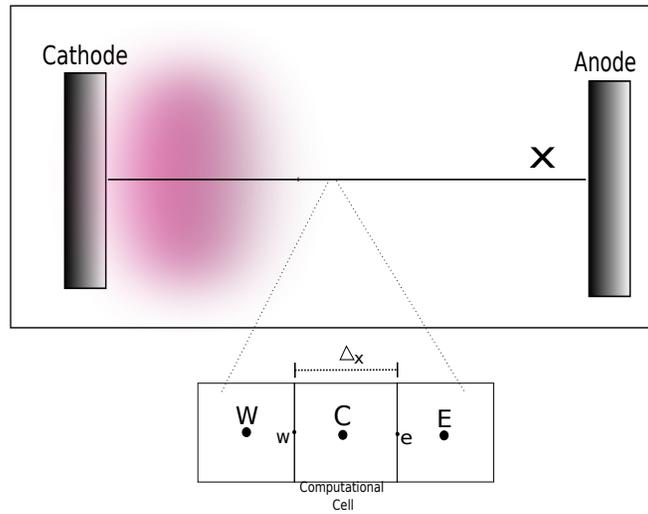
Figure 5.5: **Description of the simple gas discharge model used to test the algorithm. The domain is divided in cells of equal volume where the plasma equations are discretized using the finite volume method. In the model, uppercase subscripts refer to values taken at the cell's center while lowercase subscripts are taken at the cell interface.**

First, a simple Euler discretization in time and an implicit FV discretization in space leads to

$$\frac{n_C^{k+1} - n_C^k}{\Delta t} + \frac{J_e^{k+1} - J_w^{k+1}}{\Delta x} = S \tag{5.37}$$

where values taken at time t^k are from the current solution whereas values at time t^{k+1} are the values at the next time step, and the particle fluxes J are expressed as

$$J = bn - D\nabla n \tag{5.38}$$

with $b = Z\boldsymbol{E}\mu$ and Z is the species charge (-1 for electrons, +1 for positive ions). Subscripts have the meaning explained in Figure 5.5. A straightforward finite-volume discretization of equation 5.38 leads to numerical instabilities whenever the convective component of the flux is greater than the diffusive component [166, 168]. In other words, whenever $b > 2D/\Delta x$, the method is unstable. This condition is easily achieved at cells next to the cathode sheath where a large electric field is found.

One powerful unconditionally stable method to discretize these types of equations is the exponential scheme first described by Scharfetter and Gummel [332]. The scheme, known as the Scharfetter-Gummel model, has been extensively used in the semiconductor field to simulate the transport mechanism of holes and electrons in different semiconductor devices. The model has been adapted to the plasma fluid model in several articles [245, 248].

Using the Scharfetter-Gummel scheme, the discretization of the continuity equation leads to:

$$
\begin{aligned}
J_e &= \frac{b_e \left(n_E - n_C e\sigma_e \right)}{1 - e\sigma_e} \quad \sigma_e = \frac{b_e \Delta_x}{D_e} \\
J_w &= \frac{b_w \left(n_C - n_W e\sigma_w \right)}{1 - e\sigma_w} \quad \sigma_w = \frac{b_w \Delta_x}{D_w}
\end{aligned}
\tag{5.39}
$$

where $J_{e,w}$ is the particle flux at the cell face (e=east, w=west) according to Figure 5.5, and σ is a measure of the ratio between convective and diffusive transport.

Substituting equation 5.39 into equation 5.38 yields

$$\frac{n_C - n_C^k}{\Delta t} + n_E \left(\frac{b_e}{\Delta_x(1 - e\sigma_e)} \right) + n_W \left(\frac{b_w e\sigma_w}{\Delta_x(1 - e - \sigma_w)} \right)$$

$$- n_C \left(\frac{b_e e\sigma_e}{\Delta_x(1 - e\sigma_e)} + \frac{b_w}{\Delta_x(1 - e\sigma_w)} \right) = S_e^k \tag{5.40}$$

where indexes at $k + 1$ have been omitted for simplicity.

After some algebraic manipulation, equation 5.40 can be cast into the general expressions:

$$\alpha_C n_c + \alpha_E n_E + \alpha_W n_W = S_n \Delta t + n_i^k \tag{5.41}$$

and

$$\beta_C p_c + \beta_E p_E + \beta_W p_W = S_p \Delta t + p_i^k \tag{5.42}$$

with

$$\alpha_E = a_n \left(\frac{(\phi_E - \phi_C)}{1 - em_n(\phi_E - \phi_C)} \right) \qquad \alpha_W = a_n \left(\frac{(\phi_C - \phi_W)em_n(\phi_C - \phi_W)}{1 - em_n(\phi_C - \phi_W)} \right)$$

$$\beta_E = a_p \left(\frac{(\phi_C - \phi_E)}{1 - em_p(\phi_C - \phi_E)} \right) \qquad \beta_W = a_p \left(\frac{(\phi_W - \phi_C)em_p(\phi_W - \phi_C)}{1 - em_p(\phi_W - \phi_C)} \right)$$

and:

$$\alpha_C = 1 - a_n \left(\frac{(\phi_E - \phi_C)em_n(\phi_E - \phi_C)}{1 - em_n(\phi_E - \phi_C)} \right) - a_n \left(\frac{(\phi_C - \phi_W)}{1 - em_n(\phi_C - \phi_W)} \right)$$

$$\beta_C = 1 - a_p \left(\frac{(\phi_C - \phi_E)em_p(\phi_C - \phi_E)}{1 - em_p(\phi_C - \phi_E)} \right) - a_p \left(\frac{(\phi_W - \phi_C)}{1 - em_p(\phi_W - \phi_C)} \right)$$

where $a_n = \dfrac{\mu_n \Delta t}{\Delta x^2}$, $m_n = \dfrac{\mu_n}{D_n}$, $a_p = \dfrac{\mu_p \Delta t}{\Delta x^2}$ and $m_p = \dfrac{\mu_p}{D_p}$. In this derivation, the term $-\nabla\phi$ has been used instead of $\boldsymbol{E}$ to facilitate the Jacobian derivation.

The boundary conditions for this system of equations are

At A:

- $\phi_A = -V_0$

- $n_A = 0$

- $\dfrac{\partial p}{\partial x} = 0$

at B:

- $\phi_B = 0$

- $n_B = 0$

- $p_B = 0$

To close the system, we need to specify the initial conditions. Since the goal is to build a general method of solution, the potential field and the charge particle densities should be initialized using a global scheme. Thus, the initial densities are set to a constant low value $n_0 = 10^{12}$ and the initial potential field ϕ^0 is the solution to the equation

$$\nabla^2 \phi^0 = 0 \tag{5.43}$$

with the boundary conditions specified before.

With the discretized equations, boundary conditions and initial values already established, the next step is to numerically solve the nonlinear system of equations.

We define the nonlinear system to solve as

$$F(\phi, n, p) = \begin{bmatrix} F_\phi(\phi, n, p) \\ F_n(\phi, n, p) \\ F_p(\phi, n, p) \end{bmatrix} = 0; \tag{5.44}$$

$$\tag{5.45}$$

where

$$F_\phi = \gamma_C \phi_C - \gamma_E \phi_E - \gamma_W \phi_W - S_\phi$$
$$F_{n_e} = \alpha_C n_C - \alpha_E n_E - \alpha_W n_W - S_i$$
$$F_{n_e} = \beta_C n_C - \beta_E n_E - \beta_W n_W - S_e$$

For simplicity, I have dropped the previous notation for vectors ($\boldsymbol{u}$) and matrices ($\overline{\boldsymbol{A}}$) for a more convenient one. In the remaining of the chapter, ϕ, n_e, and p are the vectors containing the variables associated with every grid point. ϕ_i is the potential associated with the ith cell. Uppercase subscripts refer only to locations on the center of the cells (E=East,W=West) while lowercase subscripts refer to locations at the face of the cell (see Figure 5.5).

Next, I briefly describe the numerical approaches to solve the discretized plasma fluid equations.

5.6.1   Fully explicit solution

The simplest approach to deal with this particular set of equations is to solve them sequentially using a segregated solver. The procedure to solve the plasma equations is simple: first, the Poisson equation is explicitly solved using the initial conditions as starting points. Then, for a given Δt, we proceed to solve the electron and ion continuity equations keeping the electric field constant using the newly calculated potential values from the Poisson solution. With the new electron and ion densities, the Poisson equation is solved again and the process is repeated until convergence. Finally, a new Δt is chosen and the procedure is repeated until a steady state solution is found. This strategy is commonly referred as the method of successive substitutions [183] or Gummel iterations [333]. This method is stable if the variables are loosely coupled. An example of such a model is the model presented in TS2. The convergence rate is determined by the coupling between variables.

Explicit methods are easy to implement and do not require expensive matrix computations nor large amounts of computer memory. There are, however, two restrictions for the size of Δt that can be taken in the algorithm. These restrictions severely limit the application of fully-explicit methods in plasma modeling. The first limitation is the famous Courant-Friedrichs-Levy (CFL) condition [167] which is typical of time-dependent explicit methods. However, the most restrictive condition is linked to the strong coupling between the electric field and electrons. Due to this coupling, we need to choose a Δt such that

$$\Delta t < \tau_M \tag{5.46}$$

where $\tau_M = \dfrac{\epsilon_0}{\sigma}$ is the Maxwell relaxation time [248] and $\sigma = q(\mu_e n_e + \mu_p p)$ is the plasma conductivity. τ_M is a measure of the relaxation time of the electric field as a result of a change in the electron density. Since $\mu_e >> \mu_p$, the Maxwell relaxation time can be approximated as

$$\tau_M = \frac{\epsilon_0}{q\mu_e n_e} \tag{5.47}$$

For a typical magnetron plasma with $\mu_e = 1000 \text{ m}^2\text{Vs}^{-1}$ and $n_e = 1e17 \ m^{-3}$ we obtain $\tau_M \approx 5 \times 10^{-13} \ s$ which is an extremely small time step to take in a realistic simulation.

Despite the limitation, this numerical technique can be used to solve non transient problems where the coupling between variables is weak. This numerical method was used in TS2 (section 4.3.2) to solve for mass, energy and momentum transport in its original implementation. The numerical method consisted of Gauss-Seidel iterations and underrelaxation to solve for the coupling between the momentum and continuity equation. The model in TS2 is not time dependent, and the CFL condition does not apply. Gauss-Seidel iterations, and stationary methods in general, are rarely used nowadays in large scale numerical computing for their extremely slow rate of convergence.

5.6.2 Semi-explicit solution

The rate of convergence can be improved by using a semi-explicit approach where each equation is solved implicitly using a projection method such as the CG, BiConjugate Gradient (BiCG) or GMRES. The overall algorithm is similar to the explicit model described before. However, it converges faster since fast projection methods are used to solve each equation instead of using slow point iterative techniques. Unfortunately, little is gained in terms of stability or total convergence time. The CFL and the Maxwell relaxation time restrictions are still valid and a small Δt needs to be used even if a relaxation parameter is used to iteratively update the variables.

5.6.3 Semi-implicit solution

To increase Δt beyond the Maxwell relaxation time, Ventzek et al. [298] proposed a method where an estimate of the electron change at the next time level is used in the solution of the Poisson equation. The idea behind this method is to use a estimate of the charge density change in the Poisson equation. In practice, only a guess of the electron density is needed since ions are relatively heavier and have a slower response than electrons to electric field fluctuations. An example of a semi-implicit approach is the following equation

$$\nabla \cdot ((\epsilon_0 + e\Delta t \mu_n n)\nabla \phi - e\Delta t D_n \nabla n) = e(n - p) \qquad (5.48)$$

where the Poisson equation has been modified to account for an estimation of electron density change at time $t + \Delta t$. This method, or slight variants of it, has been proved successful for one and two dimensional gas discharges [233, 248, 298], allowing the use of $\Delta t \approx 10 - 100\tau_M$ without a significant effect on stability [248].

Nevertheless, it is not clear how this algorithm would perform in a three dimensional mesh. Additionally, the presence of a magnetic field complicates the calculation to estimate the electron change in Equation 5.48.

5.6.4   Fully-implicit

In the fully implicit approach, ϕ, n_e and p are solved simultaneously at every Δt. The main advantage of a fully implicit approach is that we completely avoid the CFL restriction and Δt can be significantly larger than τ_M. However, implicit methods increase the complexity of the algorithm since we require matrix computations.

The standard numerical method to solve a nonlinear equation is the Newton-Raphson [183]. The Newton-Raphson is an iterative method to solve a nonlinear equation. The procedure to solve a nonlinear equations is as follows. Suppose we want to solve $f(x) = 0$, then the Newton-Rhapson method starts with an initial guess x_0 and recursively solves

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} \tag{5.49}$$

where $f'(x)$ is the derivative of $f(x)$ and the superscript k refers to the kth Newton iteration. The iteration is stopped when $x^{k+1} - x^k = \epsilon$ where ϵ is a small number.

The extension to $\Re^N$ is straightforward

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k - \frac{f(\boldsymbol{x}^k)}{J(\boldsymbol{x}^k)} \tag{5.50}$$

where J is the derivative of $f(\boldsymbol{x})$ known as the Jacobian matrix. Therefore, the Newton-Raphson method applied to the plasma model yields

$$\begin{bmatrix} \dfrac{\partial F_\phi}{\partial \phi} & \dfrac{\partial F_\phi}{\partial n_e} & \dfrac{\partial F_\phi}{\partial p} \\ \dfrac{\partial F_{n_e}}{\partial \phi} & \dfrac{\partial F_n}{\partial n_e} & \dfrac{\partial F_{n_e}}{\partial p} \\ \dfrac{\partial F_p}{\partial \phi} & \dfrac{\partial F_p}{\partial n_e} & \dfrac{\partial F_p}{\partial p} \end{bmatrix}^k \begin{bmatrix} \delta\phi^{k+1} \\ \delta n^{k+1} \\ \delta p^{k+1} \end{bmatrix} = - \begin{bmatrix} F_\phi \\ F_n \\ F_p \end{bmatrix}^k \tag{5.51}$$

with

$$\phi^{k+1} = \phi^k + \delta\phi^{k+1}$$
$$n^{k+1} = n^k + \delta n^{k+1}$$
$$p^{k+1} = p^k + \delta p^{k+1}$$

which is iteratively solved until some convergence criteria is met. All matrices are assumed rectangular and positive definite, $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ where N is the rank of the element.

Equation 5.51 in a compacted general form reads

$$J\delta x = -f(x) \tag{5.52}$$

where $x = \{\phi, n, p\}$ is the vector formed by all cell variables, J is the Jacobian matrix and $f = (F_\phi, F_n, F_p)$ as defined before.

The solution of Equation 5.51 for a large N is computational intensive. One of the main limitations is that the rank of $\boldsymbol{J}$ is $3N$ resulting in a large matrix requiring a huge amount of memory for the solution of the linear system using projection methods. Point iterative methods are impractical due to their extremely slow convergence rate which will be even worse for such a large linear system. To overcome this limitation, and to speed up our solution, we use a classical block iteration for the solution of the kth Newton step [333]. The basic idea of the Block-Newton method is to implicitly decouple the equations by solving

$$
\begin{bmatrix}
\dfrac{\partial F_\phi}{\partial \phi} & 0 & 0 \\[2mm]
\dfrac{\partial F_n}{\partial \phi} & \dfrac{\partial F_n}{\partial n} & 0 \\[2mm]
\dfrac{\partial F_p}{\partial \phi} & \dfrac{\partial F_p}{\partial n} & \dfrac{\partial F_p}{\partial p}
\end{bmatrix}^k
\begin{bmatrix}
\delta\phi \\ \delta n \\ \delta p
\end{bmatrix}^{m+1}
= -
\begin{bmatrix}
F_\phi \\ F_n \\ F_p
\end{bmatrix}
-
\begin{bmatrix}
0 & \dfrac{\partial F_\phi}{\partial n} & \dfrac{\partial F_\phi}{\partial p} \\[2mm]
0 & 0 & \dfrac{\partial F_n}{\partial p} \\[2mm]
0 & 0 & 0
\end{bmatrix}^k
\begin{bmatrix}
\delta\phi \\ \delta n \\ \delta p
\end{bmatrix}^{m}
\tag{5.53}
$$

Since $\boldsymbol{J}$ is nonsingular, we can decouple the system into three linear equations to be solved sequentially

$$\frac{\partial F_\phi^k}{\partial \phi} \cdot \delta\phi^{km+1} = -F_\phi^k - \frac{\partial F_\phi^k}{\partial n}\dot{\delta n}^{km} - \frac{\partial F_\phi}{\partial p} \cdot \delta p^{km} \tag{5.54}$$

$$\frac{\partial F_n^k}{\partial n} \cdot \delta n^{km+1} = -F_n^k - \frac{\partial F_n^k}{\partial \phi}\dot{\delta n}^{km} - \frac{\partial F_n}{\partial p} \cdot \delta p^{km} \tag{5.55}$$

$$\frac{\partial F_p^k}{\partial p} \cdot \delta p^{km+1} = -F_p^k - \frac{\partial F_p^k}{\partial \phi}\dot{\delta n}^{km} - \frac{\partial F_p}{\partial n} \cdot \delta p^{km} \tag{5.56}$$

$$\tag{5.57}$$

The appeal of this block method is that now every equation is solved sequentially. By doing this, however, we increase the number of the iterations needed for convergence. Fortunately, it is faster to form and solve each equation separately when the number of mesh elements is large.

To put the previous equation in a more compact form, we need to substitute the series expansions of the right hand side elements. For instance, the series expansion of the right hand side of Equation 5.54 is

$$f(\phi_0, n_0 + \delta n, p_0 + \delta p) = f(\phi_0, n_0, p_0) + \delta n_0 \frac{\partial f(\phi_0, n_0, p_0)}{\partial n}$$
$$+ \delta p_0 \frac{\partial f(\phi_0, n_0, p_0)}{\partial p} + O(\delta n^2, \delta p^2) \tag{5.58}$$

If we substitute Equation 5.58 into Equation 5.54, and introduce a relaxation parameter ω, we obtain (for the three equations)

$$\frac{\partial F_\phi^k}{\partial \phi} \cdot \delta\phi^{km+1} = -\omega \cdot F_\phi\left(\phi^k, n^k + \delta n^{km}, p^k + \delta p^{km}\right) \tag{5.59}$$

$$\frac{\partial F_n^k}{\partial n} \cdot \delta n^{km+1} = -\omega \cdot F_n\left(\phi^k + \delta\phi^{km+1}, n^k, p^k + \delta p^{km}\right) \tag{5.60}$$

$$\frac{\partial F_p^k}{\partial p} \cdot \delta p^{km+1} = -\omega \cdot F_p\left(\phi^k + \delta\phi^{km+1}, n^k + \delta n^{km+1}, p^k\right) \tag{5.61}$$

which is commonly known as the block SOR Newton-Raphson method (BSNR) [183, 333].

Finally, to avoid numerical instabilities, we are required to solve for the potential and electron implicitly. This is obtained if we solve for ϕ and n simultaneously as

$$\frac{\partial F_\phi^k}{\partial \phi} \cdot \delta\phi^{km+1} = -\omega \cdot F_\phi \left(\phi^k, n^k, p^k + \delta p^{km} \right) \tag{5.62a}$$

$$\frac{\partial F_n^k}{\partial n} \cdot \delta n^{km+1} = -\omega \cdot F_n \left(\phi^k, n^k, p^k + \delta p^{km} \right) \tag{5.62b}$$

$$\frac{\partial F_p^k}{\partial p} \cdot \delta p^{km+1} = -\omega \cdot F_p \left(\phi^k + \delta\phi^{km+1}, n^k + \delta n^{km+1}, p^k \right) \tag{5.62c}$$

This approach guarantees that changes in the electric field as a consequence of a change in electron density are handled implicitly. In matrix form, Eequation 5.62 reduces to

$$\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} F_x \\ F_p \end{bmatrix} \tag{5.63}$$

where

$$\tilde{x} = \{\phi, n\} \tag{5.64}$$

This is, in essence, the numerical method used in PlaSpud2 for the solution of the magnetized plasma equations in a three dimensional grid.

5.6.5 Condition number and scaling

The condition number K of a matrix helps us to determine the sensitivity of the system to small perturbations. The condition number of a nonsingular matrix A is defined as [334]

$$K(A) := ||A|| \cdot ||A^{-1}|| \tag{5.65}$$

where $|| \cdot ||$ is any matrix norm, normally taken as the ∞ norm.

Linear systems where $K(A)$ is relatively big are considered as ill-conditioned. Ill-conditioned systems are far more complicated to solve since they are prone to numerical errors. One simple method to reduce K is row scaling which is attained by properly scaling the matrix [335]. Another important reason to reduce the condition number is that the rate of convergence of projected methods directly depends on the condition number [336].

The first step to reduce $K(A)$ has been already been taken when the variables were scaled in Equation 5.36. This was a form of column scaling, and it is the easiest recommended way to improve the condition number of a matrix [337]. Unfortunately, this scaling is not optimal, since we have not considered the value of the time step, the magnitude of the transport coefficients and the length scale of the domain. To optimally choose a scaling algorithm is still not well understood, and no single scaling algorithm can be guaranteed to always perform satisfactorily [337].

There is also the problem of estimating the condition number numerically. A close look at Equation 5.65 can explain why: while an estimation of $||A||$ is easy to compute on a per cell basis, $||A^{-1}||$ is not. Any attempt to estimate $K(A)$ to measure the reliability of the numerical method would be time consuming and counterproductive, and we have to rely on simple techniques designed to bring $K(A)$ down.

A simple method to do row and column scaling is to multiply the equations by two matrices. A good method to reduce the system $Ax - b$ is to rewrite it as [337]:

$$\hat{A}\hat{x} = \hat{b} \tag{5.66}$$

where $\hat{A} = D_1 A D_2$, $\hat{x} = D_2^{-1}$ and $\hat{b} = D_1 b$. In the above expression, D_1 and

D_2 are chosen such that the diagonal elements of $\hat{A}$ are as close as possible to 1.

A simple algorithm that recursively applies row and column scaling is the following [335]

1. $\hat{A}^{(0)} = A$, $D_1^{(0)} = I$, $D_2^{(0)} = I$
   for k=0,1,2,..., until convergence

   (a) $D_R = diag\left(\sqrt{||A^{(k)}(:,j)||_\infty}\right)$ and $D_C := diag\left(\sqrt{||A^{(k)}(i,:)||_\infty}\right)$

   (b) $\hat{A}^{(k+1)} = D_R^{-1}\hat{A}^{(k)}D_C^{-1}$

   (c) $D_1^{(k+1)} = D_1^{(k)}D_R^{-1}$ and $D_2^{(k+1)} = D_2^{(k)}D_C^{-1}$

Convergence is obtained when

$$\max_{0<j<N}\{|(1-||A(:,j)||_\infty|\} \le \varepsilon \; and \; \max_{0<j<N}\{|(1-||A(:,j)||_\infty|\} \le \varepsilon \qquad (5.67)$$

for a small value of ε.

Usually, a couple of iterations are enough to bring down $K(A)$ significantly. Unfortunately, this algorithm will most likely perform poorly in our case since $N \approx 5 \times 10^5$ and the extra burden of obtaining D_2 at every iteration and then run the algorithm is computationally demanding. To reduce the amount of extra work while maintaining the goal of reducing the condition number of our matrix, I have opted to performed only row equilibration using

$$D_1 := diag(||A(:,j)||_\infty)^{-1} \; and \; D_2 := I \qquad (5.68)$$

$$(5.69)$$

and Equation 5.66. It is argued in [337] that row equilibration is an optimal scaling strategy for the ∞ norm.

There is also another reason for choosing this scaling method: ease of implementation. In fact, D_1 can easily be obtained when the Jacobian is created.

On the other hand, D_2 is more difficult to obtain since we need first to form J and then do a column sweep to find the maximum value of a given column. In the current implementation of the code, the matrix is assumed to be diagonal dominant, thus

$$D_1 = diag(A) \qquad (5.70)$$

5.7 Method of solution

The proposed method to solve for this complex numerical solution is summarized in the following algorithm. All linear systems mentioned in the algorithm are scaled using the method described in the previous section. All matrix computations and linear solvers are implemented using the third-party sparselib++[1].

1. Start simulation, set global iteration index $i = 0$

2. Set $t_i = \Delta t$

3. Start main time loop
   DO

   (a) Form J_i and b_i

   (b) Start Block Newton-Raphson loop, set Newton index $k = 0$, set block iteration index $m = 0$.
   DO

   i. Solve Eq. 5.62a and Eq. 5.62b using the Preconditioned GMRES method

   ii. Update right hand side of Eq. 5.62c

   iii. Solve Eq. 5.62c using the preconditioned GMRES method

---

[1] http://math.nist.gov/sparselib++/

iv. Update right hand side of Eq. 5.62a and Eq. 5.62b

v. m=m+1

WHILE($||\delta\tilde{x}^{(km+1)} - \delta\tilde{x}^{(km)}||_\infty < \epsilon$ **AND**
$||\delta p^{(km+1)} - \delta p^{(km)}||_\infty < \epsilon$**)**

(c) Update variables $\tilde{x}^{k+1} = \tilde{x}^k + \delta\tilde{x}^{km+1}$ and $p^{k+1} = p^k + \delta p^{km+1}$

(d) $k = k + 1$

(e) $m = 0$

WHILE($||\delta x_k|| > \epsilon$**)**

4. $i = i + 1$ and set $t_i = t + \Delta t$

5. Check Global Convergence $(x^i - x^{i-1} < \epsilon)$. No?, return to 3.

6. End

In the previous algorithm, ϵ is a small number ($\approx 10^{-5}$). To speed convergence, Δt is adaptively increased every 5 iterations. If the new Δt introduces instabilities in the numerical algorithm, the algorithm returns to the previous densities and potential field, reduces Δt by a small factor and repeats the procedure just described until a stable Δt is found.

5.8   Results of test model

To benchmark the proposed algorithm, I have compared the one dimensional model presented in the previous section with the results of Bogaerts et al. [237] where the authors reported on the creation rates used for the solution of their fluid model. The creation rates are directly obtained from a Monte Carlo simulation of energetic electrons (see Figure 2 in [237]). An accurate description of all creation rates is key for a fair comparison between two computational models. The plasma equations are then solved using the same electrode separation and boundary conditions reported in Ref. 237.

The model presented in this chapter can be used to solve for the type of glow discharges typical of sputtering processes. The numerical algorithm presented in the previous section is applicable to a wide range of pressures and configurations only if the physics of the problem are properly captured by the transport coefficients and creation rates. For instance, while the model can be converging to a solution, if the creation rates and transport coefficients are not properly calculated, the results may not help us to relate the model to a real physical problem. Thus, the goal of this section is two fold. First, to demonstrate that the numerical algorithm proposed in this chapter to deal with a more complex glow discharge model is stable by benchmarking it using a simple one dimensional problem, and secondly, to show that the model converges to the same results when compared to a published model. The feasibility of the model for a magnetized discharge at low pressures will be covered in the next chamber.

In short, the discharge is created between two infinitely large electrodes separated by 1.5 cm. The discharge pressure is set to 100 Pa, and the electrode potentials are -1000 and 0 V respectively. The only difference between this chapter's model and that of Bogaerts et al. is the expression to calculate the reduced mobility. In Ref. 237, the authors used the Frost formula whereas the model presented in this chapter uses the Phelps equation as described in section 5.5. The simple one dimensional model was written in C++ using the sparselib++ library for all matrix computations. The program is run on an intel i7 processor running a linux operative system (Ubuntu 11.10). The program converge in 5 minutes using the Block-Newton-Rhapson method. However, the convergence time can be greatly reduced by using a simple two-stage algorithm. In the first stage of the algorithm, equations are solved using the BSNR for only a few iterations to generate a good initial guess for the subsequent stage. In the second stage, the previous solution is used as an initial condition for a nontransient version of the plasma equations (pseudo-time dependence is removed). The solver for the second stage is a fully implicit Newton-Raphson iterative scheme. Finally, and to increase the radius of convergence, a simple damped update [338] scheme with backtracking [339] is employed to increase the robustness of the algorithm. The premise of damped
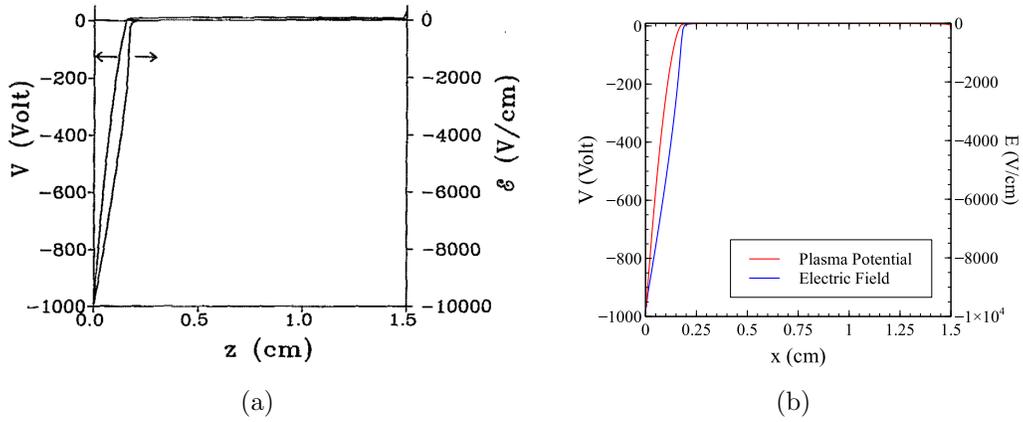
Figure 5.6: Electric field and plasma potential comparison between my simulation (a) and data from Ref. 237 (b) for a simple one dimensional discharge model

Newton-Methods [329, 338] is to generate the update

$$x^{k+1} = x^k + \xi \delta x \tag{5.71}$$

using a damping factor ξ. The correct choice of ξ is problem dependent and it is generally not easy to obtain a global sequence. For this reason, the algorithm is globalized by introducing a backtracking line search methods [340]. The backtracking algorithm returns to find a suitable ξ in case the previous value does not reduce the norm of $f(x)$. Norm reducing is a desirable characteristic of a globally convergent Newton-Rhapson algortihm [299]. For instance, if $||f(x^{k+1})|| > ||f(x^k)||$ for any given damped parameter d, the algorithm needs to backtrack and find an appropriate d for which the Newton update satisfies $||f(x^{k+1})|| < ||f(x^k)||$.

Figures 5.6 and 5.7 show the results of the simulations using the numerical algorithm presented in the previous section. In general, good agreement between the results of this section's model and that of Bogaerts et al. is observed in the figure. In Figure 5.6, both the electric field and potential distributions are practically identical.
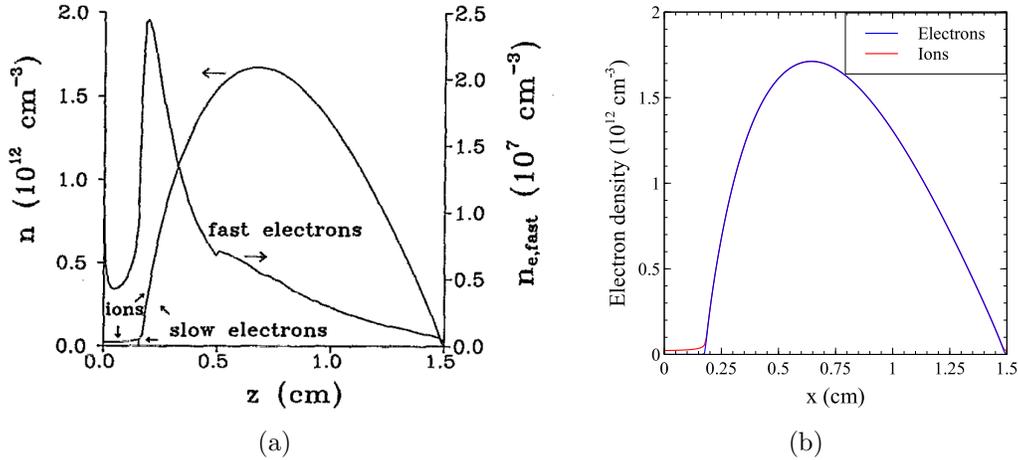
Figure 5.7: **Electron and ion density comparison between my simulation (a) and data from Ref. 237 (a) for a simple one dimensional discharge model.**

Figure 5.7 compares the densities calculated by both models. We can also observe excellent agreement between both models. It is clear from the figure that the differences introduced by using two mobility models are not enough to substantially modify the solution of the fluid plasma model. The agreement between both models confirms that the mathematical model presented in this chapter converges to the same solutions reported in Ref. 237.

Finally, the mathematical model can also be tested by running a simple numerical experiment where only one parameter is modified. If the family of solutions is contrary to the physics of the problem, the conclusion is that there is a failure in either the model of the handling of boundary conditions. Unfortunately, the contrary can not be asserted about the model in case results do agree.

One numerical experiment to test the model involves running several simulations at different pressures to observe the effect on the plasma sheath. Any solution where the plasma sheath is not reduced as pressure is increased would be indicative of a failure in the model. The simulation was performed for five

different pressures, and the potential distribution next to the cathode for each
simulation is shown in Figure 5.8. For this specific test, the electrode sepa-
ration was set to 6 cm with -290 and 0 volts as boundary conditions for the
Poisson equation. The source term where constant for all five simulations. In
reality, if we create a glow discharge in the abnormal regime at constant volt-
age, a change of pressure would result in a substantially different distribution
of sources within the electrodes. However, I have used the same source terms
only for the sake of argument.

As expected, the plasma sheath shrinks as pressure is increased as is shown in
Figure 5.8. The reduction of the plasma sheath is due to several reasons, one
of them being lower transport coefficients. The lower electron and ion diffusiv-
ity reduces diffusive losses in the discharge which increases the plasma density
between the electrodes. The higher plasma density pushes more charged parti-
cles towards the plasma sheath increasing the current collected at the cathode
J_i, since $J_i = qp_0u_0$ where p_0 is the ion density at the sheath edge, and u_0 is
the Bohm velocity defined as the velocity of ion species when they enter the
sheath region. To estimate the ion current density at the cathode, the Poisson
equation can be solved assuming that the electron and ion density are equal at
the edge of the plasma sheath. The solution to this simplified model is known
as the Child-Langmuir Law which estates [76]

$$J_i = \frac{4\epsilon_0}{9}\sqrt{2q/m_e}\frac{A|V_a|^{3/2}}{d_s^2} \tag{5.72}$$

where J is the current density, V_a is the cathode potential, A is the contact
area, and q, m_e and d_s are as previously described. It is clear in Equation
5.72 that the only viable solution for a higher current density when the ion
density is increased and V_a is held constant is to decrease d_s. Figure 5.72
indicates that the plasma model and numerical algorithm captures the physics
of a real glow discharge. The accuracy of the model will be tested in the next
chapter when the results of the model are compared against the experimental
electron densities obtained using a Langmuir probe in a magnetron sputtering
discharge.
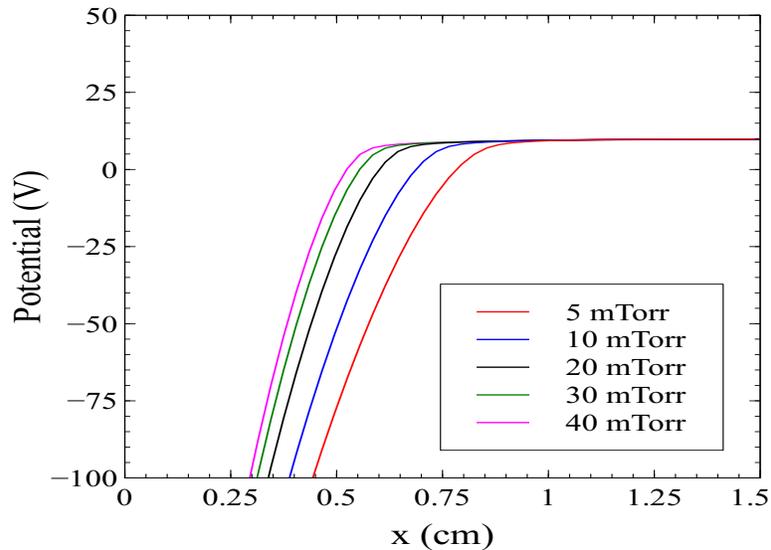
Figure 5.8: **Plasma sheath as a function of pressure. The model correctly predicts a shorter plasma sheath as presure increases.**

5.9 Summary and conclusions

In this Chapter, a hybrid numerical model for the solution of the transport of charged particles in a magnetron plasma has been presented. We started the chapter with a literature survey of mathematical and numerical models of the type of plasmas found in magnetron discharges. Next, the hybrid algorithm was presented. The method combines a Monte Carlo algorithm for the non-equilibrium component of the particle flux with a fluid model for the cold species. A model that handles the nonequilibrium component of the flux is necessary for the correct description of the magnetron discharge. The solution of the Monte Carlo algorithm was used as an input to the fluid model, providing the necessary creation rates for the latter's solution. I showed a simple derivation of the plasma fluid model using a simplified analysis and then incorporate the effect of the magnetic field in the transport equations. The obtained fluid model was globalized by the addition of a pseudo-time variable to

increase the robustness of the model. The equations thus derived were solved using a two-stage algorithm where a Block-Newton-Raphson iteration is used in the first stage to increase the stability of the numerical model. The second stage removed the pseudo-time dependence and directly solved the non linear system using a damped Newton-Raphson iterative with a backtracking line search scheme. Finally, a simple one dimensional model was build and benchmarked to test both the mathematical model and the numerical algorithm. Excellent agreement was found between results from the model presented in this chapter and those of Ref. 237.

CHAPTER 6

Integration of modules: Results and discussion

6.1 Introduction

The goal of this chapter is to describe the integration of the modules presented in the previous pages of this dissertation and to present some results from simulations using the SpudII framework as a whole. First, I describe some of the key challenges faced during the integration of the framework. I have limited the description of the module blending stage to only those aspects that are relevant to this dissertation; I have omitted in the description many other aspects of the integration which deal with parts of the framework that are less important in the context of this chapter. Following this section, I present results from simulations performed using the SpudII project as working as an integrated suite of simulators.

The last part of this chapter deals with the construction and implementation of an electric probe used for the characterization of sputtering discharges. This probe was built to obtain a two dimensional map of some of the plasma parameters of interest in the region close to the target where the plasma is denser. Results from this survey are then compared against simulation results obtained from the implementation of the computational model of the magnetron discharge presented in this dissertation. The observed discrepancies

between the model and experiment are discussed at the end of this chapter.

6.2 Integration of modules

Given the computational complexity of the SpudII framework, the final integration of all modules into a single, coherent suite of simulators was a difficult task. There are many aspects of the integration that warrant a detailed description. However, any attempt to fully detail all components that were modified during the course of the integration stage would require several pages. Therefore, I have limited the description to cover only the PS1-PS2 and the PS-TS integration. These two aspects were key to properly integrate all modules in the SpudII framework.

The self-consistent algorithm for the magnetron discharge proposed in Figure 3.3 suggests two methods of solution. In the first, the simulation is started assuming a constant voltage at the target, letting the algorithm converge to a certain current. This current would be the result of running the framework with the input conditions set in the xml input file. Alternatively, in the second method, we can hold the target current at a fixed value and iterate the modules until a voltage is found that it is in equilibrium with the input current. The method I propose in this chapter uses the constant voltage approximation. The main reason for selecting this method stems from the incomplete treatment of ions in the PS1 algorithm. This component of the PS1 algorithm is still not fully implemented and is outside of the scope of this thesis. All simulations presented in this chapter are performed assuming a constant voltage mode of operation.

6.2.1 PS1-PS2 integration

An important aspect of the PS1-PS2 integration is the treatment of the cathode sheath. In the following, I will briefly describe the cathode sheath in the context of the PS module.

The plasma sheath is a natural consequence of confining a glow discharge in a vacuum chamber [76]. As explained at the end of the previous chapter, the plasma sheath is the region of the plasma that is not quasineutral, meaning that $n \neq p$ in this region with high electric fields. In the case of typical (electropositive) discharges, the cathode sheath has a net positive charge density as discussed at the end of Chapter 5. In sputtering modeling, the cathode sheath plays a decisive role in determining most of the events occurring at the target. The physics behind plasma sheaths are beyond the scope of this thesis, and the reader is refereed to the thorough descriptions found in [76, Chapter 8] and [75, Chapter 7]. In PS2, the cathode sheath is arbitrarily defined as the location next to the target where the plasma potential is 0, as shown in Figure 6.1. Although a presheath region exists in dc sheaths, PS makes no distinction between the presheath and the bulk of the plasma. This is one of the advantages of the treating the plasma with a global model where the same numerical algorithm is used to solve for the entire region where the plasma exists. The main function of the plasma sheath in the PS module is to be the source of ions for the PS1 algorithm. These ions diffuse from the plasma bulk to the sheath edge following the process described in section 5.8.

In the computational model of the PS2 module, the sheath has its own class, and its interface can be accessed from either the PS1, PS2 or TS2 modules. The reason for the TS2 module to access the plasma sheath is briefly explained in the next section. Every cathode in the simulation is initialized with an instance of the sheath class. Two of the most important methods of the sheath class are GetIon() and GetScaleFactor(). The latter method simply returns SF_{PS1} from the PS1 module calculated from Equation 5.24. The former, on the other hand, needs to be appropriately described.

Every instance of the PS1 module needs a source of ions to start the simulation which are provide by the sheath object. These ions reach the sheath edge by diffusion and have a particular spatial distribution which is a function of the local ion density [194]. Given that, at the beginning of the simulation, the ion density is unknown, we need to determine how this ion distribution

should be initialized. This initial current density at the plasma sheath can be calculated either with a global model or by providing an estimate of the erosion profile of a real target. In the latter method, the erosion profile can be used as a good guess since it correlates with the ion flux at the target [194, 287]. Ultimately, it is this ion flux that erodes the target. This method is expected to speed up the convergence of the global algorithm. However, for cases where no initial erosion is available, the sheath needs to be initialized using a global technique. For this reason, the current density at the sheath edge is approximated using an empirical expression based on the local value of the magnetic field. Several options were investigated but the final expression was obtained after noting that an equation based on the diagonal elements of the Hal parameter (introduced later in the chapter) were closely related to the erosion track. Thus, the empirical expression to approximate the current density at the sheath edge is:

$$j_s = 1 - \frac{1 + B_z^2}{1 + |\boldsymbol{B}|^2} \tag{6.1}$$

where j_s is the estimated current density at the target. Both, the sheath approximation model proposed in this dissertation and the experimental erosion profile are shown in Figure 6.2. The approximation using expression 6.2.1 is the recommended initial condition when an accurate erosion profile is not available. In the simulations presented in this chapter, I have initialized the sheath with an experimental erosion obtained from a used target erosion in an attempt to reduce the execution time of the simulations presented in this chapter.

The nonlinear coupling between hot electron dynamics (PS1) and the potential field (PS2) requires extra care. In the PS algorithm, the coupling between both modules (PS1 and PS2) is handled as follows. First, a plasma potential is assumed to obtain an initial electric field. The initial configuration is a simple model where d_s is assigned an initial value. The bulk of the plasma is assumed
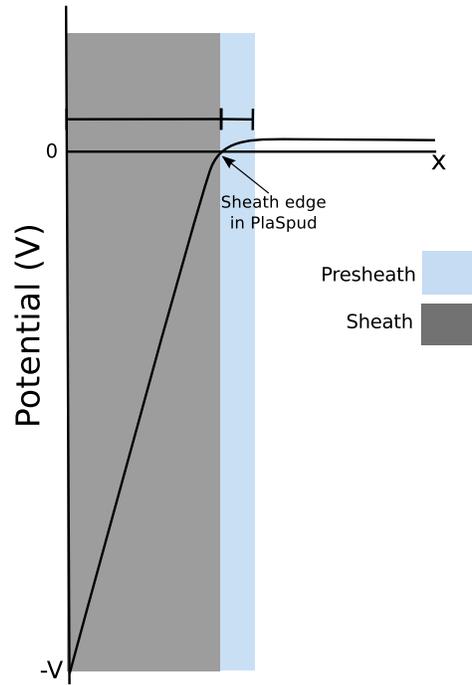
Figure 6.1: **Cathode sheath as defined in the PS2 grid. The sheath edge, the place where the plasma potential next to the target is equal to zero, is the location in the grid where ions are created to be simulated in the PS1 algorithm. The x axis is the perpendicular distance from the target.**

to be electric field free ($\phi = constant$) while the potential in the plasma sheath varies linearly between the sheath edge and electrode potential. Such a potential distribution results in a constant electric field inside the plasma sheath. This solution is not consistent with the Poisson equation which predicts a non constant electric field in the plasma sheath, but this model was found to be a good approximation for the initial stage of PS1 [2]. The initial current density at the plasma sheath is calculated using any of the methods described before. The initial electric field is used to integrate the electron trajectories in PS1 to obtain a first estimate of the source terms for the PS2 algorithm. Once PS1 has finished, PS2 solves for a new potential to obtain an updated distribution which is then used to update the plasma sheath. The new sheath is used as
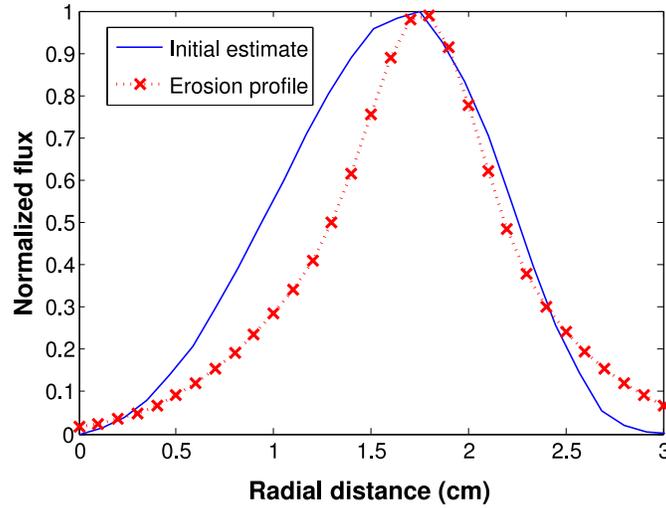
Figure 6.2: **Methods to determine the initial current density distribution in the plasma sheath in the PS module. The sheath flux can be initialized with an expression using the local value of the magnetic field or approximated with a measured erosion profile.**

a particle source for a new PS1 run which in turn modifies the source terms for PS2. This procedure is repeated during the whole PS iteration until the PS2 module converges. Convergence is defined when the potential in each cell changes no more than 1% between successive iterations.

The remaining aspects of the PS1-PS2 integration were already described when the PlasChem module and its place on the algorithm was described in the previous chapters. No further description of this module is needed.

6.2.2 PS-TS integration

Finally, the last stage to fully integrate all modules into a cohesive suite of simulators is the treatment of the PS-TS communication. The integration of the plasma and neutral modules represented another obstacle since both PS1 and PS2 constantly need the neutral density and gas temperature to

calculate most processes occurring in the algorithms described in chapter 5. However, most of the calculations pertaining either to the neutral density or the gas temperature are performed in PlasChem. For instance, all calculations during the simulation of an electron-gas collision are handled in PlasChem. Furthermore, the transport coefficients, which are also a function of the neutral density, are also calculated in PlasChem. For this reason, PS does not need to have an interface to TS: all communications requiring data from the TS grid are handled in PlasChem. TS, however, does need to extract data from the PS2 grid to solve for the argon metastable density.

The scheme that summarizes the PS-TS integration is depicted in Figure 6.3. The order of the iteration does not affect the final outcome as long as every module is run at least twice. There is, however, one aspect of the PS-TS iteration that needs to be addressed. Every time the TS1 module is run, it first clears all creation rates counters for the TS2 equations (equations 4.13). However, since the metastable creation rate was obtained from a different algorithm (PS1), it needs to be scaled with the sheath factor (SF_{PS1}) rather than the energetic particles' factor (SF_{TS1}). For this reason, when the TS1 module is executed at the beginning of the TS1-TS2 cycle, it does not clear the metastable creation rate.

There are many more aspects of the integration of the framework that are not discussed due to space constraints. The rest of the chapter is devoted to presenting results from the framework working as a single interoperable unit. However, there is only one aspect of the framework that has not been discussed and plays a key role in the simulation, that is an accurate three dimensional magnetic field.

6.2.3 Magnetic field configuration

The plasma model described in the previous chapter clearly depends on the shape and magnitude of the magnetic field. As is evident in Figure 3.3, the SpudII framework does not have a dedicated module to solve for the magnetic
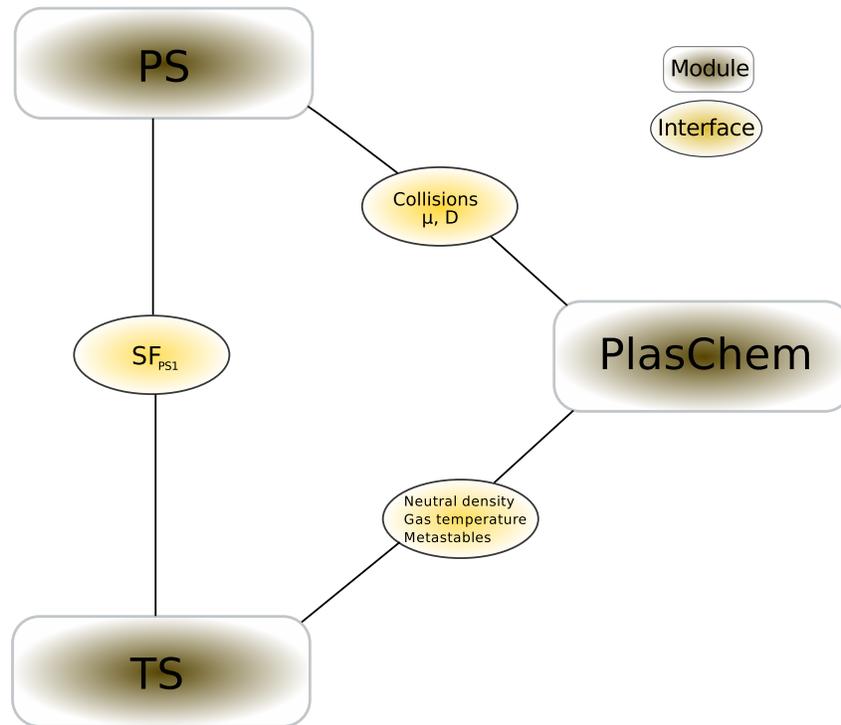
Figure 6.3: **Integration of the PS and TS modules. The TS modules need access to the plasma sheath to calculate the real creation rate to solve for the argon metastable population.**

field; therefore it depends on third party software to generate one. In this dissertation, the magnetic field was obtained using the COMSOL Multiphysics program[1]. To obtain a magnetic field as accurately as possible, the magnetic assembly and geometry of one the magnetron guns in the experimental chamber described in Chapter 4 was used as a guide to build a computational model. To calibrate the simulation, the magnetic field of the magnetron gun described in the previous chapter was measured using a LakeShore 410 Gaussmeter at several points. The simulated (after calibration) and experimental radial and axial magnetic fields are shown in Figure 6.4 while Figure 6.5 shows the magnetic field shape obtained from the simulation. Finally, an input file in the

---

[1]www.comsol.com

format used in SpudII was created from the simulation results to be used as an input to the simulator.

6.2.4 Iterative scheme

The iterative scheme for the final integration of the SpudII project is shown in Figure 6.6. In this scheme, the simulation starts with the neutral particles' transport module to obtain a neutral density and temperature grid in the chamber. The procedure described in Chapter 4 to solve for the energetic particles transport, rarefaction and gas heating is then followed. However, as mentioned before, the TS algorithm does not have at this point enough information from the plasma processes since PS has not been executed yet. Thus, the metastable argon density (in case metastables are included in the simulation) is solved using a zero creation rate everywhere in the chamber. With a new density and temperature grid, the algorithm flow goes to the PS module. Since no plasma species are created in the TS module, the PS module clears all cell counters in the PS2 grid at the beginning of the PS1-PS2 cycle. The initial conditions for this stage of the iteration are described in Section 6.4.2. The PS module solves for energetic and low energy particles' transport in the computational chamber following the procedure described in chapter 5. Once the PS2 module has finished, a new sheath is obtained using the new potential field, and the algorithm returns to the TS module to carry on with the simulation. At this point, the metastable creation rate is available from the the previous iteration of the PS module, and the TS module can solve for the metastable population. When the TS-PS cycle has converged, the program outputs all grids to output files and calculates all surface fluxes for visualization. The current implementation of the SpudII project has not yet been fully integrated with a feature scale simulator (GROFILMS in Figure 6.6).
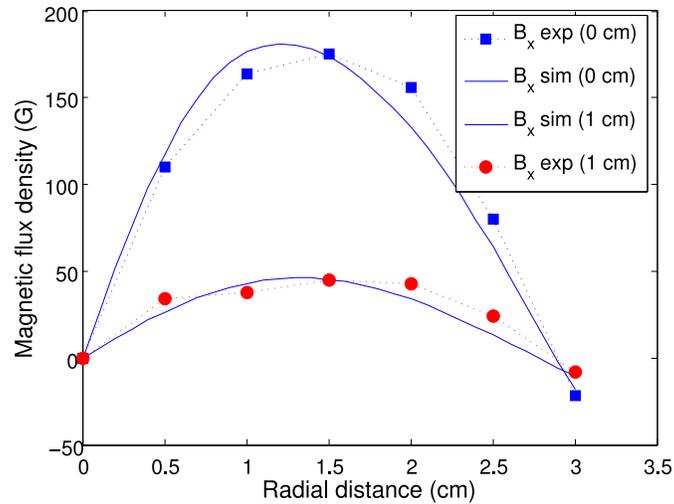
(a) B_x
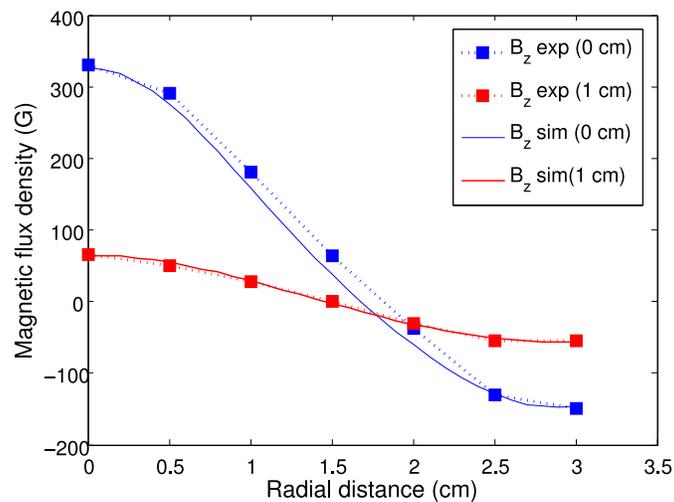
(b) B_z

Figure 6.4: **Simulated and measured magnetic fields in the region next to the target. The magnets' magnetization values were used as a fitting parameter in the COMSOL simulation to calibrate the model against experimental values. The magnetic field is an important input for an accurate description of the electron dynamics in the plasma model.**
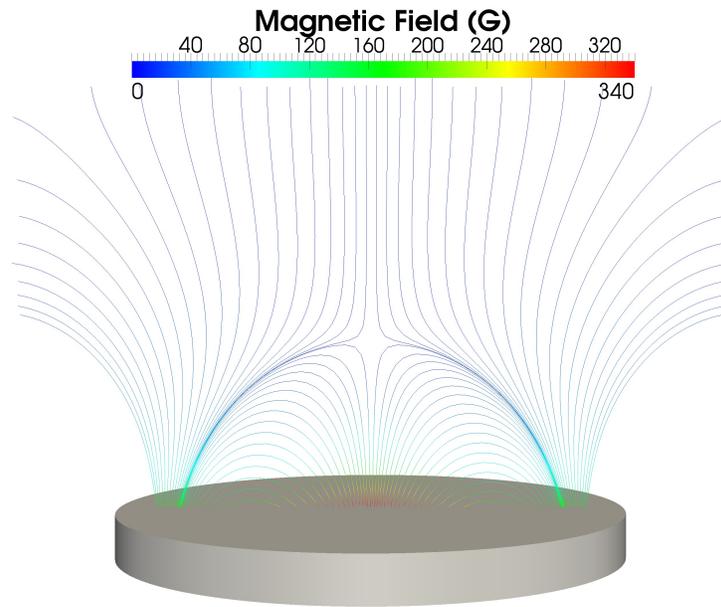
Figure 6.5: **Magnetic field shape and magnitude for all simulations presented in this chapter.**

Computational resources

The simulation results presented in this chapter were obtained using two different computer architectures. For all simulations where the PS2 module was not included in the algorithm, I used two linux clusters at the University of Alberta to perform the simulations. The first cluster is operated by the Academic Information and Communication Technologies (AICT) group and consists of 15 nodes with an Opteron 275 (6 GB RAM), 4 nodes with Opteron 275 (10 GB RAM) and 4 nodes with Opteron 280 (32 GB RAM). The second cluster is operated by the Integrated Nanosystem Research Facility (INRF). Each node consists of a dual Opteron 280 processors (4 cores @ 2.4 GHz, 8 GB RAM). The simulations involving the full iterative scheme were run on two computers each equipped with at least 12 GB of RAM running on an i7 Intel processor using linux as the operating system (Ubuntu 11.10). Each instance of the SpudII program in the full iteration mode requires at least 3 GB of
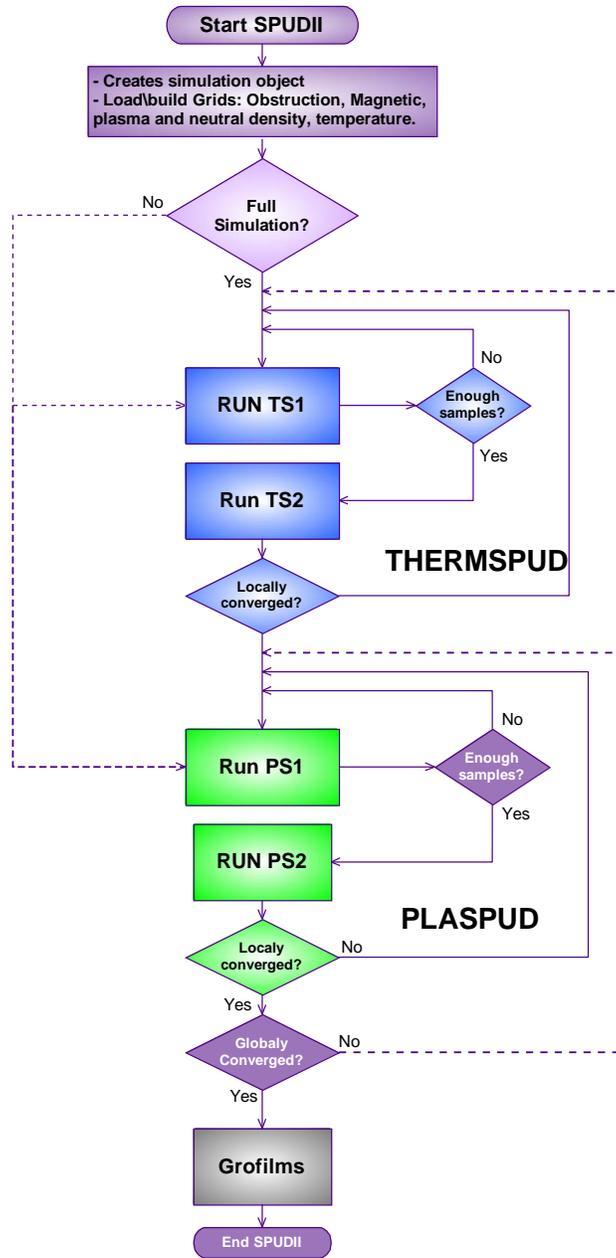
Figure 6.6: **Iterative scheme used for the integration of all modules in the SpudII framework. The model proposed in this chapter iterates through all modules until convergence.**

memory to allocate all grids and to perform all the numerical computations of the linear solvers. Typical execution times for each instance of the program range from 3-5 days when the PS2 module is not included in the simulation to about 10-15 days per simulation when the complete iterative scheme is run using the fast intel i7 architecture.

A list depicting the number of lines of code per modules/submodule of the SpudII project is shown in Table 6.1. The list only includes lines pertaining to C/C++ language.

Table 6.1: **Lines of code for each of the modules of the SpudII project obtained from the open source program cloc.**

Component	Lines of Code
PS1	3,500
PS2	10,000
TS1	6,200
TS2	6,500
PChem	2,700
TargSpud	500
ObstructionProcessor	26,500
SimulationObject	2,500
Utilities	26,000

6.3 Limitations

This section briefly discusses some of the limitations of the versions of the SpudII project reported in this dissertation.

First, the current version of the model is limited only to planar target geometries. There is, however, no technical limitation hard coded in the software that limits us from treating cylindrical targets (rotary magnetrons, post-magnetron). The components of the ObstructionProcessor module that maps input

files to simulation coordinates does not properly handle targets that are not planar surfaces and this part of the software requires a substantial modification to be able to simulate cylindrical targets.

-The limitations of the neutral module to simulate an specific target size are less limited than the ones referring to the plasma model. For instance, the neutrals module project has been tested in rectangular magnetrons as big as 15 in x 5 in. The main technical limitation associated with large magnetron configurations is the amount of available RAM for the process. For instance, to simulate a determined medium size system as the one just mentioned would require about 1×10^6 cells. Each cell may contain about 20 variables of type double. Given that each double is 64 bits(in a typical C++ compiler), the total amount of RAM required to run under this configurations is about 1.25 GB. This is just an approximations since we only considered one type of variable and only one grid. Thus, the limit of the neutrals module is dictated by the amount of RAM of the computer where the program will be run. The time required to perform such simulations will also be affected by the amount of cells in the simulation region but the simulation can always be optimized for an specific configuration.

The limitations of the plasma module are due to different factors. In the plasma module, we need to capture the smallest length scale which is typically observed at the cathode sheath. This limitations forces us to use small cells which increases the total number of equations to solve. This increase has a deep impact on the PS1 execution time. However, the execution time of the PS2 algorithm increases the most. The main reason is due to the fact that the size of the linear systems that are required to be solved in the PS2 algorithm is sensitive to the number of cells. A possible solution that can be explored to reduce the impact of this limitation is the use of mesh refinement. However, the numerical algorithm of the plasma should be modified to handle this type of grid architectures. This dissertation only deals with uniform grids and the applications to nonuniform grids.

The limits on the absolute maximum currents and voltages that can be treated

using the model presented in the previous chapter are less clear. In theory, the model should be capable of solving for any realistic combination of current and voltage. Voltages to up to 1 kV and currents up to 10 A are not a limitation of the mathematical model but this values need to be tested in the future to find a more realistic limit. However, as the current and voltage are increased, the discharge may change its operation region (from the abnormal mode of operation to the spark region). Unfortunately, the plasma model as it is does not incorporate the necessary physics to deal with this change of operation region and care should be taken when specifying the operation conditions in a simulation. As a general guideline, we should keep $V < 700$ V and $I < 1$ A to run a simulation.

6.4   Results and discussion

6.4.1   Simulation region

The simulation region for all results presented in this chapter used a cylindrical chamber (radius=5 cm, height=6 cm) similar to the simulation region from chapter 4. The smaller simulation region was required to reduce execution time and to capture the small length scale of the cathode sheath. The simulated magnetic field shown in Figure 6.5 was used as an input parameter for all simulations.

6.4.2   Initial conditions

The initial conditions greatly affect the convergence properties of the global algorithm. A poor initial estimate is normally reflected in a long execution time. Furthermore, a poor initial estimate of the plasma variables (ϕ, n and p) may cause the algorithm to diverge with catastrophic consequences. The initial conditions in the PS1 algorithm are the following. The initial sheath

current density is approximated with an experimental erosion profile using the method described before. The sheath is initialized with $d_s = 3$ mm in all cases. The initial electric field for the PS1 algorithm is initialized assuming a linearly varying potential between the cathode and the sheath edge. A constant potential is used otherwise for the plasma bulk. All simulations are run in the constant voltage mode of operation; thus the target potential for each pressure and power investigated is used as an input parameter in the xml file. During initialization, an initial current is also specified as a mean to initialize the sheath current density. The initial current density at the target is estimated using the method described in Section 6.2.1. After this initialization stage, the total current is obtained by integrating the current density at the target after each PS iteration. The I-V initial conditions for all runs presented in this chapter are summarized in Table 6.2.

Table 6.2: **I-V Initial conditions for all simulations reported in this chapter.**

Pressure (mTorr)	Current (A)	Voltage (V)	Power (W)
40	0.12	-258	30
40	0.204	-293	60
40	0.25	-300	75
20	0.11	-270	30
20	0.19	-315	60
20	0.23	-326	75
5	0.105	-280	30
5	0.167	-360	60
5	0.196	-382	75

The plasma processes included in all simulations reported in this chapter are summarized in table 6.3. The remaining parameters for the PS1 algorithm are: $E_{th} = 5.5eV$, $\gamma_e^{Al} = 0.095$, $\gamma_e^{Cu} = 0.085$, $SERC_{target} = 0.3$, $SERC_{walls} = 0$. The initial plasma density and potential for the PS2 algorithm are approximated using the same method described in the one dimensional model presented in the previous chapter ($n = p = 10^{12} \text{m} -3$ and $\phi = \phi_0$). The TS mod-

ule is initialed using the same parameters described in Chapter 4 for which an excellent agreement was found between experiment and simulations.

Table 6.3: **Plasma processes for the SpudII simulations presented in this chapter.**

Process	Type	E_{ine} (eV)
$e + Ar \rightarrow e + Ar$	Elastic collision (Ar)	0
$e + Ar \rightarrow 2e + Ar^+$	Argon ionization	15.76
$e + Ar \rightarrow e + Ar^*$	Argon excitation (total)	11.79
$e + Ar \rightarrow e + Ar^m$	Argon metastable	11.60
$e + Ar^m \rightarrow 2e + Ar^+$	Argon step ionization	4.14
$e + Al \rightarrow e + Al$	Elastic collision (Al)	0
$e + Al \rightarrow e + Al^+$	Aluminum ionization	5.99
$e + Cu \rightarrow e + Cu^+$	Copper ionization	7.73

6.4.3 Runs

Results presented in this chapter are divided into two groups. In the first group, the PS2 module was not included in the global algorithm to investigate the effect of the TS2 module on the PS1 results. For these simulations, I present results for aluminum and copper discharges. Simulations are performed at two power levels representative of low (75 W) and high (300 W) power. Only two pressures are reported: 5 and 40 mTorr.

In the second group of simulations, the global iterative scheme of Figure 6.6 was used to simulate the magnetron process comprehensively. Results are presented for three pressures representative of a low (5 mTorr), medium (20 mTorr) and high (40 mTorr) pressure regimes. This labeling in only applicable in the context of this chapter since the terms high and low are relative to the application. For each pressure, a set of three power levels (30, 60 and 75 W) is simulated. Results from the 75 W simulations will be used to benchmark the simulators against the plasma survey described in Section 6.4.10.

6.4.4   Effect of rarefaction and gas heating

The effect of rarefaction and gas heating on the source terms for the plasma fluid model was investigated with two sets of simulations each including a low (60 W) and high (300 W) power level. In the first set, the PS1 module is run with no rarefaction and with a constant gas temperature grid. However, to investigate the ionization of the metallic species, we need to generate the metal density grid. Therefore, the grid for this set of simulations was obtained from a normal TS iteration but with the rarefaction and gas heating effects removed. In the second set, the PS-TS algorithm proposed in the previous section is used to investigate the effect of rarefaction and gas heating on the plasma creation rates. Since this section concentrates on the results from the PS1 algorithm, the PS2 algorithm is not executed to reduce execution times. The results from this two sets of simulations are shown in Figure 6.7 for aluminum and in Figure 6.8 for copper. In all figures, results shown on the left of the chamber correspond to simulations where the rarefaction and gas heating have been included whereas results on the right side of the plot refer to simulations where these two effects have been neglected.

Results from Figure 6.7 indicate that the effect of rarefaction on the ion plasma creation rates is not significant for any of the conditions shown in the figure. This is indeed expected at 5 mTorr since the rarefaction levels for aluminum shown in Figure 4.14 are low and one would not expect a significant effect on the high energy electron dynamics at low pressure. However, at 40 mTorr and 300 W, where the simulated rarefaction of the process gas is about 0.6 (according to Figure 4.14), we would expect a change in the electron dynamics but the figure clearly show that the effect of rarefaction is small. Other plasma creations rates such as the low energy creation rate, the excitation rate and the metastable creation rates are also not affected by the rarefaction levels observed in the chamber configuration studied in this section for the aluminum discharge.

Figure 6.8 shows that the calculation of argon ionization rates for a copper

(a) 40 mTorr 60 W

(b) 40 mTorr 300 W

(c) 5 mTorr 60 W

(d) 5 mTorr 300 W
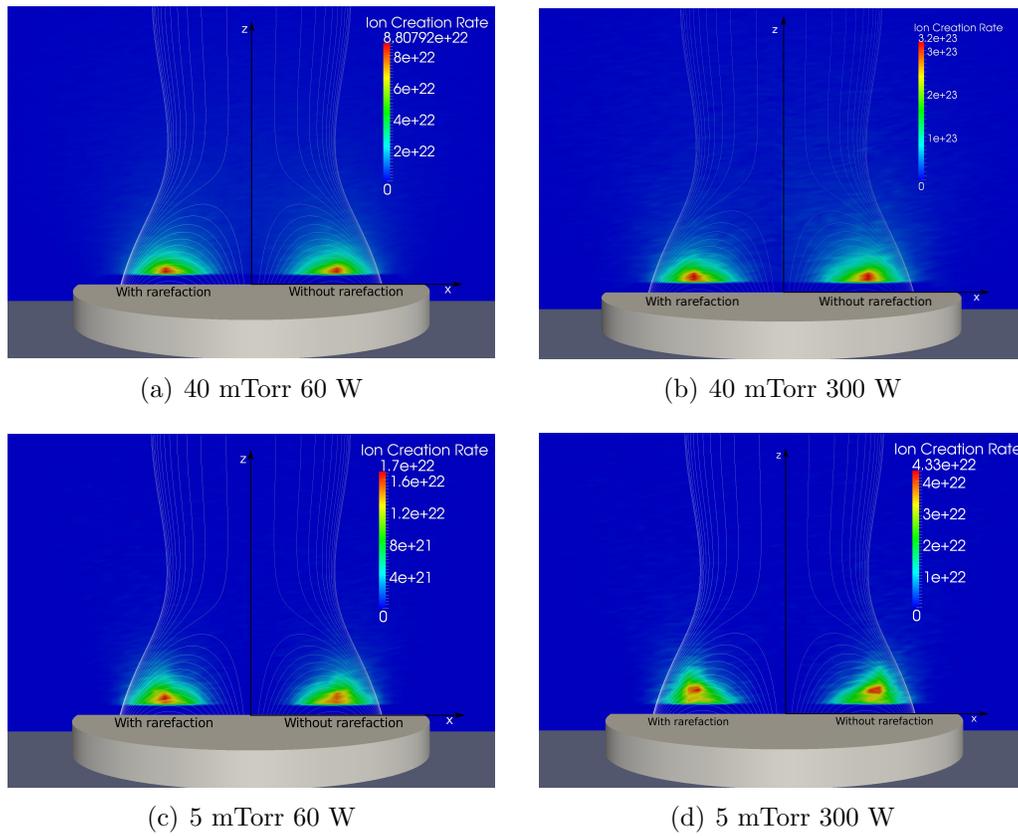
Figure 6.7: **Effect of rarefaction and gas heating on the PS1 algorithm for an aluminum simulation. Given the cylindrical geometry of the chamber, only one half of the discharge per process conditions is plotted in the figure. The left half of the plot corresponds to simulations including the rarefaction and gas heating effects whereas the results on shown on the right half side were obtained with a constant density and temperature grid.**

discharge using a rarefied grid is only affected at high pressures and powers. These results show that the impact of rarefaction on the ionization rate is bigger for Cu than for Al. This is a result of the increase ionization probability predicted in Chapter 4 where the ratio of the process gas to the metal density was significantly reduced as more rarefaction was observed in the discharge. Since electrons now can also lose energy to ionization of metal particles, they have less energy on average to ionize argon atoms during their lifetime. The small change in the figure is only an example of how the rarefaction needs to be included in calculating the electron trajectories. The rarefaction is only confined to the small volume next to the target in this example, but in bigger targets the effect is expected to be more important since the region where the electron dynamics changes is bigger.

overall

Excitation rate

The excitation rate for a 40 mTorr discharge at two powers is shown in Figure 6.9. The figure shows two halves of the discharge, each one corresponding to a power level. The plot is obtained by using a colormap with an alpha channel where low values of the excitation rates are rendered transparent in the plot. The effect observed in the figure is only for visualization purposes since we would require a better model to correlate the color and intensity of the plot to a real spectrum of a real glow discharge. The excitation rates as a function of power and pressure are shown in Figure 6.10. The plot shows the numerical values of the excitation rates along a line perpendicular to the target at the location where the excitation is the greatest. The numerical values of the excitation rates are the lowest for the low pressure simulation and rise as power or pressure are increased. This behavior of the plasma is expected in both cases for the following reasons. A higher discharge power is commonly associated with a higher current and voltage during operation. Thus, having more electrons with with higher energies coming out of the plasma sheath means that there are more collision in the plasma bulk. An excitation creation

(a) 40 mTorr 60 W

(b) 40 mTorr 300 W

(c) 5 mTorr 60 W

(d) 5 mTorr 300 W
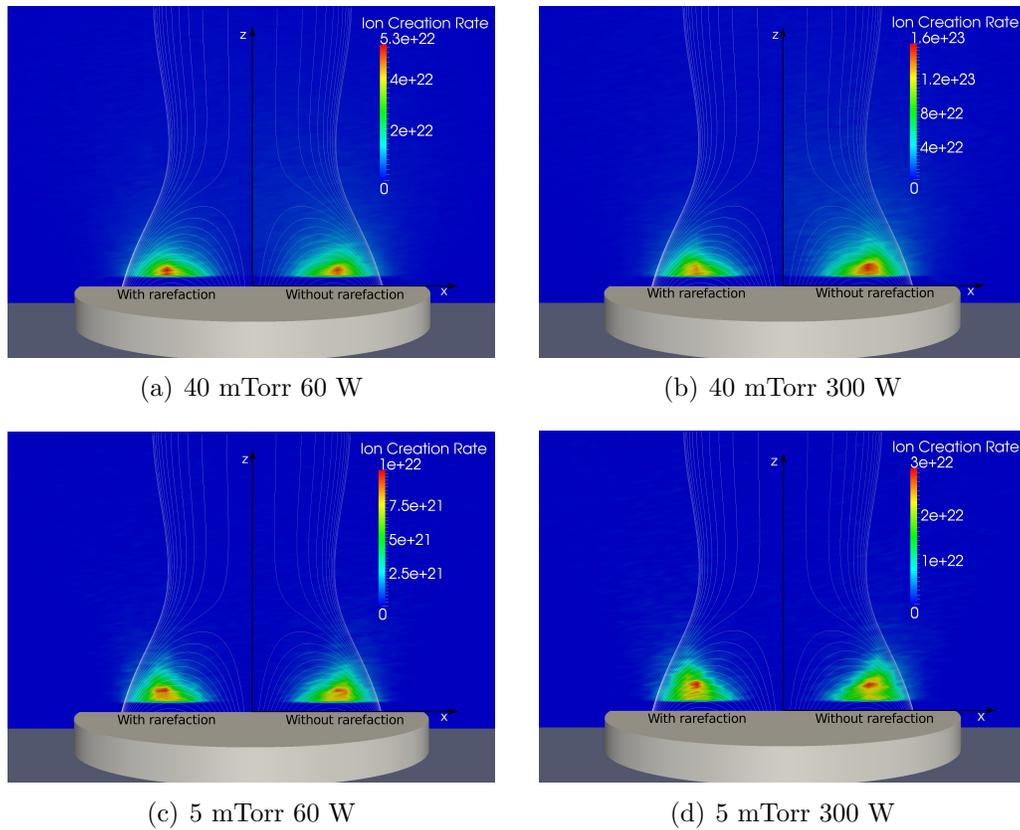
Figure 6.8: **Effect of rarefaction and gas heating on the PS1 algorithm for a copper simulation. Effect of rarefaction and gas heating on the PS1 algorithm for an aluminum simulation. In the plots, the left half corresponds to simulations with gas rarefaction and the right half plot was obtained with a constant density and temperature grid.**
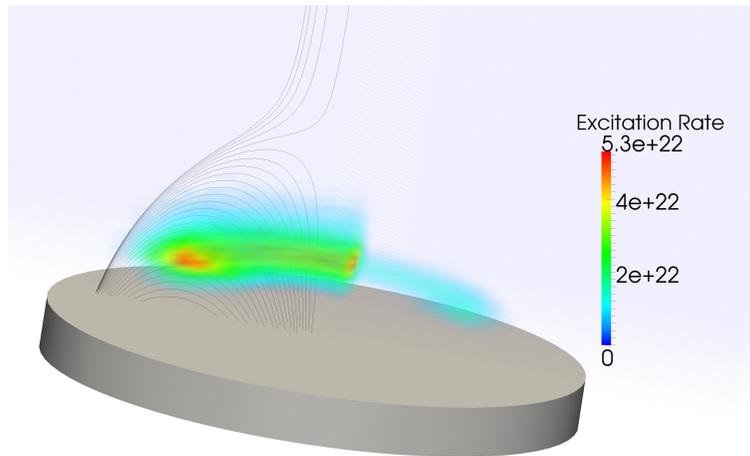
Figure 6.9: **Excitation rates for two powers for an aluminum discharge at 40 mTorr. Only excitation events from hot electrons are included in the plots. The left right half of the discharge is for 300 W while the right half is for 60 W.**

rate that increases as power is increased, on the other hand, is a consequence of a change in λ_e and related to the number of electrons or their energies. For instance, electrons at low pressures are more likely to move across the plasma region without suffering a collision, increasing the possibility of being captured by the target. This electron recapture phenomena at the target decreases the ionization efficiency of the discharge. As pressure is increased, the collision probability increases since λ_e is significantly reduced. The consequence is that electrons have more collisions in the plasma bulk, and they do not return to the target to be recaptured. It should be emphasized that the excitation rates shown in Figures 6.9 and 6.10 are only due to energetic electrons and do not include the excitation events from low energy electrons. Moreover, the plots only show the places where excitation events have occurred with no information whatsoever about the intensity of each wavelength. The current implementation of the SpudII framework does not solve for radiative transport; thus we are not able to produce a simulated spectra to be compared against

[htb]
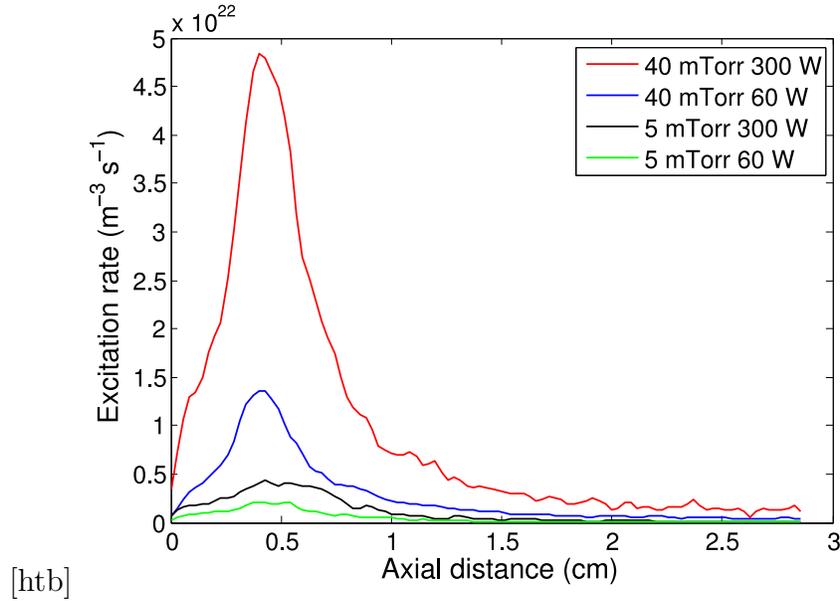
Figure 6.10: **Excitation rates for 40 and 5 mTorr and two power levels. Only excitation events from hot electrons are included in the plots.**

an experimental one. Furthermore, since all excitations were grouped into one effective excitation level, the values reported are only an estimate of the energy radiated from the plasma. The model could be extended if necessary to include a more descriptive glow discharge if all excitation levels are individually input into the xml input file. Such a model would require longer execution times but would allow us to obtain a more realistic emission from the plasma which could be correlated with experimental data. Although this could be implemented with little modifications, this aspect of the discharge is not relevant for the results presented in this dissertation.

6.4.5   Metal ionization rate

The effect of rarefaction on the ionization rate for an aluminum and copper discharge are shown in Figures 6.11 and 6.12, respectively. The simulations

were obtained for a 40 mTorr at 300 W discharge. It is clear from the figures that the rarefaction of the process gas has the effect of increasing the ionization rate in both cases. The mechanism that is responsible for the higher ionization when we include the rarefaction of the process gas was already discussed at the end of Chapter 4.
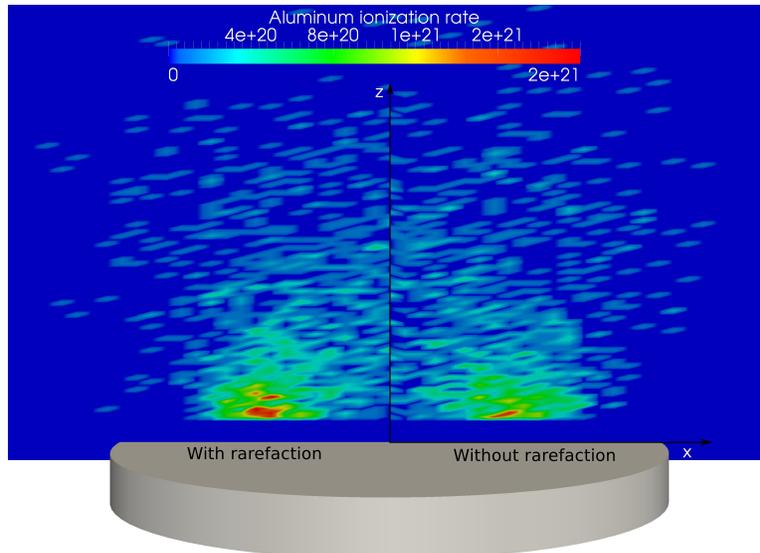
Figure 6.11: **Aluminum ionization rate in a 300 W discharge at 40 mTorr.**

Results for copper may suggest that the rarefaction of the process gas indeed contributes to the enhanced ionization of the sputtered particles in some sputtering processes. For instance, Nafarizal et al. [341] observed that Penning ionization was not sufficient to explain the observed increase in ionization of titanium. The ionization of sputtered particles has several technological implications some of which were already discussed in Chapter 2. However, it is still unclear what is the principle mechanism responsible for the enhanced ionization of the metal species in systems with no extra ionization sources. For instance, in HiPIMS, there is experimental evidence of a strong rarefaction effect in the discharge [215] which has led some researchers to suggest that rarefaction may play a significant role in increasing the ionization ef-
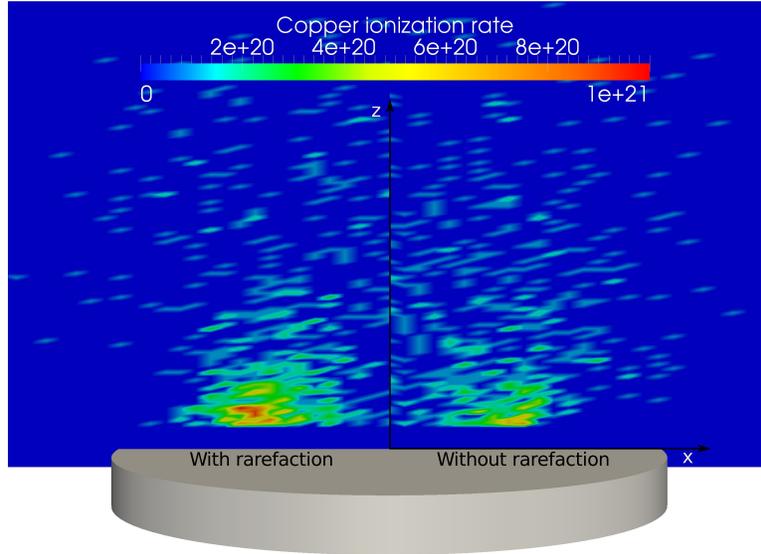
Figure 6.12: **Copper ionization rate in a 300 W discharge at 40 mTorr.**

ficiency of the sputtered material [23]. However, the extent of the effect is not easy to quantify both computationally and experimentally. HiPIMS is extremely complex to model, and only simple approaches are possible in practice [214, 342–344]; the comprehensive model described in this thesis is not sufficient for its simulation. Any attempt to modify the underlying concepts used in the SpudII framework would require significant modifications to the models described in this dissertation. For instance, the model presented in Chapters 4 and 5 would need to be modified to include the inherent time dependence of pulsed discharges using real time units (recall that PS2 used what it was labeled as 'fictitious' time units). However, since the dynamics of the plasma and the gas phase transport are now time dependent, the decoupling strategy between the high energy and low energy transport cannot longer be described using the hybrid models proposed in the SpudII framework. The main obstacle is in treating the energetic transport component. As an example, take the plasma fluid model of Chapter 5. Any attempt to use real time units would involve significant changes to the PS1 algorithm since the infor-

mation that PS1 sends to PS2 is constant through time. The same occurs with the TS1-TS2 communication. In this scheme, TS1 provides TS2 with creation rates that do not change with time. Furthermore, the model in TS2 has no implicit dependence on time.

The rest of the chapter focuses on the full iterative solution of the SpudII framework according to the algorithm of Figure 6.6.

6.4.6 Effect of pressure

In the results presented in this section, the electron density has been plotted on a line perpendicular to the surface of the target at three locations as indicated in Figure 6.13.
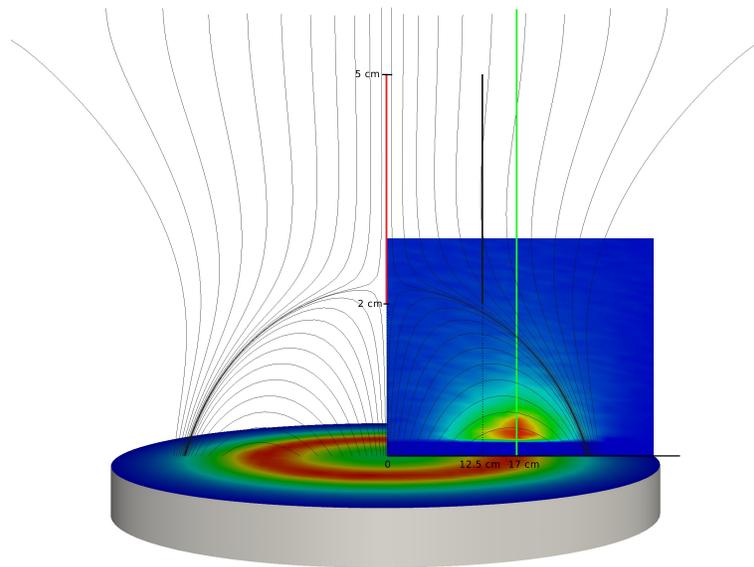
Figure 6.13: **Locations on the discharge where the electron and ion densities are plotted: center(red), erosion track (green) and max(black).**

The effect of pressure on the electron density and the plasma potential next to the target for an Al discharge are shown in Figures 6.14 and 6.15, respectively.

The discharge power was 75 W. In general, the electron density is larger for high pressures given the higher electron collision frequency of hot electrons at high pressures. This figure shows that the electron confinement is less effective at 40 mTorr than at lower pressures. For instance, in Figure 6.14(a), cold electrons are able to diffuse towards the center of the discharge following magnetic field lines on a trajectory describing an arc. These results are discussed later in the section when I compare the effect of the electron transport coefficients on the transport properties of the magnetron discharge. For instance, we will see that the electron density at 40 mTorr shown in Figure 6.14(a) is consistent with the shape of the effective diffusion coefficient.

The plasma potential effect is shown in Figure 6.15. This figure shows the plasma sheath shape and relative thickness with respect to pressure. The model correctly predicts a smaller sheath for higher pressures with a sheath thickness in the order of 1 mm for 40 mTorr and about 6 mm for 5 mTorr. Unfortunately, the grid is not as refined as necessary to capture in detail the sheath. The simulations presented in this chapter use a uniform grid to keep the simulations within a manageable amount of time. The model that treats cell refinement was not fully functional and it cannot be used in the current implementation of the algorithm. However, the results presented in this section clearly shows that the numerical algorithm for the magnetized discharge proposed in the Chapter 5 is able to solve for this very complex mathematical problem in a reasonable amount of time (about 14 days) for the full iterative scheme.

Figure 6.14 shows the effect of pressure on the transport characteristics of the low energy electron population. The electron confinement is weaker at high pressures since electrons can escape the magnetic field trap by moving perpendicularly to the magnetic field due to scattering. This can be explained in the context of single particle motion.

Recall that the electron trajectory is governed by the Lorentz force equation. In the plasma bulk, where the electric field is expected to be small (quasineutrality principle), electron trajectories are mostly determined by the magnetic
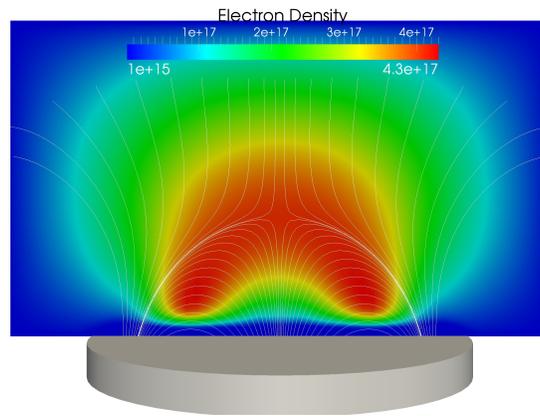
(a) 40 mTorr
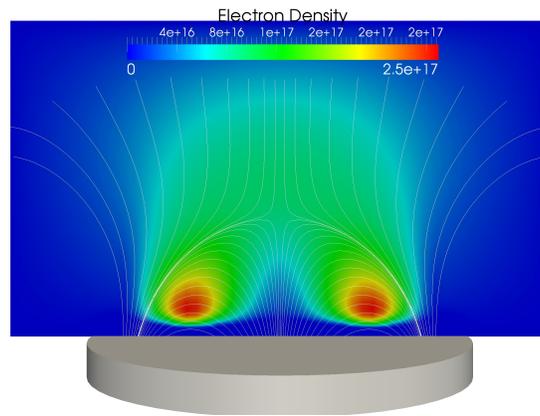
(b) 20 mTorr
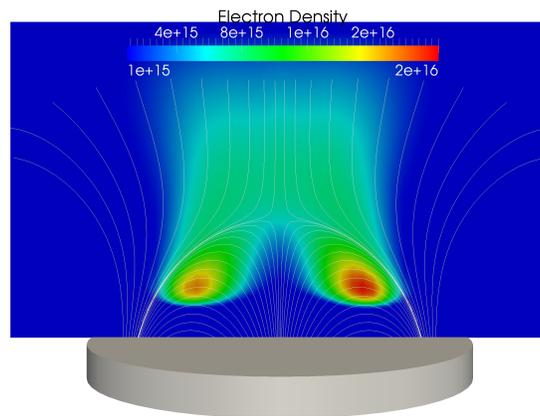
(c) 5 mTorr

Figure 6.14: **Electron density as a function of pressure. The discharge power was set to 75 W.**
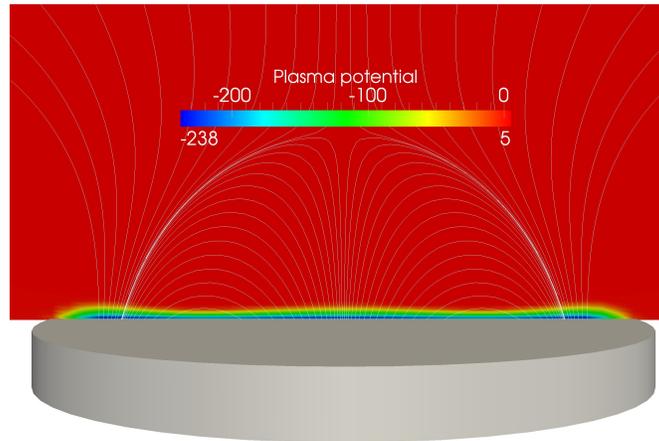
(a) 40 mTorr
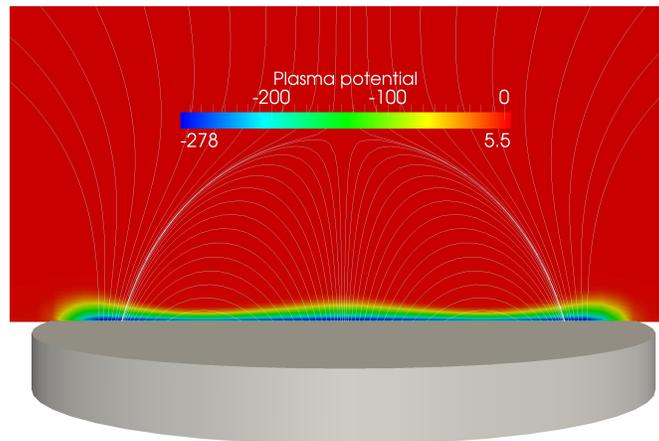
(b) 20 mTorr
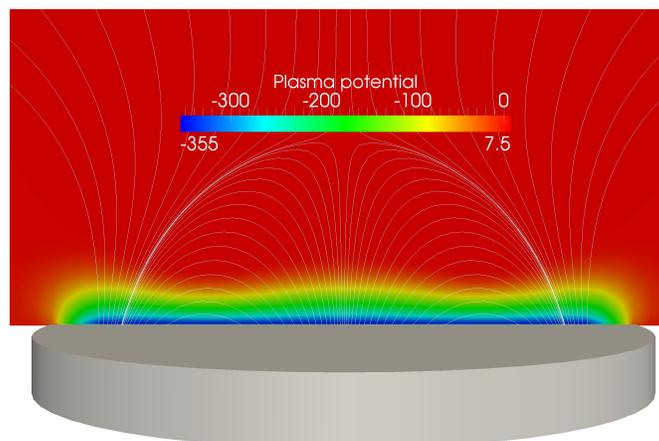
(c) 5 mTorr

Figure 6.15: **Plasma potential as a function of pressure for discharge power of 75 W.**

field shape. For the sake of argument, let us assume a zero electric field in the plasma bulk region. Thus, the solution to the Lorenz force equation for an electron trajectory results in a helical shape [75]. Electrons are thus tightly bound to the magnetic field and will tend to spiral around the field line where they happen to leave the plasma sheath. In this model, they can only escape the magnetic field through collisions (assuming that ∇B is small). It is then expected that a process wherein λ_e is small, there is a greater possibility for the electron to scape since every collision changes the motion that ties the electron to the spiral trajectory. On the other hand, at low pressures, λ_e is relatively high which increases the residence time of the electron in the magnetic field trap. Thus, at high pressures (40 mTorr), λ_e and the collision frequency increases. Collisions can be considered as a cross field diffusion coefficient whose value strongly depends on the shape of the magnetic field. The radius of the helical motion of electrons in a magnetic field is known as the gyroradius or Larmour radius r_g which can be approximated by

$$r_g = \frac{m_e v_\perp}{q|\boldsymbol{B}|} \tag{6.2}$$

where $v_\perp$ is the component of the electron velocity perpendicular to the magnetic field. Electrons will be tightly bound to the magnetic field for high values of $|\boldsymbol{B}|$ since r_g is small. In the absence of collisions, $v_\perp$ is constant throughout the electron trajectory since the electron does not gain energy in a magnetic field. However, a collision would effectively change $v_\perp$, hence affecting r_g and causing the electron to drift perpendicularly to the magnetic field. It is clear from Figure 6.13 that if an electron is moving perpendicularly to the magnetic field it is escaping the confinement region. This analysis is only a simplified model since we have not considered the inhomogeneity of the magnetic field in the discussion which also contributes to electrons effectively moving perpendicularly to the magnetic field.

The magnetic confinement of low energy electrons, which are treated in the fluid model, can be explained if we plot the diagonal elements of $\overline{\boldsymbol{L}}$ (equation

5.5.1) as a function of pressure as shown in Figure 6.16. Given the cylindrical geometry of the chamber, the radial component of the modified diffusion coefficient $\overline{\boldsymbol{D}}_e$ is just $D_{e0,0}$, while the axial component is $D_{e3,3}$. It is evident from the plots that the plasma is more efficiently confined at low pressures since electrons are more restricted to move both axially and radially. For instance, if we compare the radial component of $\overline{\boldsymbol{D}}_e$ for 5 and 40 mTorr, we can observe that electrons can only diffuse laterally in the region where the magnetic field is parallel to the target. Since it is in this region where the electron creation rate is the highest, the electrons are confined to freely move only in this small region, and their lateral transport towards the center or the walls of the chamber is limited. The same can be said if we compare the axial component of $\overline{\boldsymbol{D}}_e$ at the same pressures. The axial transport in the region where the electron creation rate is the highest is more limited at 5 than at 40 mTorr. This is only a simplified analysis but it shows how the low energy electrons are confined by the magnetic field.

6.4.7   Effect of power

The effect of power on the electron density is plotted in Figure 6.17 for 40 mTorr. The electron density next to the target increases linearly with respect to power since there are more secondary electrons being ejected from the target. This type of response is also expected since a higher power means that more energy is put into the plasma. A higher current increases the amount of ions hitting the target, thus increasing the number of secondary electrons that are ejected from the target into the bulk of the plasma. The higher secondary electron flux, if the pressure is kept the same, is invariably reflected in higher creation rates in the region next to the target.

field trap

diffusive flux

only an
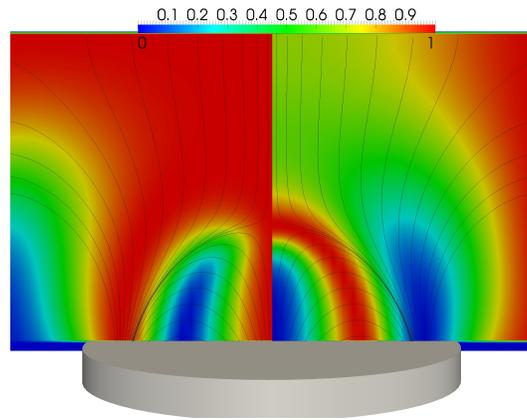
(a) 40 mTorr
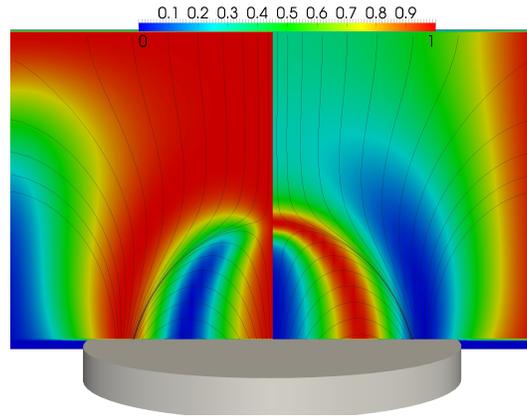
(b) 20 mTorr
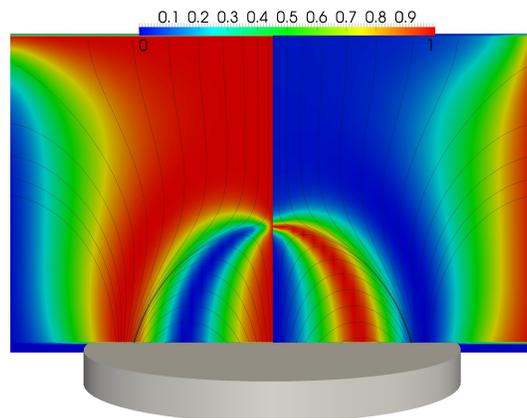
(c) 5 mTorr

Figure 6.16: **The radial and axial components of $diag(\overline{L})$ for three different pressures. In all figures, the left side of the figure depicts the axial component $(\overline{H}_{2,2})$ while the right side shows the radial component $(\overline{H}_{0,0})$.**
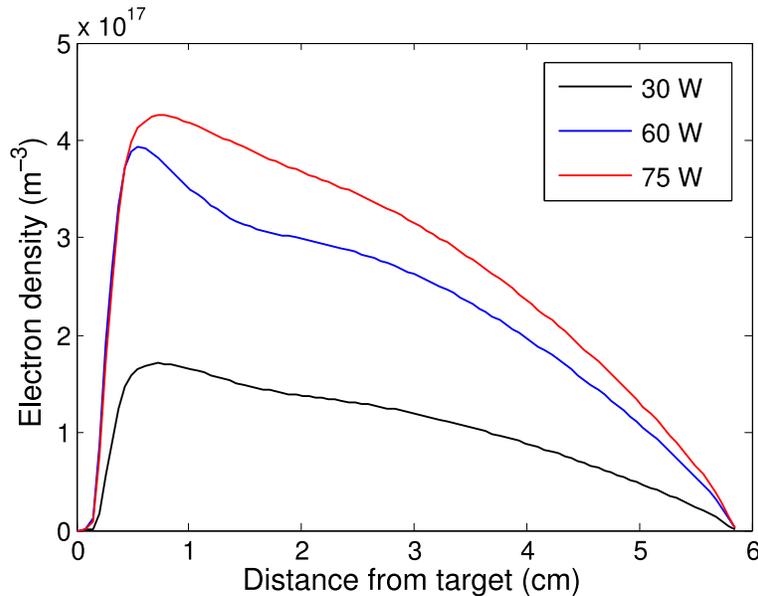
Figure 6.17: **Electron density as a function of power for 40 mTorr.**

6.4.8 Effect of magnetic field

The magnetic field is a key component in the operation of the magnetron discharge, and it is expected that variations in either the shape or the magnitude of the field can carry significant changes in the plasma dynamics. In this section, only the magnitude of the magnetic field is changed while its shape is kept constant to investigate the effect on the plasma. The reference magnetic field for all results presented in this chapter is the one depicted in Figure 6.4. The values investigated in this sections are 1.0, 1.1, 1.2 and 1.5 times the value of the reference field, $\boldsymbol{B}$. A detailed investigation of the effects of the magnetic field in a larger range of magnitudes and shapes is beyond the scope of this thesis.

The axial effect of the magnetic field magnitude on the plasma creation rate is shown in Figure 6.18. As expected, the creation rates rise by increasing the magnetic field since electrons are less likely to escape the magnetic field
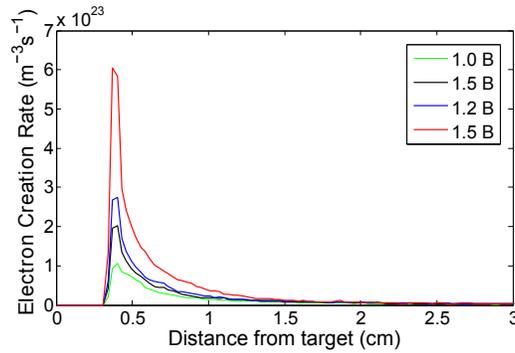
(a) 40 mTorr
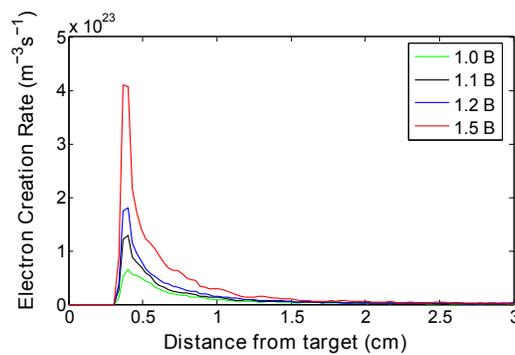
(b) 20 mTorr
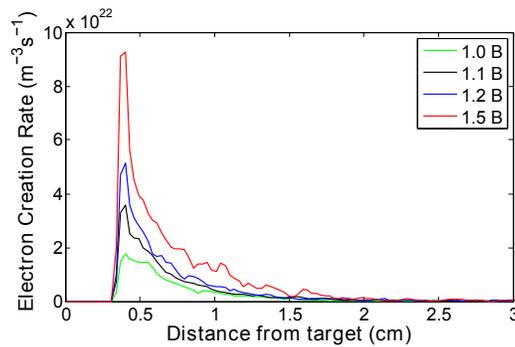
(c) 5 mTorr

Figure 6.18: **Effect of the magnetic field magnitude on the plasma creation rates for three different pressures at 60 W. The data in the plots is taken perpendicular to the target at the location where the creation rates are the maximum. B is the reference magnetic field shown in Figures 6.4 and 6.5.**
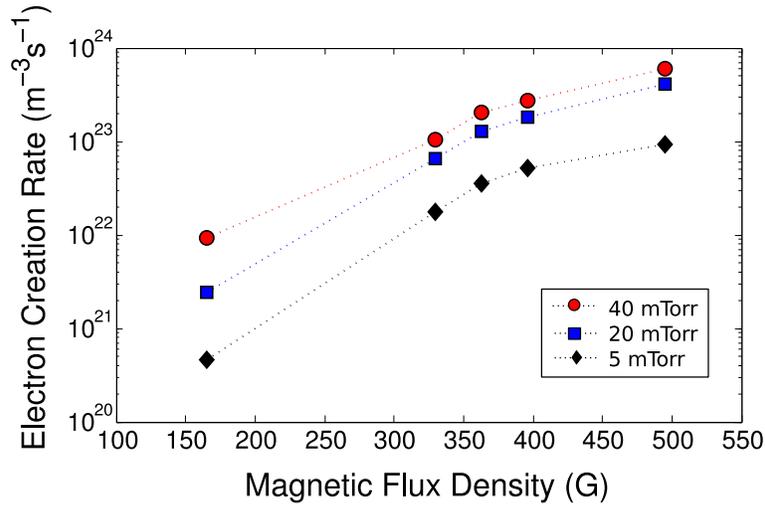
Figure 6.19: **The cold electron creation rates as function of the magnetic field magnitude.**

lines at higher fields. It is worth noting that increasing the field by 50 % increases creation rates by factors of 4-5. Figure 6.19 shows the effect of the magnetic field magnitude on the electron creation rate. The last two figures clearly shows the appeal of placing a carefully designed magnetic field in a glow discharge: ionization of the background gas is increased which leads to a more efficient process (the energy input into the system is better utilized in creating sputtered atoms leading to a better deposition rate per input unit power). However, a higher magnetic field is not always the best approach. For instance, there is experimental evidence that suggests that increasing the magnetic field does not necessarily lead to significant increases in the deposition rate at the substrate level [345]. The best compromise in efficiency-cost is usually process dependent, and it is not always easy to find.

To investigate the effect of a stronger magnetic field on the axial direction, the normalized creation rates for 40 mTorr at 60 W are plotted in Figure 6.20. From the figure, we can note how the total creation rate of low energy electrons decreases at a lower rate for the lower magnetic field ($\boldsymbol{B}$ in the figure). As the strength of the magnetic field is increased to $1.5\boldsymbol{B}$, the magnetic field is more

efficient in confining electrons in the direction perpendicular to the magnetic field. This effect is explained by the following reasoning. When the magnetic field is stronger, r_g becomes smaller according to Equation 6.4.6, implying that electrons are more tightly bound to the magnetic field. As a consequence, the effective drift perpendicular to the magnetic field due to collisions is reduced: it takes more collisions to displace an electron perpendicularly to the magnetic field when the magnetic field is stronger. Results shown in Figure 6.20 clearly indicates that there is a lower creation rate as we move away from the target for the highest magnetic field investigated.
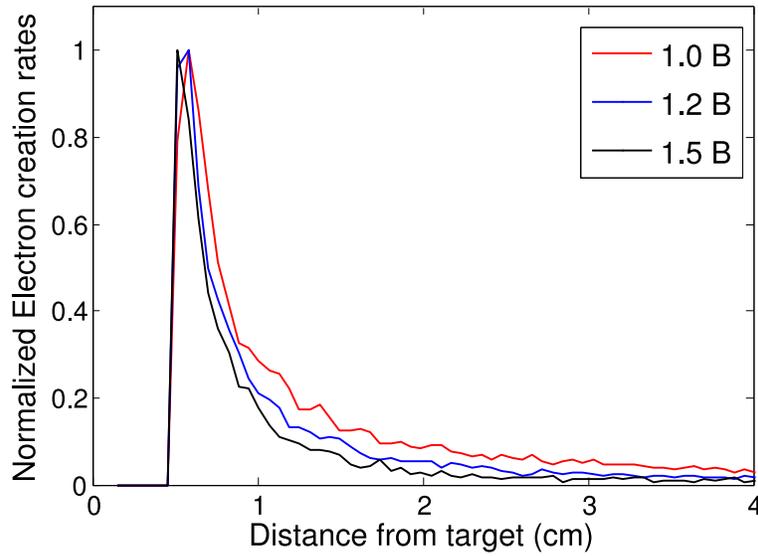
Figure 6.20: **Effect of the magnetic field magnitude on plasma creation rates at the erosion track region. Data is for simulations performed at 40 mTorr and 60 W. B is the reference magnetic field defined in figures 6.4 and 6.5. A stronger magnetic field reduces the electron creation rates in the direction perpendicular to the magnetic field.**

The effect of the magnetic field on the radial electron creation rates is shown in Figure 6.21. In this figure, the creation rates have also been normalized so as to compare them qualitatively. It is evident from the figure that the rate is just slightly affected by the magnitude only in the regions where the confinement is

weaker, i.e. at the center of the discharge and at the ends of the target (0 and 3 cm in the figure). These results suggest that the magnitude of the field does not play a significant role in determining the spatial distributions of sources next to the cathode indicating that it is the shape of the field that is the one parameter that characterizes the radial distribution of sources.
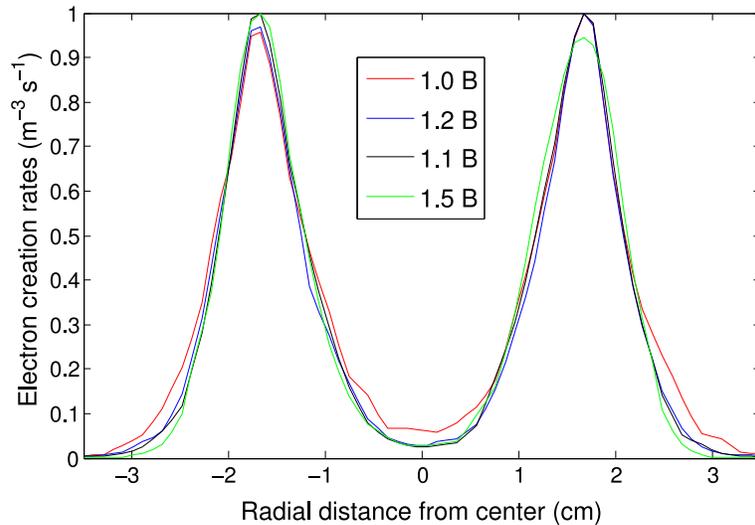
Figure 6.21: **Effect of the magnetic field magnitude on the radial electron creation rate. Data is for simulations performed at 40 mTorr and 60 W. B is the reference magnetic field defined in figures 6.4 and 6.5. In general, the magnitude of the magnetic field does not affect the shape of the electron creation rate when plotted along the radius of the chamber.**

The effect of the magnetic field on the electron density is shown in Figure 6.22 for two pressures while the normalized values for a qualitative comparison are plotted in Figure 6.23. Overall, the same trend observed in Figure 6.18 is observed in regard to the effect of the magnetic field. In these figures, the pseudo-transient simulation has been truncated to the first 50 μs of pseudo-time to reduce the execution time of the numerical experiment. The observed effect in the truncated simulation is expected to be of the same magnitude as in the complete simulation. It can be observed in figure 6.22 that a 50%

increment in the magnetic field more than doubles the electron density in the region next to the target. This density change is due mainly due to two factors.

On the one hand, as evident in Figure 6.18, a stronger magnetic field leads to a higher electron creation rate in the region next to the target where the confinement is greater. Since no other significant variable in the model has changed, it is expected that a higher source term in the plasma model results in a solution consistent with a higher electron density.

Secondly, as observed in Figure 6.16, the effective diffusivity in the direction perpendicular to the magnetic field is only zero in a small region next to the target (Figure 6.16(a)). A stronger magnetic field has the effect of spreading the confinement zone, i.e. the region in the chamber where the axial transport coefficient is zero. For instance, refer to Figure 6.16(a) where the axial component of the effective diffusion coefficient is plotted. It is observed in the plot that the effective axial diffusivity is zero only in a small region inside the magnetic trap (the blue lobe on the left in Figure 6.16(a)). In this region, low energy electrons are not allowed to diffuse perpendicularly to the field. The region where the axial component of $\boldsymbol{D}_e$ is zero extends as pressure is decreased due to a lower electron-neutral collision frequency. However, the same effect can be observed if the magnetic field is increased according to equation 5.32. Thus, a higher magnetic field also increases the region where the low energy transport coefficients are reduced.

6.4.9 Effect of γ_e

The effect of γ_e on the argon ionization creation rate shown in Figure 6.24. The effect of doubling γ_e is reflected almost linearly on the creation rates as observed in the figure. These results are consistent with the physical interpretation of γ_e. In fact, since we are not affecting significantly the overall conditions on which the PS1 grid is executed, increasing the number of simulated electrons by a factor m also increases the creation rates by about the same factor in the chamber configuration investigated in this chapter. The
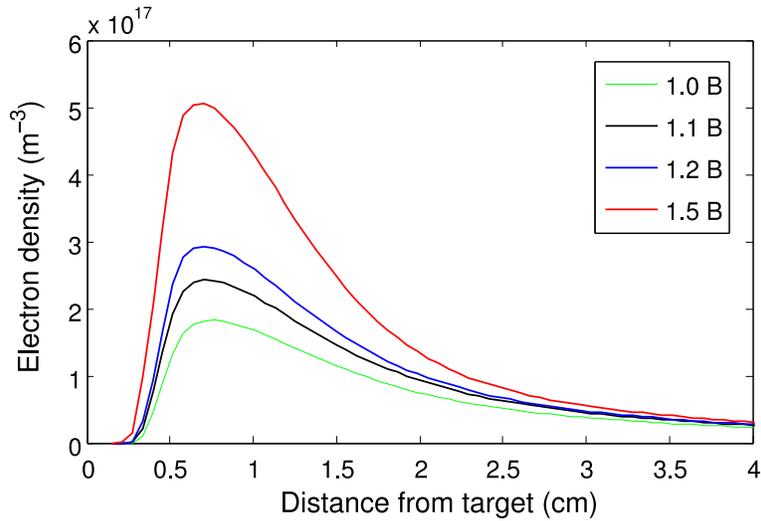
Figure 6.22: **The electron density in the direction perpendicular to the target surface for varying magnetic fields.**

increased number of secondary electrons leaving the cathode sheath leads to more collisions in the plasma region. This higher number of collisions inevitably increases the plasma creation rates since no other parameter in the model has been changed (voltage, pressure, target material).

The effect of γ_e on the electron density is shown in Figure 6.25. This plot shows data for partial simulations in a manner similar to the previous section. According to the figure, the electron density sensitivity to changes in the value of γ_e follows the same trend observed in Figure 6.24. These results suggest that the impact of γ_e in the creation rates is substantial. Given the high uncertainty in the correct value of γ_e in a real discharge, we need to treat this value as a fitting parameter.

At this point, it is worthwhile to comment about some of the challenges associated with the solution of the magnetized plasma represented in this chapter. The following lines summarizes one aspect of the model that required extra care in the algorithm.

Aside from the topics covered in the previous chapter regarding the numerical
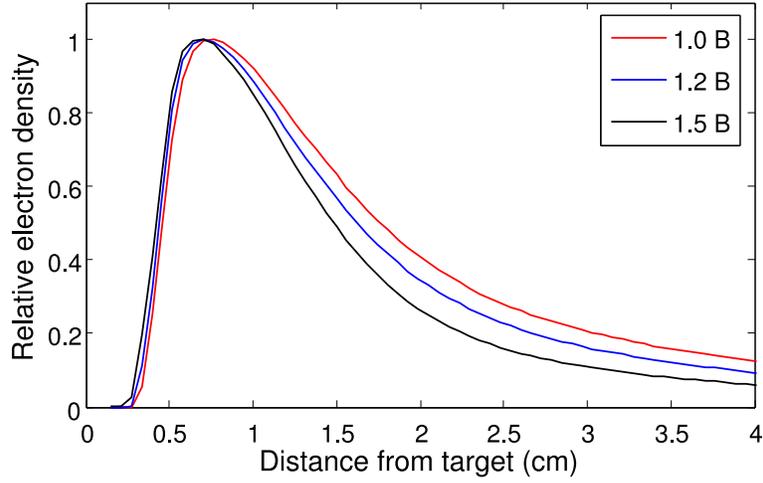
Figure 6.23: **The normalized electron density as a function of the magnetic field magnitude in the direction perpendicular to the target surface. Following the same trend observed in Figure 6.20, the normalized electron density drops faster for higher values of the magnetic field. The reduced electron diffusivity across the magnetic field lines for stronger magnetic fields causes the trend shown in the figure.**

algorithm, the difficulties faced when solving for the magnetized electron flux are evident if we take a close look at the expression to calculate the total electron flux. For the sake of simplicity, we will only look into the x component of this flux. Thus, from Equation 5.31, the divergence of the total electron flux (continuity equation with zero sources) in the x direction is

$$\nabla \cdot \mathcal{J}_x = \nabla \cdot \left(\tilde{J}_x + \hat{J}_x \right) \tag{6.3}$$

where $\tilde{J}_x$ and $\hat{J}_x$ are defined by

$$\tilde{J}_x = L_{0,0} J_x \tag{6.4a}$$

$$\hat{J}_x = L_{0,1} J_y + L_{0,2} J_z \tag{6.4b}$$
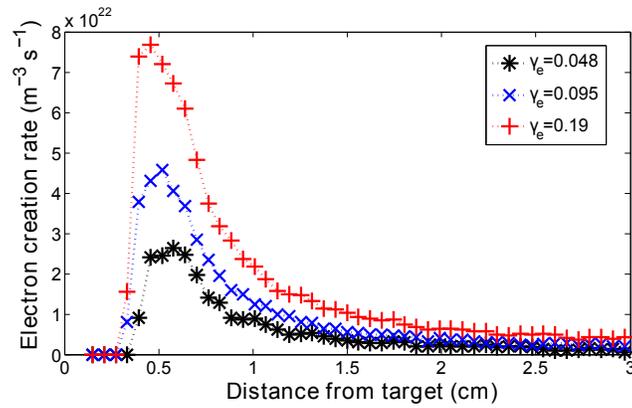
(a) 40 mTorr
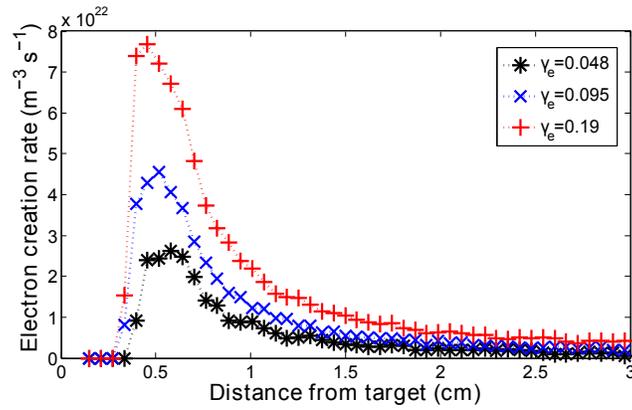
(b) 20 mTorr

Figure 6.24: **Effect of γ_e on the electron creation rate for 20 and 40 mTorr. The discharge power was 60 W. The electron creation rates scales linearly with respect to changes in γ_e.**
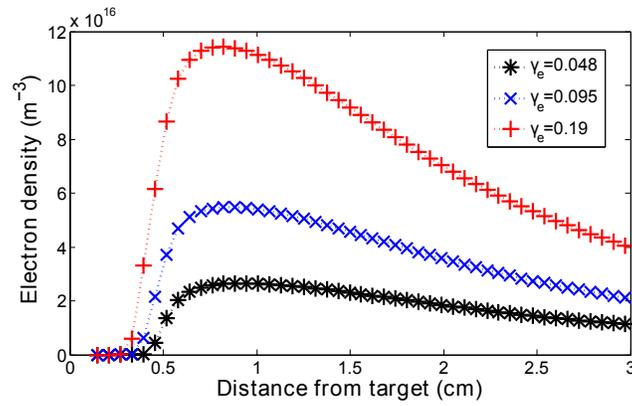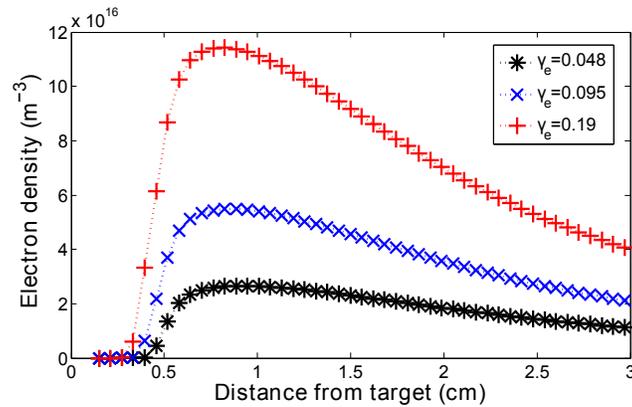
(a) 40 mTorr

(b) 20 mTorr

Figure 6.25: **Effect of γ_e on the electron density. Data taken at an early stage of the PS2 algorithm.**

with $L_{0,0}$, $L_{0,1}$ and $L_{0,2}$ obtained from Equation 5.5.1. In the above equation, J_x, J_y and J_z refer to the non-magnetized electron fluxes in the x, y and z directions, respectively.

The term $\nabla \cdot \tilde{J}_x$ is relatively easy to approximate using the finite volume method and the Scharfetter-Gummel approximation for the flux as already explained in section 5.5. In practice, we can use the same model to solve for the magnetized plasma by substituting the classical transport coefficients with the modified ones which include the effect of the magnetic field. However, to approximate the term $\nabla \cdot \hat{J}_e$ using the same finite volume formulation, we need to obtain partial cross derivatives for which we need to know the components of J_y and J_z at the edge of the cell. Given the geometry of the grid in the PS2 module, these quantities are not easy to obtain implicitly using a simple model. For this reason, all terms involving partial cross derivatives are sent to the right hand side of the continuity equations and treated explicitly in the solution algorithm. With this approach, the cross derivatives are calculated once we get an estimate of electron density at the current time step. Recall that, as mentioned in the previous chapter, the difference between an implicit and explicit approach is that the terms on the left for the discretized equations are calculated at a time $t+\Delta t$, while all terms on the right hand side are discretized at time t and calculated once a solution to the equations is obtained.

The challenge now is to approximate the cross derivatives and to find the proper numerical algorithm that does not break the numerical stability of the algorithm proposed in Chapter 5. The approach taken in this dissertation to calculate $\hat{J}_x$ is the following.

First, the magnetized plasma model is solved for time $t + \Delta t$ using the numerical algorithm proposed in Chapter 5 with $\hat{\boldsymbol{J}} = 0$. With n, p and ϕ calculated at time $t + \Delta t$, an estimate of $\boldsymbol{J}_x$, $\boldsymbol{J}_y$ and $\boldsymbol{J}_z$ is computed and $\nabla \cdot \hat{J}_e$ is approximated using a simple finite difference scheme. Then, an updated estimate of the right hand side of Equation 5.33b is obtained for the next time step iteration. This procedure, however, was numerically unstable during the early stages of the algorithm. To improve the stability of the algorithm, the

term $\hat{\boldsymbol{J}}$ was multiplied by a relaxation factor ϖ whose numerical value was incremented every few iterations.

Recall that the modified plasma model to solve, the one including the effect of the magnetic field, is described by Equation 5.33. Thus, including ϖ, the modified plasma model to solve in PS2 is

$$\nabla^2\phi = -\frac{q}{\epsilon_0}(n_i - n_e - n_{hot}) \tag{6.5a}$$

$$\nabla \cdot \tilde{\boldsymbol{J_e}} = S_e - \nabla \cdot \varpi\hat{\boldsymbol{J}}_e \tag{6.5b}$$

$$\nabla \cdot \boldsymbol{J_i} = S_i \tag{6.5c}$$

The numerical scheme used in PS2 is then the following. Equations 6.4.9 are solved using the algorithm described in Section 5.6. During the first iterations ($t < 1$ μs), the model was solved using $\varpi = 0.1$. Then, the value ϖ was successively incremented every iteration until it reached its final value ($\varpi = 1$).

The influence of $\hat{\boldsymbol{J}}$ in the plasma model is shown in Figure 6.26 which compares the solution of the full iterative method with $\varpi = 0$ and when $\varpi = 1$. It is clear from the figure that by neglecting $\hat{\boldsymbol{J}}$ in the model, the electron density next to the target does not have a localized peak. In this case, the cross field diffusivity next to the target is not enough to constrict the transport of the low electron population. If this term is not included in the solution, the electron density close to the target is underestimated. However, including $\hat{\boldsymbol{J}}$ in the solution algorithm greatly decreases the convergence properties of the numerical algorithm. In this case, the correct choice of ϖ is found empirically. The method employed in this thesis, as explained before, is to successively increment ϖ since if the algorithm starts with a high value ($\varpi > 0.2$) the numerical algorithm is unstable. The impact of $\hat{\boldsymbol{J}}$ on the results presented in Figure 6.26 is greater at low pressures as observed in the plots.
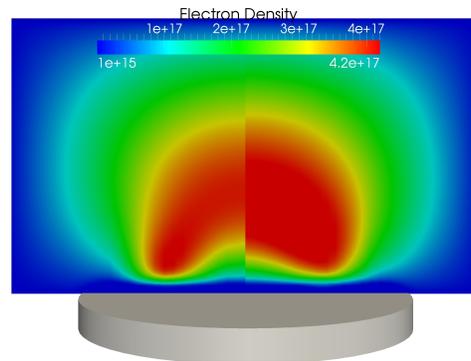
(a) 40 mTorr
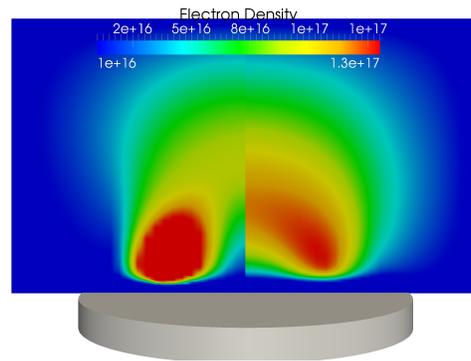
(b) 20 mTorr
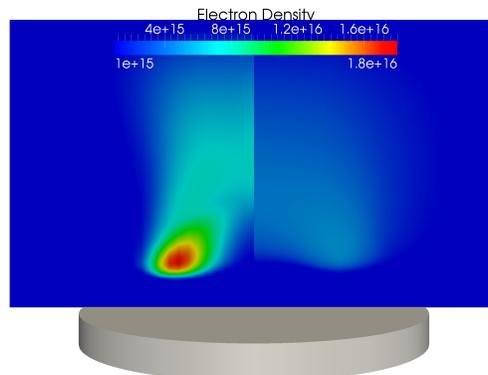
(c) 5 mTorr

Figure 6.26: **Effect of $\hat{\boldsymbol{J}}$ in the solution of the electron density for the magnetized discharge. Left: results from the magnetized model when $\hat{\boldsymbol{J}}$ is not included in the solution ($\varpi = 0$ in equation 6.5b). Right: solution of the complete magnetized model ($\varpi = 1$ in Equation 6.5b) .**

6.4.10 Plasma survey and comparison with simulations

Langmuir probe

A Langmuir probe is a small metal electrode immersed in the bulk of the plasma from which charged particles are extracted [346]. This type of probe was popularized by I. R. Langmuir and L. Tonks in the early decades of the 20th century as an effective tool to characterize a cold plasma [76]. This type of probe has been extensively used to measure electron density and temperature in magnetron discharges [347–359]. The theoretical framework of this type of probes is well documented in the literature [75, 346].

A typical Langmuir probe consists of a small conductor placed in the bulk of the plasma. Being an intrusive method, care should be taken to handle and analyze the data obtained from the probe. In brief, a Langmuir probe works as follows. When the probe is placed in the region where a plasma exists, it starts collecting electrons and/or ions depending on the probe's potential with respect to the plasma. Let us assume for the sake of the argument that the probe is floating. As soon as the probe is in contact with the plasma, it starts being bombarded by ions and electrons. However, since electrons are more mobile, the electron flux will be initially higher than the ion flux. This process rapidly charges the probe with a negative potential. The magnitude of the potential rises, and as a consequence the probe starts attracting more ions from the plasma. Eventually, a balance between the electron and ion flux is reached. At steady state, both electron and ion fluxes to the probe are the same and the measured electrical current is zero. The potential at which the probe is with respect to the plasma at steady state is known as the plasma potential. However, there is little information that can be obtained from a floating probe. The solution therefore is to connect the probe to an external voltage so the probe can draw current from the plasma, and this current can be related to the plasma parameters. This procedure is repeated for many voltage-current combinations to obtain an I-V curve which is typical of the type of the probe.
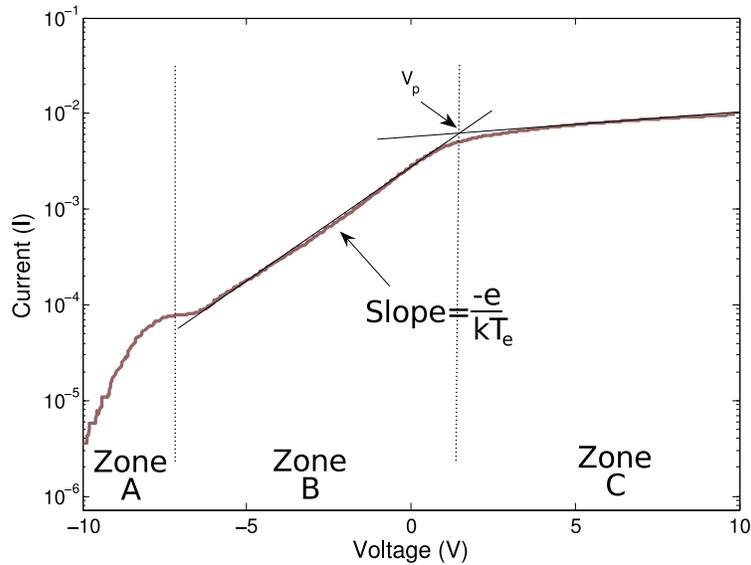
Figure 6.27: **Example of a typical curved obtained from a cylindrical Langmuir probe. Data taken from the experiments reported in this section.**

A typical I-V curve obtained from a linearly varying potential is shown on Figure 6.27. The plot can be divided in three regions. If the probe voltage is too negative (zone A) with respect to the plasma, the probe will tend to repel all electrons and only those in the high tail of the distribution function are able to surpass the potential barrier. The measured current in this region is then due only to ions. On the other hand, if the potential is too positive with respect to the plasma (zone C), the probe collects only electrons and the obtained current is a function of the local density of electrons in the vicinity of the probe. In between these two zones (zone B), electrons and ions are being collected by the probe in different proportions, and the total current is due to both species.

The Langmuir probe used for the measurements reported in this section was built using the design proposed by Field [2]. The Langmuir probe consists of two concentric alumina tubes attached together using a high vacuum epoxy

(TorrSeal). Figure 6.28 depicts the dimensions and an example of one of the probes used in the survey. The outer tube (outer diameter 6 mm and internal diameter 3 mm) works as a mechanical support for the probe while the inner tube (outer diameter 3 mm and internal diameter 1.5 mm) holds the metal probe in place (0.25 mm diameter W cable). Tungsten is the material of choice for the probe given its low secondary electron coefficient and high melting point. A low γ_e is needed to minimize the interaction of the plasma with the probe while a high melting point prevents the probe from evaporating. The total length of the probe is about 4 cm.

The probe shown in Figure 6.28 shows that the electrode tip is turned 90° with respect to its original direction. This bending is to minimize the effect of the magnetic field on the measurements. This is explained using a very simple example. Let us assume that we introduce the probe in a plasma with no magnetic field. In this case, the transport coefficients are isotropic and we should, ideally, measure the same values irrespective to the orientation of the probe. However, when there is a magnetic field, the transport coefficients are clearly anisotropic and the measured values depend on the orientation of the probe. Since the electron transport parallel to the magnetic field is expected to be higher for reasons explained before, the probe needs to be oriented such that its length is perpendicular to the direction of the flow. For the results presented in this section, I oriented the probe such that the length of the probe was perpendicular to the direction of the magnetic field in all cases.

The electronic circuit used is shown in Figure 6.29, and it is the same reported in Ref. 2.

In practice, the I-V analysis of Langmuir probe data is more complicated than the simplified model described above. The analysis was performed using the following procedure. First, a digital oscilloscope (Tektronics TDS210) was used to extract the I-V information from the circuit shown in Figure 6.29. The oscilloscope was used as an interface to a personal computer. Matlab was used to collect and anlalyze the data. For this, a simple script was written to connect the Matlab interface with the oscilloscope using the serial port.
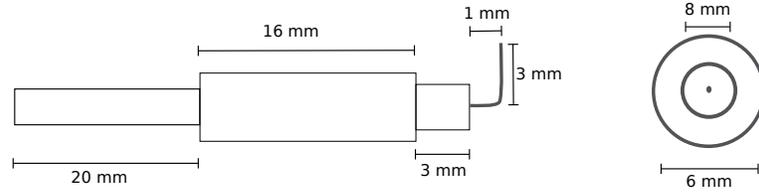
(a) Schematic

(b) Example

Figure 6.28: **a) Schematic of the Langmuir probe built for the electron density measurements reported in this section. b) Photographs of the Langmuir probe used in the experiments.**
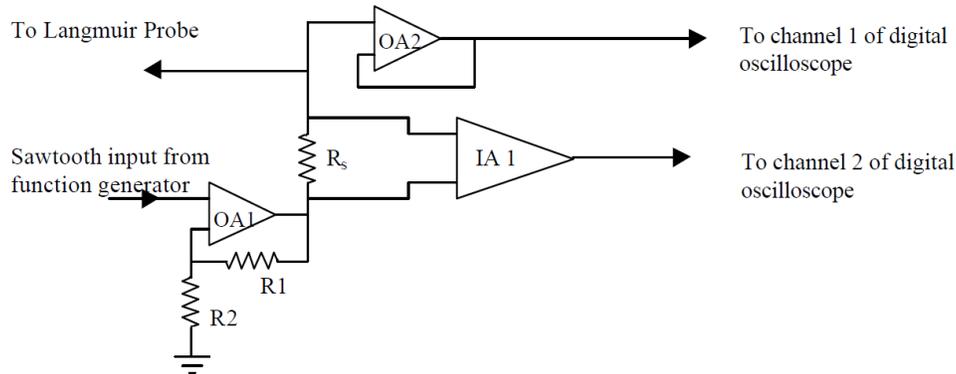
Figure 6.29: **Electrical circuit used for capturing the electrical signal delivered by the Langmuir probe when inserted in the plasma region.**

Once all measurements were taken, the data was analyzed and the plasma parameters were extracted from the I-V curves using the model described in Ref. [2]. The process is summarized below.

First, each collected waveform is filtered using a Savitzky-Golay procedure to reduce noise [299]. This filter is readily available as internal function in MATLAB. The filtered waveform is then plotted on an semi-log scale. The ion current is approximated and extracted from the data and a simple algorithm searches for the best fit. If the fit is correct (checked by visual inspection), the script saves the calculated plasma parameters, and the next waveform is analyzed. The uncertainty in the measurement is approximated using the same procedure described in Ref. 2.

Past analysis of Langmuir probe data in magnetron plasmas has shown that there are regions in the discharge where bi-Maxwellian distributions have been observed [349, 352, 355]. The exact origin of these distributions has not been elucidated, but some theories have been put forward to explain them. Field [352] speculated about the mechanisms to create a two-temperature distribution and attributed the effect to electric confinement of low energy electrons in some parts of the discharge. A detailed investigation of the ori-

gin and location of these distributions is beyond the scope of this chapter. Since bi-Mawelilian distributions are normally characterized by two electron populations, each with its own characteristic temperature, the data has been weighted with respect to each population to obtain an effective electron density and temperature. The procedure to obtain these quantities is described in Ref. 2. The values herein reported correspond then only to effective quantities.

Using this probe, the radial dependence of n_e and T_e at a distance of 2 cm from the target surface is shown in Figure 6.30. The discharge power was set to 75 W for all measurements. The electron density at this distance shows a peak at about 12 mm from the center of the target. However, the peak of the erosion rate taken at the target surface is about 17.5 mm from the center, indicating that the location of the electron peaks shifts towards the center as we move away from the target. This behavior is consistent with the Hall parameter shown in Figure 5.32. For instance, it can be observed in all three plots that the radial transport next to the target coincided with the location of the highest erosion rate at the target. As we move upwards, the constrained region gradually moves towards the center. The axial transport also plays a role. Figure 6.16(a) shows that at 40 mTorr, the axial transport is less constrained since the location where electrons cannot axially diffuse and only extends a couple of centimeters from the target. This region is indicated by the blue lobe on the left extending from the target surface to a few mm above the target. As pressure is increased, the region where the electron diffusivity along the axial direction is zero increases. This is clearly evident if we compare Figure 6.16(a) and 6.16(c).

The axial dependence of n_e is shown in Figure 6.31 at two locations above the target surface. The effect of the Hall parameter and the cross field diffusion as a function of the discharge pressure is observed if we compare both figures. For instance, take Figure 6.31(a) where we can observed how the electron density is higher in the region close to the target. We observed in Section 6.4.6 that the localized peak in the region is mainly due to the term $\hat{\boldsymbol{J}}\boldsymbol{e}$. This was a conclusion based on the fact that the simulation run with no cross
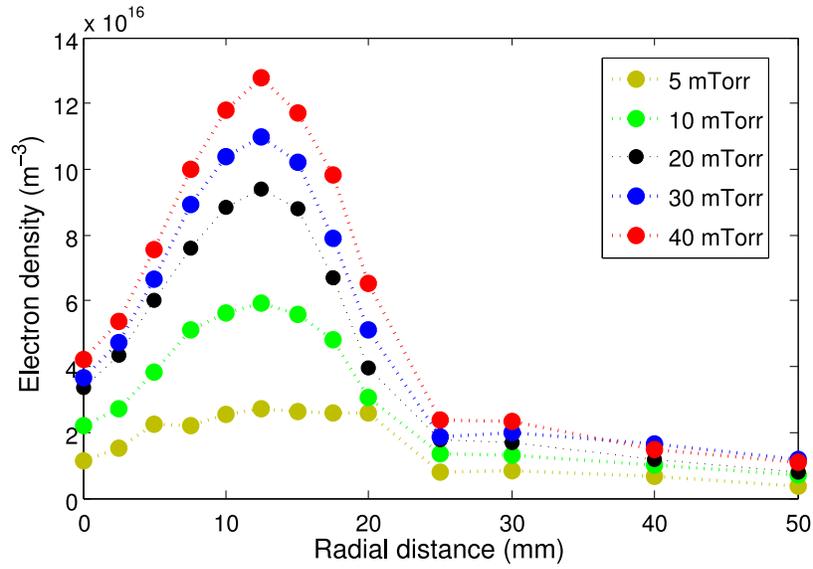
(a) n_e
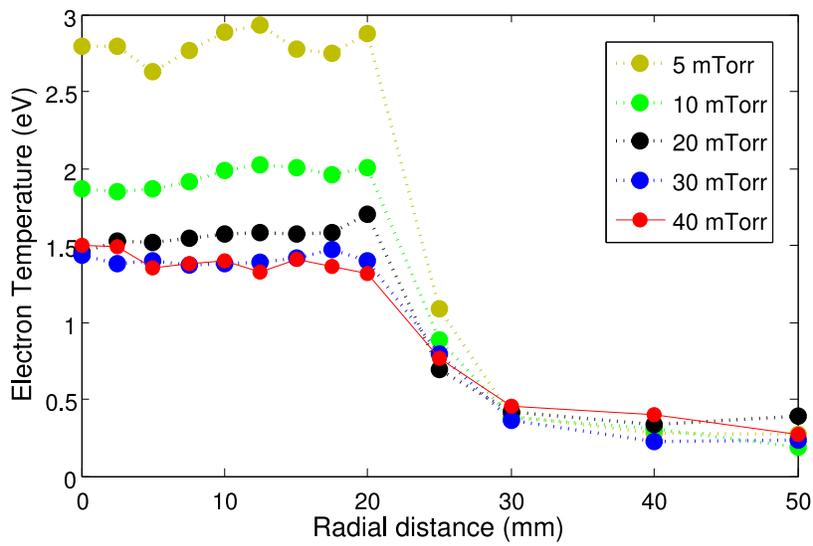
(b) T_e

Figure 6.30: **Radial electron density and temperature in the magnetron discharge at 2 cm from the target. Discharge power was 75 W.**

field did not show a pronounced peak of the electron density in the magnetic confinement region.

On the other hand, the plasma density at the center of the discharge is mainly governed by the effective diffusivity. This is clear if we compare the results shown in Figure 6.31(b) for all five pressures. At 5 and 10 mTorr, the plasma density is almost constant in the direction away from the target. However, as pressure is increased, the electron density along the center of the chamber increases as we move away from the target. This is more evident for 40 mTorr where the electron density increases from about $4.2 \times 10^{16} m^{-3}$ at 2 cm to about $6.8 \times 10^{16} m^{-3}$ at 3 cm away from the target. This plot is experimental evidence suggesting that the electron diffusion in the region next to the target is mainly determined by the effective diffusivity as shown in Figure 6.16. The mechanisms by which the Hall parameter dictates the electron transport characteristics were already explained in Section 6.4.6.

The complete survey of electron density and temperature is shown in Figures 6.32 and 6.33, respectively. At the power level these measurements were taken (75 W), the transition from high (40 mTorr) to low (5 mTorr) pressures depicted in Figure 6.32 shows the trend. The electron density shows at about 12.5 mm from the center of the cylinder. This clearly indicates that the location where the electron density shows a maxima moves towards the center of the cylinder as we move away from the target. For instance, next to the target, the electron density peak is observed at about 17.5 mm, consistent with the erosion profile observed experimentally. As we move away from the target, the electron density peak tends to shift towards the center. This shift is consistent with the shape of the anisotropic transport coefficients shown in Figure 6.16.

6.4.11 Experimental comparison

The comparisons between experiment and simulations for 40, 20 and 5 mTorr are shown in Figures 6.34, 6.35 and 6.36, respectively. The discharge power for all these plots was set to 75 W. Overall, there is some qualitative agreement
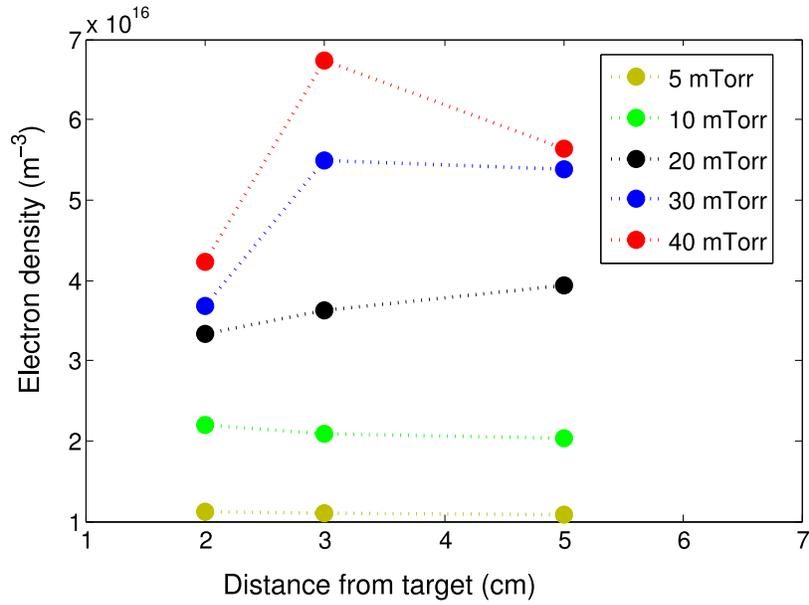
(a) Center
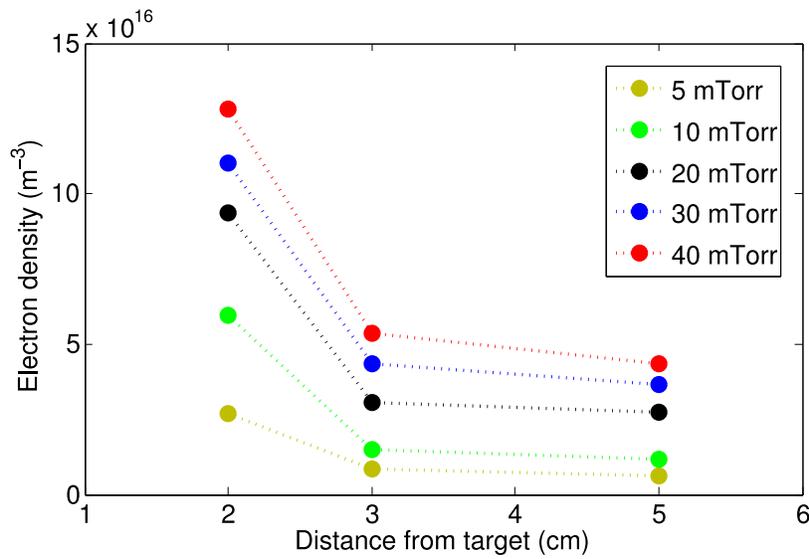
(b) Max

Figure 6.31: **Axial electron density measured at two locations: a) at the center of the cylinder and b) at the location of the maximum density as shown in Figure 6.30**
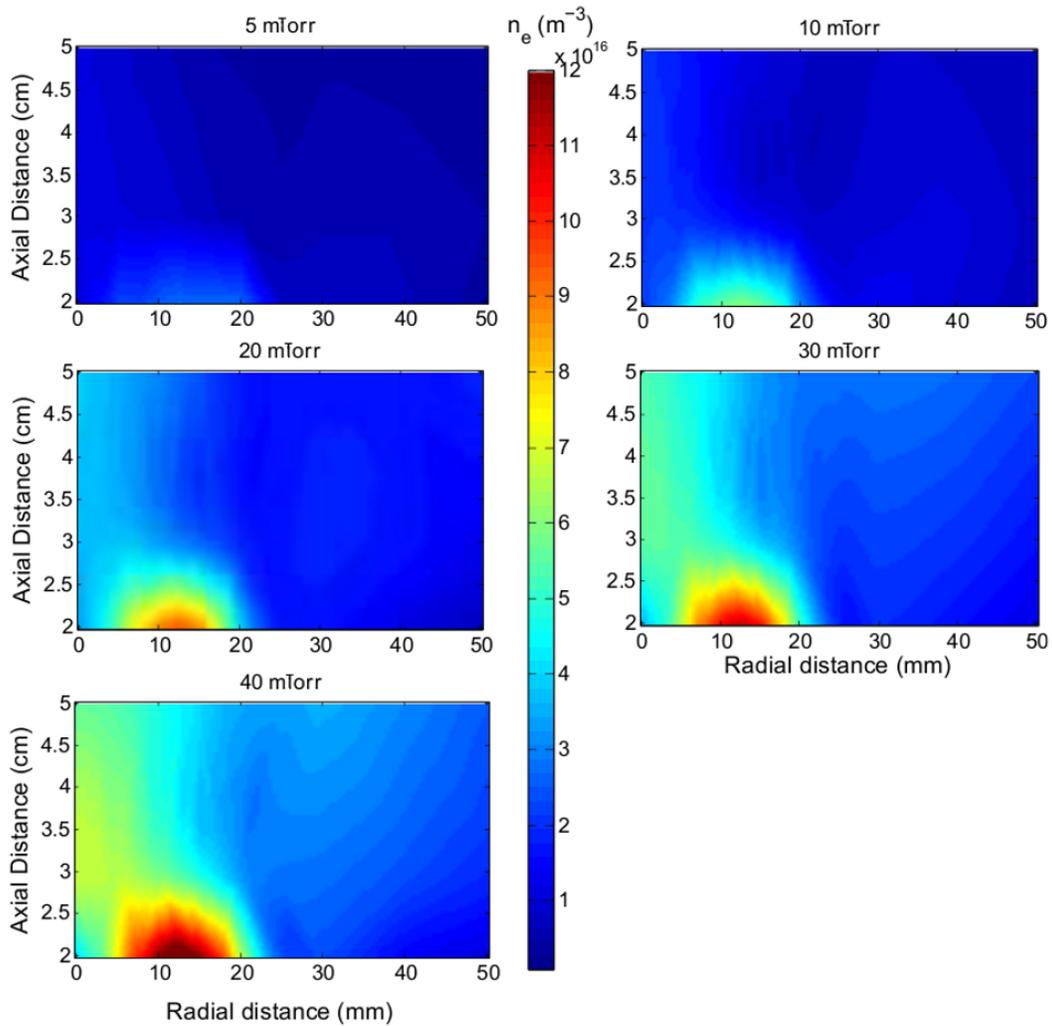
Figure 6.32: **Electron density as a function of position in the chamber. The figure only shows an estimate of the real electron density distribution since only three data points were available to interpolate the electron temperature in the axial direction.**
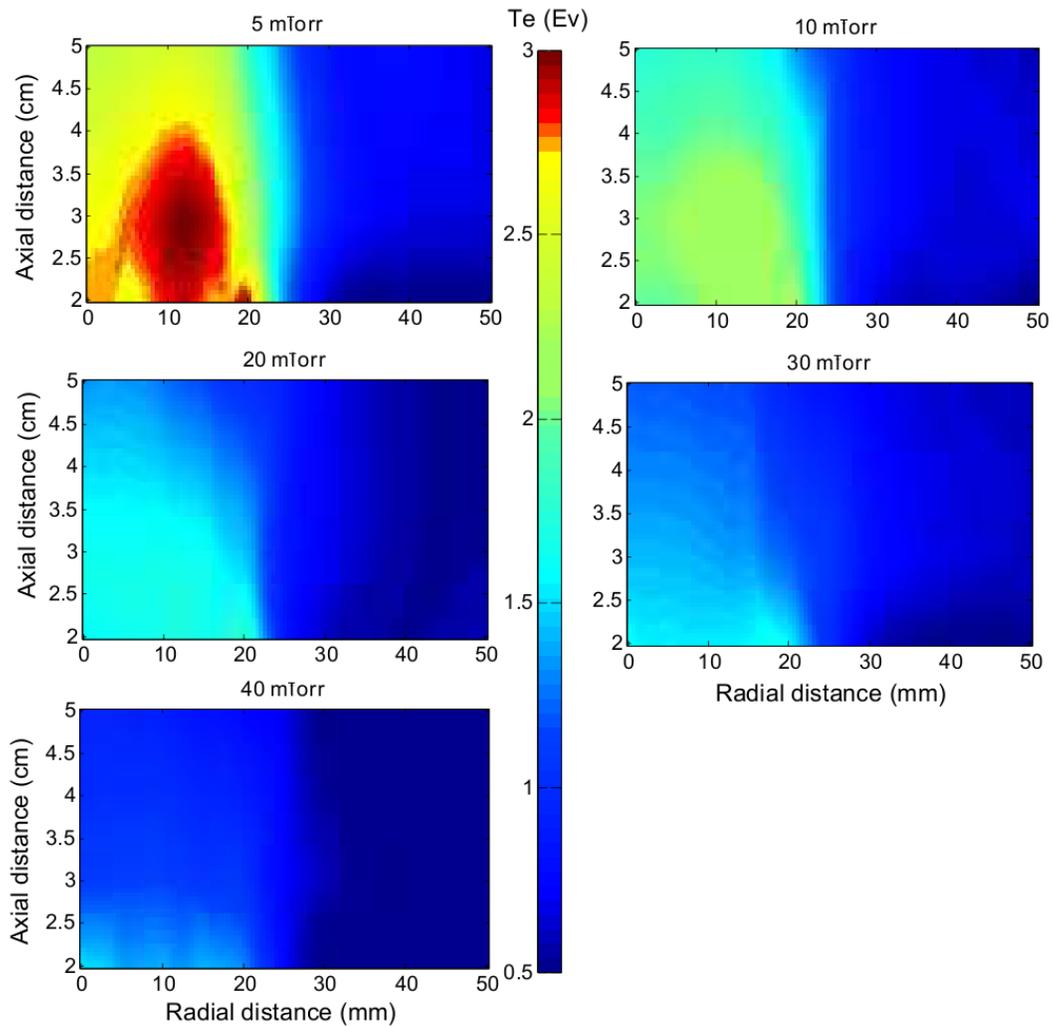
Figure 6.33: **Electron temperature as a function of position in the chamber. The figure only shows an estimate of the real electron temperature distribution since only three data points were available to interpolate the electron temperature in the axial direction.**

in the plots with important quantitative differences. This type of result is expected from a simulation of a complex problem if the transport coefficients are not properly adjusted.

The results of Figure 6.34 suggests that the electron density falls at a faster rate in the simulation than in the experiment. This may be indicative of a higher effective diffusivity in the direction parallel to the field in the simulation.
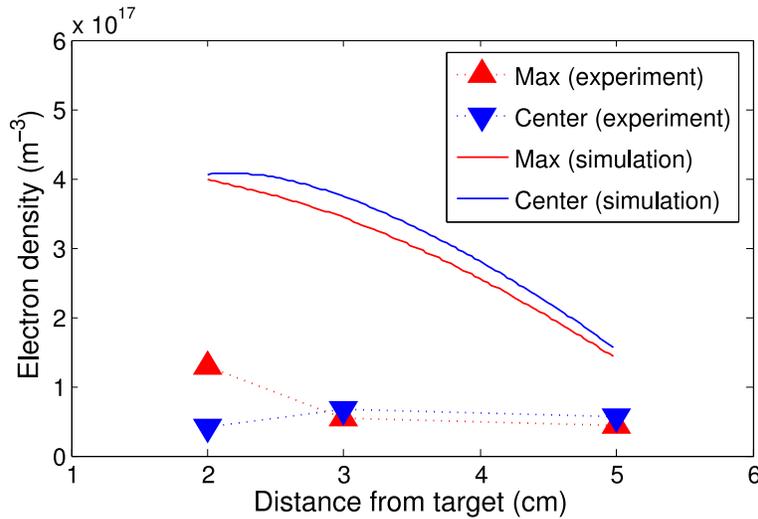
Figure 6.34: **Simulated versus experimental electron density for 40 mTorr at 75 W.**

Results from Figure 6.35 for a 20 mTorr simulation are in better quantitative agreement than those of Figure 6.34 for 40 mTorr. At this pressure, the electron confinement by the magnetic field is stronger, and a small difference in the effective diffusion coefficient has a smaller impact on the transport perpendicularly to the magnetic field lines. However, both simulation and experiment show the same trend: the electron density is higher at the max location at 2 cm from the target but as we move away from the target both densities tend to about the same value. The fact that the electron density in the simulation decays at a slower rate than the rate observed in the experiments may indicate that the electron confinement in the simulations is weaker than it should be.
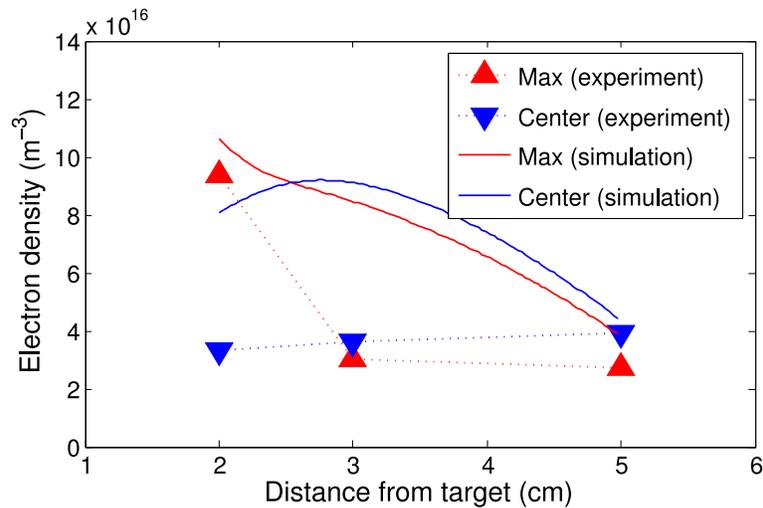
Figure 6.35: **Simulated versus experimental electron density for 20 mTorr at 75 W.**

The constant electron temperature approximation can also affect the simulation results in another way. For instance, using a constant electron temperature in the whole discharge has as a consequence that the electron diffusivity is underestimated in the region where the electron temperature is high and is overestimated where the electron temperature is low. The error in the approximation for the electron transport coefficient is even greater at low pressures where the electron temperature can be as high as 3 eV (according to Figure 6.33).

In an attempt to obtain a more accurate electron temperature distribution, I tried to implement an equation describing the electron energy transport. However, the solution to this equation resulted in an extremely complicated numerical problem. The solution to this equation was then left then for a further improvement of the model.
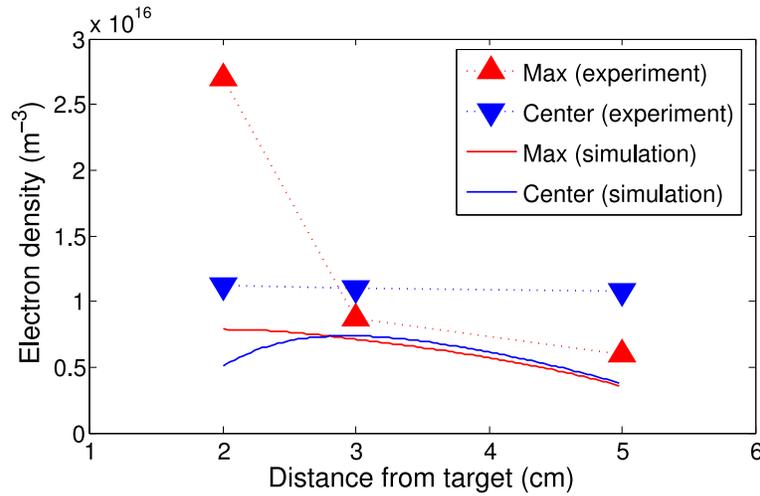
Figure 6.36: **Simulated versus experimental electron density for 5 mTorr at 75 W.**

6.5   Summary and conclusions

This chapter has presented results from a comprehensive sputter simulation tool that incorporates all relevant aspects of the process in the solution strategy. The results presented in this chapter reflect the contribution of all modules presented in this thesis. The plasma simulation model presented in Chapter 5 has been coupled with the neutralized particles' module of Chapter 4 using a strategy where the nonlinear coupling between all phenomena is solved using an iterative approach. This computational model has been used to investigate the effect of several process conditions on the plasma characteristics of a typical magnetron process. It was found that the effective diffusion coefficient is key in explaining the confinement of low energy electrons in the region next to the target. It was observed that the confinement is stronger at low pressures since the electron transport across magnetic field lines is significantly reduced under these conditions.

A Langmuir probe was built to benchmark the simulation against experimental data. A series of measurements were taken in a conventional cylindrical

magnetron for an Al deposition process. The data collected includes five different pressures at one power level. Results from the Langmuir probe are in qualitative agreement with similar measurements found in the literature. It was suggested that one of the reasons for the observed differences with the simulation results was due to one of the simplifications of the discharge model presented in Chapter 5.

CHAPTER 7

Design of a heat flux measurement device for magnetron sputtering processes

7.1 Introduction

This chapter describes the construction and implementation of a device designed to measure the heat flux in an industrial sputter coater. The chapter starts with a short review on the methods to measure this important processes parameter in low pressure discharges. The chapter continues with a brief description of the heat flux measurement device especially built for its use on sputtering reactors. The device was used to measure the heat flux at the substrate level in two types of rectangular magnetrons. These magnetrons are a scaled-down version of an industrial coater used for the production of nanocrystalline silver coatings for antimicrobial applications.

7.2 Heat flux measurement in plasma equipment

During the coating of heat sensitive substrates such as plastics and polymer webs, the energy flux at the substrate is one of the most important parameters to consider [360]. In choosing the optimal parameters for the process, we need to consider not only the deposition rate, but also the total energy input to the substrate during deposition. However, the effect of all system variables on the total energy flux at the target is not always easy to determine. In addition to the total energy, the energy per depositing particle also plays an important role in determining the growth dynamics of the film [361].

Another concern in industrial coaters is the uniformity of the heat flux at the substrate region. For heat sensitive substrates, any nonuniformity on the flux may severely damage the film by locally increasing the temperature of the substrate beyond its failure point. Furthermore, a highly nonuniform flux may impose very strict limitations on the maximum deposition rate that can be obtained for an specific process. The problem of determining the failure point for the substrate is key for optimizing the deposition process. The economics of the deposition process are dictated by several factors such as the cost per square meter. One alternative to reduce this quantity is to maximize the total throughput by tweaking the process conditions of the deposition. However, the maximum throughput is normally limited by the amount of heat the substrate can withstand before failure.

There are several techniques to estimate the heat flux, but only a few can be implemented in a low pressure environment. We can broadly divide these approaches into two methods: direct and indirect. This division is strictly based on the method to extract the total heat flux from the experimental data.

In an indirect method, the total heat flux impinging on the substrate is inferred from the heating and/or cooling curves of a material exposed to the source of

heat. The properties of the material must be well known in order to perform an accurate measurement. These properties have been used for estimating heat fluxes in glow discharges processes. For instance, Ball [362] used a temperature sensitive probe consisting on a tantalum disk connected to a thermally isolated thermocouple on a ceramic base. The probe was used to measure the heat flux in a diode sputtering system to estimate that about 40% of the power delivered to the substrate was carried out by secondary electrons. He used relatively short times to calculate the initial rate of change of temperature to estimate the heat flux since he was interested in reducing the radiation from the target. Thornton [363] used a similar probe made of stainless steel to measure the heat flux in a cylindrical post magnetron configuration. The heat flux was obtained from the heating and cooling curves of the probe similarly to Ball's process. Using this method, Thornton estimated the plasma radiation using a simple model whereby half of the plasma radiation ends at the at the substrate. According to his measurements and calculations, the percentage of heat flux per atom due to radiation ranged from about 13% for Cd to 42% for Vd. Thornton also found a relatively weak dependence of the energy per deposited atom as a function of power. Kersten et al. [364] improved the Thornton design to measure the heat flux in an magnetron system used to deposit molybdenum [364], copper [120], tungsten [120] and highly amorphous hydrogenated carbon films [365]. Their measurements were in agreement with a simplified model that considered all heat sources during the process. Similar probes have been used for measuring heat fluxes in HiPIMS. Another example of indirect probes used for sputtering is found in Ref. 366. Ekpe and Dew [361, 367, 368] used the change of resistivity of a thin polysilicon array to build an integrated heat flux sensor which has been successfully implemented to measure heat fluxes in different chamber configurations. According to the authors' calculations, the plasma accounts for as much as 65% of the total heat flux at the substrate [361], while sputtered atoms were found to contribute only about 23% of the total flux when aluminum is sputtered at 5 mTorr in a small circular planar magnetron [361]. A big disadvantage of indirect methods is the long measuring times associated with the heating and cooling curve of

the material. In some cases, it may take one hour to get a singe point of a typical heat flux curve.

In contrast, in a direct sensing method, a transducer is used to convert heat flux into a measurable electrical signal. This type of transducer has been used to measure the contributions of different sources for the total heat flux. Thomann et al. [369] used an active sensor to estimate the component of the heat flux due to the plasma bombardment in an argon plasma. The authors found a linear heat flux dependence with respect to the rf power. Within the same research group, Bedra et al. [370] measured the contribution of sputtered particles to the total heat flux. They found that the sputtered particles contribute to at most 5% of the total power with the plasma species responsible for most of the heat load to the sensor. However, their setup was not representative of a magnetron discharge and these results should be analyzed in the context of the experiment. Other examples of heat flux measurements using an active sensor in a glow discharge include [369, 371, 372]. Disadvantages associated with direct methods are the high price of the sensing device and the electronics needed to handle the small electrical signal. Furthermore, the range and sensitivity needed for accurate measurements requires small sensors. Such small sensors are more sensitive to damage and extra care should be taken when placing them in a sputtering chamber. For instance, the sensor should be fully protected from any metal flux to protect it from irreversible damage.

7.3   Design of a heat flux measurement device

The device presented in this section is based on an active sensor. The short acquisition times commonly associated with this type of sensors allows the experimenter to perform a comprehensive investigation of the effect of several process parameters on the heat flux load at the substrate level. A fast measurement is required to span a wide range of the control variables, keeping the length of the experiment within a reasonable duration. In a manufacturing

environment, there is only a limited amount of time available for such experiments and having a fast measurement technique is essential for obtaining as much information as possible with minimal impact on production.

The active sensor element is a thin thermopile (1 in x 0.5 in) element manufactured by RdF corporation (pn 27134-3). In brief, the sensor is an array of several thermocouples connected in series forming a thermopile. This array produces a voltage that is proportional to the gradient of temperature between the two faces of the sensor. This proportionality factor is characterized by a sensitivity which is then used to convert the small electrical signal delivered by the sensor to real heat flux units. The sensor is mounted on a water cooled copper base to keep one face of the sensor at a constant temperature. To protect the sensor from the metal flux, a thin metal cover was placed on top of the sensor. The thin aluminum cover as the shielding material offered several advantages over other alternatives investigated. The choice of aluminum was based on availability, price and malleability. All surfaces are glued together using an insulating thermal adhesive (bond-Ply-100) given its high thermal conductivity and electrical isolation properties.

The voltage delivered by the sensor is connected to a USB data acquisition system (DAQ DI-718) which is equipped with an industrial amplifier module (8B-30) both from dataQ corp [1]. The DAQ is connected to a PC running a Windows operating system to capture the signal using the software provided by dataQ. The small signal delivered by the sensor is then converted to real units using the expression

$$H = \frac{S_s h}{a} \tag{7.1}$$

where H is the total heat flux in mW/cm^2, S_s is the sensitivity of the sensor as provided by the manufacturer in mV/W/cm^2, h is the small signal delivered by the sensor in mV and a is a proportionality factor which is a function of

---

[1]www.dataq.com

the area ratio between the sensor and the sensor cap (taken as 1.75 in these experiments).

One disadvantage of this device is the low sensitivity to the plasma irradiation. Since the sensor is in direct contact with the metallic flux, the sensor shield is rapidly coated with a thin silver film within seconds of igniting the plasma. This silver film (silver is highly reflective to thermal radiation) reflects part of the spectrum associated with thermal radiation from hot surfaces and some radiation from the plasma. The sensor may be collecting some UV radiation from argon resonant states since the reflectivity of silver significantly drops for radiation with a wavelength below 400 nm. Therefore, values herein reported should be considered as representative of the deposition conditions. The process to quantify the radiation from the plasma is beyond the scope of this chapter.

Typical response times for the device as a whole range from 10 to 25 seconds per measurement. All measurements reported in this chapter were obtained using a grounded sensor cap.

7.4 Results and discussion

7.4.1 Experimental setup

Three different magnetrons were used for the experiments reported in this chapter. The first magnetron, hereafter referred as Magnetron A, is a custom-made rectangular magnetron manufactured by Sierra Applied Science [2] with a patented magnetic field arrangement. The second magnetron, hereafter referred as Magnetron B, is an in-house rectangular magnetron specially designed to modify the magnetic field by moving the plate where the magnets are mounted away from the target surface. A schematic of magnetron B is shown in Figure 7.1 and is fully described elsewhere [345]. Both magnetrons

---

[2]www.sierraapplied.com

house a 36.48 cm x 10.88 cm x 0.6 cm silver target (99.99%) which was used as the sputtering material. Lastly, the third magnetron configuration was the 3″ gun and vacuum chamber fully described in Chapter 4.
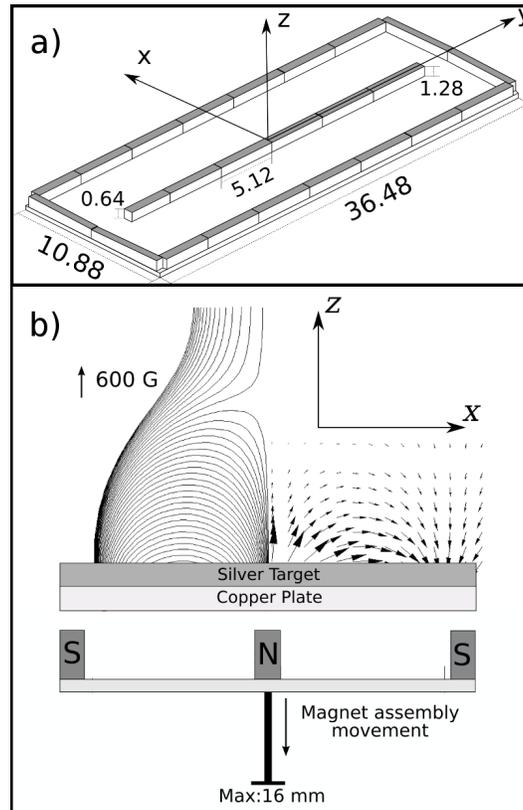
Figure 7.1: **In-house rectangular magnetron with movable magnetic field (Magnetron B)**

The experiments using magnetrons A and B were carried out using a 177 L semi-cylindrical stainless steel vacuum chamber which was evacuated to a base pressure $< 5 \times 10^{-6}$ Torr. Argon was introduced into the chamber and its flow rate was adjusted to 300 sccm. The chamber pressure was regulated at 32 mTorr by throttling the high vacuum valve. All walls in the chamber were water cooled and kept constant at ≈ 15 °C. The substrate of the experiment was placed parallel to the target surface and the throw distance adjusted according to the experiment. For the experiments using magnetron A, the throw

distance was varied in the range 5-12 cm, while for the other two magnetrons it was kept constant at 10 cm.

A typical measurement taken with the heat flux device at the center of the substrate is shown in Figure 7.2. Initially, the heat flux is close to zero when the plasma is off (point A). This small value ($\approx$ 1.5) could be attributed to the small temperature difference between the sensor and the walls of the system. A low signal has also been observed in other experiments using a similar sensor [369] and attributed to the thermal radiation from the walls of the vacuum chamber. However, in contrast to the experimental setup of Ref. 369, the experiments reported in this chapter used water cooled walls thus minimizing the thermal radiation signal. As soon as the plasma is ignited (point B), the signal rises with a time constant that depends on the device design. Although the sensor is relatively fast (0.5 s to reach 63% of a step function according to the data sheet), the response time of the device as a whole is increased given its higher thermal mass. After about 20 s, the signal reaches steady state: it is at this point that a measurement is taken (point C). Next, the discharge current is increased which results in an instantaneous change in the value measured by the heat flux sensor (point D) until a steady state associated with the new discharge conditions is reached. If the discharge is extinguished, the signal from the sensor drops rapidly (point E). During this process the temperature of the sensor is monitored to account for any change in the sensitivity of the sensor as a function of temperature (also provided by the manufacturer).

Deposition rates were measured using a crystal microbalance at the same locations on the substrate where the heat flux measurements were taken.

7.4.2 Magnetron A

Magnetron A was used to measure the effect of current and throw distance on the total heat flux at the substrate. Results from these measurements are plotted in Figure 7.3. As expected, the heat flux increases as discharge power
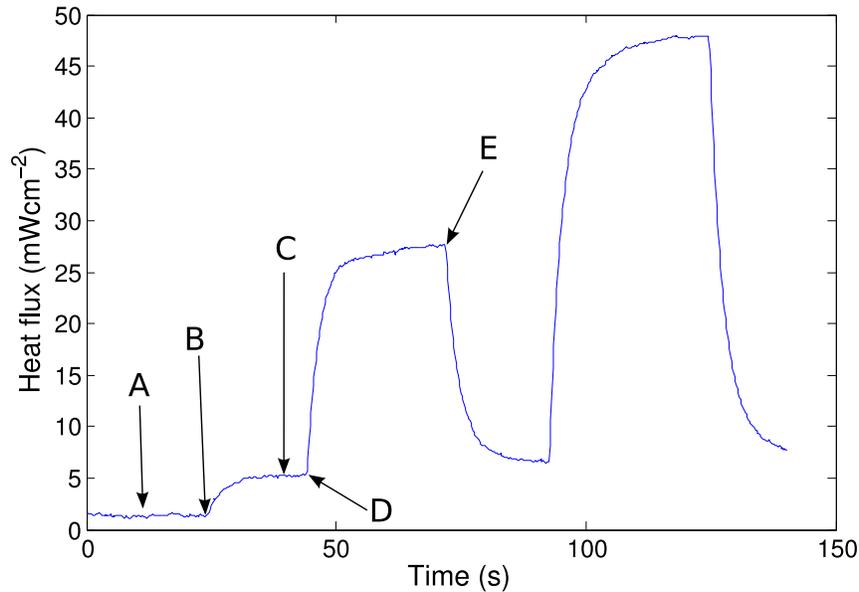
Figure 7.2: **Typical waveform of an experiment.**

is increased due to the higher plasma density and particle flux associated with a higher power process.

Another important parameter of the process is the total energy per deposited atom (EPA) which is a measure of the total energy available per particle at the substrate as a results of all contributing heat sources. To obtain this value, I measured the deposition rate for all points plotted in the previous figure using the same location where the heat flux was measured. Results from these measurements are shown in Figure 7.4.

Figures 7.5 and 7.6 show the EPA as a function of power and throw distance resulting from combining combining Figures 7.3 and 7.4. When the EPA is plotted as a function of power as in Figure 7.5, the EPA is reduced for increasing power. At low currents, there is a discrepancy of about 20 eV between the maximum and minimum registered values but this difference is reduced as power is increased. The reduction of the EPA observed in the figure is consistent with results presented in Ref. 373 and 361 for a cylindrical planar magnetron. The fact that this trend is observed in both cylindrical and
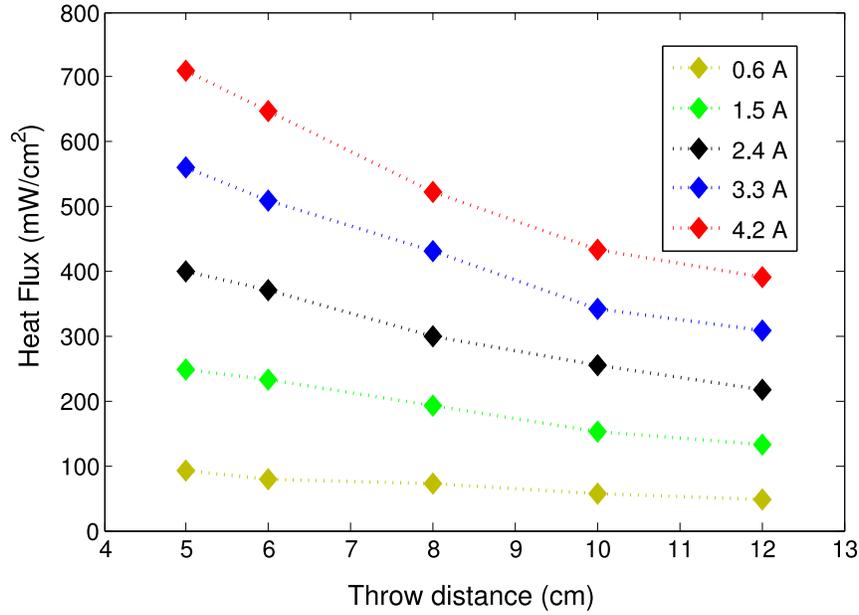
Figure 7.3: **Heat flux as a function of the throw distance for 5 discharge currents. The heat flux diminishes almost linearly in the range 5-8 cm and shows a slower rate in the range 8-10 cm.**

rectangular planar magnetrons for different materials may suggest that this is a characteristic of discharges using planar targets. However, it remains to be determined if the trend is also observed in nonplanar configurations.

Results from Figure 7.5 imply that the particle flux changes at a different rate than the energy flux as a result of a change in the discharge current. The different rates of change may imply that sputtered particles do not contribute significantly to the total energy input for this process conditions. A simple calculation shows that this is indeed the case. Sputtered particles contribute to the total energy at the substrate through the heat flux due to their kinetic energy J_{KE} and to their enthalpy of condensation, J_{HC}. These two contributions can be roughly estimated as $P_{KE} \cong J_n E_{ave}$ and $P_{HC} \approx J_n \Delta H_{Ag}$ where J_n is the particle flux reaching the substrate and ΔH_{Ag} is the enthalpy of condensation. If we assume that silver particles are fully thermalized at this
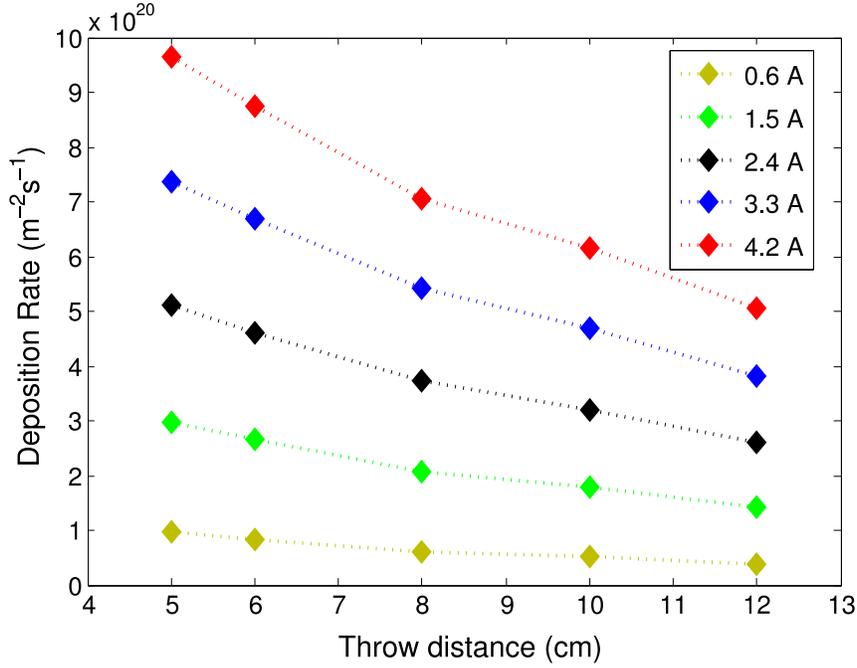
Figure 7.4: **Deposition rate measurements at the center of the substrate. These values are then used to calculate the energy per deposited atom as a function of pressure and power.**

pressure, then $E_{ave} \approx 0.03$ eV. Considering for the sake of the argument a deposition process at 32 mTorr, discharge current of 1.5 A and throw distance of 10 cm, we obtain from Figure 7.4 $P_n \approx 2 \times 10^{20}$ m^{-2}s^{-1} yielding $P_{KE} < 1$ and $P_{HC} \approx 8.32$ both in mWcm^{-2}. For the latter calculation, $\Delta H_{Ag} \approx 2.6eV$ [374]. From Figure 7.3 we see that the total heat flux is about 180 mWcm^{-2} indicating that the sputtered particles have a negligible contribution to the total energy. In this simple analysis, reflected neutrals have been assumed to be fully thermalized due to the following. At 32 mTorr, $n_{Ar} = 1 \times 10^{21}$ m^{-3}, $\lambda_{Ar} = 0.1/\sigma_{Ar}$ for σ_{Ar} in Å^2 (using the ZBL interatomic potential described in Chapter 4). For an average energy in the range 10-20 eV (see Chapter 4), $\sigma = 20 - 30$ Å^2 yielding $\lambda_{Ar} = 3 - 5$ mm. Since $\lambda_{Ar} << 10$ cm, reflected neutrals are fully thermalized and do not contribute to the heat flux at this pressure. A very similar analysis for silver atoms would show that their mean
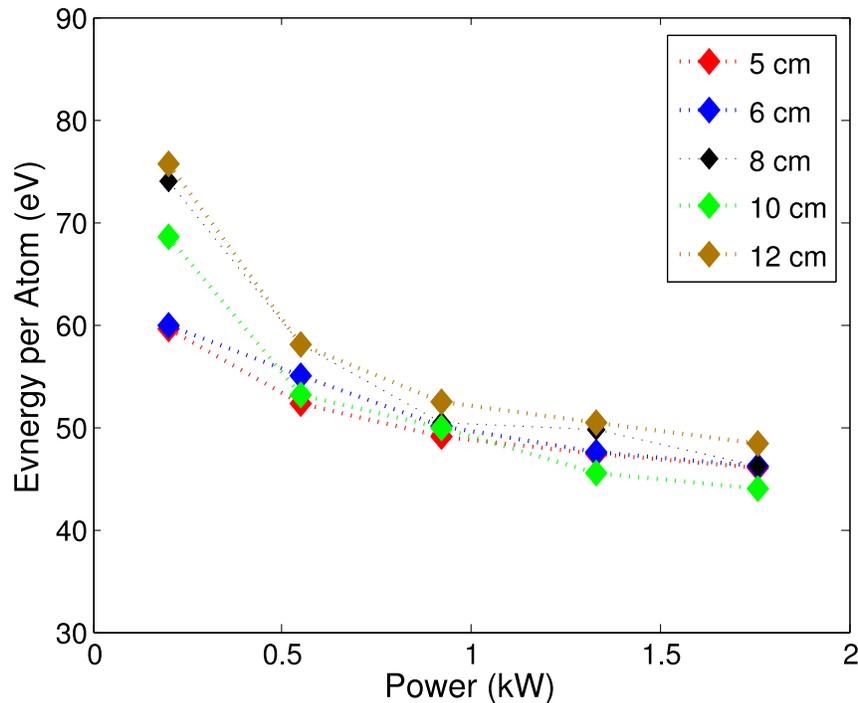
Figure 7.5: **Energy per deposited atom as a function of power. The experiment shows that the total available energy diminishes as power in increased. The consequence is that different crystal structures can be forming in the conditions are right as a result of increasing power.**

free path is also very small compared with the throw distance. This means that the total heat flux is mostly determined by the plasma and that the deposition rate plays a minor role in the total heat flux.

Figure 7.6 shows the change in the energy per atom as a function of the throw distance. The figure shows how the EPA, on average, is almost constant for all plots with the exception of the low current plot which shows a small fluctuation. The plot indicates that both the energy and particle flux change at the same rate as the substrate is brought closer to the target. The plot is more difficult to explain using a simple model since in this case we would need to consider the change in the plasma density and temperature as the substrate

is moved towards the region where the plasma density is greater.
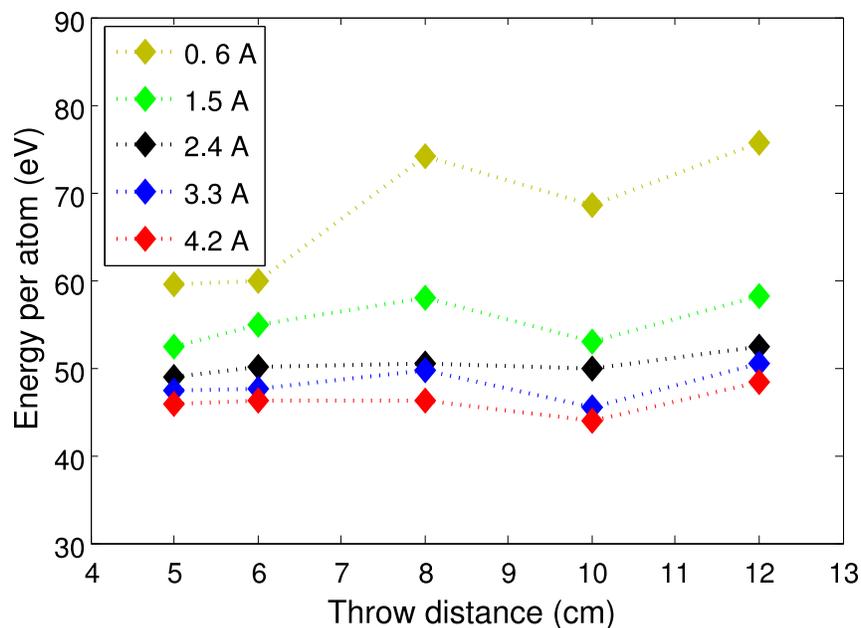
Figure 7.6: **Energy per deposited atom as function of throw distance for five discharge currents.**

The change of the EPA for different discharge characteristics has important implications in the growing dynamics of the film. Unfortunately, the analysis of the as deposited silver films may not be a representative example since we may need to significantly change one process variable to observe a significant change in the EPA that would affect the growing of the films in a quantitative manner. A process more suitable to describe the effect of the EPA on the growth dynamics is reactive sputtering. In this method, a small change in one of the process variables can induce a significant modification of the growing process at the substrate. During a typical reactive sputtering process, reactions at the substrate are promoted by introducing a gas to react with depositing flux. For instance, aluminum nitride films can be deposited in a mixture of argon and nitrogen gases, and the composition and crystal structure of the films can be controlled by selectively choosing the appropriate process conditions. To investigate the effect of EPA on the growing dynamics

of reactively deposited AlN films, the chamber described in Chapter 4 was used to deposit AlN films on Si (100) substrates [375] using a mixture of argon (99.9999) and N_2 (99.995) as process gases. The substrate was placed 10 cm above the magnetron gun which was operated in the constant power mode. The deposition pressure was kept constant at 5 m Torr by throttling a high vacuum vacuum valve for a particular $Ar - N_2$ combination. The Ar/N_2 proportion was controlled by changing their respective flow rates.

The film composition analysis was performed using an electron probe microanalyzer, model JEOL 8900, equipped with five wavelength dispersive spectrometers, an energy dispersive spectrometer and cathode luminescence detector. The surface morphology and the cross-section of the film were analyzed using a JEOL 6302F field emission scanning electron microscope (SEM). The thicknesses of the deposited films were measured from the cross-sectional SEM images. The type and structural orientation of the as-deposited AlN film were obtained with a Rigaku rotary anode x-ray diffraction (XRD) using Co K_α radiation.

The discharge voltage and stoichiometry of the as-deposited AlN films is shown in Figure 7.7. In Figure 7.7(a), the discharge voltage as a function of the N_2 flow clearly indicates a negligible hysteresis effect in the process [376]. Hysteresis is observed in some reactive magnetrons, and it is commonly attributed to several factors such as 'poisoning' of the target. The discharge voltage follows the expected trend for this target-gas combination [197]: the discharge voltage drops as more N_2 is introduced into the system. The voltage discharge response to changes in the reactive gas flow into the system (O_2 or N_2) is commonly divided into two groups: those for which the voltage increases and those for which the opposite occurs. This type of behavior is attributed to the different γ_e for some materials when a thin oxide/nitride film appears on top of the target [197]. For the AlN films deposited using this setup, the voltage discharge shown in Figure 7.7(a) can be explained by a higher γ_e of the thin AlN formed at the target by the following reasoning. For a discharge in the constant power mode of operation, a higher γ_e means more secondary electrons

entering the discharge which increases the total current collected at the target. The power supply then reacts by lowering the voltage delivered to the cathode to keep the power constant. As more N_2 enters the system, the nitride layer at the target grows thicker and γ_e also grows which in turn reduces the discharge voltage even further until the discharge enters the fully nitride mode at about 30%. A this point the nitridation of the target is in equilibrium with the flux of ions eroding the surface layer.

Figure 7.7(b) shows the composition of the AlN as a function of the N_2 proportion obtained by EDS. At low N_2 flows, the aluminum flux reaching the substrate is larger than the amount of available N_2 to react and form stoichiometric AlN. However, as more N_2 introduced into the system, the amount of available N_2 increases and stoichiometric AlN starts to grow. The difference between the 100 W and 200 W plots is due to the dissimilar metal fluxes reaching the substrate at these two powers. For instance, at 0.2 N_2, doubling the power from 100 to 200 W means that there is a higher aluminum flux reaching the substrate; however the flux of N_2 is not expected to significantly change thus lowering the relative N concentration in the film.

The X-ray diffraction spectra of films grown at three different N_2 proportions is shown in Figure 7.8. At 10% N_2 (Figure 7.8a), the film shows a mixture of metallic Al(111), cubic AlN(101) and AlN(111). Visually, films growth under this condition have a black tarnished appearance. As the N_2 flow is increased to 20%, the stoichiometry of the film is improved as Figure 7.8b shows. The crystal phase having the fastest growing rate under this process conditions is the (002). As more N_2 is put into the system, other planes' growth rates improve, in particular the (100) and (101) planes until a fully stoichiometric polycrystalline AlN films is grown. Visually, these films are transparent to the naked eye.

The different crystal orientations for 20 and 50% could be explained in terms of the energy per deposited atom. Figure 7.9 shows both the total energy and the EPA for the 100 W films. The total energy includes the energy from sputtered particles, plasma and surface reactions at the substrate. Data from
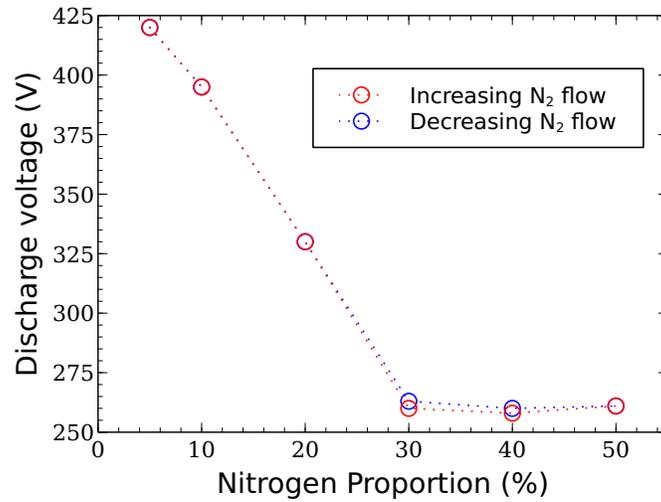
(a) Discharge voltage
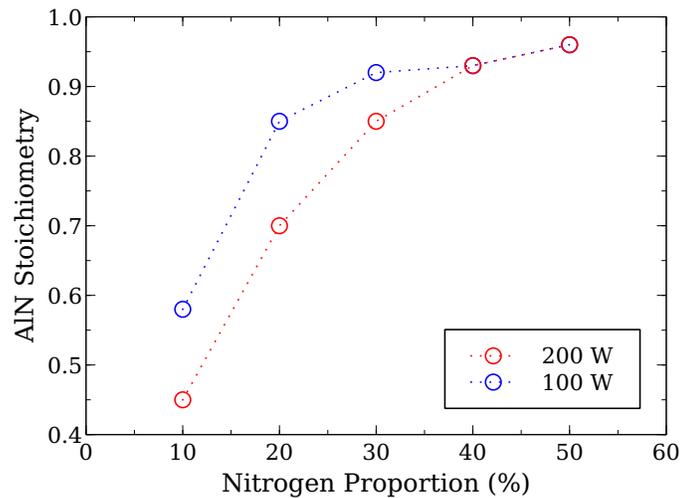
(b) AlN stochiometry

Figure 7.7: **AlN films deposited at 5 mTorr. a) The discharge voltage as a function of N_2 proportion showing the negligible hysteresis effect during the experiment. b) Stoichiometry of AlN films as a function of N_2 proportion for 100 W and 200 W.**
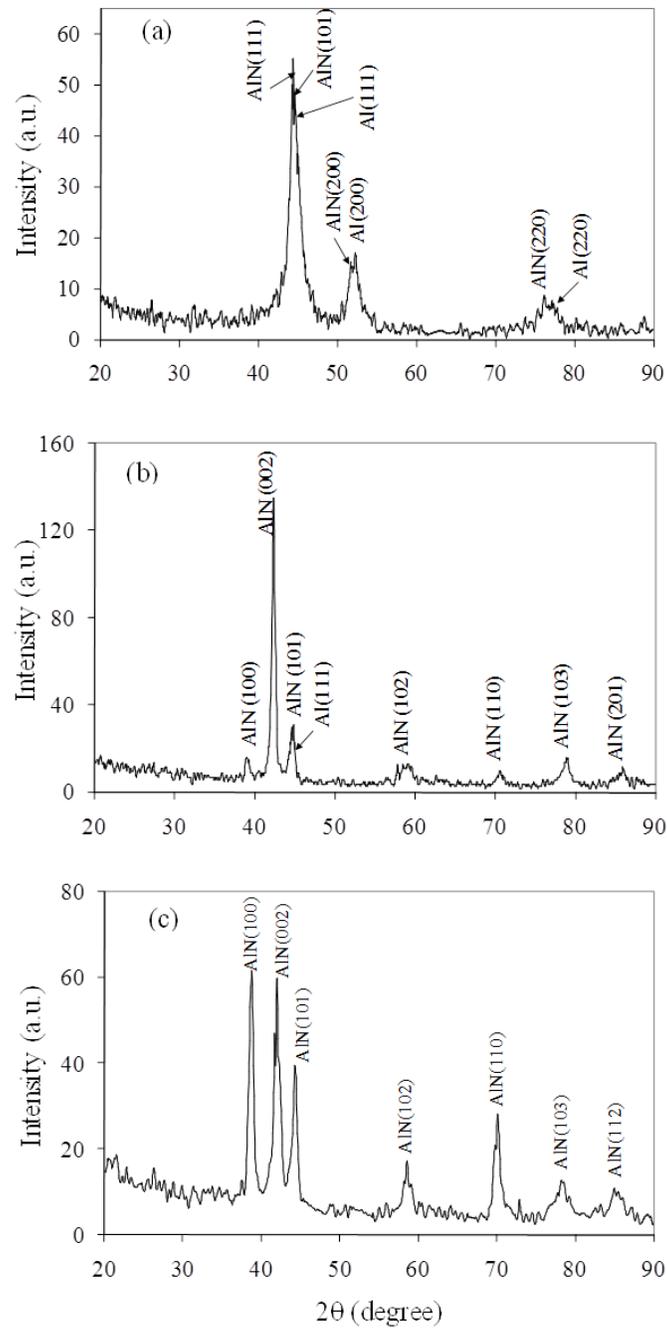
Figure 7.8: **XRD spectra showing a) 10% , b) 20% and c) 50% N_2 proportion of the gas (5 mTorr total pressure).**

Figure 7.9 shows that the total energy grows as more N_2 is introduced into the system. However, as the N_2 proportion reaches the point where stoichiometric AlN is grown, the total energy flux drops. On the other hand, the EPA does not show the same transition and continues to monotonically increase as more N_2 is put into the system. The main reason for the EPA to increase despite the total energy trend is due to a significant reduction of the deposition rate as the discharge enters the fully poisoned mode (not shown). At the low N_2 composition, the high kinetic energy of sputtered particles and the relatively low N_2 flux at the substrate seems to favor the growth of metallic Al(111) and surface planes with a high formation energy such as the AlN(101) [377]. When the N_2 proportion is increased, the peaks associated with the metallic Al and the AlN(101) planes strongly diminish and the preferred growth orientation changes to AlN(002). The c-axis of such planes is oriented perpendicular to the substrate suggesting that the high kinetic energy of the sputtered particles and a relatively low EPA promotes the growth of this plane. The same transition has also been observed in other works [378–381]. Finally, further increase in N_2 flow changes the crystal planes observed in the XRD spectra. Since the EPA continues to grow, coupled with the fact that sputtered particles arrive at the substrate with a lower kinetic energy (note how the discharge voltage has substantially been reduced from the non reactive mode), suggests that the growth dynamics enters a different regime. The appearance of the AlN(100) and (100) peaks at 50% N_2 suggests that this plane is promoted by a low deposition rate with a high EPA. Since the c-axis of the (100) plane is parallel to the the surface, the higher EPA may be promoting surface diffusion. Some physical properties of the film depend in the crystal orientation of the films, and it is expected that films grown at 20 and 50 % N_2 have different properties.

7.4.3   Magnetron B

To investigate the effect of a varying magnetic field, magnetron B was used to obtain a two dimensional map of the heat flux at the substrate. The data collection process was assisted by an xy PC-controlled probe positioning sys-
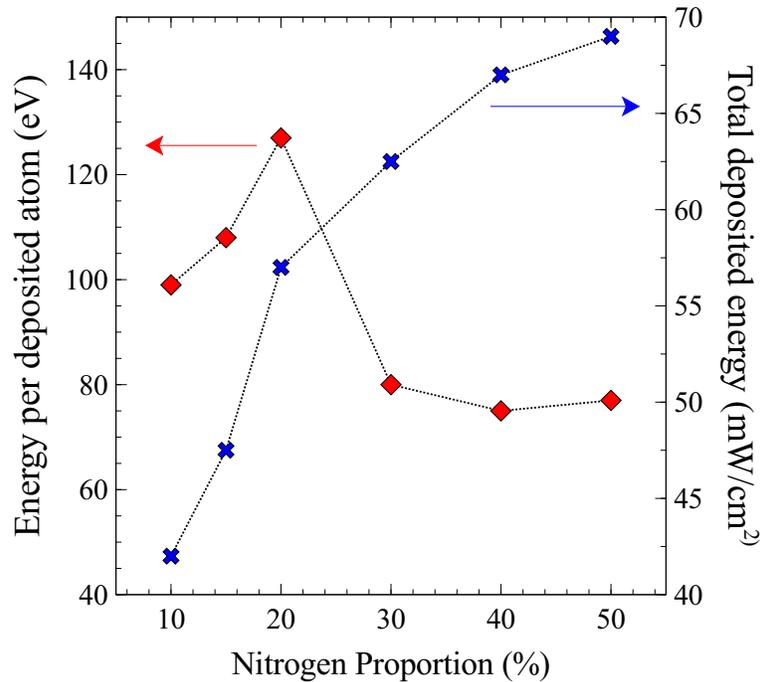
Figure 7.9: **Energy flux (total and per atom) as a function of N$_2$ proportion in the gas (5 mTorr total pressure).**

tem that was placed above the substrate. With this arrangement, the location of the sensor was externally modified without the need to vent the chamber. This setup significantly reduced the time needed to perform the experiments.

The magnitude of the magnetic field was changed by modifying the distance between the magnets and the surface of the target as indicated in Figure 7.1. The whole magnet assembly was moved downwards by 2 mm steps producing a set of 4 different magnetic fields. Hereafter, the magnetic fields labeled as B1, B2, B3 and B4 represent a displacement of 0, 2, 4 and 6 mm in the z direction according to Figure 7.1. The magnetic fields along the x axis for all four configurations are plotted in Figure 7.10. The magnetic fields were obtained using the COMSOL Multiphysics suite[3]. Only the x component of

---

[3]www.comsol.com

the vector field is plotted for ease of visualization.
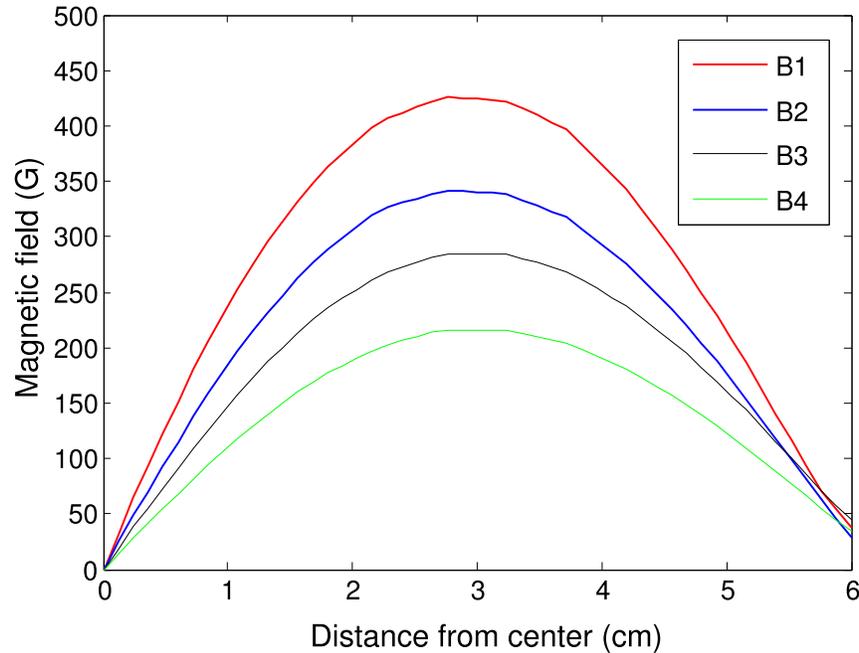
Figure 7.10: **Calculated magnetic fields used to investigate the effect on the total heat flux at the substrate.**

The heat flux as a function of the magnetic field for the same rectangular magnetron has been previously determined [345] using the integrated sensor described in Ref. 368. The measurements reported in Ref. [345] were taken only at two locations on the substrate. A direct comparison between results obtained using the micro-electro-mechanical sensor and the device presented in this chapter is shown in Figure 7.11.

The difference between both measurements, although small, is of about 15%. The discrepancy on the measurements could be explained if we consider differences between the measurement techniques. The first is due to the uncertainty on the correct proportionality factor a in Equation 7.3. It is possible that this value is in reality smaller than the estimated value I used for the experiment. This value was calculated using the active area of the sensor cap divided by the sensor area and considering the bending at the edges of the sensor cap.
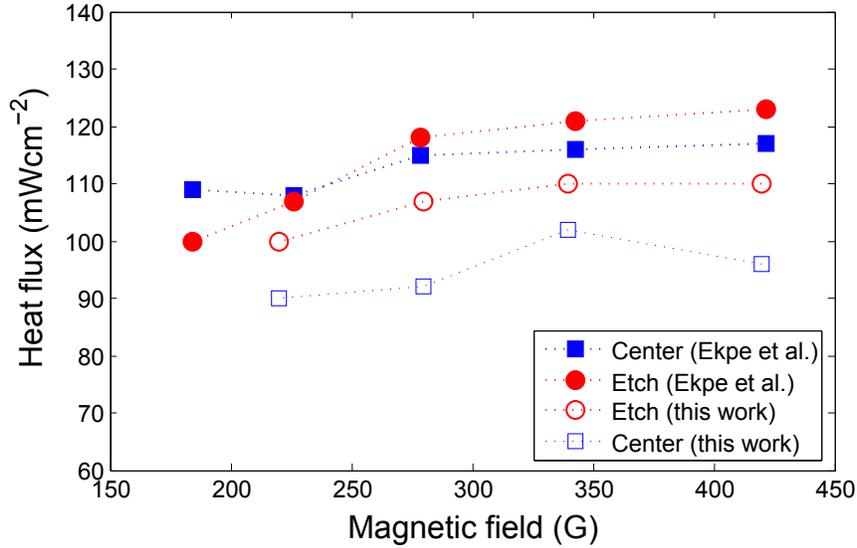
Figure 7.11: **Sensor comparison with a previous, alternating design by Ekpe et al. [368].**

The bending angle for all edges of the cap was estimated to be about $60°$. However, even a small change of $10°$ (bending angle of $70°$) would be enough to explain a 10% change in the value measured by the sensor since a would change from 1.75 to 1.53. Another source of error could be attributed to the cooling of the sensor. It was noted during the experiment that the cooling water was not constant during the measurements, presenting a peculiar and periodic fluctuation of about $\pm 3°C$ with a period of about 4 minutes. The temperature signal presented zones of relatively slow temperature changes followed by zones where the temperature changed faster. This temperature fluctuation was enough to change the output of the sensor by about 10 mWcm^{-2} when the sensor was resting in the chamber with no heat input. Although some measures were taken to minimize the effect of the cooling water, it could be the case that some measurements present an offset of $\pm$ 5 mWcm^{-2}. These two effects, when added, could explain the small offset between both measurement techniques observed in Figure 7.11.

The effect of the changes on the magnetic field on the total heat flux measured

along the width of the substrate (x in Figure 7.1) is plotted in Figure 7.12 for three discharge currents. For the four magnetic fields investigated, the difference in total heat flux is relatively small but qualitative differences can be observed in the plots. In general, results for magnetic field B1 show a higher and less uniform heat flux along the width of the substrate. The nonuniformity grows larger as current is increased as can be observed if we compare results for magnetic field B1 in Figures 7.12(a), 7.12(b) and 7.12(c). In contrast, the shape of all plots for magnetic field B4 are almost uniform for all three currents in the range $0-6$ cm. These results suggest that a lower magnetic field is preferred for a more uniform heat flux at the substrate. As shown in Figure 6.4(b), the magnetic field B4 is almost half of B1.

The heat flux was measured at several locations on the surface to investigate lateral and longitudinal variations of the heat flux. Figure 7.13 depicts the results from these measurements. It is clear that for B1 the energy flux shows a strong nonuniformity which gradually disappears as the magnetic field is reduced to B4. This bulge in the heat flux can cause local damage to the substrate. In general, the observation of a less uniform heat flux for the higher magnetic field is confirmed by Figure 7.13.

To obtain the EPA for all four magnetic fields, the deposition rate was was measured and is shown in Figure 7.14. This figure confirms the experimental results reported in Ref. [345] where it was observed that the deposition rate was improved by lowering the magnitude of the magnetic field.

Finally, the energy per deposited atom as a function of the magnetic field for all four magnetic field investigated in this chapter is shown in Figure 7.15. Based on the results of Figures 7.13 and 7.14, we conclude that for this particular geometry and operation pressure, the use of a relatively low magnetic field results in similar deposition rates and a more uniform heat load to the substrate when compared to a magnetic field magnitude typical of magnetron sputtering configurations (300-400 G).
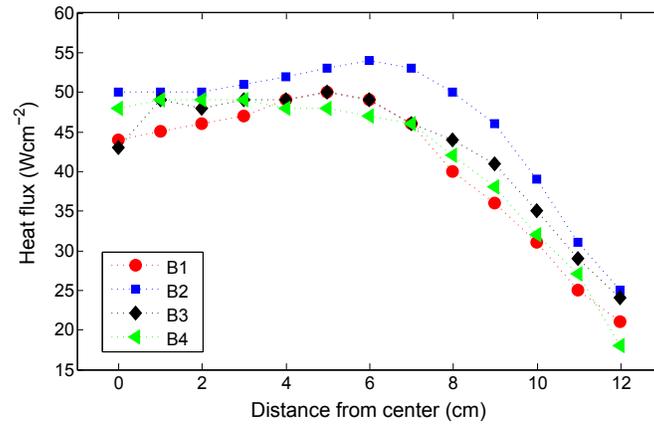
(a) 0.5 A
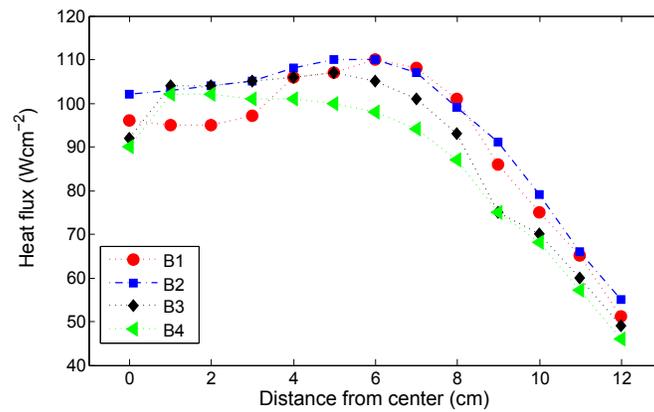
(b) 1 A
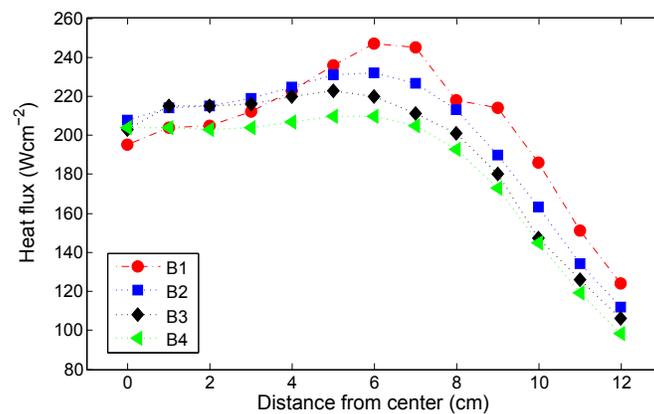
(c) 2 A

Figure 7.12: **Heat flux along the width of the substrate for four magnetic fields. The operating pressure was set to 32 mTorr with a throw distance of 10 cm currents of a) 0.5 A, b) 1 A and c) 2 A were used.**
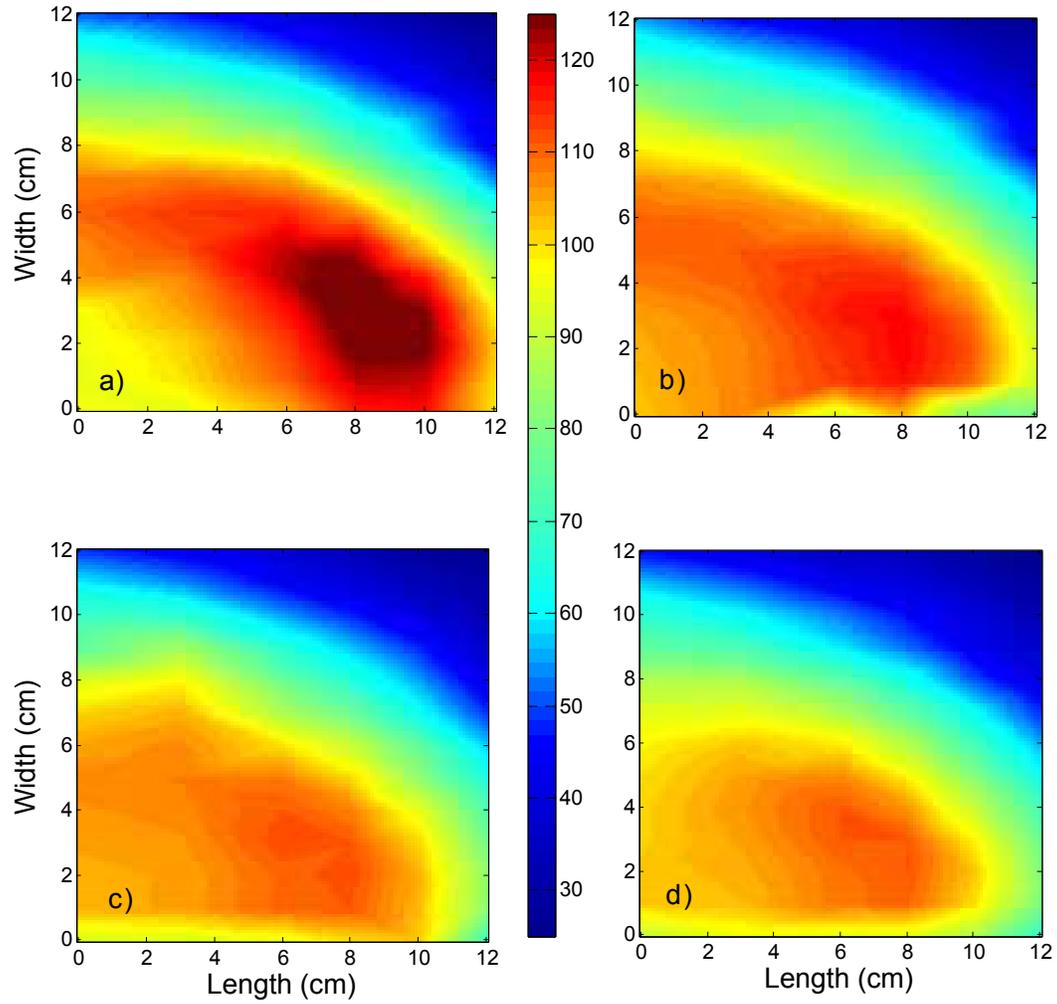
Figure 7.13: **Heat flux measurements (in W/cm^2) for four magnetic fields at 32 mTorr and 1A. The throw distance was kept constant at 10 cm.**
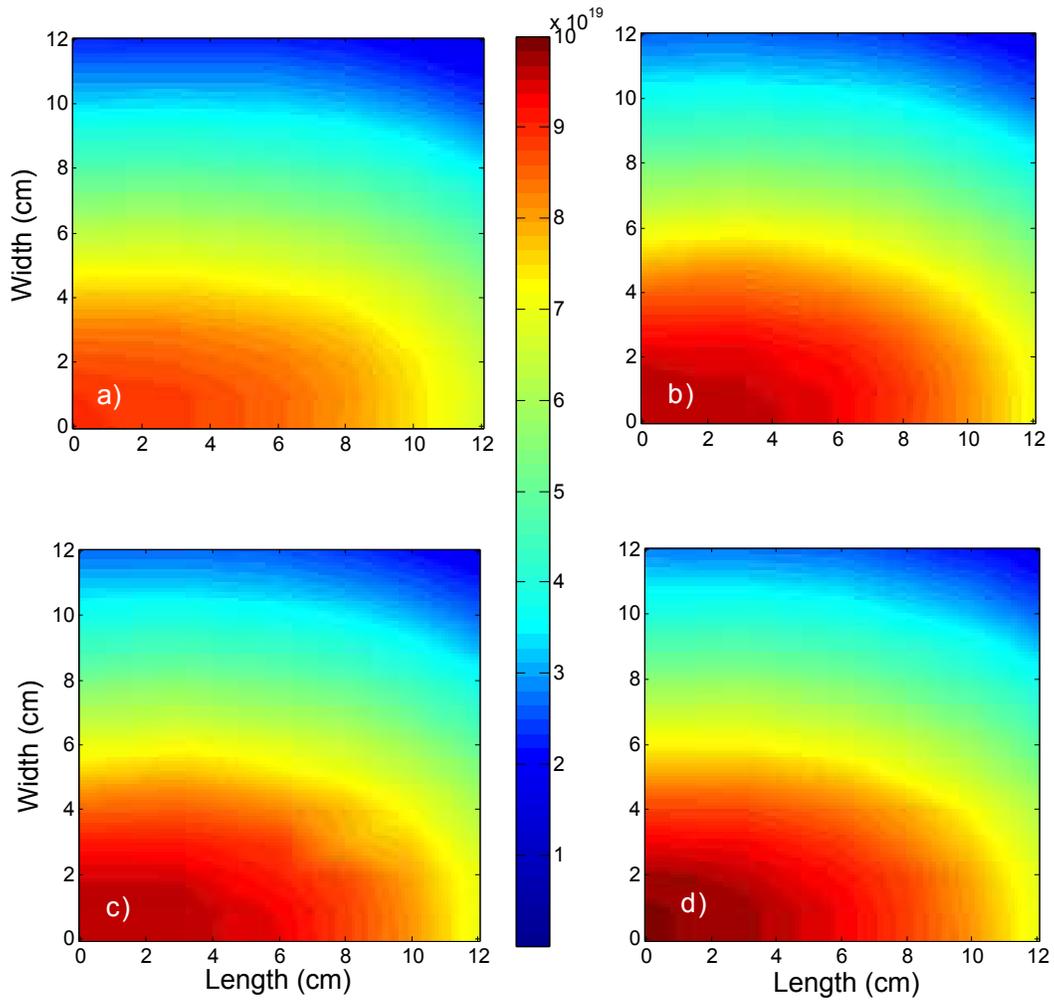
Figure 7.14: **Deposition rate (in** atoms m^{-2}s^{-1}**) for four magnetic fields at 32 mTorr and 1A. The throw distance was kept constant at 10 cm. Magnet configurations are a) B1, b) B2, c) B3 and d) B4.**
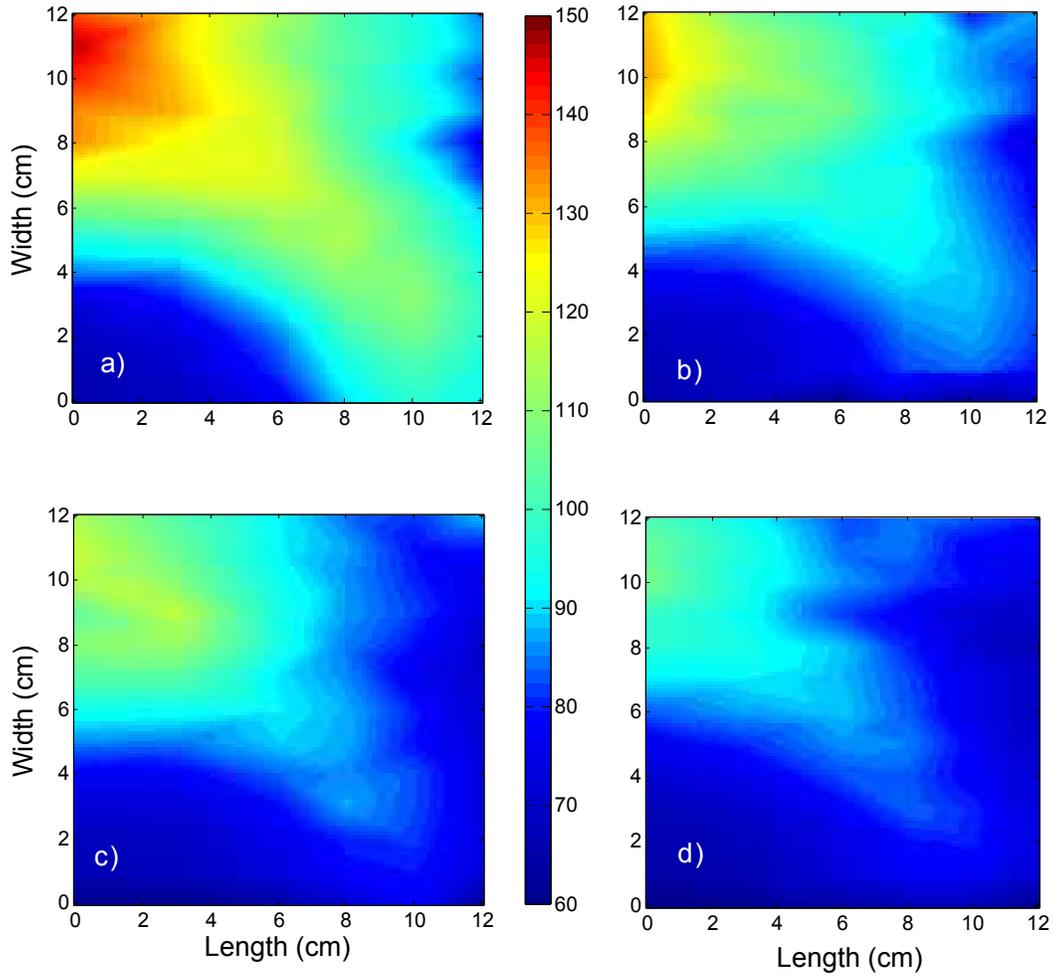
Figure 7.15: **Energy per deposited atom (in eV) for four magnetic fields at 32 mTorr and 1 A. The throw distance was kept constant at 10 cm. Magnet configurations are a) B1, b) B2, c) B3 and d) B4.**

7.5 Summary and conclusions

In this chapter, I described the design and implementation of a heat flux sensor amenable for low pressure gas discharge experiments. The device was used to measure the heat flux at the substrate level for two types of rectangular magnetrons. The effect of current, throw distance and magnetic field strength were investigated for a wide range of the control parameters.

The heat flux was found to change linearly as a function of the discharge current, while a quasilinear response was observed when plotted as a function of the throw distance. The energy per deposited particle was found to diminish as a function of increasing power when the discharge is operated at 32 mTorr. The same was also observed at low pressures (5 and 10 mTorr) on a circular planar target suggesting that this is indeed a characteristic of processes using planar magnetrons. Changes in the plasma as power is increased were found to be the main mechanisms for the observed reduction of the EPA. Sputtered particles were found to have a low contribution to the total energy flux using a simple calculation. Finally, the magnetic field was found to affect the heat flux uniformity at the substrate in such a way that the least uniform flux was for the higher magnetic field investigated. By reducing the magnetic field, the heat flux uniformity was improved while the deposition rate increased when all measurements are compared at the same current. These results suggest there is room for improvement in designing more efficient magnetic fields for practical applications. Finally, the energy per deposited atoms was also found to have a better uniformity at the substrate also for the smallest magnetic field of the study.

CHAPTER 8

Conclusions and Recommendations

8.1   Summary

This dissertation presented a three-dimensional comprehensive model of the magnetron sputtering discharge that incorporates all aspects of the process in a computational framework. To my knowledge, this is the first such comprehensive model. The presented model couples most phenomena occurring in the discharge using a scheme that breaks down the process into several parts and abstracts them into highly specialized and independent modules. Each module, working as an individual component, is in itself a fully featured simulator. The communication and operation between these single units, and the development of one of the key modules in the framework, were at the focus of the work presented in this dissertation. The ultimate goal was to build a virtual environment to perform numerical experiments and have a better understating of the complex interactions between process variables occurring during a typical operation of the magnetron sputtering discharge.

In Chapter 2, I presented a thorough description of the underlaying principles behind the operation of a typical magnetron sputtering discharge. The chapter covered most aspects of the process placing emphasis on the modeling strategies to capture the phenomena at hand. The computational framework

presented in this dissertation was briefly summarized in Chapter 3. This chapter also covered the basics of the computational architecture of the framework. Two of the most important modules of the project were described in Chapters 4 and 5. In Chapter 4, a detailed description of the neutrals module was presented together with experimental validation of the numerical model. Chapter 5 presented a mathematical description of a magnetized glow discharge and included a description of the robust numerical algorithm developed in this work to solve for this complex problem. Chapter 6 presented simulation results from the computational model of the magnetron discharge presented in this dissertation. Finally, Chapter 7 covered the design and implementation of a heat flux sensor developed for the measurement of this process parameter in industrial deposition systems.

8.1.1 Neutrals module

Results from Chapter 4 indicated that the simulation and experiment were in excellent agreement with experimental data. In particular, gas temperature and metal species densities in the reactor were found to agree with experimental data found in the literature. Predicted deposition rates for a wide range of pressures and powers were found to match those measured in the laboratory. The simulation study presented in this chapter covered a wide range of pressures, powers and target materials for which the model was shown to describe the transport of neutral particles.

8.1.2 Plasma module

The difficulties associated with the numerical solution of the magnetized discharge were discussed in Chapter 5. This complex problem was solved using a robust numerical algorithm applicable for a wide range of pressures. One important aspect of the algorithm presented in Chapter 5 was the partitioning of the magnetized electron flux into two components, each handled differently

in the solver. This aspect of the model was key to the stabilization of the numerical algorithm. In a like manner, the Block-Newton-Raphson solver implemented for the solution of the large system of linear equations in the plasma model was an important component of the solver. This solver allows the PS2 module to handle a grid with a large number of cells in an efficient manner.

8.1.3 Integration

The integration of all modules into a suite of simulators was a complex task to accomplish. This stage not only demanded an extensive knowledge of the computational architecture of the framework and a good understanding of basic software engineering concepts, but also a sound understanding of the fundamental physics involved during a sputter deposition process.

The iteration scheme proposed in Chapter 6 was proved to be sufficiently efficient to iterate all modules in a consistent manner. The comprehensive discharge model presented in this dissertation was able to correctly predict the localized electron density peaks in the region next to the target otherwise not seen using simplified models. Results from the comprehensive discharge model showed that the electron flux should be treated in three dimensions for a correct description of the process. Some disadvantages of the proposed model include long execution times, the need for experimental calibration and the complex interoperability of the software components. In particular, long execution times are a big obstacle to perform parameterized studies and sensitivity analysis. However, the model correctly captures the transport mechanisms of low energy electrons in a glow discharge subject to an inhomogeneous three dimensional magnetic field. I should mention that there were several aspects of the integration that were omitted in this document. Topics such as debugging, testing and optimization of modules were an integral part of the process and demanded a large amount of time. However, their description add little to the quality of the document.

8.1.4   Heat flux measurement

Another important contribution of this dissertation was the development and implementation of a heat flux measurement device for its application in industrial coaters. This device, described in Chapter 7, has been successfully implemented to estimate the heat flux impinging on the substrate in several industrial setups. The study reported in Chapter 7 reports only on measurements using a scaled-down system. The effect of several process conditions on the total heat flux was investigated using two types of rectangular magnetrons. In the first study, the effect of the throw distance, power and pressure is reported. The energy per deposited atoms was found to play a key role in the growth dynamics of the films. This parameter's significance in a reactive processes allowed us to prove its impact on growth dynamics of the film.

8.2   Recommendations for further work

As pointed out in Chapter 5, the mathematical description of the plasma using the fluid model is only a good approximation. More work needs to be done to calibrate the simulation against experimental data to increase the accuracy of the results presented in Chapter 6. Furthermore, a sensitivity analysis with respect to some of the many variables affecting the results predicted by the model would also be required to increase our understanding of the discharge. Such a study, given the computer resources available at the time this dissertation was written, would probably take months to complete using the current implementation of the SpudII program.

The mathematical model described in Chapter 5 has room for improvement. One such improvement involves including a model to obtain a more accurate description of the electron temperature. This model, which could be based on the solution of an electron energy equation or on a direct solution to the Boltzmann equation, could increase the accuracy of the results presented in chapter 6. This modification to the current model, however, comes with a

high price: the complexity of the numerical algorithm would be substantially increased. Ultimately, a compromise must be made. It is uncertain how the numerical model proposed in this chapter would behave when the electron temperature is nonuniform. This topic should be explored with care to avoid breaking the robustness of the numerical algorithm presented in chapter 6.

Another improvement would be to include real time units in the models presented in Chapters 4 and 5 to account for pulsed and radio-frequency discharges. Pulsed discharges have been recently the focus of several research groups, and efforts have been put into understanding the dynamics of the process. Unfortunately, to include real time in the current mode, we would need to modify some of the fundamental concepts on which the PS and TS modules are built upon. Particularly, the coupling between the high and low energy transport modules would need to be substantially revised. However, it is feasible, to a certain extent, to build a simplified model using the available tools that would allow us to use the framework to simulate simple pulsed discharges.

Finally, the long execution times of the iterative scheme proposed in this dissertation could be reduced. There are two modification that would significantly improve the convergence time of the overall iterative scheme. Both changes involve the plasma model, which is by far the most demanding component of the framework.

The first change would involve a modification to the plasma model that allows the PS2 module to use a refined grid in places where it requires better spatial resolution while keeping a coarse grid everywhere else. The current implementation of the SpudII framework does allow for cell refinement, and a numerical scheme that can handle those cells in the interface between a coarse and a refined grid is already implemented in the TS2 module. However, the same model cannot be applied *verbatim* to the plasma model since the exponential scheme used to discretize the plasma equations demands a different, more intricate approach.

The second improvement to improve the execution time involves the implementation of a faster numerical library for all linear solvers, especially those used

in the PS2 module. These faster libraries would most probably rely on parallelization techniques. In the current implementation of the SpudII project, all linear solvers are implemented using the SparseLib++ library which, unfortunately, does not take advantage of multi-core architectures. Libraries such as pARMS and PETSc would be great alternatives to explore. These libraries, in theory, could be integrated in the SpudII project although it is not clear if the integration can be as smooth as with the SparseLib++ library. A different, newer approach to the problem involves the use of the graphical processing unit (GPU) which is found in every desktop computer. The use of GPUs for scientific applications is a recent topic of research and represents a paradigm shift in the field. Unfortunately, this topic is in the early days of development, and there are not many open libraries available at the moment. This limitation forces us to use closed libraries such as CUDA (from NVIDIA) which, unfortunately, are tied to the hardware. There is, however, a recent trend to develop numerical libraries that can run on heterogeneous systems (GPU + multi-core), and projects such as the ViennaCL or MAGMA look promising for implementing the necessary changes into the SpudII framework to bring down the execution times.

Bibliography

[1] Sputtering Targets and Sputtered Films: Technology and Markets (2011).

[2] D. J. Field. *Study of electrons in magnetron sputtering systems for novel thin film applications.* PhD thesis, University of Alberta (2005).

[3] D. Depla, X. Y. Li, S. Mahieu, and R. D. Gryse, J. Phys. D: Appl. Phys. **41**, 202003 (2008).

[4] W. D. Westwood, *Sputter deposition*, AVS (2003).

[5] R. Behrisch and W. Eckstein, *Sputtering by Particle Bombardment: Experiments and Computer Calculations from Threshold to MeV Energies*, Springer, 1 edition (2007).

[6] K. H. Kingdon and I. Langmuir, Phys. Rev. **22**, 148 (1923).

[7] D. M. Mattox, *The foundations of vacuum coating technology*, Noyes publications (2003).

[8] W. R. Grove, Philos. Trans. Roy. Soc. London **142**, 87 (1852).

[9] M. W. Thompson, J. S. Collington, and R. Smith, Philos. Trans. Roy. Soc. London, Ser. A **362** (2004).

[10] M. Faraday, Philos. Trans. Roy. Soc. London **147**, 145 (1857).

[11] A. Wright, Am. J. Sci. Arts **13**, 49 (1877).

[12] T. A. Edison. Art of plating one material with another (1894).

[13] R. L. Boxman, IEEE T. Plasma. Sci. **29**, 759 (2001).

[14] D. M. Mattox. Decorative coating. In D. M. Mattox and V. H. Mattox, editors, *50 Years of Vacuum Coating Technology*, chapter 2, p. 198. Society of Vacuum Coaters (2007).

[15] J. S. Chapin. Sputtering process and apparatus (1979).

[16] R. J. Hill, S. Nadel, and M. Andreasen. History of large area coatings. In D. M. Mattox and V. H. Mattox, editors, *50 years of vacuum coating technology*, chapter 4. Society of Vacuum Coaters (2007).

[17] P. D. D. Davidse and L. Maissel, J. Appl. Phys. **37**, 574 (1966).

[18] M. Wright, J. Vac. Sci. Technol., A **4**, 388 (1986).

[19] B. Window and N. Savvides, J. Vac. Sci. Technol., A **4**, 453 (1986).

[20] S. K. Dew, J. Appl. Phys. **76**, 4857 (1994).

[21] S. M. Rossnagel, J. Vac. Sci. Technol., B **12**, 449 (1994).

[22] V. Kouznetsov, K. Macak, J. M. Schneider, U. Helmersson, and I. Petrov, Surf. Coat. Technol. **122**, 290 (1999).

[23] A. Anders, Surf. Coat. Technol. **205**, S1 (2011).

[24] A. Vetushka and A. P. Ehiasarian, J. Phys. D: Appl. Phys. **41**, 015204 (2008).

[25] W.-D. Münz, M. Schenkel, S. Kunkel, J. Paulitsch, and K. Bewilogua, J. Phys. Conf. Ser. **100**, 2001 (2008).

[26] K. Wasa and S. Hayakawa, *Handbook of sputter deposition system*, Noyes Publications (1992).

[27] J. Lindhard, V. Nielsen, and M. Scharff, Kgl. Dan. Vidensk. Selsk., Mat.-Fys. Medd. **36**, 1 (1968).

[28] W. Moller. Fundamentals of Ion-Surface Interaction. Technical report, Technische Universitat Dresden (2004).

[29] P. Sigmund, Phys. Rev. **184**, 383 (1969).

[30] I. S. T. Tsong and D. J. Barber, J. Mater. Sci. **8**, 123 (1973).

[31] H. Oechsner, Appl. Phys. **8**, 185 (1975).

[32] J.-C. Pivin, J. Mater. Sci. **18**, 1267 (1983).

[33] G. Betz and K. Wien, Int. J. Mass Spectrom. Ion Processes **140**, 1 (1994).

[34] V. S. Smentkowski, Prog. Surf. Sci. **64**, 1 (2000).

[35] G. K. Wehner, D. Rosenberg, and D. Wehner, G. K. and Rosenberg, J. Appl. Phys. **31**, 128 (1960).

[36] J. Bohdansky, Nucl. Instrum. Methods Phys. Res., Sect. B **2**, 587 (1984).

[37] C. E. Carlston, G. D. Magnuson, A. Comeaux, and P. Mahadevan, Phys. Rev. **138**, A759 (1965).

[38] R. Behrisch and W. Eckstein, Nucl. Instrum. Methods **82**, 255 (1993).

[39] D. Mercs, F. Perry, and A. Billard, Surf. Coat. Technol. **201**, 2276 (2006).

[40] N. Matsunami, Y. Yamamura, Y. Itikawa, N. Itoh, Y. Kazamuta, S. Miyagawa, K. Morita, R. Shimizu, and H. Tawara, At. Data Nucl. Data Tables **31** (1984).

[41] C. Garcia-Rosales, W. Eckstein, and J. Roth, J. Nucl. Mater. **218**, 8 (1994).

[42] Y. Yamamura, N. Matsunami, and N. Itoh, Radiat. Eff. Defects Solids **71**, 65 (1983).

[43] Y. Yamamura, Radiat. Eff. Defects Solids **55**, 49 (1981).

[44] W. Eckstein and R. Preuss, J. Nucl. Mater. **320**, 209 (2003).

[45] Y. Yamamura and H. Tawara, At. Data Nucl. Data Tables **62**, 149 (1995).

[46] M. Stepanova. Personal Communication (2006).

[47] W. Eckstein, Vacuum **82**, 930 (2008).

[48] M. Stepanova and S. K. Dew, J. Appl. Phys. **92**, 1699 (2002).

[49] M. Stepanova, Nucl. Instrum. Methods Phys. Res., Sect. B **103**, 33 (1995).

[50] N. Laegreid and G. K. Wehner, J. Appl. Phys. **32**, 365 (1961).

[51] M. T. Robinson and I. M. Torrens, Phys. Rev. B: Condens. Matter Mater. Phys. **9**, 5008 (1974).

[52] M. A. Nastasi, J. W. Mayer, and J. K. Hirvonen, *Ion-solid interactions: fundamentals and applications*, Cambridge University Press (1996).

[53] H. M. Urbassek, Nucl. Instrum. Methods Phys. Res., Sect. B **122**, 427 (1997).

[54] C. F. Abrams and D. B. Graves, IEEE Trans. Plasma Sci. **27**, 1426 (1999).

[55] G. Kornich, G. Betz, V. Zaporojtchenko, A. Bazhin, and F. Faupel, Nucl. Instrum. Methods Phys. Res., Sect. B **227**, 261 (2005).

[56] J. P. Biersack and L. G. Haggmark, Nucl. Instrum. Methods **174**, 257 (1980).

[57] J. P. Biersack and W. Eckstein, Appl. Phys. A: Mater. Sci. Process. **34**, 73 (1984).

[58] K. Wittmaack, J. Appl. Phys. **96**, 2632 (2004).

[59] M. W. Thompson, Philos. Mag. **18**, 377 (1968).

[60] P. Sigmund, Nucl. Instrum. Methods Phys. Res. **B27**, 1 (1987).

[61] A. Brizzolara, Robert, C. B. Cooper, and K. Olson, Nucl. Instrum. Methods Phys. Res., Sect. B **35**, 36 (1988).

[62] A. Goehlich, N. Niemöller, and H. Döbele, J. Nucl. Mater. **266-269**, 501 (1999).

[63] A. Goehlich, D. Gillmann, and H. F. Döbele, Nucl. Instrum. Methods Phys. Res., Sect. B **164-165**, 834 (2000).

[64] L. Zhang and Z. L. Zhang, Radiat. Eff. Defects Solids **160**, 337 (2005).

[65] M. Stepanova and S. K. Dew, J. Vac. Sci. Technol., A **19**, 2805 (2001).

[66] A. Wucher and W. Reuter, J. Vac. Sci. Technol., A **6**, 2316 (1988).

[67] T. Aoyama, M. Tanemura, and F. Okuyama, Appl. Surf. Sci. **100-101**, 351 (1996).

[68] W. Zhenxia, P. Jisheng, Z. Jiping, T. Zhenlan, W. Wenmin, Z. Huiming, and Z. Yuewu, Vacuum **46**, 247 (1995).

[69] S. Bhattacharjee, J. Zhang, V. Shutthanandan, P. K. Ray, N. R. Shivaparan, and R. J. Smith, Nucl. Instrum. Methods Phys. Res., Sect. B **129**, 123 (1997).

[70] B. Rubin, J. L. Topper, C. C. Farnell, and A. P. Yalin, Rev. Sci. Instrum. **80**, 103506 (2009).

[71] H. Tsuge and S. Esho, J. Appl. Phys. **52**, 4391 (1981).

[72] R. E. Jones, IBM J. Res. Dev. **16**, 27 (1972).

[73] M. Horkel, K. V. Aeken, C. Eisenmenger-Sittner, D. Depla, S. Mahieu, and W. P. Leroy, J. Phys. D: Appl. Phys. **43**, 075302 (2010).

[74] K. V. Aeken, S. Mahieu, and D. Depla, J. Phys. D: Appl. Phys. **41**, 205307 (2008).

[75] A. Piel, *Plasma physics*, Springer-Verlag (2010).

[76] F. F. Chen, *Introduction to plasma physics and controled fusion. Volume 1: Plasma Physics*, Springer (2001).

[77] W. D. Sproul, D. M. Mattox, and V. H. Mattox. PVD Processes for Depositing Hard Tribological Coatings. In D. M. Mattox and V. H. Mattox, editors, *50 years of vacuum coating technology*, chapter 6, p. 35. Society of Vacuum Coaters (2007).

[78] A. Brudnik, H. Czternastek, K. Zakrzewska, and M. Jachimowski, Thin Solid Films **199**, 45 (1991).

[79] I. Safi, Surf. Coat. Technol. **127**, 203 (2000).

[80] R. A. Swady and M. K. Jasim, Adv. Appl. Mech. **5**, 1 (2012).

[81] J. Musil, P. Baroch, J. Vlcek, K. Nam, and J. Han, Thin Solid Films **475**, 208 (2005).

[82] M. Ohring, *Materials science of thin films: Deposition and structure*, Academin Press, 2nd edition (2002).

[83] J. A. Thornton, J. Vac. Sci. Technol., A **11**, 666 (1974).

[84] B. A. Movchan and A. V. Demchishin, Phys. Met. Metallogr. **28**, 83 (1969).

[85] P. J. Kelly and R. D. Arnell, J. Vac. Sci. Technol., A **16**, 2858 (1998).

[86] R. Messier, A. P. Giri, and R. A. Roy, J. Vac. Sci. Technol., A **2**, 500 (1984).

[87] S. Mahieu, G. Buyle, D. Depla, S. Heirwegh, G. P, and R. De Gryse, Nucl. Instrum. Methods Phys. Res., Sect. B **243**, 313 (2006).

[88] H. M. Choi, S. K. Choi, O. Anderson, and K. Bange, Thin Solid Films **358**, 202 (2000).

[89] Y. Fu, H. Du, and S. Zhang, Surf. Coat. Technol. **167**, 120 (2003).

[90] Y. G. Yang, X. W. Zhou, R. A. Johnson, and H. N. G. Wadley, Acta Mater. **49**, 3321 (2001).

[91] T. J. Vink, M. A. J. Somers, J. L. C. Daams, and A. G. Dirks, J. Appl. Phys. **70**, 4301 (1991).

[92] K.-H. Muller, J. Appl. Phys. **62**, 1796 (1987).

[93] E. Winsberg, *Science in the Age of Computer Simulation*, The university of Chicago Press, Chicago, 1st edition (2010).

[94] T. Schulze, P. Smereka, and E. Weinan, J. Comput. Phys. **189**, 197 (2003).

[95] C. Ratsch and J. A. Venables, J. Vac. Sci. Technol., A **21**, S96 (2003).

[96] M. Pellicione, *Evolution of thin film morphology*, volume 478, Springer-Verlag, Berlin (2008).

[97] J. A. Venables, G. D. T. Spiller, and M. Hanbucken, Rep. Prog. Phys. **47**, 399 (1984).

[98] W. W. Mullins, J. Appl. Phys. **28**, 333 (1957).

[99] V. G. Dubrovskii, M. A. Kazansky, M. V. Nazarenko, and L. T. Adzhemyan, J. Chem. Phys. **134**, 094507 (2011).

[100] J. Haider, M. Rahman, B. Corcoran, and M. Hashmi, J. Mater. Process. Technol. **168**, 36 (2005).

[101] K. T. Kuwata, R. I. Erickson, and J. R. Doyle, Nucl. Instrum. Methods Phys. Res., Sect. B **201**, 566 (2003).

[102] X.-Y. Liu, F. Ercolessi, and J. B. Adams, Modelling Simul. Mater. Sci. Eng. **12**, 665 (2004).

[103] M. Karolewski, Nucl. Instrum. Methods Phys. Res., Sect. B **256**, 354 (2007).

[104] K. A. Fichthorn, Y. Tiwary, T. Hammerschmidt, P. Kratzer, and M. Scheffler, Phys. Rev. B: Condens. Matter Mater. Phys. **83**, 195328 (2011).

[105] K. A. Fichthorn and M. Scheffler, Phys. Rev. Lett. **84**, 5371 (2000).

[106] P. Brault and G. Moebs, Eur. J. Appl. Phys. **28**, 43 (2004).

[107] D. B. Graves and P. Brault, J. Phys. D: Appl. Phys. **42**, 194011 (2009).

[108] H. Rafii-Tabar, Physics Reports **325**, 239 (2000).

[109] V. Georgieva, M. Saraiva, N. Jehanathan, O. I. Lebelev, D. Depla, and A. Bogaerts, J. Phys. D: Appl. Phys. **42**, 065107 (2009).

[110] I. T. Dimov, *Monte Carlo methods for applied scientists*, World Scientific Publishing Co. (2008).

[111] A. K. Bhuiyan, S. K. Dew, and M. Stepanova, J. Phys. Chem. C **115**, 19557 (2011).

[112] Z. Jiang and C. Ebner, Phys. Rev. B: Condens. Matter Mater. Phys. **2**, 33 (1989).

[113] P. Ramanlal and L. Sander, Phys. Rev. Lett. **54**, 1828 (1985).

[114] M. J. Brett, J. Mater. Sci. Lett. **8**, 415 (1989).

[115] P. Vogl, U. Hansen, and V. Fiorentini, Comp. Matter. Sci. **24**, 58 (2002).

[116] Gill, P. Spencer, and A. Cocks, Mater. Sci. Eng., A **365**, 66 (2004).

[117] G. Russo, L. M. Sander, P. Smereka, and Russo, Phys. Rev. B: Condens. Matter Mater. Phys. **69**, 1 (2004).

[118] L. J. Friedrich, S. K. Dew, M. J. Brett, and T. Smy, IEEE T. Semiconduct. M. **12**, 353 (1999).

[119] H. Basagaoglu, C. T. Green, P. Meakin, B. J. McCoy, and H. Baaaolu, J. Chem. Phys. **121**, 7987 (2004).

[120] M. Stahl, T. Trottenberg, and H. Kersten, Rev. Sci. Instrum. **81**, 023504 (2010).

[121] J. G. Han, J. Phys. D: Appl. Phys. **42**, 043001 (2009).

[122] S. Leonard. *A framework for reactor-scale PVD simulations.* PhD thesis, University of Alberta (2002).

[123] S. K. Dew. *Processes and simulation for advanced integrated circuit metallization.* PhD thesis, University of Alberta (1992).

[124] M. Fayad and D. C. Schmidt, Communications of the ACM **40**, 32 (1997).

[125] D. Riehle. *Framework design: A role modeling approach.* PhD thesis, University of Hamburg (2000).

[126] F. J. Jimenez. *Modeling transport of thermalized neutrals in a sputter deposition system.* PhD thesis, University of Alberta (2005).

[127] B. Chapman, *Glow discharge processes*, Wiley publications (1980).

[128] G. Ecker and K. G. Emeleus, Proc. Phys. Soc. London, Sect. B **67**, 546 (1954).

[129] W. D. Westwood, J. Vac. Sci. Technol., A **15**, 1 (1978).

[130] J. H. Keller and R. G. Simmons, IBM J. Res. Dev. **23**, 24 (1979).

[131] J. B. Webb, J. Appl. Phys. **53**, 9043 (1982).

[132] T. P. Drüsedau, M. Lohmann, B. Garke, and T. P. Drusedau, J. Vac. Sci. Technol., A **16**, 2728 (1998).

[133] H. Dieker and M. Wuttig, Thin Solid Films **478**, 248 (2005).

[134] O. Kappertz, R. Drese, J. Ngaruiya, and M. Wuttig, Thin Solid Films **484**, 64 (2005).

[135] I. Knittel, M. Gothe, and U. Hartmann, J. Vac. Sci. Technol., A **23**, 1714 (2005).

[136] C. Salinga, O. Kappertz, and M. Wuttig, Thin Solid Films **515**, 2760 (2006).

[137] L.-W. Lai and C.-T. Lee, Mater. Chem. Phys. **110**, 393 (2008).

[138] A. Palmero, E. D. van Hattum, H. Rudolph, and F. H. P. M. Habraken, Thin Solid Films **494**, 18 (2006).

[139] S. D. Ekpe, L. W. Bezuidenhout, and S. K. Dew, Thin Solid Films **474**, 330 (2005).

[140] G. Bondarenko, O. Bonk, and V. Kristya, Vacuum **81**, 771 (2007).

[141] V. Kristya, J. Surf. Invest. **2**, 203 (2008).

[142] F. J. Cadieu and N. Chencinski, IEEE T. Magn. **MAG-11**, 227 (1975).

[143] A. Gras-Marti and J. A. Valles-Abarca, J. Appl. Phys. **54**, 1071 (1983).

[144] J. A. Valles-Abarca and A. Gras-Marti, J. Appl. Phys. **55**, 1370 (1984).

[145] G. Gonzalez-Diaz, I. Martil, F. Sanchez-Quesada, M. Rodriguez-Vidal, A. Gras-Marti, and J. A. Valles-Abarca, J. Vac. Sci. Technol., A **1**, 1394 (1983).

[146] T. W. Tuer and G. S. Springer, Computers & Fluids **1**, 399 (1973).

[147] T. Motohiro and Y. Taga, Surf. Sci. **134**, L494 (1983).

[148] R. E. Somekh, J. Vac. Sci. Technol., A **2**, 1285 (1984).

[149] G. M. Turner, I. S. Falconer, B. W. James, D. R. Mckenzie, and R. W. James, J. Appl. Phys. **65**, 3671 (1989).

[150] G. M. Turner, J. Vac. Sci. Technol., A **13**, 2161 (1995).

[151] V. V. Serikov and K. Nanbu, J. Appl. Phys. **82**, 5948 (1997).

[152] A. M. Myers, J. R. Doyle, J. R. Abelson, and D. N. Ruzic, J. Vac. Sci. Technol., A **9**, 614 (1991).

[153] A. M. Myers, J. R. Doyle, and D. N. Ruzic, J. Appl. Phys. **72**, 3064 (1992).

[154] A. Bogaerts, R. Gijbels, and V. V. Serikov, J. Appl. Phys. **87**, 8334 (2000).

[155] V. Vyas and M. J. Kushner, J. Vac. Sci. Technol., A **24**, 1955 (2006).

[156] I. Kolev and A. Bogaerts, IEEE Trans. Plasma Sci. **34**, 886 (2006).

[157] I. Kolev and A. Bogaerts, Plasma Processes Polym. **3**, 127 (2006).

[158] A. Malaurie and A. Bessaudou, Thin Solid Films **286**, 305 (1996).

[159] S. Senthil-Nathan, G. Mohan Rao, and S. Mohan, J. Appl. Phys. **84**, 564 (1998).

[160] O. Yamazaki, K. Iyanagi, N. Shigeyuki, and K. Nanbu, Jpn. J. Appl. Phys. **41**, 1230 (2002).

[161] E. Hollmann, V. Vol[']pyas, and R. Wördenweber, Physica C: Superconductivity **425**, 101 (2005).

[162] R. Sobbia, P. K. Browning, and J. W. Bradley, J. Vac. Sci. Technol., A **26**, 103 (2008).

[163] A. Settaouti and L. Settaouti, Appl. Surf. Sci. **254**, 5750 (2008).

[164] G. Bird, *Molecular Gas Dynamics*, Oxford University Press (1976).

[165] K. Nanbu, J. Phys. Soc. Jpn. **49**, 2050 (1980).

[166] S. V. Patankar, *Numerical heat transfer and fluid flow*, McGraw-Hill (1980).

[167] J. D. Anderson, *Computational fluid dynamics: The basics with applications*, Mc Graw-Hill Inc., 1st edition (1995).

[168] H. K. Versteeg and W. Malalasekera, *An introduction to Commputational Fluid Dynamics: the finite volume method*, Prentice Hall, 2nd edition (2007).

[169] J. Stache, J. Vac. Sci. Technol., A **12**, 2867 (1994).

[170] T. Nakano and S. Baba, Thin Solid Film **343-344**, 24 (1999).

[171] V. V. Serikov, S. Kawamoto, and K. Nanbu, IEEE Trans. Plasma Sci. **27**, 1389 (1999).

[172] S. H. Choi, U. H. Kwon, W. J. Lee, and Y. H. Park, Thin Solid Films **475**, 17 (2005).

[173] L. Devroye, *Non-uniform random variate generation*, Springer-Verlag (1986).

[174] J. F. Ziegler, J. P. Birsack, and J. Littmark, *The Stopping and Range of Ions in Solids*, Pergamon Press (1985).

[175] R. S. Robinson, J. Vac. Sci. Technol. **16**, 185 (1979).

[176] A. Bogaerts. *Mathematical modeling of a direct current glow discharge in argon.* PhD thesis (1996).

[177] T. Motohiro, J. Vac. Sci. Technol., A **4**, 189 (1986).

[178] K. Macak, P. Macak, and U. Helmersson, Comput. Phys. Commun. **120**, 238 (1999).

[179] E. H. Kennard, *Kinetic theory of gases*, McGraw Hill (1934).

[180] Y. Yamamura and M. Ishida, J. Vac. Sci. Technol., A **13**, 101 (1995).

[181] M. Benzi, G. H. Golub, and J. Liesen, Acta Numerica **14**, 1 (2005).

[182] N. Mctaggart, R. Zagorski, X. Bonnin, a. Runov, and R. Schneider, Comput. Phys. Commun. **164**, 318 (2004).

[183] J. M. Ortega and R. W. C, *Iterative solution of nonlinear equations in several variables*, Academi Press (1970).

[184] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM (1995).

[185] M. Hestenes and E. Stiefel, J. Res. Nat. Bur. Stand. **49**, 409 (1952).

[186] Y. Saad and M. H. Schultz, SIAM Journal on Scientific and Statistical Computing **7**, 856 (1986).

[187] J. Dongarra, A. Lumsdaine, X. Niu, R. Pozo, and K. A. Remington. A Sparse Matrix Library in C++ for High Performance Architectures. In *Proceedings of the Second Object Oriented Numerics Conference* (1994).

[188] W. Eckstein. Sputtering Yields. In *Sputtering by particle bombardment*, pp. 33–187 (2007).

[189] C. Vitelaru, L. de Poucques, T. M. Minea, and G. Popa, Plasma Sources Sci. Technol. **20**, 045020 (2011).

[190] W. D. Davis and T. A. Vanderslice, Phys. Rev. **131**, 219 (1963).

[191] J. Rickards, Vacuum **34**, 559 (1984).

[192] S. Hamaguchi, R. T. Farouki, and M. Dalvie, Phys. Rev. A: At., Mol., Opt. Phys. **44**, 3804 (1991).

[193] C. V. Budtz-Jorgensen, J. Bottiger, and P. Kringhoj, Vacuum **56**, 9 (2000).

[194] M. J. Goeckner, J. A. Goree, and T. E. Sheridan, IEEE T. Plasma. Sci. **19**, 301 (1991).

[195] C. H. Shon, J. K. Lee, H. J. Lee, Y. Yang, and T. H. Chung, IEEE T. Plasma. Sci. **26**, 1635 (1998).

[196] R. A. Baragiola, E. V. Alonso, J. Ferron, and A. Oliva-Florio, Surf. Sci. **90**, 240 (1979).

[197] D. Depla, S. Mahieu, and R. De Gryse, Thin Solid Films **517**, 2825 (2009).

[198] S. Kadlec, C. Quaeyhaegens, G. Knuyt, and L. M. Stals, Surf. Coat. Technol. **97**, 633 (1997).

[199] V. Abhilash, R. Balu, S. Balaji, S. Senthil Nathan, and S. Mohan, Comp. Matter. Sci. **30**, 523 (2004).

[200] A. Kersch, C. Werner, and W. Murokoff, J. Appl. Phys. **75**, 2278 (1994).

[201] C. Eisenmenger-Sittner, R. Beyerknecht, A. Bergauer, W. Bauer, and G. Betz, J. Vac. Sci. Technol., A **13**, 2435 (1995).

[202] T. Heberlein, G. Krautheim, and W. Wuttke, Vacuum **42**, 47 (1991).

[203] D. G. Coronell, E. W. Egan, G. Hamilton, A. Jain, R. Venkatraman, and B. Weitzman, Thin Solid Films **333**, 77 (1998).

[204] V. Efimova, A. Derzsi, A. Zlotorowicz, V. Hoffmann, Z. Donkó, and J. Eckert, Spectrochim. Acta, Part B **65**, 311 (2010).

[205] N. Britun, M. Gaillard, Y. Kim, K. Kim, and J.-G. Han, Metals and Materials International **13**, 483 (2007).

[206] N. Britun, M. Gaillard, S.-G. Oh, and J. G. Han, J. Phys. D: Appl. Phys. **40**, 5098 (2007).

[207] M. Wolter, H. T. Do, H. Steffen, and R. Hippler, J. Phys. D: Appl. Phys. **38**, 2390 (2005).

[208] R. Ramos, G. Cunge, M. Touzeau, and N. Sadeghi, J. Phys. D: Appl. Phys. **41**, 152003 (2008).

[209] A. P. Yalin and V. Surla, Opt. Lett. **30**, 3219 (2005).

[210] M. Gaillard, N. Britun, Y. M. Kim, and J. G. Han, J. Phys. D: Appl. Phys. **40**, 809 (2007).

[211] D. W. Hoffman, J. Vac. Sci. Technol., A **3**, 561 (1985).

[212] S. M. Rossnagel, J. Vac. Sci. Technol., A **6**, 19 (1988).

[213] F. J. Jimenez, S. D. Ekpe, and S. K. Dew, J. Vac. Sci. Technol., A **24**, 1630 (2006).

[214] S. Kadlec, Plasma Processes Polym. **4**, S419 (2007).

[215] D. Horwat and A. Anders, J. Appl. Phys. **108**, 123306 (2010).

[216] M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing*, Wiley-Interscience, 1 edition (1994).

[217] A. Bogaerts and R. Gijbels, J. Appl. Phys. **84**, 121 (1998).

[218] M. Dickson, F. Qian, and J. Hopwood, J. Vac. Sci. Technol., A **15**, 340 (1997).

[219] J. Bohlmark, J. Alami, C. Christou, A. P. Ehiasarian, and U. Helmersson, J. Vac. Sci. Technol., A **23**, 18 (2005).

[220] C. Christou and Z. H. Barber, J. Vac. Sci. Technol., A **18**, 2897 (2000).

[221] J. Hopwood and I. Technology, Phys. Plasmas **5**, 1624 (1998).

[222] Z. Donko, P. Hartmann, and K. Kutasi, Plasma Sources Sci. Technol. **15**, 178 (2006).

[223] V. E. Golant, A. P. Zhilinsky, and I. E. Sakharov, *Fundamentals of plasma physics*, John Wiley & Sons, Inc. (1980).

[224] J. W. Bradley, Plasma Sources Sci. Technol. **5**, 662 (1996).

[225] J. W. Bradley and G. Lister, Plasma Sources Sci. Technol. **6**, 524 (1997).

[226] J. W. Bradley, Plasma Sources Sci. Technol. **7**, 572 (1998).

[227] N. F. Cramer, J. Phys. D: Appl. Phys. **30**, 2573 (1997).

[228] a. Palmero, E. van Hattum, W. Arnoldbik, and F. Habraken, Surf. Coat. Technol. **188-189**, 392 (2004).

[229] A. Bogaerts, E. Bultinck, I. Kolev, L. Schwaederle, K. V. Aeken, G. Buyle, and D. Depla, J. Phys. D: Appl. Phys. **42**, 194018 (2009).

[230] S. T. Surzhikov and J. S. Shang, J. Comput. Phys. **199**, 437 (2004).

[231] C. Costin, L. Marques, G. Popa, and G. Gousset, Plasma Sources Sci. Technol. **14**, 168 (2005).

[232] C. Costin, G. Popa, and G. Gousset, J. Optoelectron. Adv. Mater. **7**, 2465 (2005).

[233] G. J. M. Hagelaar, Plasma Sources Sci. Technol. **16**, S57 (2007).

[234] G. Leray, P. Chabert, a. J. Lichtenberg, and M. A. Lieberman, J. Phys. D: Appl. Phys. **42**, 194020 (2009).

[235] J.-P. Boeuf, J. Appl. Phys. **63**, 1342 (1988).

[236] S. Hashiguchi, J. Vac. Sci. Technol., A **10**, 1339 (1991).

[237] A. Bogaerts, R. Gijbels, and W. J. Goedheer, J. Appl. Phys. **78**, 2233 (1995).

[238] K. Kutasi, Solid State Physics **33**, 1081 (2000).

[239] V. A. Zhovtyansky, Ukranian Journal of Physics **53**, 497 (2008).

[240] K. Kutasi, Solid State Physics **34**, 3368 (2001).

[241] G. Bánó, P. Hartmann, K. Kutasi, P. Horváth, R. Plašil, P. Hlavenka, J. Glosík, and Z. Donkó, Plasma Sources Sci. Technol. **16**, 492 (2007).

[242] P. S. Kothnur and L. L. Raja, J. Appl. Phys. **97**, 043305 (2005).

[243] M. Dalvie, M. Surendra, and G. S. Selwyn, **62**, 3207 (1993).

[244] M. S. Barnes, T. J. Colter, and M. E. Elta, J. Appl. Phys. **61**, 81 (1986).

[245] J.-P. Boeuf, Phys. Rev. A: At., Mol., Opt. Phys. **36**, 2782 (1987).

[246] J.-P. Boeuf and L. C. Pitchford, Phys. Rev. **51**, 1376 (1995).

[247] J. D. P. Passchier and W. J. Goedheer, J. Appl. Phys. **74**, 3744 (1993).

[248] A. Salabas, G. Gousset, and L. L. Alves, Plasma Sources Science Technology **11**, 448 (2002).

[249] A. D. Richards, B. E. Thompson, and H. H. Sawin, Appl. Phys. Lett. **50**, 128 (1987).

[250] J. D. Bukowski, D. B. Graves, and P. Vitello, J. Appl. Phys. **80**, 2614 (1996).

[251] D. Hash, D. Bose, T. R. Govindan, and M. Meyyappan, J. Appl. Phys. **93**, 6284 (2003).

[252] G. J. Nienhuis and W. J. Goedheer, J. Appl. Phys. **82**, 2060 (1997).

[253] S. K. Nam, C. B. Shin, and D. J. Economou, Mater. Sci. Semicond. Process. **2**, 271 (1999).

[254] S. Sfikas, E. Amanatides, D. Mataras, and D. Rapakoulias, IEEE Trans. Plasma Sci. **35**, 1420 (2007).

[255] T. E. Sheridan, M. J. Goeckner, and J. A. Goree, J. Vac. Sci. Technol., A **8**, 30 (1989).

[256] J. E. Miranda, M. J. Goeckner, J. A. Goree, and T. E. Sheridan, J. Vac. Sci. Technol., A **8**, 1627 (1990).

[257] S. Ido and K. Nakamura, Jpn. J. Appl. Phys. **32** (1993).

[258] M. Holik, J. Bradley, V. Bellido-Gonzalez, and D. Monaghan, Plasma Processes Polym. **6**, S789 (2009).

[259] J. Sun, Y. Gong, and D. Wang, J. Phys. D: Appl. Phys. **26**, 436 (1993).

[260] M. Yousfi, A. Hennad, and A. Alkaa, Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys. **49**, 3264 (1994).

[261] A. Bogaerts and R. Gijbels, J. Appl. Phys. **78**, 6427 (1995).

[262] J. Musschoot, D. Depla, G. Buyle, J. Haemers, and R. D. Gryse, J. Phys. D: Appl. Phys. **39**, 3989 (2006).

[263] V. Zhakhovskii and K. Nishihara, physics/0507076 pp. 1–8 (2008).

[264] C. K. Birdsall and L. Fellow, IEEE T. Plasma. Sci. **19**, 65 (1991).

[265] K. Nanbu, S. Segawa, and S. Kondo, Vacuum **47**, 1013 (1996).

[266] S. Kondo and K. Nanbu, IEEE Trans. Plasma Sci. **27**, 92 (1999).

[267] T. M. Minea, J. Bretagne, G. Gousset, L. Magne, D. Pagnon, and M. Touzeau, Science **119**, 558 (1999).

[268] K. Nanbu, K. Mitsui, and S. Kondo, J. Phys. D: Appl. Phys. **33**, 2274 (2000).

[269] S. Yonemura and K. Nanbu, IEEE T. Plasma. Sci. **31**, 479 (2003).

[270] V. Vahedi, Comput. Phys. Commun. **87**, 179 (1995).

[271] C. H. Shon, J. S. Park, B. K. Kang, and J. K. Lee, Jpn. J. Appl. Phys. **38**, 4440 (1999).

[272] S. Kondo and K. Nanbu, J. Vac. Sci. Technol., A **19**, 830 (2001).

[273] E. Bultinck, I. Kolev, and A. Bogaerts, pp. 483–486 (2007).

[274] S. Kuroiwa, T. Mine, T. Yagisawa, and T. Makabe, J Vac Sci Technol B **23**, 2218 (2005).

[275] E. Bultinck and A. Bogaerts, Plasma Sources Sci. Technol. **20**, 045013 (2011).

[276] I. Kolev and A. Bogaerts, J. Appl. Phys. **104**, 093301 (2008).

[277] M. Yusupov, E. Bultinck, D. Depla, and A. Bogaerts, New Journal of Physics **13**, 033018 (2011).

[278] E. Bultinck and A. Bogaerts, J. Phys. D: Appl. Phys. **41**, 202007 (2008).

[279] F. Sigeneger, Z. Donkó, D. Loffhagen, and Z. Donko, Eur. J. Appl. Phys. **38**, 161 (2007).

[280] R. K. Porteous and D. B. Graves, IEEE T. Plasma. Sci. **19**, 204 (1991).

[281] K. Grapperhaus, M. J. Kushner, and M. J. Grapperhaus, J. Appl. Phys. **83**, 6 (1998).

[282] R. L. Kinder and M. J. Kushner, J. Vac. Sci. Technol., A **19**, 76 (2001).

[283] A. V. Vasenkov, J. Appl. Phys. **95**, 834 (2004).

[284] S. Rauf and M. J. Kushner, J. Appl. Phys. **81**, 5966 (1997).

[285] E. Shidoji, H. Ohtake, N. Nakano, and T. Makabe, Jpn. J. Appl. Phys. **38**, 2131 (1999).

[286] E. Shidoji, E. Ando, and T. Makabe, Plasma Sources Sci. Technol. **8**, 621 (2001).

[287] E. Shidoji, K. Ness, and T. Makabe, Vacuum **60**, 299 (2001).

[288] E. Shidoji and T. Makabe, Thin Solid Film **443**, 27 (2003).

[289] M. Rondanini, C. Cavallotti, D. Ricci, D. Chrastina, G. Isella, T. Moiseev, and H. von Kanel, J. Appl. Phys. **104**, 013304 (2008).

[290] I. Kolev and A. Bogaerts, Contrib. Plasma Phys. **44**, 582 (2004).

[291] A. Bogaerts, R. D. Guenard, B. W. Smith, J. D. Winefordner, W. W. Harrison, and R. Gijbels, Spectrochim. Acta, Part B **52**, 219 (1997).

[292] K. H. Schoenbach and H. Chen, J. Appl. Phys. **67**, 154 (1990).

[293] M. Surendra, D. B. Graves, and G. M. Jellum, Phys. Rev. A: At., Mol., Opt. Phys. **41**, 1112 (1990).

[294] D. P. Lymberopoulos and D. J. Economou, J. Appl. Phys. **73**, 3668 (1993).

[295] A. Fiala, L. C. Pitchford, and J.-P. Boeuf, Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys. **49**, 5607 (1994).

[296] E. Suetomi, M. Tanaka, S. Kamiya, S. Hayashi, S. Ikeda, H. Sugawara, and Y. Sakai, Comput. Phys. Commun. **125**, 60 (2000).

[297] A. Bogaerts, I. Kolev, and G. Buyle. Modeling of the magnetron discharge. In *Reactive Sputter Deposition*, pp. 61–130. Springer-Verlag, 1st edition (2008).

[298] P. L. G. Ventzek, T. J. Sommerer, R. J. Hoekstra, M. J. Kushnerb, and S. T. J, Appl. Phys. Lett. **63**, 605 (1993).

[299] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd edition (1992).

[300] K. Nanbu, Jpn. J. Appl. Phys. **33**, 4752 (1994).

[301] A. B. Phelps. http://jila.colorado.edu/˜avp/collision_data/ electronneutral/ELECTRON.TXT (2008).

[302] A. B. Phelps, Plasma Sources Sci. Technol. **10**, 329 (2001).

[303] A. B. Phelps, J. Phys. Chem. Ref. Data **20**, 557 (1991).

[304] W. C. Fon, K. A. Berrington, P. G. Burke, and A. Hibbert, J. Phys. B: At., Mol. Opt. Phys. **16**, 307 (1983).

[305] S. N. Nahar, Wadehra, and J. M, Phys. Rev. A: At., Mol., Opt. Phys. **35**, 2051 (1987).

[306] F. J. de Heer, R. H. Jansen, and W. van der Kaay, J. Phys. B: At., Mol. Opt. Phys. **12**, 979 (1979).

[307] M. Hayashi. http://jila.colorado.edu/˜avp/collision_data/ electronneutral/hayashi.txt (1992).

[308] M. Kotera, K. Murata, and K. Nagami, J. Appl. Phys. **52**, 7403 (1981).

[309] A. Jablonski, F. Salvat, and C. J. Powell, J. Phys. Chem. Ref. Data **33**, 409 (2004).

[310] I. Adesida, R. Shimizu, and T. E. Everhart, J. Appl. Phys. **51**, 5962 (1980).

[311] A. H. Futch and F. A. Grant, Phys. Rev. **104**, 356 (1956).

[312] A. Bogaerts, Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys. **52** (1995).

[313] N. Baguer, A. Bogaerts, Z. Donko, R. Gijbels, and N. Sadeghi, J. Appl. Phys. **97**, 123305 (2005).

[314] R. J. Carman, J. Phys. D: Appl. Phys. **22**, 55 (1989).

[315] H. Tawara and T. Kato, At. Data Nucl. Data Tables **36**, 167 (1987).

[316] R. S. Freund, R. C. Wetzel, R. J. Shul, and T. R. Hayes, Phys. Rev. A: At., Mol., Opt. Phys. **41**, 3375 (1990).

[317] W. Lotz, Zeitschrift für Physik A Hadrons and Nuclei **216**, 241 (1968).

[318] W. J. M. Brok. *Modelling of Transient Phenomena in Gas Discharges.* PhD thesis, Technishe Universiteit Eindhoven (2005).

[319] L. M. Kishinevsky, Radiat. Eff. Defects Solids **19**, 23 (1973).

[320] A. B. Phelps and Z. L. Petrovic, Plasma Sources Sci. Technol. **8**, R21 (1999).

[321] D. Depla, G. Buyle, J. Haemers, and R. De Gryse, Surf. Coat. Technol. **200**, 4329 (2006).

[322] G. Buyle, W. De Bosshcer, D. Depla, K. Eufinger, J. Haemers, and R. De Gryse, Vacuum **70**, 29 (2003).

[323] L. S. Frost, Phys. Rev. **105**, 354 (1957).

[324] A. L. Ward, J. Appl. Phys. **33**, 2789 (1962).

[325] H. W. Ellis, R. Y. Pai, E. W. Mcdaniel, E. A. Mason, and L. A. Viehland, At. Data Nucl. Data Tables **17**, 177 (1976).

[326] G. J. M. Hagelaar and L. C. Pitchford, Plasma Sources Sci. Technol. **14**, 722 (2005).

[327] G. J. M. Hagelaar and G. M. W. Kroesen, J. Comput. Phys. **159**, 1 (2000).

[328] Ron S. Dembo, S. C. Eisenstat, T. Steihaug, and R. S. Dembo, SIAM J. Numer. Anal. **19**, 400 (1982).

[329] R. E. Bank and D. J. Rose, SIAM J. Numer. Anal. **17**, 806 (1980).

[330] P. R. McHugh and D. A. Knoll, Int. J. Numer. Methods Fluids **19**, 439 (1994).

[331] D. Knoll, J. Comput. Phys. **193**, 357 (2004).

[332] D. L. Scharfetter and H. K. Gummel, IEEE T. Electron. Dev. **16**, 64 (1969).

[333] S. Selberherr, *Analysis and simulation of semiconductor devices*, Springer-Verlag Wien-New York (1984).

[334] R. J. Schilling, *Applied numerical methods for engineers* (2000).

[335] P. Amestoy, I. Duff, D. Ruiz, and B. Uçar. A Parallel Matrix Scaling Algorithm. Technical report (2008).

[336] C. Doucet, I. Charpentier, J.-L. Coulomb, C. Rin, Y. L. Floch, G. Meunier, and C. Guérin, COMPEL **27**, 897 (2008).

[337] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM: Society for Industrial and Applied Mathematics, 2nd edition (1996).

[338] R. E. Bank and D. J. Rose, Numerische Mathematik **37**, 279 (1981).

[339] S. C. Eisenstat and H. F. Walker, SIAM Journal on Optimization **4**, 393 (1994).

[340] R. S. Tuminaro, H. F. Walker, and J. N. Shadid, J. Comput. Phys. **180**, 549 (2002).

[341] N. Nafarizal, N. Takada, K. Shibagaki, K. Nakamura, Y. Sago, and K. Sasaki, Jpn. J. Appl. Phys. **44**, L737 (2005).

[342] N. Brenning, I. Axnäs, M. A. Raadu, D. Lundin, and U. Helmersson, Plasma Sources Sci. Technol. **17**, 045009 (2008).

[343] J. Vlček and K. Burcalová, Plasma Sources Sci. Technol. **19**, 065010 (2010).

[344] M. a. Raadu, I. Axnäs, J. T. Gudmundsson, C. Huo, and N. Brenning, Plasma Sources Sci. Technol. **20**, 065007 (2011).

[345] S. D. Ekpe, F. J. Jimenez, D. J. Field, M. J. Davis, and S. K. Dew, J. Vac. Sci. Technol., A **27**, 1275 (2009).

[346] I. H. Hutchinson, *Principles of plasma diagnostics*, 2nd edition (2002).

[347] I. Petrov, I. Ivanov, V. Orlinov, J. Kourtev, and I. Ivaxov, Contrib. Plasma Phys. **30**, 223 (1990).

[348] S. M. Rossnagel and H. R. Kaufman, J. Vac. Sci. Technol., A **4**, 1822 (1986).

[349] T. E. Sheridan, M. J. Goeckner, and J. A. Goree, J. Vac. Sci. Technol., A **9**, 688 (1991).

[350] M. J. Goeckner, J. A. Goree, and T. E. Sheridan, Physics of Fluids B: Plasma Physics **4**, 1663 (1992).

[351] J. W. Bradley, S. Thompson, and Y. A. Gonzalvo, Plasma Sources Sci. Technol. **10**, 490 (2001).

[352] D. J. Field, S. K. Dew, and R. E. Burrell, J. Vac. Sci. Technol., A **20**, 2032 (2002).

[353] S. Z. Wu, J. Appl. Phys. **98**, 083301 (2005).

[354] D. Desideri, M. Bagatin, M. Spolaore, V. Antoni, R. Cavazzana, E. Martines, G. Serianni, and M. Zuin, COMPEL **24**, 261 (2005).

[355] S.-H. Seo, J.-H. In, H.-Y. Chang, and J.-G. Han, Phys. Plasmas **11**, 4796 (2004).

[356] S.-H. Seo, J.-H. In, and H.-Y. Chang, Plasma Sources Sci. Technol. **15**, 256 (2006).

[357] A. D. Pajdarova, J. Vlcek, P. Kudlacek, and J. Lukas, Plasma Sources Sci. Technol. **18**, 025008 (2009).

[358] J. H. In, B. K. Na, S. H. Seo, H. Y. Chang, and J. G. Han, Plasma Sources Sci. Technol. **18**, 045029 (2009).

[359] P. Poolcharuansin and J. W. Bradley, Plasma Sources Sci. Technol. **19**, 025010 (2010).

[360] C. Bishop, *Vacuum Deposition onto Webs, Films and Foils*, William Andrew Publishing (2006).

[361] S. D. Ekpe and S. K. Dew, J. Vac. Sci. Technol., A **20**, 1877 (2002).

[362] D. Ball, J. Appl. Phys. **43**, 3047 (1972).

[363] J. A. Thornton, J. Vac. Sci. Technol., A **15**, 171 (1978).

[364] H. Kersten, D. Rohde, J. Berndt, H. Deutsch, and R. Hippler, Thin Solid Films **377-378**, 585 (2000).

[365] D. Rohde, P. Pecher, H. Kersten, W. Jacob, and R. Hippler, Surf. Coat. Technol. **149**, 206 (2002).

[366] M. Andritschky, F. Guimardes, V. Teixeira, and F. Guimaraes, Vacuum **44**, 809 (1993).

[367] S. D. Ekpe and S. K. Dew, J. Vac. Sci. Technol., A **21**, 476 (2003).

[368] S. D. Ekpe and S. K. Dew, J. Vac. Sci. Technol., A **22**, 1420 (2004).

[369] A.-L. Thomann, N. Semmar, R. Dussart, J. Mathias, and V. Lang, Rev. Sci. Instrum. **77**, 033501 (2006).

[370] L. Bedra, A. L. Thomann, N. Semmar, R. Dussart, and J. Mathias, J. Phys. D: Appl. Phys. **43**, 065202 (2010).

[371] P.-A. Cormier, M. Stahl, A.-L. Thomann, R. Dussart, M. Wolter, N. Semmar, J. Mathias, and H. Kersten, J. Phys. D: Appl. Phys. **43**, 465201 (2010).

[372] R. Dussart, a. L. Thomann, L. E. Pichon, L. Bedra, N. Semmar, P. Lefaucheux, J. Mathias, and Y. Tessier, Appl. Phys. Lett. **93**, 131502 (2008).

[373] T. P. Drusedau, T. Bock, T.-M. John, F. Klabunde, and W. Eckstein, J. Vac. Sci. Technol., A **17**, 2896 (1999).

[374] D. R. Lide, *Handbook of chemistry and physics*, 72nd edition (1991).

[375] S. D. Ekpe, F. J. Jimenez, and S. K. Dew, J. Vac. Sci. Technol., A **28**, 1210 (2010).

[376] S. Berg, T. Nyberg, and T. Kubart. Modelling of reactive sputtering processes. In D. Depla and S. Mahieu, editors, *Reactive Sputter deposition*, pp. 131–152. 1st edition (2008).

[377] H. Cheng, Surf. Coat. Technol. **167**, 297 (2003).

[378] F. S. Ohuchi and P. E. Russell, J. Vac. Sci. Technol., A **5**, 1630 (1987).

[379] H. Cheng, J. Cryst. Growth **254**, 46 (2003).

[380] H. Cheng, Thin Solid Films **434**, 112 (2003).

[381] S. Venkataraj, D. Severin, R. Drese, F. Koerfer, and M. Wuttig, Thin Solid Films **502**, 235 (2006).

APPENDIX A

Derivation of discrete magnetized plasma equations

A.1 Plasma under the influence of a mangetic field

To derive equation 5.31, we start with the electron force equation under the influence of a magnetic field [75, 216, 223] which reads

$$nm \left[\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} \right] = qn(\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B}) - \nabla p \qquad \text{(A.1)}$$

Neglecting inertia [216, 223] and using the approximation $nm\partial u/\partial t \approx \nu nm$ [223], we obtain

$$T\nabla n + mn\nu\boldsymbol{u} = qn\boldsymbol{E} + qn(\boldsymbol{u} \times \boldsymbol{B}) \qquad \text{(A.2)}$$

This equation is more difficult to solve for $n\boldsymbol{u}$ (see Chapter 5) and requires a more elaborated process. The first step is to divide the flux into two components: one parallel and another perpendicular to the magnetic field in the form

$$\boldsymbol{u} = \boldsymbol{u}_\perp + \boldsymbol{u}_\parallel \tag{A.3}$$

We proceed then to solve for each component separately. Solving for $\boldsymbol{u}_\parallel$, and noticing that $\boldsymbol{u} \times \boldsymbol{B} = 0$, we obtain

$$n\boldsymbol{u}_\parallel = \frac{q}{m\nu} n\boldsymbol{E}_\parallel - \frac{T}{m\nu}(\nabla n)_\parallel \tag{A.4}$$

which has a form similar to the plasma equation without a magnetic field (equation 5.6).

Next, we solve for $\boldsymbol{u}_\perp$ from

$$m\nu n\boldsymbol{u}_\perp = qB(\boldsymbol{u}_\perp \times \boldsymbol{b}) + qn\boldsymbol{E}_\perp - T(\nabla n)_\perp \tag{A.5}$$

where $\boldsymbol{b}$ is the unit vector in the direction of the magnetic field $\boldsymbol{B}$ and B is its magnitude. To solve this equation, I follow the procedure described in Ref. 223. First, we multiply vectorialy equation A.5 by $\boldsymbol{b}$ to obtain

$$n(\boldsymbol{u}_\perp \times \boldsymbol{b}) = qB\boldsymbol{u}_\perp + qn(\boldsymbol{E} \times \boldsymbol{b}) + \boldsymbol{b} \times (T\nabla n)_\perp \tag{A.6}$$

Eliminating the vector product $(\boldsymbol{u}_\perp \times \boldsymbol{b})$ from equations A.5 and A.6 yields [223]

$$n\boldsymbol{u}_\perp = \frac{(\boldsymbol{E} \times \boldsymbol{b})}{B(1 + 1/\Omega^2)} + \frac{\boldsymbol{b} \times (T\nabla n)}{qB(1 + 1/\Omega^2)}$$
$$\frac{qm\nu\boldsymbol{E}_\perp}{1 + \Omega^2} + \frac{m\nu T(\nabla n)_\perp}{1 + \Omega^2} \tag{A.7}$$

where Ω is the Hall Parameter defined by

$$\Omega = \frac{|q\boldsymbol{B}|}{m\nu} \tag{A.8}$$

Finally, we obtain

$$nu = n(u_\parallel + u_\perp) = \mu_\parallel n E_\parallel + \mu_\perp n E_\perp + \mu_\alpha n[E \times b]$$
$$-D_\parallel (\nabla n)_\parallel - D_\perp (\nabla n)_\perp + D_\alpha [b \times \nabla n] \tag{A.9}$$

with

$$\mu_\parallel = \frac{q}{m\nu} = \mu \tag{A.10}$$

$$D_\parallel = \frac{T}{m\nu} = D \tag{A.11}$$

$$\mu_\perp = \frac{\mu}{1 + \Omega^2} \tag{A.12}$$

$$D_\perp = \frac{D}{1 + \Omega^2} \tag{A.13}$$

$$\mu_\alpha = \frac{1}{B(1 + 1/\Omega^2)} \tag{A.14}$$

$$D_\alpha = \frac{T}{qB(1 + 1/\Omega^2)} \tag{A.15}$$

The terms $E_\parallel$, $E_\perp$, $E \times b$, $(\nabla n)_\parallel$, $(\nabla n)_\perp$ and $b \times \nabla n$ in equation A.9 are still unknown and need to be expanded. The following expressions are used to derive these terms.

$$E_\parallel = (E \cdot b)b = (E_x b_x + E_y b_y + E_z b_z)(b_x \hat{i} + b_y \hat{j} + b_z \hat{k}) =$$
$$(E_x b_x^2 + E_y b_x b_y + E_z b_x b_z)\hat{i} + (E_x b_x b_y + E_y b_y^2 + E_z b_y b_z)\hat{j}$$
$$+(E_x b_x b_z + E_y b_y b_z + E_z b_z^2)\hat{k} \tag{A.16}$$

$$E_\perp = E - E_\parallel = (E_x \hat{i} + E_y \hat{j} + E_z \hat{z}) - E_\parallel$$
$$= (E_x(1 - b_x^2) - E_y b_x b_y - E_z b_x b_z)\hat{i} + (-E_x b_x b_y + E_y(1 - b_y^2) - E_z b_y b_z)\hat{j}$$
$$+(-E_x b_x b_z - E_y b_y b_z + E_z(1 - b_z^2))\hat{k} \tag{A.17}$$

$$
\begin{aligned}
(\nabla n)_\parallel =& (\nabla n \cdot \boldsymbol{b})\boldsymbol{b} = \left(\frac{\partial n}{\partial x}b_x + \frac{\partial n}{\partial y}b_y + \frac{\partial n}{\partial n}b_z \right) (h_x\hat{i} + h_y\hat{j} + h_z\hat{k}) \\
=& \left(\frac{\partial n}{\partial x}b_x^2 + \frac{\partial n}{\partial y}b_xb_y + \frac{\partial n}{\partial z}b_xb_z \right)\hat{\boldsymbol{i}} + \left(\frac{\partial n}{\partial x}b_xb_y + \frac{\partial n}{\partial y}b_y^2 + \frac{\partial n}{\partial z}b_yb_z \right)\hat{\boldsymbol{j}} \\
&+ \left(\frac{\partial n}{\partial x}b_xb_z + \frac{\partial n}{\partial y}b_yb_z + \frac{\partial n}{\partial z}b_z^2 \right)\hat{\boldsymbol{k}}
\end{aligned} \tag{A.18}
$$

$$
\begin{aligned}
(\nabla n)_\perp = \nabla n - (\nabla n)_\parallel =& \left(\frac{\partial n}{\partial x}(1 - b_x^2) - \frac{\partial n}{\partial y}b_xb_y - \frac{\partial n}{\partial z}b_xb_z \right)\hat{\boldsymbol{i}} \\
&+ \left(-\frac{\partial n}{\partial x}b_xb_y - \frac{\partial n}{\partial y}(1 - b_y^2) - \frac{\partial n}{\partial z}b_yb_z \right)\hat{\boldsymbol{j}} \\
&+ \left(\frac{\partial n}{\partial x}b_xb_z - \frac{\partial n}{\partial y}b_yb_z - \frac{\partial n}{\partial z}(1 - b_z^2) \right)\hat{\boldsymbol{k}}
\end{aligned} \tag{A.19}
$$

$$
\begin{aligned}
\boldsymbol{b} \times \nabla n =& \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ b_x & b_y & b_z \\ \dfrac{\partial n}{\partial x} & \dfrac{\partial n}{\partial y} & \dfrac{\partial n}{\partial z} \end{vmatrix} \\
=& \left(b_y\frac{\partial n}{\partial z} - b_z\frac{\partial n}{\partial y} \right)\hat{\boldsymbol{i}} + \left(b_z\frac{\partial n}{\partial x} - b_x\frac{\partial n}{\partial z} \right)\hat{\boldsymbol{j}} + \left(b_x\frac{\partial n}{\partial y} - b_y\frac{\partial n}{\partial x} \right)\hat{\boldsymbol{k}}
\end{aligned} \tag{A.20}
$$

$$
\begin{aligned}
\boldsymbol{E} \times \boldsymbol{b} =& \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ E_x & E_y & E_z \\ h_x & h_y & hz \end{vmatrix} \\
=& (E_yb_z - E_zb_y)\hat{\boldsymbol{i}} + (E_zb_x - E_xb_z)\hat{\boldsymbol{j}} + (E_xb_y - E_yb_x)\hat{\boldsymbol{k}}
\end{aligned} \tag{A.21}
$$

Substituting equations A.16, A.17, A.18, A.19, A.20 and A.21 into equation A.9 yields

$$nu_x = \mu_\parallel (E_x b_x^2 + E_y b_x b_y + E_z b_x b_z) + \mu_\perp (E_x(1 - b_x^2) - E_y b_x b_y - E_z b_x b_z)$$

$$+ \mu_\alpha (E_y b_z - E_z b_y) - D_\parallel \left(\frac{\partial n}{\partial x} b_x^2 + \frac{\partial n}{\partial y} b_x b_y + \frac{\partial n}{\partial z} b_x b_z \right)$$

$$- D_\perp \left(\frac{\partial n}{\partial x}(1 - b_x^2) - \frac{\partial n}{\partial y} b_x b_y - \frac{\partial n}{\partial z} b_x b_z \right) + D_\alpha \left(\frac{\partial n}{\partial z} b_y - \frac{\partial n}{\partial y} b_z \right) \quad \text{(A.22a)}$$

$$nu_y = \mu_\parallel (E_x b_x b_y + E_y b_y^2 + E_z b_y b_z) + \mu_\perp (-E_x b_x b_y + E_y(1 - b_y^2) - E_z b_y b_z)$$

$$+ \mu_\alpha (E_z b_x - E_x b_z) - D_\parallel \left(\frac{\partial n}{\partial x} b_x b_y + \frac{\partial n}{\partial y} b_y^2 + \frac{\partial n}{\partial z} b_y b_z \right)$$

$$- D_\perp \left(-\frac{\partial n}{\partial x} b_x b_y + \frac{\partial n}{\partial y}(1 - b_y^2) - \frac{\partial n}{\partial z} b_y b_z \right) + D_\alpha \left(\frac{\partial n}{\partial x} b_z - \frac{\partial n}{\partial z} b_x \right)$$

$$\text{(A.22b)}$$

$$nu_z = \mu_\parallel (E_x b_x b_z + E_y b_y b_z + E_z b_z^2) + \mu_\perp (-E_x b_x b_z - E_y b_y b_z + E_z(1 - b_z^2))$$

$$+ \mu_\alpha (E_z b_x - E_x b_z) - D_\parallel \left(\frac{\partial n}{\partial x} b_x b_z + \frac{\partial n}{\partial y} b_y b_z + \frac{\partial n}{\partial z} b_z^2 \right)$$

$$- D_\perp \left(-\frac{\partial n}{\partial x} b_x b_z - \frac{\partial n}{\partial y} b_y b_z + \frac{\partial n}{\partial z}(1 - b_z^2) \right) + D_\alpha \left(\frac{\partial n}{\partial y} b_x - \frac{\partial n}{\partial x} b_y \right)$$

$$\text{(A.22c)}$$

Lastly, to obtain the form that is used in this dissertation, I modify equation A.22 in the following form. For the sake of simplicity, I will only derive the modified equations for the x component.

First, the terms involving the $\dfrac{\partial n}{\partial x}$ are grouped to obtain

$$- \left(D_\parallel b_x^2 + D_\perp(1 - b_x^2) \right) \frac{\partial n}{\partial x} = D b_x^2 + \frac{D}{1 + \Omega^2}(1 - b_x^2) \frac{\partial n}{\partial x} = \frac{(1 + \Omega^2) D b_x^2 + D(1 + b_x^2)}{1 + \Omega^2} \frac{\partial n}{\partial x}$$

$$= D(1 + \Omega^2 b_x^2) \frac{\partial n}{\partial x}$$

Similarly, grouping the terms involving E_x we obtain

$$\left(\mu_\parallel b_x^2 + \mu_\perp(1 - b_x^2) \right) E_x = \mu b_x^2 + \frac{\mu}{1 + \Omega^2}(1 - b_x^2) E_x = \frac{(1 + \Omega^2)\mu b_x^2 + \mu(1 + b_x^2)}{1 + \Omega^2} \boldsymbol{E} =$$

$$\mu(1 + \Omega^2 b_x^2) \boldsymbol{E}$$

Next, grouping the terms involving the term $\dfrac{\partial n}{\partial y}$ yields

$$-\left(D_\parallel b_x b_y + D_\perp b_x b_y + D_\alpha b_z\right)\frac{\partial n}{\partial y} = Db_x b_y - \frac{b_x b_y}{1 + \Omega^2} + \frac{Db_z}{1 + \Omega^2}\frac{\partial n}{\partial y}$$
$$=\frac{\Omega_x \Omega_y + \Omega_z}{1 + \Omega^2}\frac{\partial n}{\partial y}$$

Similarly, for the terms involving E_y

$$\left(\mu_\parallel b_x b_y + \mu_\perp b_x b_y + D_\alpha h_z\right) E_y = Db_x b_y - \frac{b_x b_y}{1 + \Omega^2} + \frac{Db_z}{1 + \Omega^2}E_y$$
$$=\frac{\Omega_x \Omega_y + \Omega_z}{1 + \Omega^2}\frac{\partial n}{\partial y}$$

Finally, the terms involving E_z are obtained analogously. Thus, the total nu_x term can be rewritten as

$$nu_x =\frac{1 + \Omega_x^2}{1 + \Omega^2}\left[\mu n E_x - D\frac{\partial n}{\partial x}\right] + \frac{\Omega_x \Omega_y + \Omega_z}{1 + \Omega^2}\left[\mu n E_y - D\frac{\partial n}{\partial y}\right]$$
$$\frac{\Omega_x \Omega_z - \Omega_y}{1 + \Omega^2}\left[\mu n E_z - D\frac{\partial n}{\partial z}\right]$$

and

$$nu_y =\frac{1 + \Omega_y^2}{1 + \Omega^2}\left[\mu n E_y - D\frac{\partial n}{\partial y}\right] + \frac{\Omega_y \Omega_x - \Omega_z}{1 + \Omega^2}\left[\mu n E_x - D\frac{\partial x}{\partial x}\right]$$
$$\frac{\Omega_y \Omega_z + \Omega_x}{1 + \Omega^2}\left[\mu n E_z - D\frac{\partial n}{\partial z}\right]$$
$$nu_z =\frac{1 + \Omega_z^2}{1 + \Omega^2}\left[\mu n E_z - D\frac{\partial n}{\partial z}\right] + \frac{\Omega_z \Omega_x + \Omega_y}{1 + \Omega^2}\left[\mu n E_x - D\frac{\partial n}{\partial x}\right]$$
$$\frac{\Omega_z \Omega_y - \Omega_x}{1 + \Omega^2}\left[\mu n E_y - D\frac{\partial n}{\partial y}\right]$$

where $\boldsymbol{\Omega} = \Omega\boldsymbol{B}/B$.

APPENDIX B

Derivation of Jacobian matrix

The terms in equation 5.51 are straightforward to calculate with the exception of the following

$$\frac{\partial F_n}{\partial \phi_C} = n_C \frac{\partial \alpha_C}{\partial \phi_C} + n_E \frac{\partial \alpha_E}{\partial \phi_C} + n_W \frac{\partial \alpha_W}{\partial \phi_C}$$

$$\frac{\partial F_n}{\partial \phi_E} = n_C \frac{\partial \alpha_C}{\partial \phi_E} + n_E \frac{\partial \alpha_E}{\partial \phi_E}$$

$$\frac{\partial F_n}{\partial \phi_W} = n_C \frac{\partial \alpha_C}{\partial \phi_W} + n_W \frac{\partial \alpha_W}{\partial \phi_W}$$

Some of this terms involve differentiation of exponential terms. If we use the following template

$$\frac{\partial}{\partial b} \left(\frac{a(b-c)}{1 - e^{m(b-c)}} \right) = \frac{-ae^{m(b-c)} \left[1 - m(b-c) \right] + a}{\left(1 - e^{m(b-c)} \right)^2}$$

$$\frac{\partial}{\partial c} \left(\frac{a(b-c)}{1 - e^{m(b-c)}} \right) = \frac{ae^{m(b-c)} \left[1 - m(b-c) \right] - a}{\left(1 - e^{m(b-c)} \right)^2}$$

$$\frac{\partial}{\partial b} \left(\frac{a(b-c)e^{m(b-c)}}{1 - e^{m(b-c)}} \right) = \frac{(1 - e^{m(b-c)})(ae^{m(b-c)})[m(b-c)+1] + \left(am(b-c)e^{2m(b-c)} \right)}{\left(1 - e^{m(b-c)} \right)^2}$$

$$\frac{\partial}{\partial c} \left(\frac{a(b-c)e^{m(b-c)}}{1 - e^{m(b-c)}} \right) = \frac{-(1 - e^{m(b-c)})(ae^{m(b-c)})[m(b-c)+1] - \left(am(b-c)e^{2m(b-c)} \right)}{\left(1 - e^{m(b-c)} \right)^2}$$

to obtain the partial derivatives whenever there is an exponential term, we
have, for electrons

$$\frac{\partial \alpha_E}{\partial \phi_C} = \frac{ae^{m(\phi_E - \phi_C)}(1 - m(\phi_E - \phi_C)) - a}{(1 - e^{m(\phi_E - \phi_C)})^2}$$

$$\frac{\partial \alpha_E}{\partial \phi_E} = \frac{-ae^{m(\phi_E - \phi_C)}(1 - m(\phi_E - \phi_C)) + a}{(1 - e^{m(\phi_E - \phi_C)})^2}$$

$$\frac{\partial \alpha_W}{\partial \phi_C} = \frac{(1 - e^{m(\phi_C - \phi_W)})(ae^{m(\phi_C - \phi_W)})(1 + m(\phi_C - \phi_W)) + \left(am\left(\phi_C - \phi_W\right)e^{2m(\phi_C - \phi_W)}\right)}{(1 - e^{m(\phi_C - m\phi_W)})^2}$$

$$\frac{\partial \alpha_W}{\partial \phi_W} = \frac{-(1 - e^{m(\phi_C - \phi_W)})(ae^{m(\phi_C - \phi_W)})(1 + m(\phi_C - \phi_W)) - \left(am(\phi_C - \phi_W)e^{2m(\phi_C - \phi_W)}\right)}{(1 - e^{m(\phi_C - m\phi_W)})^2}$$

$$\frac{\partial \alpha_C}{\partial \phi_E} = \frac{-(1 - e^{m(\phi_E - \phi_C)})(ae^{m(\phi_E - \phi_C)})(1 + m(\phi_E - \phi_C)) - \left(am(\phi_E - \phi_C)e^{2m(\phi_E - \phi_C)}\right)}{(1 - e^{m(\phi_E - m\phi_C)})^2}$$

$$\frac{\partial \alpha_C}{\partial \phi_W} = \frac{-ae^{m(\phi_C - \phi_W)}(1 - m(\phi_C - \phi_W)) + a}{(1 - e^{m(\phi_C - m\phi_W)})^2}$$

$$\frac{\partial \alpha_C}{\partial \phi_C} = \frac{(1 - e^{-m(\phi_E - \phi_C)})(ae^{-m(\phi_E - \phi_C)})(1 + m(\phi_E - \phi_C)) + \left(am(\phi_E - \phi_C)e^{2m(\phi_E - \phi_C)}\right)}{(1 - e^{m(\phi_E - m\phi_C)})^2}$$
$$+ \frac{ae^{m(\phi_C - \phi_W)}(1 - m(\phi_C - \phi_W)) - a}{(1 - e^{m(\phi_C - m\phi_W)})^2}$$

In a similar way, we obtain for ions

$$\frac{\partial \beta_E}{\partial \phi_C} = \frac{-ae^{m(\phi_C - \phi_E)}(1 - m(\phi_C - \phi_E)) + a}{(1 - e^{m(\phi_C - \phi_E)})^2}$$

$$\frac{\partial \beta_E}{\partial \phi_E} = \frac{ae^{m(\phi_C - \phi_E)}(1 - m(\phi_C - \phi_E)) - a}{(1 - e^{m(\phi_C - \phi_E)})^2}$$

$$\frac{\partial \beta_W}{\partial \phi_C} = \frac{-(1 - e^{m(\phi_W - \phi_C)})(ae^{m(\phi_W - \phi_C)})(1 + m(\phi_W - \phi_C)) - \left(am\,(\phi_W - \phi_C)\,e^{2m(\phi_W - \phi_C)}\right)}{\left(1 - e^{m(\phi_W - m\phi_C)}\right)^2}$$

$$\frac{\partial \beta_W}{\partial \phi_W} = \frac{(1 - e^{m(\phi_W - \phi_C)})(ae^{m(\phi_W - \phi_C)})(1 + m(\phi_W - \phi_C)) + \left(am(\phi_W - \phi_C)e^{2m(\phi_W - \phi_C)}\right)}{\left(1 - e^{m(\phi_W - m\phi_C)}\right)^2}$$

$$\frac{\partial \beta_C}{\partial \phi_E} = \frac{(1 - e^{m(\phi_C - \phi_E)})(ae^{m(\phi_C - \phi_E)})(1 + m(\phi_C - \phi_E)) + \left(am(\phi_C - \phi_E)e^{2m(\phi_C - \phi_E)}\right)}{\left(1 - e^{m(\phi_C - m\phi_E)}\right)^2}$$

$$\frac{\partial \beta_C}{\partial \phi_W} = \frac{ae^{m(\phi_W - \phi_C)}(1 - m(\phi_W - \phi_C)) - a}{\left(1 - e^{m(\phi_W - m\phi_C)}\right)^2}$$

$$\frac{\partial \beta_C}{\partial \phi_C} = \frac{-(1 - e^{-m(\phi_C - \phi_E)})(ae^{-m(\phi_C - \phi_E)})(1 + m(\phi_C - \phi_E)) - \left(am(\phi_C - \phi_E)e^{2m(\phi_C - \phi_E)}\right)}{\left(1 - e^{m(\phi_C - m\phi_E)}\right)^2}$$
$$+ \; \frac{-ae^{m(\phi_W - \phi_C)}(1 - m(\phi_W - \phi_C)) + a}{\left(1 - e^{m(\phi_W - m\phi_C)}\right)^2}$$

The extension of the fluid equations to 3D is straightforward (only shown for electrons)

$$\frac{n_C^{k+1} - n_C^k}{\Delta t} + \frac{J_e^{k+1} - J_w^{k+1}}{\Delta x} + \frac{J_n^{k+1} - J_s^{k+1}}{\Delta y} + \frac{J_f^{k+1} - J_b^{k+1}}{\Delta z} = S \qquad \text{(B.1)}$$

$$\alpha_C n_c + \alpha_E n_E + \alpha_W n_W + \alpha_N n_N + \alpha_S n_S + \alpha_F n_F + \alpha_B n_B = S\Delta t + n_i^k \quad \text{(B.2)}$$

where

$$\alpha_E = a_e \left(\frac{(\phi_E - \phi_C)}{1 - e^{m_e(\phi_E - \phi_C)}} \right) \quad \alpha_W = a_w \left(\frac{(\phi_C - \phi_W)e^{m_w(\phi_C - \phi_W)}}{1 - e^{m_w(\phi_C - \phi_W)}} \right)$$

$$\alpha_N = a_n \left(\frac{(\phi_N - \phi_C)}{1 - e^{m_n(\phi_N - \phi_C)}} \right) \quad \alpha_S = a_s \left(\frac{(\phi_C - \phi_S)e^{m_s(\phi_C - \phi_S)}}{1 - e^{m_s(\phi_C - \phi_S)}} \right)$$

$$\alpha_F = a_f \left(\frac{(\phi_F - \phi_C)}{1 - e^{m_f(\phi_F - \phi_C)}} \right) \quad \alpha_B = a_b \left(\frac{(\phi_C - \phi_B)e^{m_b(\phi_C - \phi_B)}}{1 - e^{m_b(\phi_C - \phi_B)}} \right)$$

$$\alpha_C = V_{cell} - a_e \left(\frac{(\phi_E - \phi_C)e^{m_e(\phi_E - \phi_C)}}{1 - e^{m_e(\phi_E - \phi_C)}} \right) - a_w \left(\frac{(\phi_C - \phi_W)}{1 - e^{m_w(\phi_C - \phi_W)}} \right) -$$

$$a_n \left(\frac{(\phi_N - \phi_C)e^{m_n(\phi_N - \phi_C)}}{1 - e^{m_n(\phi_N - \phi_C)}} \right) - a_s \left(\frac{(\phi_C - \phi_S)}{1 - e^{m_s(\phi_C - \phi_S)}} \right) -$$

$$a_f \left(\frac{(\phi_F - \phi_C)e^{m_f(\phi_F - \phi_C)}}{1 - e^{m_f(\phi_F - \phi_C)}} \right) - a_b \left(\frac{(\phi_C - \phi_B)}{1 - e^{m_b(\phi_C - \phi_B)}} \right)$$

where

$$a_e = \frac{A_e \mu_e \Delta t}{\Delta x} \quad a_w = \frac{A_w \mu_w \Delta t}{\Delta x} \quad a_n = \frac{A_n \mu_n \Delta t}{\Delta y}$$

$$a_s = \frac{A_s \mu_s \Delta t}{\Delta y} \quad a_f = \frac{A_f \mu_f \Delta t}{\Delta z} \quad a_b = \frac{A_b \mu_b \Delta t}{\Delta z}$$

$$m_e = \frac{\mu_e}{D_e} \quad m_w = \frac{\mu_w}{D_w} \quad m_n = \frac{Z \mu_n}{D_n}$$

$$m_s = \frac{\mu_s}{D_s} \quad m_f = \frac{\mu_f}{D_f} \quad m_b = \frac{Z \mu_b}{D_b}$$

and A stands for the contact area between cells.