Advances in ultra-high dose rate radiation therapy

Julien Bancheri

Department of Physics
McGill University
Montréal, Québec, Canada

November 19, 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Acknowledgements

First, I would like to express my utmost gratitude to my supervisor Dr. Jan Seuntjens. Your constant support and confidence in me have enabled me to keep going through uncertain times and keep going until the completion of this work. Your wisdom allowed this work and by extension myself to succeed scientifically and professionally. I look forward to how our paths will cross as I pass into my professional life.

I am deeply thankful to Dr. Anatoly Krasnykh for his invaluable mentorship during my visit to SLAC, without which none of the pulser work would have been possible. Your deep knowledge on all aspects of the DWA and DSRDs and our many discussions have made the work possible and greatly increased its depth and competency.

I would like to thank my PhD supervisory committee, Dr. Shirin A. Enger, Dr. Malcolm McEwen, Dr. Anatoly Krasnykh, Dr. Hugo Palmans and Dr. Harald Paganetti for their expertise and wisdom. It has greatly aided me in steering this research in the right direction.

I would like to acknowledge NSERC and FRQNT for providing funding for this work.

There are several collaborators at the Princess Margaret Cancer Centre that I would like to thank. Andrew Currell for assembling the MOSFET pulser and testing it. I have really enjoyed all of our discussions and you have taught me a lot about electronics. Bern Norrlinger for your help in assembling the MST pulser and the many discussions. Jake Broske

and Matthew Filleti of the machine shop for the excellent items they have machined and fulfilling my countless requests.

I must acknowledge Margery Knewstubb and Tatjana Nisic for always dealing with my many administrative questions and requests.

Many thanks to the faculty and staff of the McGill Medical Physics Unit as well as towards my friends, colleagues and fellow students; Chris, Morgan, Veng, Alaina, Jason, James, Liam, Santiago, Felix and Bryan. You all have made this journey so much more enjoyable and I will always remember our times together fondly.

Finally, I express my greatest thanks to my family. Ayleigh, who carried me through every moment and kept me going, Mom, Dad, Andreas and Kristian for constant support and love. Cam, Lynn, Courtney and Robbie for continual interest and friendship.

Abstract

Ultra-high dose rate (UHDR) radiation therapy (RT) has recently gained significant interest. Studies, mostly carried out in proton beams, show that UHDR reduces healthy tissue damage in comparison to conventional dose rates. The rise in UHDR-RT has highlighted the need for cost-effective proton acceleration methods. Proton RT is still relatively inaccessible due to high costs and the large size of the proton accelerators and gantries. One solution to address these issues is the development of novel proton acceleration methods. In terms of UHDR dosimetry, the measurement of absorbed dose in UHDRs beams is compromised by ion recombination in commercial ionization chambers. Hence, it is relevant to revisit the theory of ion recombination to expand its applicability to the UHDR regime. Previous theoretical investigations and methods used to measure the ion recombination correction factor in the clinic do not hold above 10 mGy dose-per-pulse (DPP). The main goal of this thesis is to address these two needs that arise from the introduction of UHDR radiation therapy, including proton therapy.

The Dielectric Wall Accelerator (DWA) is a compact, cost-effective proton acceleration method. The DWA accelerates protons with high accelerating gradients that are formed by injecting high voltage (HV) pulses into radial waveguides. The HV pulses are formed by a pulser with a semiconductor switch. The pulse width must be on the order of ns and should operate above 1 kHz to achieve UHDRs. The drift step recovery diode (DSRD) is a Silicon opening switch that can withstand HV and operate in the ns scale, which makes it a suitable candidate for the switch. In order to determine if a DSRD-based pulser is suitable

for the DWA in terms of the pulse requirements, two pulser designs were built and tested; a magnetic saturation transformer (MST)-DSRD-based pulser and a MOSFET-DSRD-based multi-module pulser. The MOSFET-DSRD-based pulser consistently generated pulses with shorter widths, the shortest being 3 ns. The largest amplitude pulse (10.9 kV, rise time 1.66 ns, pulse width 3.48 ns) was also generated by the MOSFET-DSRD-based pulser with six modules. Both pulsers were able to be run at 1 kHz and up to 10 kHz for the MOSFET-DSRD-based pulser. The pulse stability of the MOSFET-DSRD-based pulser was within 1% up to 10 kHz. It was shown that parasitic coupling between the modules of the MOSFET-DSRD-based pulser affects the pulse through the MOSFET on-time. Simulations with Sentaurus TCAD by Synopsys® studied the influence of the DSRD doping profile and material on the pulse. To reduce the pedestal effect, the width of the p doped region should be larger than the width of the n doped region. Silicon carbide DSRDs result in lower amplitude pulses and longer pulse widths, due to incomplete ionization.

In order to solve the partial differential equations (PDEs) describing the charge carrier transport and space charge, a semi-analytical solution method based on the homotopy perturbation method (HPM) is employed. The resulting analytical expression for the ion recombination correction factor was used in a fit procedure with published measured data. The estimate of the charge collection efficiency obtained with this fit procedure were within 6% of the published values for (1) values below 1 Gy DPP for a 1 mm plate separation parallel-plate chamber and (2) below 3 Gy DPP for a 0.5 mm plate separation. A fit based on a general equation consisting of a constant term and a $1/V^2$ term (V is the chamber voltage) performed just as well.

The work carried out in this thesis addresses two needs brought about by UHDR radiation therapy. The investigations into the MOSFET-DSRD-based pulser and its driving conditions pave the way for the development of the DWA system and achieving low-cost, accessible

UHDR proton RT. The method presented for determining the ion recombination correction factor allows for its determination in UHDR beams in the clinical setting.

Abrégé

La radiothérapie à ultra-haut débit de dose (UHDR) est devenue un sujet d'intérêt majeur en raison de son potentiel pour réduire les dommages aux tissus sains par rapport aux débits de dose conventionnels. Cette approche, étudiée principalement avec des faisceaux de protons, a suscité un besoin urgent de technologies rentables et compactes pour l'accélération des protons. En effet, la protonthérapie reste inaccessible à grande échelle en raison des coûts élevés des infrastructures, notamment la taille et la complexité des accélérateurs et des portiques. L'une des solutions à ces problèmes consiste à mettre au point de nouvelles méthodes d'accélération des protons. En termes de dosimétrie UHDR, la mesure de la dose absorbée dans les faisceaux UHDR est compromise par la recombinaison des ions dans les chambres d'ionisation commerciales. Il est donc important de revoir la théorie de la recombinaison des ions afin d'étendre son applicabilité au régime UHDR. Les études théoriques précédentes et les méthodes utilisées pour mesurer le facteur de correction de la recombinaison des ions en clinique ne sont pas valables au-delà de 10 mGy de dose par impulsion (DPP). L'objectif principal de cette thèse est de répondre à ces deux besoins qui découlent de l'introduction de la radiothérapie UHDR, y compris la protonthérapie.

Le DWA est une méthode d'accélération innovante qui exploite des gradients d'accélération élevés, générés par des impulsions haute tension (HV) injectées dans des guides d'ondes radiaux. Ces impulsions, formées par un pulseur doté d'un commutateur à semi-conducteur, doivent présenter une largeur temporelle de quelques nanosecondes et une fréquence d'au moins 1 kHz pour répondre aux critères UHDR. La diode de récupération de l'étape de

dérive (DSRD), un commutateur silicium, se distingue par sa capacité à fonctionner à haute tension et à l'échelle du nanoseconde, faisant d'elle un choix idéal. Deux prototypes de pulseurs basés sur la DSRD ont été conçus et testés : un à transformateur magnétique à saturation (MST) et un multi-modules MOSFET-DSRD. Les tests ont révélé que le modèle MOSFET-DSRD générait systématiquement des impulsions plus courtes, avec une largeur minimale de 3 ns, et des amplitudes plus élevées, atteignant 10,9 kV avec un temps de montée de 1,66 ns. Ce prototype a fonctionné de manière stable jusqu'à 10 kHz, avec une variation de l'ordre de 1%. Les simulations TCAD ont permis d'optimiser le profil de dopage des DSRD et d'évaluer les performances des matériaux alternatifs. Les résultats ont montré que l'élargissement de la région dopée en p réduisait l'effet de piédestal. Les DSRD en carbure de silicium généraient des impulsions plus longues et moins intenses, en raison d'une ionisation incomplète.

Une méthode semi-analytique, basée sur la perturbation homotopique (HPM), a été développée pour résoudre les équations différentielles décrivant le transport des charges et la charge d'espace. L'expression analytique obtenue pour le facteur de correction de la recombinaison des ions a été confrontée à des données expérimentales publiées. Cette approche a démontré une précision de l'ordre de 6%, pour des doses inférieures à 1 Gy DPP dans une chambre à plaques parallèles avec une séparation de 1 mm. Une méthode alternative, basée sur une équation générale comprenant un terme constant et un terme en $1/V^2$ (où V est la tension de la chambre), a produit des résultats similaires. Ces avancées permettent une application fiable des chambres d'ionisation en clinique, même dans le régime UHDR.

Les travaux présentés dans cette thèse apportent des contributions significatives à deux besoins majeurs de la radiothérapie UHDR. Le développement de pulseurs basés sur la DSRD ouvre la voie à des systèmes DWA compacts et économiques pour une protonthérapie UHDR plus accessible. Parallèlement, l'amélioration des techniques de dosimétrie garantit

une meilleure précision dans les environnements cliniques exigeants.

Contents

List of Figures

List of Tables

Abbreviations

Linac	Linear accelerator
PSDL	Primary standards dosimetry laboratory
DPP	Dose-per-pulse
TVM	Two-voltage method
SOBP	Spread-out Bragg peak
UHDR	Ultra-high dose rate
DWA	Dielectric wall accelerator
HGI	High gradient insulator
DSRD	Drift step recovery diode
HV	High voltage
LET	Linear energy transfer
PDD	Percent depth dose distribution
RF	Radiofrequency
LIA	Linear induction accelerator
MST	Magnetic saturation transformer
HPM	Homotopy perturbation method

Contribution to Original Knowledge

This thesis contains original research carried out in order to address current needs that arise from the increasing prevalence of ultra-high dose rate radiotherapy and has resulted in two published manuscripts. The first manuscript is part of a concerted effort to develop a compact DWA system for cost effective and accessible proton therapy. Two high voltage DSRD-based pulsers were built and evaluated in terms of the output pulses. Multiple circuit parameters are varied in order to observe how the output pulses change and how far they can be driven, in terms of pulse width, rise time and magnitude. Simulations of each pulser with different DSRD doping profiles were also carried out. To the authors' knowledge, this is the only comparative study of DSRD-based pulsers and DSRD doping profiles. The performance of a MOSFET-DSRD-based pulser with up to six modules is also presented for the first time, leading to higher pulse amplitudes and shorter pulse widths. Additionally, this is the only study that provides such an in-depth study of how the DSRD doping profile influences the output pulses. This is necessary to determine the suitability of a pulser and DSRD doping profile for the DWA. The use of SiC DSRDs with the MOSFET-DSRD-based pulser has also not yet been studied. The second manuscript describes the use of a semi-analytical solution method to solve the charge carrier transport PDEs in UHDR beams and obtain an analytical expression for the ion recombination correction factor. To the authors' knowledge, there is no analytical expression for the correction factor in UHDR beams or extrapolation method in UHDR beams that can be used in the clinic.

Contribution of Authors

Chapters 1-6 contain the introduction and a review on relevant theory and literature such as radiation physics and dosimetry, ion recombination, particle acceleration methods and semiconductor physics and simulations. Chapter 9 is the conclusion that summarizes the work done in this thesis and provides future research direction. These chapters were written by myself and proofread by Jan Seuntjens. Chapters 7 and 8 are manuscripts for which the contributions are as follows:

1. Julien Bancheri, Andrew Currell, Anatoly Krasnykh, Morgan Maher, Christopher Lund, Alaina Giang Bui and Jan Seuntjens. Evaluation of DSRD-based pulsers for a Dielectric Wall Accelerator. Submitted to NIM A. 2024.

   I performed the assembly and all of the measurements with the MST-DSRD-based pulser and performed the DSRD simulations. The machine shop at Princess Margaret Cancer Centre, Toronto machined the board and the plastic holders. Andrew Currell assembled and tested the MOSFET-DSRD-based pulser. Anatoly Krasnykh is part of the DWA project and provided technical support and advice. Morgan Maher, Christopher Lund and Alaina Giang Bui are part of the DWA project and provided additional information on the DWA radial waveguides and beam optics. Jan Seuntjens provided oversight, regular discussion and support of the project. All authors commented on the paper contents as part of the initial version and revised version.

2. Julien Bancheri and Jan Seuntjens. A semi-analytical procedure to determine the ion recombination correction factor in high dose-per-pulse beams. Med Phys. 2024; 51: 4458–4471. https://doi.org/10.1002/mp.17005

I carried out all of the theoretical calculations and analysis. Jan Seuntjens provided oversight, regular discussion and support of the project. Both authors commented on the paper contents and reviewed the submitted and final version of the article.

1

Introduction

1.1 Cancer statistics and treatments

In 2022, approximately 20 million new cancer cases and 10 million cancer deaths were reported globally [1]. For most developed countries, the incidence rate (per 100000 people) for all cancer types is slowing or decreasing and the mortality rate (per 100000 people) is decreasing [1]. In Canada, it is estimated that there will be 247000 new cancer cases and 88100 cancer deaths in 2024 [2], up from 2023. It is currently the leading cause of death in Canada with the incidence and mortality rates decreasing overall. This can be attributed to increasing cancer screening and prevention efforts. The most common types of cancers found among Canadian males and females are lung and colorectal, breast (females) and prostate (males), making up almost half of new cancer cases (47%). While the incidence rates of these common cancers are decreasing, the rates of rarer types of cancer such as melanoma, kidney

and head and neck are increasing [2]. Focusing on age-specific sub-populations, the incidence rate in the children age group (0-14) is increasing while that for young adults (15-29) and adults (30+) is decreasing for males and slowing for females [3]. As the burden of cancer increases on Canadian society due to an aging population and relatively long lifespans, it is crucial to continue any cancer awareness and screening programs as well as research into novel treatment modalities and techniques.

There are several treatment modalities that are currently available to cancer patients. While the ultimate goal of cancer treatment is to completely remove the cancer from the patient, these treatments are also used to limit the growth of tumours and the spread of cancer. The most common treatment method is surgery, whereby the tumour volume and some surrounding healthy tissue and lymph nodes are physically removed from the patient. The lymph nodes are sometimes removed because they contribute to the spread of cancer in the body. If the tumour cannot be fully removed from the patient for some reason, such as potential damage to an adjacent organ, part of the the tumour may be removed and the rest treated with a different modality. Chemotherapy involves the use of a drug or a combination of drugs to kill the cancer cells and limit their growth and division. Several types of chemotherapy drugs are available depending on the type and stage of the cancer as well as the patient health and age and any other additional cancer treatments the patient may undergo. Radiation therapy is another common form of cancer treatment whereby ionizing radiation is used to kill the cancer cells. While these three treatment modalities are the most common, there are other available treatments, some of which are still in the research phase. Immunotherapy involves the use of drugs to aid the immune system's response to the cancer. Targeted therapy employs drugs that target the cancer proteins that aid in the cell's division and viability. While primarily in the research phase, gene therapy involves the mutation of the cancer cell genes [4].

1.2 Radiation therapy and its effect on cancer

The primary cancer treatment modality this thesis focuses on is radiation therapy. Radiation therapy only employs *ionizing radiation*, that is, radiation capable of ionizing atoms and liberating orbital electrons. *Non-ionizing radiation* is another class of radiation but is not employed in radiation therapy. Therefore, for the remainder of this thesis, radiation will refer solely to ionizing radiation. The ionizing radiation used in radiation therapy can be further sub-divided into two classes. One class is *directly ionizing radiation*, which consists of charged particles such as electrons, protons and light ions (carbon, helium, oxygen, etc). A charged particle interacts with an atom through the Coulomb force and can either excite or ionize it. In the case of ionization, an orbital electron is freed which can further ionize other atoms and liberate other electrons. As a charged particle traverses through a medium, it loses kinetic energy through these interactions and radiative processes (see Section 2.2). This energy loss (except that from radiative processes) can be considered as energy deposited to a local region of the medium. From this, a definition of absorbed dose D can be developed,

$$D = \frac{\mathrm{d}\bar{\epsilon}}{\mathrm{d}m}, \tag{1.1}$$

where $d\bar{\epsilon}$ is the mean energy deposited by the radiation in a local region of mass dm. The units of absorbed dose is J/kg or Gray (Gy). In the case of ion beams, after a collision with a nucleus, both the projectile and the nucleus may fragment. The daughter fragments of the projectile still maintain a high velocity and contribute to the absorbed dose further downstream. *Indirectly ionizing radiation* consists of neutral particles such as photons and neutrons. Energy and thus absorbed dose is deposited through a two-step process. For photons, the photon interacts with an orbital electron and frees it, transferring to it some or all of its kinetic energy. This electron is known as a *secondary electron*. Then, this secondary electron undergoes its own interactions in the same manner as those mentioned above. In contrast to charged particles, a neutral particle may only undergo a few interactions as it tra-

verses a medium. Neutrons interact mainly with the nucleus and their interaction produces secondary nucleons. It is these secondary nucleons that fully contribute to the absorbed dose.

Radiation therapy is an effective treatment method because these radiation-matter interactions occur with the DNA of the cancer cells, causing DNA damage. The DNA damage affects the cell's ability to repair itself, undergo mitosis and can even cause cell death (known as lethal damage) [5]. If the damage to the cell does not induce cell death, the damage is known as sub-lethal. The damage to the cellular DNA occurs when a single or both strands of the DNA is broken. These are known as single and double strand breaks, respectively. Macroscopically, this leads to the elimination or reduction of cancer in the patient body. When radiation interacts directly with one of the constituent DNA molecules, it is known as *direct action*. Direct action occurs more frequently for radiation types that deposit a large amount of energy locally, such as neutrons, protons and light ions. The more common mechanism of cell DNA damage occurs through *indirect action*. In this case, the radiation interacts with other targets, mainly water molecules (a cell is comprised of about 80% water). The excitation of the water molecule or the subsequent interactions of the secondary electron generate so-called *free radicals*, such as aqueous electrons e^-_{aq}, hydroxyl radicals $\bullet$OH and ionized hydrogen H^+. These free radicals are highly reactive with the molecules of the cell DNA and can cause DNA damage through the breakage of chemical bonds. Indirect action is more common for high energy electrons and photons (typical in conventional radiation therapy). There are several factors which influence a cell's radiosensitivity. For example, during the cell life cycle, the cell is most radiosensitive during the mitosis (M) and second gap (G2) phase and most radioresistant during the DNA synthesis (S) stage. The presence of oxygen also increases radiosensitivity.

Radiation-matter interactions make no distinction between cancer cells and healthy cells. Therefore, healthy tissues is also damaged by radiation exposure. Fortunately, most mam-

malian normal tissues are much less sensitive to radiation than cancer cells [6]. To improve the sparing of healthy tissue and the killing of cancer cells, *fractionation* is employed. Fractionation is when the total radiation dose to the patient is spread out over time, usually days and weeks. Fractionation takes advantage of five biological processes that occur after cellular irradiation: (1) radiosensitivity, where healthy tissues are less radiosensitive, (2) repair, where sublethal damage to healthy cells is repaired in between fractions, (3) repopulation, where the healthy cell population grows in between fractions, (4) redistribution, where cells redistribute within the cell cycle and increase the amount of cells in the radiosensitive stages and (5) reoxygentation, where oxygen is reintroduced into the cell population after a fraction.

1.3 External beam radiation therapy

Radiation therapy most often employs photon or electron beams that are generated externally to the patient and aimed towards them. This type of radiation therapy is known as *external beam radiation therapy*. Low and medium energy photon beams in the 40 - 300 keV range are generated by x-ray tubes. They are not considered in this thesis and will not be discussed further. High energy photon or electron beams in the MeV range are generated by particle accelerators known as a medical *linear accelerators* or *linacs* for short. These linacs can accelerate electrons up to energies of 22 MeV. A diagram of a medical linac and its components is shown in Figure 1.1. There are several components of a linac involved in the radiation beam generation. The *injection system* is the source of the electrons and often called the electron gun. It is comprised of a filament (cathode), grid and anode. Electrons are thermionically ejected from the heated filament which is kept at a constant potential of -25 kV. When the beam is "off", the grid is kept at a lower potential than the cathode to keep the electrons from leaving the gun. When the beam is "on", the grid is switched to a positive potential by the *pulsed modulator* and the electrons are accelerated towards the anode and enter the *accelerating waveguide*. The waveguide is the main accelerating structure

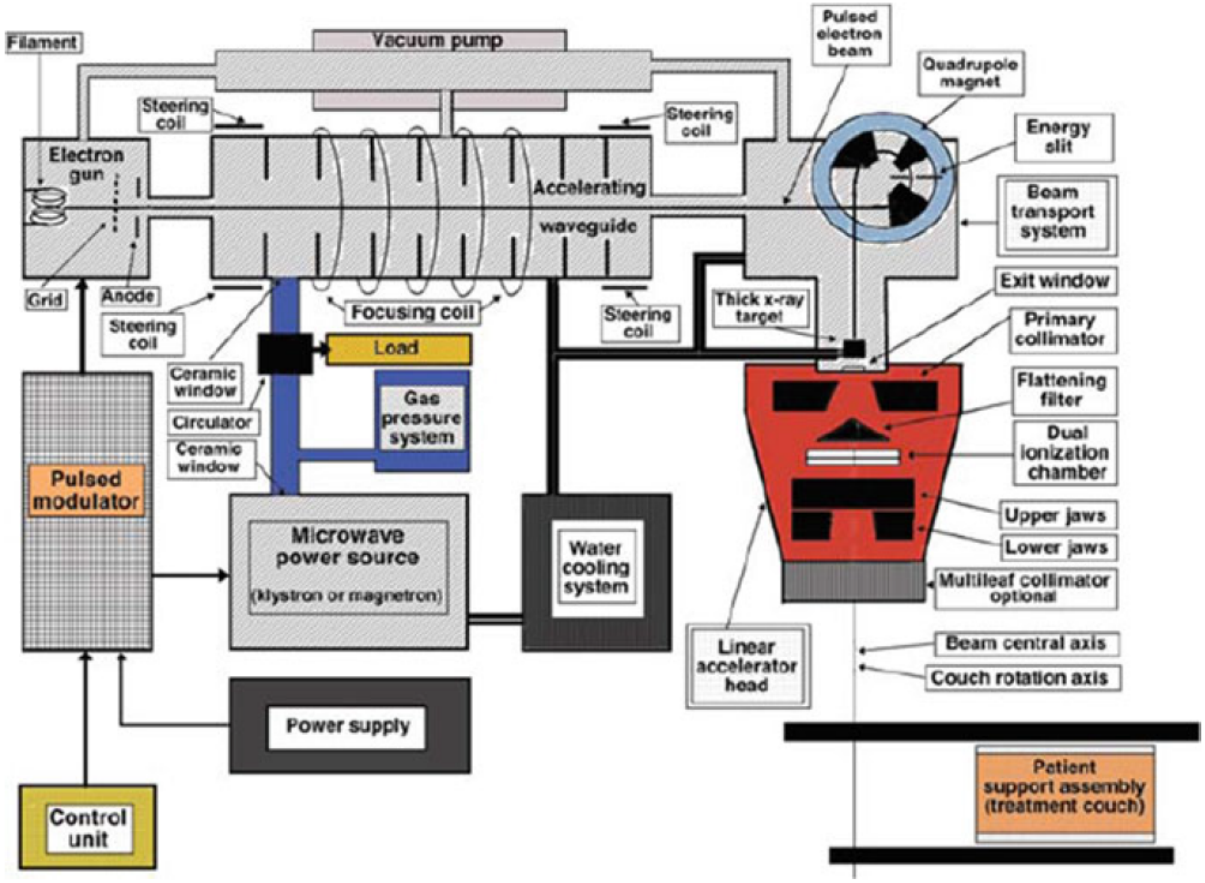


Figure 1.1: A diagram of a medical linear accelerator (linac) and its components. Reproduced with permission from [7].

of the linac and is fed microwave RF power from a RF power source, either a magnetron or klystron. In modern high energy medical linacs, the klystron is most common. The pulsed modulator also powers the RF source and keeps it and the injection system in-sync. Due to the pulsed nature of the modulator and the klystron, the final photon or electron beam is pulsed in nature. The waveguide is in fact a disk-loaded waveguide where disks separate the waveguide into individual cavities. The electric field established in these cavities accelerates the electrons. The theory of electron acceleration with this type of waveguide is discussed in detail in Section 5.1. The waveguide is kept under vacuum. Steering coils establish magnetic fields that are used to steer and focus the electron beam. The *beam transport system* bends the beam towards the exit window and includes a slit that removes any electrons at undesired

energies. A *dose monitoring system*, comprising of two independent ionization chambers (see Section 3.3), known as monitor chambers, is used to ensure the linac is delivering an appropriate amount of radiation dose. These monitor chambers measure monitor units (MU), which correspond to a specific absorbed dose at a specific point in water under reference conditions (distance, field size, temperature and pressure). This will be further discussed in Section 1.3.3. The portion of the linac that is able to rotate is known as the gantry and rotates about a point known as isocentre, as shown in Figure 1.2. This is done so that radiation can be delivered to the patient from multiple angles. Irradiation from multiple angles improves dose coverage of the tumor and healthy tissue sparing. A movable treatment couch where the patient lies during treatment is placed downstream from the radiation beam.

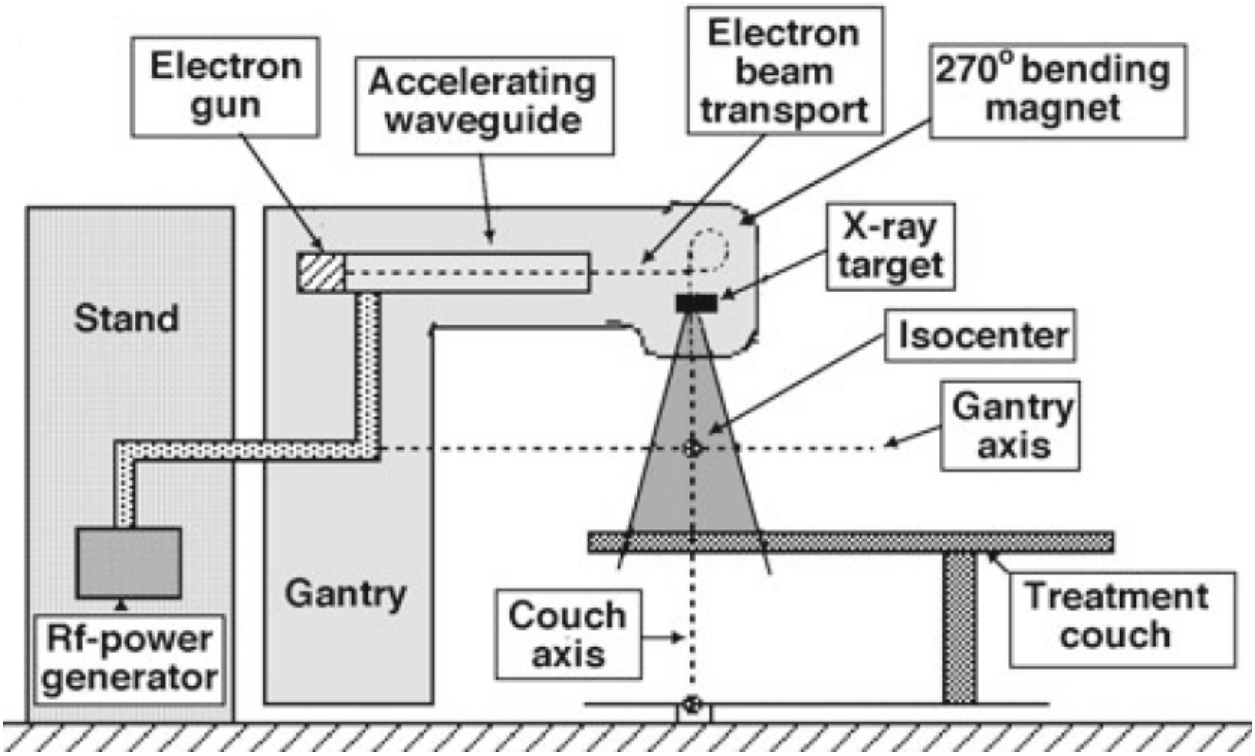


Figure 1.2: A diagram of a medical linear accelerator gantry that rotates about the isocentre. Reproduced with permission from [7].

1.3.1 High energy photon beams

Photon beams are generated by placing an *x-ray target* in the path of the electron beam. Electron interactions with the target atoms result in bremsstrahlung (see Section 2.2). This x-ray target is usually made of Tungsten (W) due to its high Z value, increasing the bremsstrahlung yield, and high melting point. The resulting photon beam then passes through a *flattening filter* which flattens the beam profile in the lateral direction (perpendicular to the beam axis). Since the angular distribution of megavoltage bremsstrahlung emission is forward directed, there is a larger fraction of photons in the forward direction compared to the angles away from the beam axis. The flattening filter then removes these photons and also removes any low energy photons through the photoelectric effect (see Section 2.1.1). Currently, there are linacs that also allow for flattening filter-free (FFF) operation. The photon beam then passes through two sets of *collimators*. The primary collimator removes any photons that scatter beyond the maximum possible field size (usually 40x40 cm^2). The secondary collimator then shapes the photon beam to a prescribed square field size, XxY cm^2. The secondary collimator is made up of two independently moving jaws, one of which defines the X dimension of the field size and the other the Y dimension. The *multi-leaf collimator* or MLC consists of 120 leaves (60 pairs) that can be individually positioned so as to form irregularly shaped fields. MLCs are crucial for various modern radiation therapy treatment techniques such as intensity-modulated radiation therapy (IMRT) and volumetric arc therapy (VMAT). Photon beams are specified by the maximum electron energy that strikes the x-ray target (divided by the electron charge e) and are called megavoltage (MV) photon beams. For example, for a 6 MV photon beam, the maximum electron energy striking the target is 6 MeV. The actual photon energy is not used because the bremsstrahlung has an energy spectrum. Modern medical linacs have photon energies from 6 MV to 15 MV available.

1.3.2 High energy electron beams

Electron beams are produced by removing the x-ray target and replacing the flattening filter with a scattering foil. This enables a widening of the electron beam through scattering. The scattering foil is made of high-Z material such as Tantalum (Ta) or Gold (Au). Often a secondary scattering foil made of Aluminum (Al) is inserted to remove any bremsstrahlung from the interaction between the electron beam and the primary foil. Field sizes for electron beams are established with the primary and secondary collimators along with a so-called *applicator*. An applicator attaches to the end of the linac head and collimates the electron beam to a specific field size near the patient surface. This is done because without the applicator, the electrons would scatter significantly in air and broaden the beam. Electron beams are specified by the final energy of the electrons, for example 6 MeV. Modern medical linacs have electron energies from 6 MeV to 22 MeV available.

1.3.3 Calibration of linear accelerators

A patient undergoing radiation therapy will receive radiation dose from the photon or electron beam generated by the linac. It is therefore essential to ensure that the linac is delivering the dose that is prescribed to the patient. This is the purpose of *reference dosimetry*. *Radiation dosimetry* is the field that concerns the measurement and quantification of radiation absorbed dose. The definition of absorbed dose was discussed in Section 1.2. Dosimetry is performed with various types of radiation-measuring devices, known as *dosimeters*. Reference dosimetry concerns the quantification of absorbed dose at a reference point in water under reference conditions. Water is chosen because of its similarity to human tissue, which is mainly comprised of water. The reference conditions concern the distance between the beam source and the surface of the water (known as the source-to-surface distance, SSD), the field size at the surface, the depth in water and the ambient pressure and temperature. One monitor unit (1 MU) is then defined to be 1 cGy at the reference point in water under

the same reference conditions. The point in water may not be the same as that used in the reference dosimetry measurement. Reference dosimetry measurements take place in a large water-filled tanks of size 30x30x30 cm^3, called a water phantom.

In the clinic, *ionization chambers* are the dosimeters of choice for reference dosimetry. They contain an air-filled collecting volume and operate by collecting the charges generated in this volume by the radiation beam through the application of a voltage across the two electrodes. A detailed discussion of ionization chambers can be found in Section 3.3. Ionization chambers, and any dosimeter for that matter, can not directly measure absorbed dose. For example, ionization chambers measure charge or current. A conversion factor, known as a *absorbed dose to water calibration coefficient*, is then required to convert from the measured quantity to absorbed dose to water. The calibration coefficient of a chamber is determined at a *primary standards dosimetry laboratory (PSDL)* with an *absorbed dose primary standard*. An absorbed dose primary standard is a measurement device that is able to determine absorbed dose with low uncertainty and is independent of any other measurements. Water and graphite calorimeters are common absorbed dose primary standards [8]. Through this framework of ionization chamber calibrations, traceability to PSDLs is continually maintained. Ionization chambers should be calibrated every two years [9].

1.4 Ionization chamber measurements for reference dosimetry

Codes of practice (CoP) for MV photon, MeV electron and proton reference dosimetry are offered by various international and national agencies for standardized reporting of results across various medical institutions and uniformity in practice. Such CoPs are the TG-51 and TG-51 addendum from the American Association of Physicists in Medicine (AAPM) [9, 10] and the TRS-398 and its update from the International Atomic Energy Agency (IAEA) [11,

12]. Both of these CoPs are based on in-clinic ionization chamber measurements of collected charge and absorbed dose to water primary standards. In general, the absorbed dose to water of a beam of quality Q (see Section 3.1.5 for further detail) under reference conditions at the reference point, $D_{\mathrm{w,Q}}$, is obtained by multiplying the corrected chamber reading M_{Q} at the reference point and the absorbed dose to water calibration coefficient $N_{\mathrm{D,w,Q}}$, that is,

$$D_{\mathrm{w,Q}} = M_{\mathrm{Q}} N_{\mathrm{D,w,Q}}. \tag{1.2}$$

The ionization chamber is generally used in charge collection mode, so the units of M_{Q} is Coulombs (C). The absorbed dose to water calibration coefficient then has units of Gy/C. The corrected chamber reading is used because, in practice, the raw chamber measurement is not done at the reference pressure and temperature and polarity effects and ion recombination are present. All of these are corrected for through correction factors. Each of these effects will be discussed in Section 3.3.4.

1.4.1 Ion recombination

When a raw ionization chamber reading is taken, the charge value is less than the initial value produced by the radiation beam. The radiation interacts with the air molecules of the chamber's sensitive volume and produces free electrons and positive ions. Due to the electronegativity of oxygen molecules, the electrons attach to the molecules and form negative ions. The chamber collects positive and negative ions and free electrons that do not attach. Some positive and negative ions are able to recombine and form neutral particles that are not collected. The ionization chamber then has a charge collection efficiency f less than 100%. The collection efficiency is defined as the ratio of the charge collected to the charge initially created by the radiation beam, prior to any recombination, $\frac{Q}{Q_0}$. Were recombination ignored, the absorbed dose to water would be underestimated (through Eq. 1.2) and the linac would not be calibrated properly (the patient will be over-dosed because more MUs will be deliv-

ered than necessary to reach a desired dose). Ion recombination is taken into consideration by multiplying the raw ionization chamber reading M_{raw} by an *ion recombination correction factor* P_{ion} or k_{s}, which is equivalent to $1/f$.

Ion recombination is classified into *initial* and *general* recombination. Initial recombination occurs when a positive and a negative ion formed in the same radiation track recombine. Initial recombination is more significant for densely-ionizing radiation, such as protons and light ions. Since initial recombination takes place in a single beam track, it is dose-rate independent. Initial recombination depends on the chamber geometry, the applied voltage and the radiation type and energy. General (or volume) recombination concerns the recombination of positive and negative ions from separate tracks. The electric field in the chamber accelerates the ions away from their parent track. Since the track density depends on dose rate, general recombination is dose-rate dependent. In conventional MV photon and MeV electron beams, general recombination is the dominant recombination mechanism. General recombination is also dependent on the chamber geometry, applied voltage and the radiation energy and type. General recombination also depends on the time structure of the radiation beam, whether it is pulsed or continuous. An in-depth discussion of both initial and general recombination will be given in Chapter 4.

While the free electron collection is negligible at the dose-per-pulse (DPP) levels found in conventional MV photon and MeV electron beams ($<$ 1 mGy/pulse), it can not be ignored in the intermediate (10-100 mGy/pulse), high (100-1000 mGy/pulse) and ultra-high ($>$ 1 Gy/pulse) DPP regimes. Additionally, a large free electron fraction results in a large positive space charge that distorts the electric field within the ionization chamber. This leads to the breakdown of almost all previous theoretical analyses that do not consider the free electron fraction and space charge. The usual methods to determine the ion recombination correction factor in the clinic, such as the two-voltage method (TVM) and Jaffé plot extrapolation, also

breakdown because they are based on theoretical studies. This will be discussed in greater detail in Section 4.5.

1.5 Proton therapy

Proton therapy is the use of protons for the purposes of radiation therapy. The potential benefits of proton therapy were first reported in 1946 by Wilson [13]. The rationale behind using protons for radiation therapy is their unique percent depth dose distribution (PDD). In brief, a PDD presents the distribution of dose from a radiation beam as a function of depth in water on the central beam axis for a given field size and SSD. It is normalized to the maximum dose. A full discussion of PDDs can be found in Section 3.1.4. The PDD of a 6 MV photon beam and a 190 MeV proton beam can be seen in Figure 1.3. The PDD of a MV photon beam is characterized by a low surface dose, a rise to a maximum and then a slow, exponential decrease. While a tumour volume can receive a large dose, the shallow depth of the maximum dose and slow decrease results in a high dose to adjacent healthy tissues and organs. A monoenergetic proton beam is characterized by a low surface dose as well as a low dose plateau for a considerable depth within water. At some deeper depth, the dose rapidly rises to a maximum and then falls off rapidly to almost zero. This peak is known as the *Bragg peak*. Given a properly positioned beam, the Bragg peak can be entirely localized within a tumour volume. The tumour volume receives a high dose while the healthy tissue is spared to a greater extent than with a conventional MV photon beam. Several monoenergetic proton beams can be delivered during a single treatment, resulting in a *spread-out Bragg peak (SOBP)*, shown in Figure 1.4. Due to all of the Bragg peaks being in close vicinity to one another, their sum results in a high, uniform dose across a certain region and a low dose everywhere else. With a SOBP, an entire tumour volume can be completely covered. There are several cases where proton therapy is thought to be superior to other radiation types. Pediatric cancer cases are thought to greatly benefit from proton therapy.

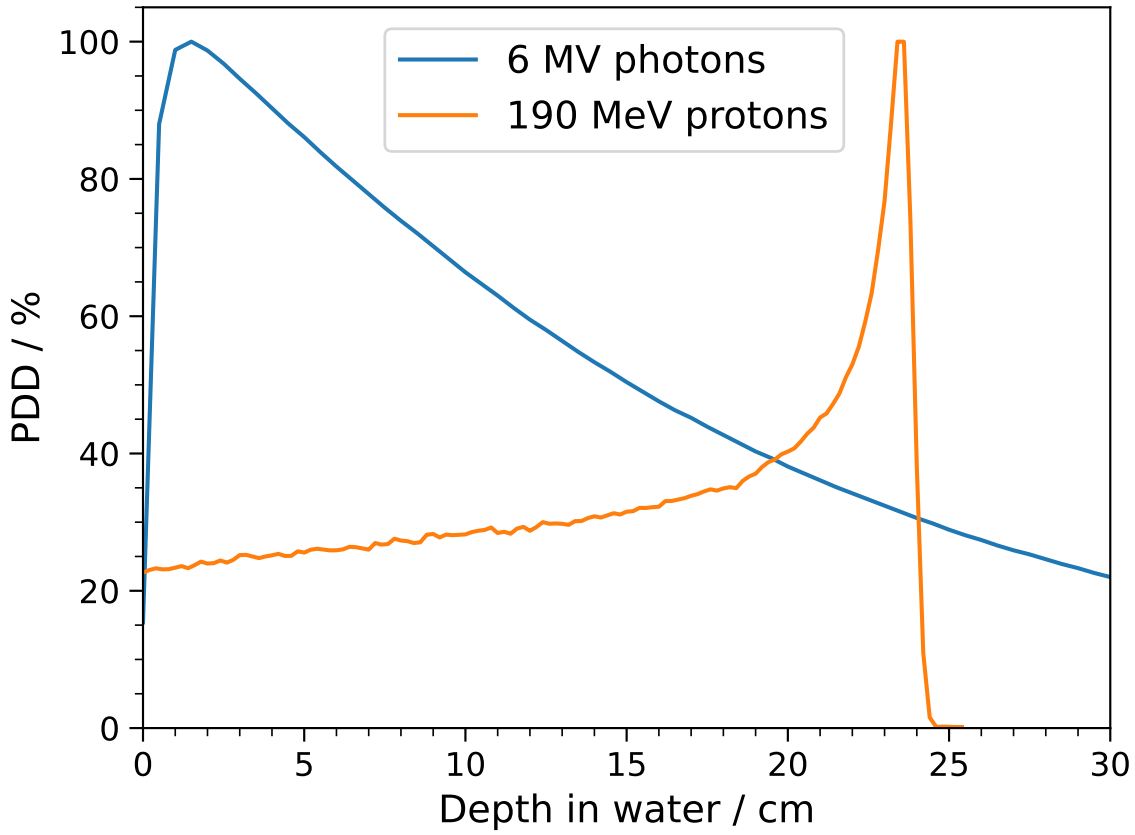


Figure 1.3: Percent depth dose distributions (PDDs) of a 6 MV photon beam and a 190 MeV proton beam. The proton PDD was calculated with the GEANT4 Monte Carlo toolkit, version 11.2. Courtesy of Alaina Giang Bui.

This is due to the greater healthy tissue sparing of proton beams than MV photon beams. This fact can aid in reducing adverse late term effects such as secondary tumours. Cancers of the eye and at the skull base are also thought to benefit from proton therapy, due to the close proximity of many organs at risk (OARs). Due to the steep dose gradients of the Bragg peak, the dose can be delivered to the tumour without a large dose being delivered to the OARs [14].

Medical proton beams are currently generated by cyclotrons or synchrotrons [15]. These two acceleration methods will be discussed further in Chapter 5. These accelerators can

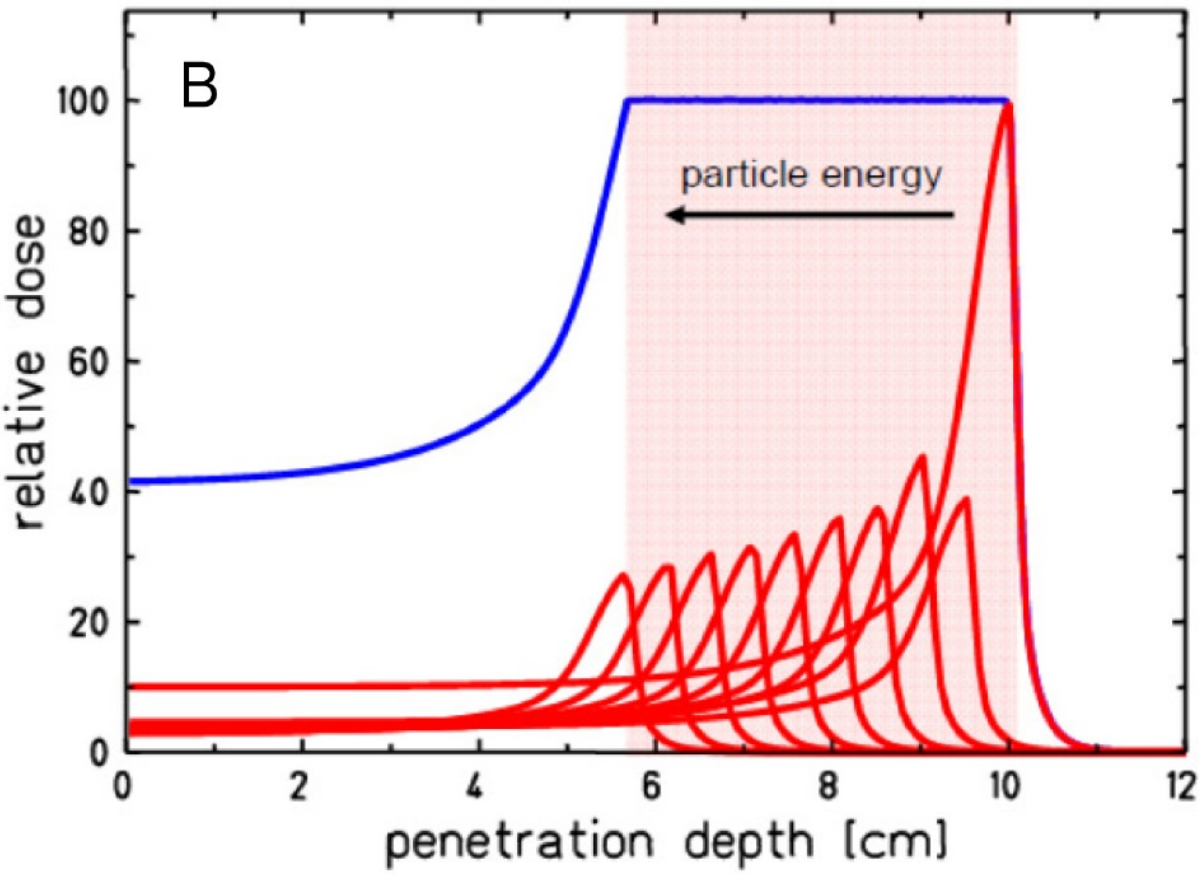


Figure 1.4: A spread-out Bragg Peak (SOBP) consists of the sum of many monoenergetic proton beams. Reproduced with permission from [14].

accelerate protons up to energies of 250 MeV and higher. In either case, the beam can be delivered to multiple treatment rooms. In each room is a treatment head, known as a *nozzle*. The treatment head is mounted on a gantry that can rotate around the patient. Two methods are used to deliver the proton beam, shown in Figure 1.5. *Passive scattering* uses a rotating range modulation wheel to modulate the proton energies and form the SOBP. The wheel is made of ridges of various thickness. Two scatterers of high Z material scatter the beam laterally and flatten the dose profile, much like a flattening filter in conventional linacs. A collimator limits the lateral spread of the beam and a *range compensater* is used to further conform the beam to the tumour shape. The other method, *active scanning*, uses magnets to steer the beam towards a spot or sets of spots to cover the tumour. A raster in the x-y plane of the tumour is delivered, known as scanning, and the depth varied by

changing the beam energy or adding additional absorbers in the beam path.

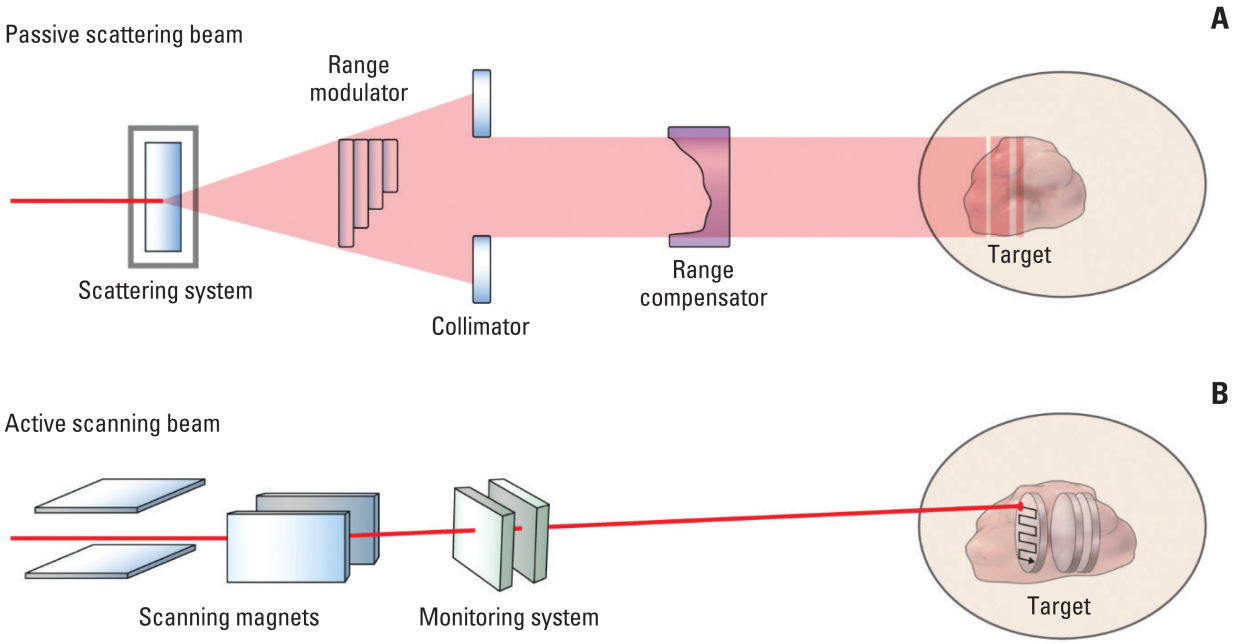


Figure 1.5: A: A schematic of a passively-scattered proton beam. B: A schematic of a magnetically actively-scanned proton beam. Reproduced under the terms of the Creative Commons Attribution Non-Commercial License from [16].

1.6 Ultra-high dose rate radiation therapy

Conventional MV photon and MeV electron radiation therapy take place within the mean dose rate range of 30-100 mGy/s [17]. In 2014, Favaudon et. al. [18] demonstrated that an ultra-high dose rate (UHDR) electron beam ($\geq$ 40 Gy/s) resulted in the repression of lung cancer growth in mice, reduced late term effects and less damage to healthy tissue, in comparison to a conventional dose rate beam of 30 mGy/s. The use of UHDRs in radiation therapy has come to be known as *FLASH*. With FLASH-capable beams, dose rates up to 1000 Gy/s can be achieved. Such high dose rates and very short pulse times on the order of 2 μs enable ultra-high dose per pulses, almost 10 Gy/pulse. The FLASH effect is the reduction in healthy tissue damage under UHDR radiation. The mechanism or mechanisms behind the FLASH effect is still not entirely settled. In general, oxygen consumption is considered to be a crucial mechanism [17]. This can be attributed to the rapid deposition of dose and

the inability for oxygen to be resupplied to the cells in FLASH beams, in comparison to conventional beams. Currently, most FLASH beams are electron beams produced by dedicated linear accelerators. Of course, combining the benefits of proton beams and FLASH is of major interest [19].

1.7 Thesis motivation and objectives

The potential benefits of proton therapy makes it a desirable treatment modality for many cancer patients. However, less than 2% of patients undergoing radiation therapy receive proton therapy [20]. This is in stark contrast to the estimate that 15% or more of patients undergoing radiation therapy may benefit from proton therapy. At the current rate of new proton therapy facilities in operation per year vs. patient load per year, this gap will only be closed in about 40 years. There are two main reasons for this discrepancy. The first concerns the high costs of the main proton accelerator, the cyclotron or the synchrotron, and the gantry. For a single-room proton therapy facility, the cost is between US $30 to $50 million. The second reason is the size of the equipment involved in proton radiotherapy. The gantry itself cannot fit inside a conventional bunker made for photon radiation therapy. The proton accelerator and beam line which transports the beam from the accelerator to the gantry must also be located somewhere outside the treatment room. The use of superconducting magnets with the accelerator have aided in decreasing their size but also significantly add to the cost. Efforts have been made to use alternative acceleration methods, such as synchrocyclotrons (used in the Mevion S250 and Proteus ONE) and laser-driven acceleration, to decrease the size and cost of the accelerator [20]. In terms of reducing the size of the gantry, one promising development is to have it fixed at one angle and treat the patient in an upright position (such as sitting) who rotates around the beam [21].

Proton therapy is also the radiation therapy modality that is currently most suitable for

UHDR radiation therapy. Concerning electron beams at conventional energies (< 16 MeV), UHDR-capable linacs do exist [22] but the small range of electrons in water (less than 10 cm, see section 3.1.4) limits their applicability. Very high energy electron (VHEE) beams between the range of 50-250 MeV are able to overcome these range issues but are not clinically available [23]. X-ray tubes are able to achieve UHDRs (< 140 Gy/s) but only near the tube exit window and the depth is limited to 2 mm [24]. Sampayan et. al. [25] argued that high energy photon beams at UHDRs are able to be produced with linear induction accelerators (see section 5.4). Again, LIAs are not available for clinical use.

The cost and size concerns, along with the recent interest in UHDR radiation therapy, underpin the recent global drive to develop novel proton acceleration methods. One potential solution is the Dielectric Wall Accelerator (DWA). The DWA is a structure of stacked modules in the direction of the beam axis [26]. At the beam end of each module, the accelerating gradient is formed which accelerates the passing proton bunches. Each module contains a pulser with a fast, high voltage (HV) switch and a radial waveguide. The pulsers generate HV pulses through the operation of the switch. The HV pulses are then injected into the radial waveguide. The injection method allows for time-varying pulse profiles which aid in longitudinal beam stability [27]. The radial waveguides amplify the pulses and transport them towards the beam pipe. The beam pipe is a high gradient insulator (HGI) structure whose design prevents surface flashover dielectric breakdown at high field strengths. Due to the large accelerating gradient, the DWA is able to accelerate protons to clinically relevant energies in a few meters. The materials involved are also relatively low cost. Such a compact accelerator at a lower cost is able to directly address the cost and size concerns of common proton therapy units and improve proton therapy accessibility.

The pulsers of the DWA must be able to generate HV pulses in order to achieve large accelerating gradients. Such HV pulsers generally contain an inductive element that can

store energy from a HV source. This energy is then commuted to a load (in the case of the DWA, the waveguide) through the operation of a HV switch. This is when the HV pulse is formed. The pulse width should be on the order of 1 ns. This is due to the breakdown threshold of the HGI being inversely proportional to the pulse width. At a pulse width on the order of 1 ns, an HGI is able to sustain an accelerating gradient up to 100 MV/m [26]. The pulse should also have a fast rise time, minimizing the pulse distortion as it travels down the waveguide. Finally, the pulser and switch should be able to operate at sufficiently high frequencies, above 1 kHz, so that the DWA can achieve UHDRs. To see why this should be, consider the time-averaged dose rate $\dot{D}$ of a pulsed monoenergetic proton beam, given by (see section 3.2)

$$\dot{D} = \left(\frac{S_{\text{el}}}{\rho}\right)\frac{\dot{N}}{A} \tag{1.3}$$

$$= \left(\frac{S_{\text{el}}}{\rho}\right)\frac{I \cdot \text{PW} \cdot f}{e}\frac{1}{A}, \tag{1.4}$$

where $\frac{S_{\text{el}}}{\rho}$ is the proton stopping power (defined in section 2.2), I is the current in a single pulse, PW is the pulse width, f is the operating frequency of the DWA, e is the electronic charge and A is the beam spot area. $\dot{N}$ is the time averaged rate of protons in the beam. For a frequency of 1 kHz, a beam spot diameter of 8 mm (that of 200 MeV proton beam in water from a Mevion S250 [28]), a current of 0.1 mA (D-Pace ISV.F-40) and a pulse width of 1 ns, the time averaged dose rate at the Bragg peak in liquid water is 164 Gy/s and 164 mGy/pulse. The dose rate and DPP can be increased by decreasing the beam spot size or increasing the current. The dose rate can also be increased by increasing the frequency.

Common semiconductor switches such as MOSFETs and diodes are not able to meet all of these requirements. Additionally, these are all *closing* switches ('on' when closed) which are less efficient than opening switches ('on' when open). Drift step recovery diodes (DSRDs) are Silicon (Si)-based opening switches that may be able to meet these requirements [29].

The operation of the DSRD is discussed in detail in section 7.4.1.

The arrival of UHDR radiation therapy in the clinic also necessitates the review of common dosimetric concepts that are affected by the high dose-per-pulse or dose rate. This is the case with ion recombination. Under UHDR, the results of previous theoretical studies do not hold and the methods used in reference dosimetry to determine the ion recombination correction factor (TVM and and Jaffé plot extrapolation) which are based on these studies also do not work. Metrological-based solutions have yet to enter the clinic and numerical solutions are not practical in a clinical setting. Analytical solutions based on theoretical studies, on the other hand, have the benefit of being quick to evaluate and all parameters involved retain a physical meaning. Quick evaluation methods such as the TVM can also be developed from analytical solutions. So far, there has been no investigation into this area.

The main goal of this thesis is to address two needs that arise from the introduction of UHDR radiation therapy, including proton therapy. The development of the DWA and more specifically its pulsers aims to address the first need; the need for less costly and compact proton acceleration methods. Theoretical investigations into ion recombination in the UHDR regime aims to address the second need; the need for analytical expressions of the ion recombination correction factor. In the context of these two needs, the objectives set out for this thesis are as follows;

1. **Establish the suitability of a DSRD-based pulser within the context of the DWA and the pulse requirements.**

   Any DSRD-based pulser must meet the above mentioned pulse requirements. The extent of the capabilities of a pulser to meet these requirements must therefore be systematically studied. Additionally, there has been no direct comparison between pulser designs while keeping the DSRD parameters consistent. This is necessary if one is to

be considered for the DWA and the optimal driving conditions determined to generate suitable pulses. For the purposes of this thesis, two pulser designs are investigated; a magnetic saturation transformer (MST)-DSRD-based pulser and a MOSFET-DSRD-based multi-module pulser.

2. **Study the influences of the DSRD parameters on the output pulse.**

   The DSRD doping profile and its material are the main parameters that determine the DSRD operation. Establishing the DRSD profile and material which yield the optimal pulse in terms of the pulse requirements, given a constant pulser, is crucial to informing future DSRD design for the DWA. The increasing prominence and affordability of silicon carbide (SiC) also motivates this thesis objective.

3. **Obtain an analytical expression for the ion recombination correction factor in UHDR beams through a mathematical procedure.**

   So far, there have been no theoretical investigations into ion recombination in UHDR beams. This has also resulted in a lack of analytical expressions for the ion recombination correction factor. Analytical expressions have an advantage over other solution methods as they are quick to evaluate, which is advantageous in the clinical setting, and all the parameters retain a physical meaning.

4. **Develop an experimental procedure to determine the ion correction factor in a clinical setting.**

   This would enable the evaluation of the ion recombination correction factor in the clinic, simplifying the reference dosimetry workflow.

1.8 Thesis outline

Chapter 2 is a review of the fundamental interactions and quantities concerning radiation physics. These are crucial in understanding how radiation dosimetric quantities such as absorbed dose arise. Chapter 3 is a review of radiation dosimetry and how absorbed dose can be

measured. Ionization chambers, which are the dosimeters most commonly used in the clinic, are also reviewed in this chapter. Chapter 4 is a review of the theory of ion recombination and the methods used to measure it in the clinic for reference dosimetry. The issues in UHDR beams are also discussed. Chapter 5 is a review of the particle acceleration methods used in clinical accelerators. The DWA and DSRDs are also discussed int his chapter. Chapter 6 is a quick review of semiconductor physics and the models used in semiconductor simulations. The semiconductor simulation software, Sentaurus TCAD by Synopsys®, is also introduced, as it is used in Chapter 7. Chapter 7 is an article submitted to *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* investigating the MST and MOSFET-DSRD-based pulsers for the DWA. Simulations of the DSRD with different materials and doping profiles are presented. Chapter 8 is an article published in the journal *Medical Physics* describing a semi-analytical solution method to solve the PDEs describing ion recombination in UHDR beams. A method to determine the ion recombination correction factor with ionization chamber experiments is also provided. The conclusion reviews the original research carried out for this thesis and discusses future research avenues.

References

[1] The Global Cancer Observatory. "World fact sheet." (2024), [Online]. Available: `https://gco.iarc.fr/today/en` (visited on 2024).

[2] Canadian Cancer Statistics Advisory Committee. "Canadian cancer statistics." (2024), [Online]. Available: `https://cancer.ca/en/research/cancer-statistics/canadian-cancer-statistics` (visited on 2024).

[3] The Global Cancer Observatory. "Cancer over time." (2024), [Online]. Available: `https://gco.iarc.who.int/overtime/en` (visited on 2024).

[4] S. K. Das *et al.*, "Gene therapies for cancer: Strategies, challenges and successes," *Journal of Cellular Physiology*, vol. 230, no. 2, pp. 259–271, 2015.

[5] International Atomic Energy Agency, *Radiation Biology: A Handbook for Teachers and Students*. Vienna, Austria: IAEA, 2010.

[6] N. Suntharalingam, E. B. Podgorsak, and J. H. Hendry, "Basic radiobiology," in *Review of Radiation Oncology Physics: A Handbook for Teachers and Students*, E. B. Podgorsak, Ed., Vienna: International Atmoic Energy Agency, 2005, ch. 14, pp. 397–412.

[7] E. B. Podgorsak, *Radiation Physics for Medical Physicists*. Springer International Publishing, 2016.

[8] J. Renaud, H. Palmans, A. Sarfehnia, and J. Seuntjens, "Absorbed dose calorimetry," *Physics in Medicine & Biology*, vol. 65, no. 5, 05TR02, Mar. 2020.

[9] P. R. Almond *et al.*, "AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon and electron beams," *Medical Physics*, vol. 26, no. 9, pp. 1847–1870, 1999.

[10] M. McEwen *et al.*, "Addendum to the AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon beams," *Medical Physics*, vol. 41, no. 4, p. 041 501, 2014.

[11] *Absorbed Dose Determination in External Beam Radiotherapy* (Technical Reports Series 398). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2001.

[12] *Absorbed Dose Determination in External Beam Radiotherapy* (Technical Reports Series 398 (Rev. 1)). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2024.

[13] R. R. Wilson, "Radiological use of fast protons," *Radiology*, vol. 47, no. 5, pp. 487–491, 1946, PMID: 20274616.

[14] M. Durante and H. Paganetti, "Nuclear physics in particle therapy: A review," *Reports on Progress in Physics*, vol. 79, no. 9, p. 096 702, Aug. 2016.

[15] R. Mohan, "A review of proton therapy – current status and future directions," *Precision Radiation Oncology*, vol. 6, no. 2, pp. 164–176, Apr. 2022.

[16] H. K. Byun *et al.*, "Physical and biological characteristics of particle therapy for oncologists," *Cancer Res Treat*, vol. 53, no. 3, pp. 611–620, 2021.

[17] B. Lin *et al.*, "Flash radiotherapy: History and future," *Frontiers in Oncology*, vol. 11, May 2021.

[18] V. Favaudon *et al.*, "Ultrahigh dose-rate flash irradiation increases the differential response between normal and tumor tissue in mice," *Science Translational Medicine*, vol. 6, no. 245, 245ra93–245ra93, 2014.

[19] A. E. Mascia *et al.*, "Proton FLASH Radiotherapy for the Treatment of Symptomatic Bone Metastases: The FAST-01 Nonrandomized Trial," *JAMA Oncology*, vol. 9, no. 1, pp. 62–69, Jan. 2023.

[20] S. Yan, T. A. Ngoma, W. Ngwa, and T. R. Bortfeld, "Global democratisation of proton radiotherapy," en, *The Lancet Oncology*, vol. 24, no. 6, e245–e254, Jun. 2023.

[21] S. Yan, H.-M. Lu, J. Flanz, J. Adams, A. Trofimov, and T. Bortfeld, "Reassessment of the necessity of the proton gantry: Analysis of beam orientations from 4332 treatments at the massachusetts general hospital proton center over the past 10 years," *International Journal of Radiation Oncology*Biology*Physics*, vol. 95, no. 1, pp. 224–233, 2016, Particle Therapy Special Edition.

[22] M. Rahman *et al.*, "Electron flash delivery at treatment room isocenter for efficient reversible conversion of a clinical linac," *International Journal of Radiation Oncology*Biology*Physics*, vol. 110, no. 3, pp. 872–882, Jul. 2021.

[23] M. G. Ronga *et al.*, "Back to the future: Very high-energy electrons (vhees) and their potential application in radiation therapy," *Cancers*, vol. 13, no. 19, p. 4942, Sep. 2021.

[24] M. Bazalova-Carter and N. Esplen, "On the capabilities of conventional x-ray tubes to deliver ultra-high (FLASH) dose rates," *Medical Physics*, vol. 46, no. 12, pp. 5690–5695, 2019.

[25] S. E. Sampayan *et al.*, "Megavolt bremsstrahlung measurements from linear induction accelerators demonstrate possible use as a FLASH radiotherapy source to reduce acute toxicity," *Scientific Reports*, vol. 11, no. 1, Aug. 2021.

[26] G. J. Caporaso, Y.-J. Chen, and S. E. Sampayan, "The Dielectric Wall Accelerator," *Reviews of Accelerator Science and Technology*, vol. 2, no. 1, 2009.

[27] C. Lund *et al.*, "A novel approach to longitudinal beam stability in a compact dielectric wall accelerator for proton therapy," *Radiotherapy and Oncology*, vol. 186, S15–S16, Sep. 2023.

[28] G. Vilches-Freixas *et al.*, "Beam commissioning of the first compact proton therapy system with spot scanning and dynamic field collimation," *The British Journal of Radiology*, vol. 93, no. 1107, Dec. 2019.

[29] I. Grekhov, V. Efanov, A. Kardo-Sysoev, and S. Shenderey, "Power drift step recovery diodes (DSRD)," *Solid-State Electronics*, vol. 28, no. 6, pp. 597–599, 1985.

2

Radiation Physics

2.1 Photon interactions

Photons undergo a variety of interactions as they travel through a medium and interact with atoms. The probability of a certain interaction occurring is given by the interaction cross section σ. A cross section has units of m^2. The definition of a cross section is [1]

$$\sigma = \frac{\text{\# of scattered particle/unit time}}{\text{\# of incident particle/unit time/unit area}}. \tag{2.1}$$

The angular distribution of scattered particles is given by the differential cross section (DCS), $\frac{d\sigma}{d\Omega}$, where Ω is solid angle. The (total) cross section can be obtained through integration of the differential cross section, that is,

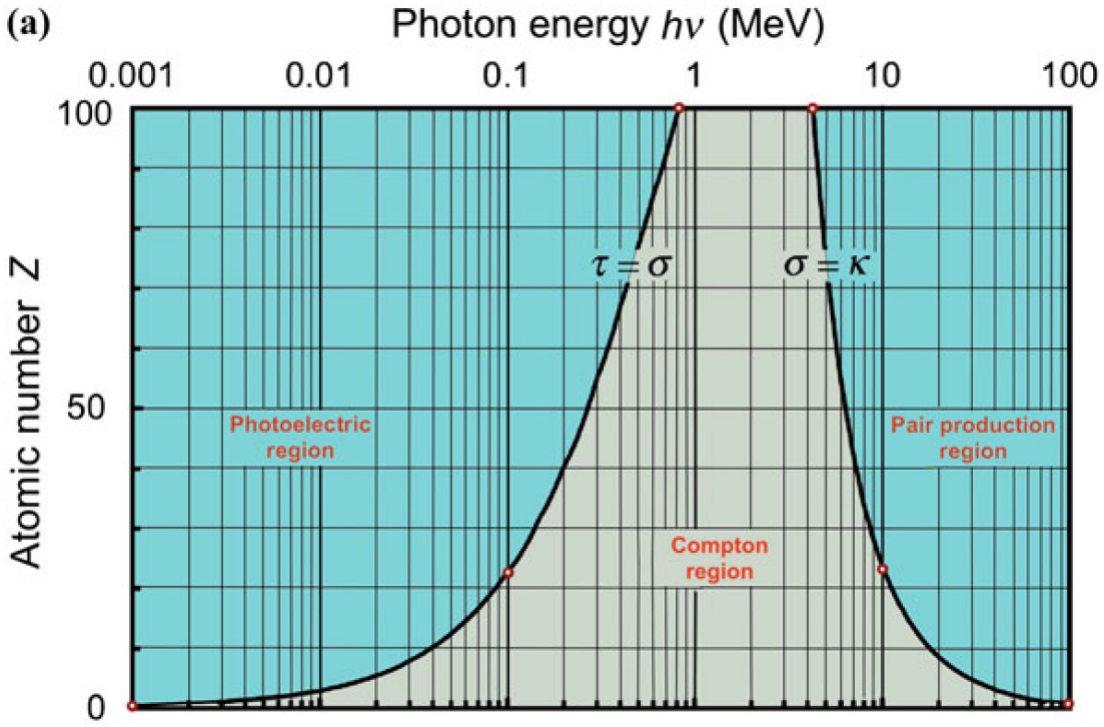


Figure 2.1: The photon energy ranges in which a certain photon interaction is most dominant. The black lines indicate where two adjacent photon interaction cross section are equal. As the atomic number increases, the photoelectric effect increasingly covers higher energies and pair production increasingly covers lower energies. Reproduced with permission from [2].

$$\sigma = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\, d\Omega = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\, \sin\theta d\theta d\phi. \tag{2.2}$$

The individual interaction cross sections depend on the incident photon energy and the atomic number Z of the atom. Therefore, the photon energy range under which a certain interaction is more dominant (larger cross section) than others depends on the material, as shown in Figure 2.1.

2.1.1 Photoelectric effect

The photoelectric effect is the most common type of photon interaction at low energies, below 100 keV for low atomic numbers [3]. This includes water, graphite and air, materials commonly found in medical physics. The photoelectric effect was first mathematically described by Albert Einstein in 1905 [4]. A diagram of the photoelectric effect is shown in

Figure 2.2. In the photoelectric effect, the incident photon interacts with the electromagnetic field of the atom and is absorbed. An orbital electron is then ejected. The kinetic energy of this electron is the photon energy minus the binding energy of the shell and the kinetic energy of the nuclear recoil, though this is small due to the large nuclear mass. The ejected electron is generally ejected close to the forward direction, becoming less forward as the photon energy decreases. Following the ejection of the electron, relaxation processes occur and characteristic x-rays or Auger electrons are emitted. It has been found that the photoelectric cross section is proportional to Z^n where n varies from 4-5 and $k^{-3.5}$ where k is the incident photon energy [3]. Near the absorption edges, these relations do not hold. Due to the shell structure of the atom, each shell has its own photoelectric cross section and the K-shell's is the largest. The total photoelectric cross section of an atom is then the sum of the individual shell photoelectric cross sections.

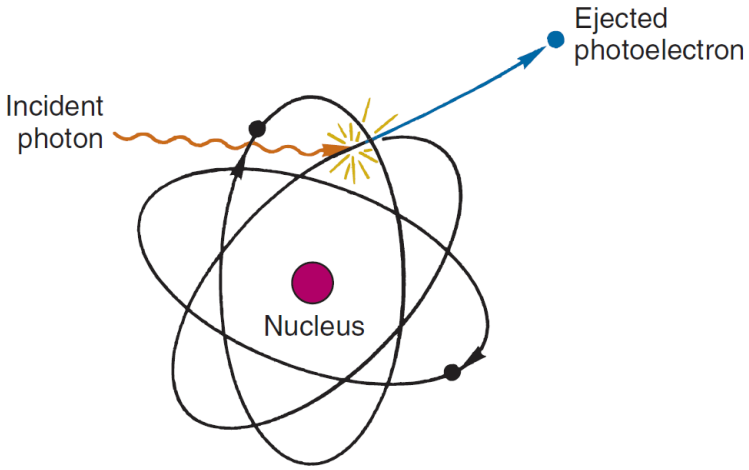


Figure 2.2: A diagram of the photoelectric effect. Reproduced with permission from [5].

2.1.2 Compton scattering

Due the photon energies found in conventional MV photon beams, around 1 MeV, and the materials encountered (water, air and graphite), Compton scattering is by far the most common photon interaction. Compton scattering was first observed by Arthur H. Compton in 1923 [6]. A diagram of the Compton scattering interaction is shown in Figure 2.3. An

incident photon interacts with an outer shell orbital electron, transfers some of its energy to the electron and is then scattered to a different angle with less energy. The electron is also ejected with a kinetic energy equal to the energy transferred from the photon minus the binding energy. The scattered photon and the ejected electron are generally scattered near the forward direction, becoming more forward as the incident photon energy increases. Relaxation processes also occur after the ionization event. Klein and Nishina [7] derived the

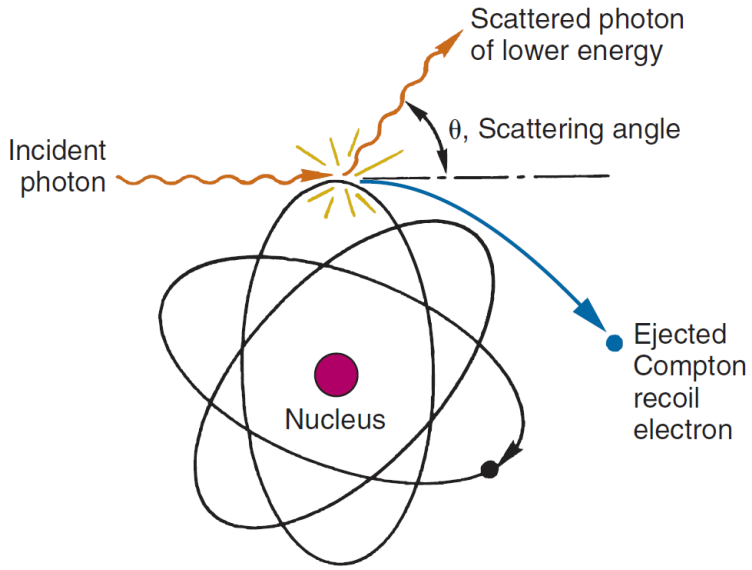


Figure 2.3: A diagram of Compton scattering. Reproduced with permission from [5].

now famous DCS for Compton scattering. In their derivation, the orbital electron is considered to be at rest and unbound. A more complete theory describing Compton scattering is the Relativistic Impulse Approximation (RIA) [8]. The RIA incorporates the bound nature of the electrons and that they are not at rest. The electrons of a given shell are assumed to move within a momentum distribution. The double differential cross section (DDCS) in the RIA formalism is then given by

$$\frac{\partial^2\sigma_{\mathrm{C}}}{\partial k'\partial\Omega} = \frac{r_{\mathrm{e}}^2}{2}\frac{k'}{k}\frac{m_{\mathrm{e}}}{q}\left[1+\left(\frac{p_{\mathrm{z}}}{m_{\mathrm{e}}c}\right)^2\right]^{-1/2} XJ(p_{\mathrm{z}}), \tag{2.3}$$

where

$$J(p_z) = \iint \rho(\mathbf{p}) dp_{\mathrm{x}} dp_{\mathrm{y}} \tag{2.4}$$

is the Compton profile and is the integration of the momentum distribution $\rho(\mathbf{p})$ over the x and y momentum components. In Eq. 2.3, k is the incident photon energy, k' is the scattered photon energy, r_e is the classical electron radius, q is momentum transferred to the electron by the incident photon, p_z is the projection of q on the initial electron momentum vector, X is a function of all of these variables and scattering angle. Each shell of the atom has a Compton profile and they must be summed over to obtain the total Compton cross section. The Compton cross section is proportional to k^{-1} above 100 keV and increases with photon energy at lower energies.

2.1.3 Rayleigh scattering

Rayleigh scattering concerns the elastic scattering of a photon by an atom. It was first described by Lord Rayleigh (John William Strutt) in 1871 [9]. Therefore, the emitted photon has the same energy as the incident photon but is scattered in another direction. Since all of the orbital electrons are involved in the absorption and emission of the photon, the constructive and destructive interference of all outgoing waves must be considered. This is done through the atomic form factor $F(\mathbf{q}, Z)$, where $\mathbf{q}$ is the difference between the incident and the scattered photon momentum vectors. As the photon energy increases, the scattering angle is more forward directed. The Rayleigh DCS is given by

$$\frac{\mathrm{d}\sigma_{\mathrm{R}}}{\mathrm{d}\Omega} = r_{\mathrm{e}}^2 \frac{1 + \cos^2\theta}{2} \left| F(\mathbf{q}, Z) + f' + if'' \right|^2 , \tag{2.5}$$

where θ is the photon scattering angle and $F(\mathbf{q}, Z)$ is the atomic form factor with $\mathbf{q}$ being the photon momentum transfer vector. The quantity $f' + if''$ is known as the anomalous scattering factor and accounts for scattering near the absorption edges. The Rayleigh scattering cross section is larger than the Compton cross section for photon energies below 10

keV, for low atomic numbers [3].

2.1.4 Pair and triplet production

Pair production occurs when a photon interacts with the electromagnetic field of the atomic nucleus. The photon can decay into a electron and positron pair. In order to conserve energy and momentum, there must also be a non-zero kinetic energy from the nuclear recoil. Pair production was first observed by Blackett and Occhialini in 1933 [10]. Triplet production occurs when the photon interacts with the electromagnetic field of an orbital electron instead. Along with the electron and positron pair being produced, the orbital electron is also ejected with a non-zero kinetic energy. For low Z materials, triplet production is very small in comparison to pair production.

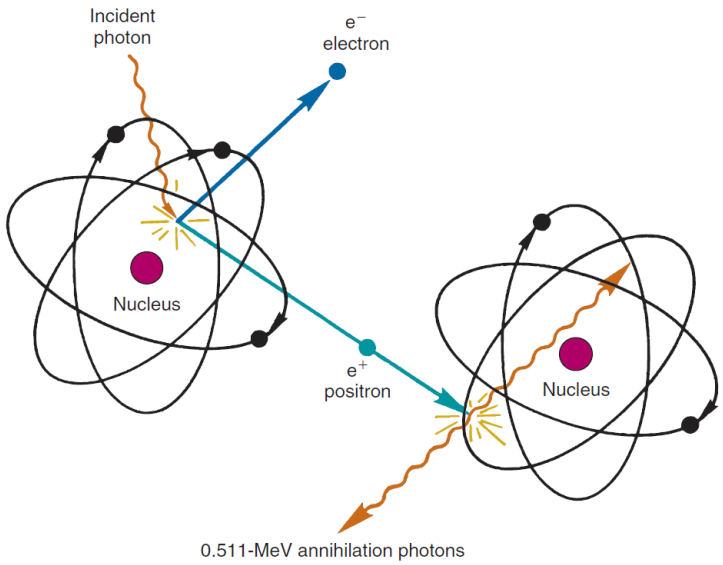


Figure 2.4: A diagram of pair production. The positron then annihilates with an orbital electron of a different atom and produces two annihilation photons. Reproduced from with permission [5].

2.1.5 Photon interaction coefficients

The total cross section σ of a medium is given by the sum of the individual interaction cross sections

$$\sigma = \sigma_{\mathrm{R}} + \sigma_{\mathrm{PE}} + \sigma_{\mathrm{C}} + \sigma_{\mathrm{PP}} + \sigma_{\mathrm{TP}}, \tag{2.6}$$

where R stands for Rayleigh, PE for photoelectric, C for Compton, PP for pair production and TP for triplet production. The cross section only concerns a single atom of a medium. It is a microscopic quantity. The macroscopic equivalent is the photon mass attenuation coefficient, $\frac{\mu}{\rho}$, given by

$$\frac{\mu}{\rho} = \frac{N_{\mathrm{A}}}{A}\left[\sigma_{\mathrm{PE}} + \sigma_{\mathrm{R}} + \sigma_{\mathrm{C}} + \sigma_{\mathrm{PP}} + \sigma_{\mathrm{TP}}\right] = \frac{\mu_{\mathrm{PE}}}{\rho} + \frac{\mu_{\mathrm{C}}}{\rho} + \frac{\mu_{\mathrm{R}}}{\rho} + \frac{\mu_{\mathrm{PP}}}{\rho} + \frac{\mu_{\mathrm{TP}}}{\rho}, \tag{2.7}$$

where ρ is the element's physical density, N_{A} is Avogradro's number and A is the molar mass of the element. The mass attenuation coefficient has units of m^2/kg but is often given in cm^2/g.

Photons transfer some or all of their initial energy to the kinetic energy of the electrons (and positrons in the case of pair and triplet production) upon interaction. This energy transfer occurs for all of the photon interactions except Rayleigh scattering, due to its elastic nature. The energy transfer coefficient for an individual interaction $\bar{f}_i$ is defined as the fraction of the initial photon energy transferred to the kinetic energy of the charged particle. The photon energy-transfer coefficient, $\frac{\mu_{\mathrm{tr}}}{\rho}$ is then defined to be

$$\frac{\mu_{\mathrm{tr}}}{\rho} = \frac{\mu_{\mathrm{PE}}}{\rho}\bar{f}_{\mathrm{tr,PE}} + \frac{\mu_{\mathrm{C}}}{\rho}\bar{f}_{\mathrm{tr,C}} + \frac{\mu_{\mathrm{PP}}}{\rho}\bar{f}_{\mathrm{tr,PP}} + \frac{\mu_{\mathrm{TP}}}{\rho}\bar{f}_{\mathrm{tr,TP}} \tag{2.8}$$

$$= \frac{\mu_{\mathrm{tr,PE}}}{\rho} + \frac{\mu_{\mathrm{tr,C}}}{\rho} + \frac{\mu_{\mathrm{tr,PP}}}{\rho} + \frac{\mu_{\mathrm{tr,TP}}}{\rho}. \tag{2.9}$$

As discussed in Section 1.2, charged particles deposit kinetic energy locally through multiple interactions. Any kinetic energy of the charged particle that is ultimately converted to radiative losses is not deposited as dose locally. Radiative losses take the form of bremsstrahlung,

positron in-flight annihilation and fluorescence following relaxation due to impact ionization. These radiative losses are incorporated in the $\bar{g}$ factor. The photon energy-absorption coefficient, $\frac{\mu_{\text{en}}}{\rho}$ is then defined to be

$$\frac{\mu_{\text{en}}}{\rho} = \frac{\mu_{\text{tr}}}{\rho}(1 - \bar{g}). \tag{2.10}$$

An example of how the mass attenuation coefficient and the energy-absorption coefficient differ for different materials and over photon energies is shown in Figure 2.5.

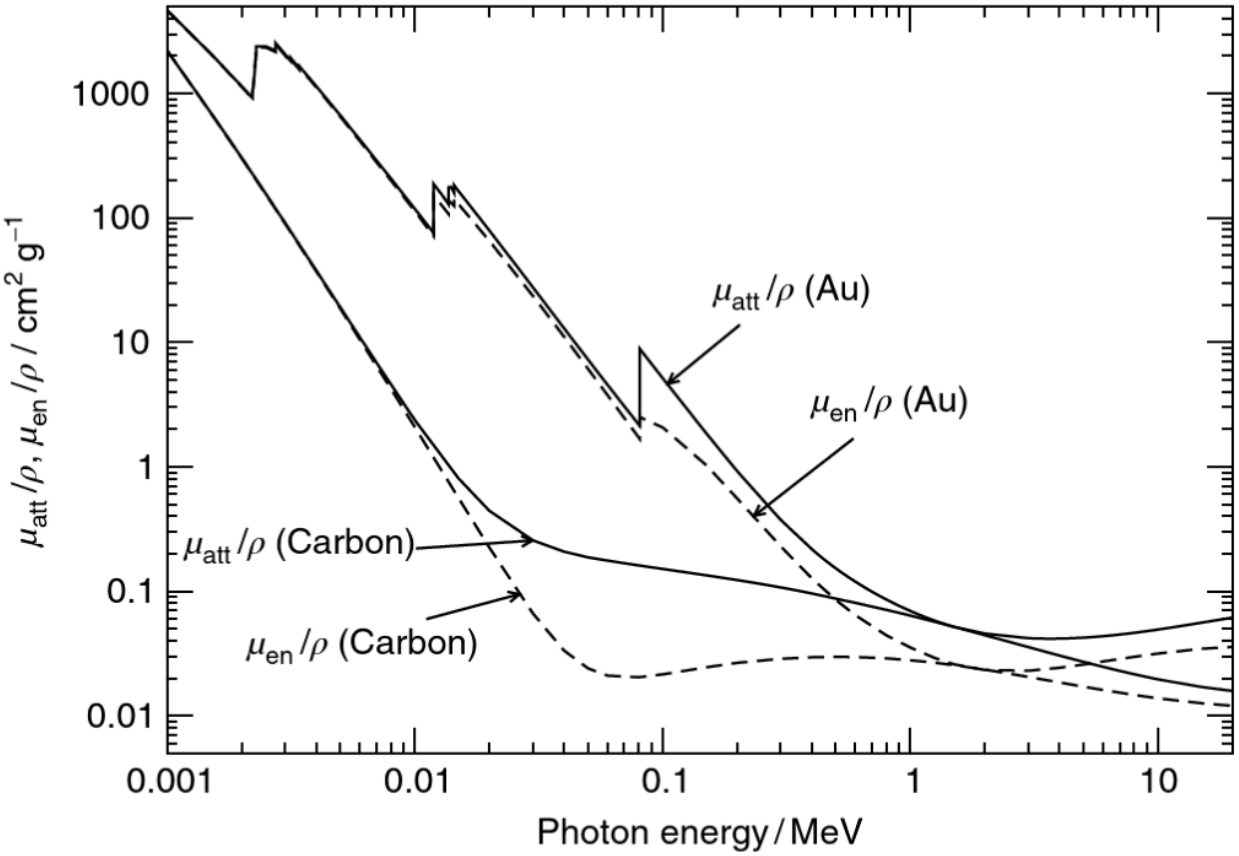


Figure 2.5: The mass attenuation coefficient and the energy absorption coefficient for Carbon and Gold (Au). Reproduced with permission from [3].

2.2 Charged particle interactions

Charged particles lose kinetic energy as they travel through a medium and interact with atoms. The rate of kinetic energy loss is given by the mass stopping power, $\frac{S}{\rho}$, given by

$$\frac{S}{\rho} = -\frac{1}{\rho}\frac{dE}{dx}, \tag{2.11}$$

where dE is the kinetic energy lost as the charged particle travels a distance dx. The types of interactions can be classified into elastic, inelastic and radiative interactions. In elastic interactions, the charged particle is scattered in a different direction by the electromagnetic field of the nucleus of the atom. There is of course some kinetic energy transferred to the nuclear recoil, but this is very small due to the nuclear mass. This type of collision occurs when the charged particle passes close to the nucleus. There are two types of inelastic collisions. The first type, "soft" inelastic collisions, occur when the charged particle is of considerable distance from the atom. Most of the time, the atom is excited. The second type, "hard" inelastic collisions, occur when the distance between the charged particle and the atom is about the atomic radius. In this case, the charged particle carries enough kinetic energy to eject an inner shell orbital electron. This ejected electron is known as a delta ray (δ) or knock-on electron. It is able to further liberate other electrons and deposit energy. The kinetic energy loss rate of the initial charged particle through inelastic collisions is known as the mass electronic stopping power, $\frac{S_{\text{el}}}{\rho}$. The radiative interaction also occurs when a charged particle passes close to the nucleus. If the charged particle is accelerated or decelerated by the electromagnetic field of the nucleus or orbital electrons, the charged particle emits photons. This is known as breaking radiation or bremsstrahlung. The amount of bremsstrahlung produced is given by the (non-relativistic) Larmor formula

$$P = \mu_0 \frac{q^2 a^2}{6\pi c} \tag{2.12}$$

where P is the power emitted in the form of bremsstrahlung, q is the charge of the particle and a is the acceleration of the particle. Due to the a^2 term, an accelerating or decelerating

particle produces bremsstrahlung. Because of Newton's second law,

$$a \propto \frac{qQ}{m} \tag{2.13}$$

where Q is the charge of the target and m is the mass of the particle. Therefore most of the bremsstrahlung produced comes from interactions with the nucleus in comparison to any orbital electrons. Due to the inverse square mass term, heavy particles produce much less bremsstrahlung. In the case of protons vs. electrons, protons produce 1 part in 1 million less bremsstrahlung than electrons. Thus, bremsstrahlung is of little concern in proton beams and proton radiation therapy. The energy loss rate of the initial charged particle through radiative interactions is known as the mass radiative stopping power, $\frac{S_{\text{rad}}}{\rho}$. The total mass stopping power is their sum, i.e.,

$$\frac{S}{\rho} = \frac{S_{\text{el}}}{\rho} + \frac{S_{\text{rad}}}{\rho}. \tag{2.14}$$

In general, the electronic stopping power decreases with kinetic energy until about 1 MeV, where it stabilizes, increasing slightly with energy. The radiative stopping power increases monotonically with kinetic energy. The radiative stopping power becomes larger than the electronic stopping power at kinetic energies between 10-100 MeV. For low Z materials, like water, air and graphite, it is closer to 100 MeV.

A closely related concept is linear energy transfer (LET). Whereas stopping power deals with the energy loss of a charged particle due to inelastic collisions, LET concerns the energy deposited to the medium along its track. The energy deposited is in the form of inelastic collisions that result in secondary electrons with kinetic energies below a certain cutoff level.

References

[1] "The international commission on radiation units and measurements," *Journal of the ICRU*, vol. 11, no. 1, pp. 5–6, 2011.

[2] E. B. Podgorsak, *Radiation Physics for Medical Physicists*. Springer International Publishing, 2016.

[3] P. Andreo, D. T. Burns, A. E. Nahum, J. Suentjens, and F. H. Attix, *Fundamentals of Ionizing Radiation Dosimetry*. Wiley-VCH, 2017.

[4] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," *Annalen der Physik*, vol. 322, no. 6, pp. 132–148, 1905.

[5] S. R. Cherry, J. A. Sorenson, and M. E. Phelps, *Physics in Nuclear Medicine*, 4th ed. Philadelphia, PA, USA: Elsevier, 2012.

[6] A. H. Compton, "A Quantum Theory of the Scattering of X-rays by Light Elements," *Phys. Rev.*, vol. 21, pp. 483–502, 5 May 1923.

[7] O. Klein and Y. Nishina, "Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac," *Zeitschrift für Physik*, vol. 52(11), no. 11, pp. 853–868, 1929.

[8] R. Ribberfors, "Relationship of the relativistic Compton cross section to the momentum distribution of bound electron states," *Phys. Rev. B*, vol. 12, pp. 2067–2074, 1975.

[9] J. Strutt, "XV. On the light from the sky, its polarization and colour," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 41, no. 271, pp. 107–120, 1871.

[10] P. M. S. Blackett, G. P. S. Occhialini, and E. Rutherford, "Some photographs of the tracks of penetrating radiation," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 139, no. 839, pp. 699–726, 1933.

3

Radiation Dosimetry

In the previous chapter, photon and charged particle interactions were discussed. These interactions take place at the atomic level and are singular events. In radiation dosimetry, macroscopic quantities are measured in materials where millions of interactions take place. In this chapter, we will see how the quantities discussed in the previous chapter give rise to radiation dosimetric quantities that are regularly measured.

3.1 Dosimetric quantities

3.1.1 Fluence and energy fluence

Given a monoenergetic beam of particles of energy E, the amount of particles dN traversing a sphere of cross-sectional area da is given by the fluence Φ,

$$\Phi = \frac{\mathrm{d}N}{\mathrm{d}a}. \tag{3.1}$$

The amount of energy crossing this sphere is given by the energy fluence, Ψ,

$$\Psi = E\Phi. \tag{3.2}$$

A fluence (and energy fluence) rate can also be defined. If the beam consists of a spectrum of energies, a fluence spectrum, Φ_E, can be defined as

$$\Phi_{\mathrm{E}} = \frac{\mathrm{d}\Phi}{\mathrm{d}E}. \tag{3.3}$$

The energy fluence spectrum, Ψ_{E}, is defined as

$$\Psi_{\mathrm{E}} = E\Phi_{\mathrm{E}}. \tag{3.4}$$

The total fluence (and energy fluence) can be obtained by integrating the spectrum over all energies present in the beam.

3.1.2 Kerma

Photons and other uncharged particles transfer kinetic energy to charged particles upon interaction. The kinetic energy released per unit mass, or kerma, K, is defined as

$$K = \frac{\mathrm{d}E_{\mathrm{tr}}}{\mathrm{d}m}, \tag{3.5}$$

where dE_{tr} is the initial kinetic energy transferred to all charged particles by uncharged particles in a region of mass dm. For a spectrum of energies, the kerma can be written as

$$K = \int_0^{E_{\mathrm{max}}} \Psi_{\mathrm{E}} \frac{\mu_{\mathrm{tr}}(E)}{\rho} dE. \tag{3.6}$$

where E_{max} is the maximum kinetic energy in the spectrum. The units of kerma is Gray (Gy). The component of kerma that results only in local dose deposition is known as electronic kerma, K_{el}, given by

$$K_{\mathrm{el}} = K(1-\bar{g}) = \int_0^{E_{\max}} \Psi_{\mathrm{E}} \frac{\mu_{\mathrm{en}}(E)}{\rho} dE. \tag{3.7}$$

3.1.3 Charged particle equilibrium

The absorbed dose D at a point in a medium can be equated to the electronic kerma at the same point when the conditions for charged particle equilibrium (CPE) exist. First, consider a medium homogeneous in density and atomic number that is irradiated by a uniform beam of uncharged particles. Now consider a point in the medium (far away from any boundaries) and an infinitesimal sphere around the point. Because of the uniform medium and beam, any charged particle (created inside the sphere) of a given energy and direction exiting the sphere is replaced by an identical charged particle (created outside the sphere) entering the sphere. Additionally, any radiative loss occurring outside of the sphere is replaced by an identical one occurring inside the sphere. It can now be said that CPE exists at this point and all other points of the medium since they are identical. Therefore, under CPE,

$$D \overset{\mathrm{CPE}}{=} K_{\mathrm{el}}. \tag{3.8}$$

In reality, CPE does not exist. This is because uncharged particle beams are attenuated as they traverse a distance comparable to the maximum secondary electron range. Thus, the secondary electrons produced up-stream of the beam are not identical to those produced down-stream. Past the point of maximum dose deposition, the absorbed dose at one point is greater than the electronic kerma at the same point, but they are proportional to one another. Under these conditions, transient (or partial) CPE exists and

$$D \overset{\mathrm{TCPE}}{=} \beta K_{\mathrm{el}}, \tag{3.9}$$

where β is a constant of proportionality greater than unity.

3.1.4 Percent depth dose

The percent depth dose distribution (PDD) describes how absorbed dose varies with depth in a medium on the central beam axis, for a given field size and SSD. Generally, the medium is water. A PDD is normalized to the maximum dose. Figure 3.1 presents the PDD of a 6 and 10 MV photon beam, a 9 and 16 MeV electron beam and a 50 and 190 MeV proton beam. A MV photon beam is characterized by a low surface dose, a buildup region to the depth of maximum dose and then an exponential fall off. The buildup region is due to the photons travelling some distance before interacting. A MeV electron beam is characterized by a large surface dose, a shallow depth of maximum dose and then a fast decrease. The higher surface dose in comparison to MV photon beams is due to electrons being able to deposit dose on the spot, unlike photons. A low dose region is found deep in the medium and is due to the bremsstrahlung produced in the medium and linac treatment head. It is about 5-15% of the maximum dose. A proton PDD is characterized by a low surface dose and a plateau region for some distance into the medium. The dose then rises rapidly, peaks and then falls off quickly. This is known as the Bragg peak. It is attributed to the rapid rise in proton stopping power as it reaches the end of its range, where it has low kinetic energy.

3.1.5 Beam quality

The quality Q of a beam is a specifier of a given radiation beam. A beam quality specifier should be easily measurable in-clinic. This is in contrast to specifying the beam through the more descriptive, but much more complex position, energy and angular distribution of each particle type at all points in the medium.

For MV photon beams, the TG-51 [1] beam quality specifier is the value of the photon component of the PDD at 10 cm depth in water, $\text{PDD}(10)_x$, where x denotes the photon component. The field size is 10x10 cm^2 at the phantom surface and the SSD is 100 cm. Only the photon component is desired because electrons generated in the linac head are included

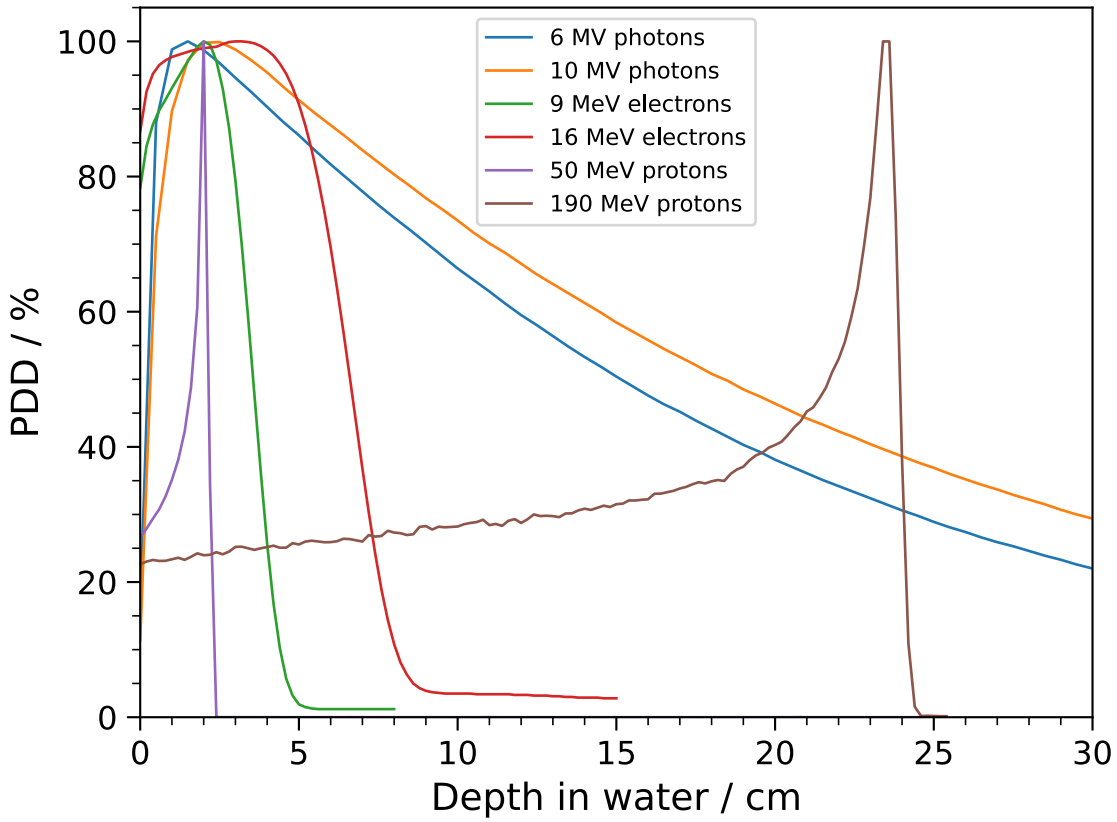


Figure 3.1: PDDs of a 6 and 10 MV photon beam, a 9 and 16 MeV electron beam and a 50 and 190 MeV proton beam. The field size is 10x10 cm^2 for the photon and electron beams and 4x4 cm^2 for the proton beams. The SSD is 100 cm for the photon and electron beams and 10 cm for the proton beams. Proton PDDs courtesy of Alaina Giang Bui and was calculated with the GEANT4 Monte Carlo toolkit, version 11.2.

in the total PDD.

For MeV electron beams, the TG-51 [1] beam quality specifier is the depth in water at which the PDD reaches 50% of the maximum, R_{50}. The SSD is 100 cm and the field size is 10x10 cm^2 at the phantom surface for R_{50} <8.5 cm and 20x20 cm^2 for R_{50} >8.5 cm.

For proton beams, the TRS-398 [2] beam quality specifier is the residual range R_{res}, given by

$$R_{\text{res}} = R_{\text{p}} - z_{\text{ref}}, \tag{3.10}$$

where R_{p} is the practical range, where the PDD reaches 10% of the maximum dose and the z_{ref} is the reference depth of the measurement. For SOBP beams, z_{ref} is taken at the centre of the SOBP. In a monoenergetic proton beam, z_{ref} is a depth of 1 or 2 cm.

3.2 Cavity theory

Radiation dosimetry consists of measurements to determine absorbed dose. In order to perform the measurement, a radiation dosimeter must be inserted in the medium, typically an air-filled ionization chamber is inserted into a water phantom. For this chapter, dosimeter and detector will be used synonymously. The sensitive, charge collecting volume of the dosimeter is not the same material as that of the medium. Therefore, a relation between the dose delivered to the sensitive volume of the chamber D_{det}, and the dose to the medium in the absence of the chamber, D_{med}, must be established. Cavity theory attempts to determine these relationships. The word cavity is used because the detector can be thought of as a cavity of a different material within the medium. It is noted that for a spectrum of charged particles, the absorbed dose to the medium can be written as [3]

$$D_{\text{med}} = \int_0^{E_{\max}} \Phi_{\text{med,E}} \left(\frac{S_{\text{el}}(E)}{\rho} \right)_{\text{med}} dE, \tag{3.11}$$

where $\Phi_{\text{med,E}}$ is the charged particle fluence spectrum. Eq 3.11 applies when there is δ-CPE, CPE for delta rays.

3.2.1 Bragg-Gray cavity theory

Consider a medium with a small detector or "cavity" placed within it. The detector sensitive volume is made of low-density material, such as air. For MV photon and MeV electron beams,

the secondary electrons are energetic enough so that they can cross the cavity without the secondary electron fluence spectrum being disturbed by the cavity. The photon fluence spectrum is also undisturbed. Thus, the cavity is essentially "sensing" the electrons that cross it. In this scenario, the cavity can be considered as small compared to the electron ranges. Note that the usage of "small" depends on the beam energy and the material of the cavity. Due to the secondary electron fluence spectrum remaining undisturbed, Eq 3.11 can be used to determine the ratio $\frac{D_{\text{med}}}{D_{\text{det}}}$, that is,

$$\frac{D_{\text{med}}}{D_{\text{det}}} = \frac{\int_0^{E_{\max}} \Phi_{\text{med,E}}^{\text{e,prim}} \left(\frac{S_{\text{el}}(E)}{\rho}\right)_{\text{med}} dE}{\int_0^{E_{\max}} \Phi_{\text{med,E}}^{\text{e,prim}} \left(\frac{S_{\text{el}}(E)}{\rho}\right)_{\text{det}} dE}, \tag{3.12}$$

where $\Phi_{E,\text{med}}^{\text{e,prim}}$ is the "primary" electron fluence (only the secondary electrons initially liberated by photons) in the medium.

3.2.2 Spencer-Attix cavity theory

As discussed in Section 3.1.3, CPE is never attained in reality. This applies to δ-CPE as well. This is because high energy delta rays produced inside the cavity lose energy as they cross the cavity and the remaining energy they carry away is not necessarily equivalent to that deposited in the cavity by incoming delta rays. Therefore, Bragg-Gray cavity theory is only an approximation. Spencer and Attix [4] incorporated these high energy delta rays by considering two energy ranges, separated by an energy cut-off Δ. This energy cut-off is chosen for an electron energy that is just able to cross the cavity. For air, Δ is 10-15 keV. Delta rays with energies above Δ are added to the total electron fluence spectrum, $\Phi_{E,\text{med}}^{\text{e,tot}}$. The total electron fluence spectrum remains undisturbed by the cavity. Delta rays with energies below Δ are assumed to deposit their energy on the spot. The ratio $\frac{D_{\text{med}}}{D_{\text{det}}}$ can then be written as

$$\frac{D_{\text{med}}}{D_{\text{det}}} = \frac{\int_{\Delta}^{E_{\max}} \Phi_{\text{med,E}}^{\text{e,tot}} \left(\frac{L_{\Delta}(E)}{\rho}\right)_{\text{med}} dE + TE_{\text{med}}}{\int_{\Delta}^{\text{E}_{\max}} \Phi_{\text{med,E}}^{\text{e,tot}} \left(\frac{L_{\Delta}(E)}{\rho}\right)_{\text{det}} dE + TE_{\text{det}}}, \tag{3.13}$$

where $(L_\Delta(E)/\rho)_{\text{med}}$ and $(L_\Delta(E)/\rho)_{\text{det}}$ are the *restricted* mass electronic stopping powers in the medium and detector, respectively. Restricted mass electronic stopping powers only consider charged particle interactions that result in an energy transfer of energy Δ or less. The track end terms

$$TE_{\text{med}} = \Phi^{\text{e,tot}}_{\text{E,med}}(\Delta) \left(\frac{S_{\text{el}}(\Delta)}{\rho} \right)_{\text{med}} \Delta, \tag{3.14}$$

and

$$TE_{\text{det}} = \Phi^{\text{e,tot}}_{\text{E,med}}(\Delta) \left(\frac{S_{\text{el}}(\Delta)}{\rho} \right)_{\text{det}} \Delta, \tag{3.15}$$

were added by Nahum [5] in order to incorporate the dose deposition of electrons with energies below Δ.

3.3 Ionization chambers

3.3.1 Cylindrical chambers

Cylindrical ionization chambers are the main dosimeter used in cancer clinics for MV photon and MeV electron beam dosimetry. The IAEA TRS-398 also recommends cylindrical chambers in proton SOBP beams if $R_{res} > 0.5$ cm and in single energy beams if $R_{res} > 15$ cm. They are cylindrical in shape and air-filled. A schematic of a Farmer-type, cylindrical, air-filled ionization chamber is shown in Figure 3.2. The sensitive, charge collection volume is surrounded by a wall and filled with dry air (no humidity). The air volume is vented so that the air does not remain at constant temperature and pressure. Typically, the volume of the collecting volume is 0.6 cm^3. The outer wall is generally made of graphite or an air-equivalent plastic such as C552. The wall is made thick enough to maintain TCPE without disturbing the photon fluence (in a photon beam) [3]. This allows the application of Spencer-Attix cavity theory (Section 3.2.2). With graphite, the wall thickness is typically

0.5 cm. In the centre of the collecting volume is a cylindrical electrode, about 1 mm in diameter. It is typically made of aluminum or an air-equivalent plastic such as C552. During operation, a bias voltage is applied between the central electrode and the wall, which act as two electrodes. The region apart from the collecting volume is known as the stem. An insulator separates the wall from the central electrode in the stem so that no current leaks between the wall and central electrode. The guard electrode ensures electric field uniformity when the voltage is applied and also redirects any leakage current to ground so that it is not collected. The charges created inside the sensitive volume are collected by the electrodes

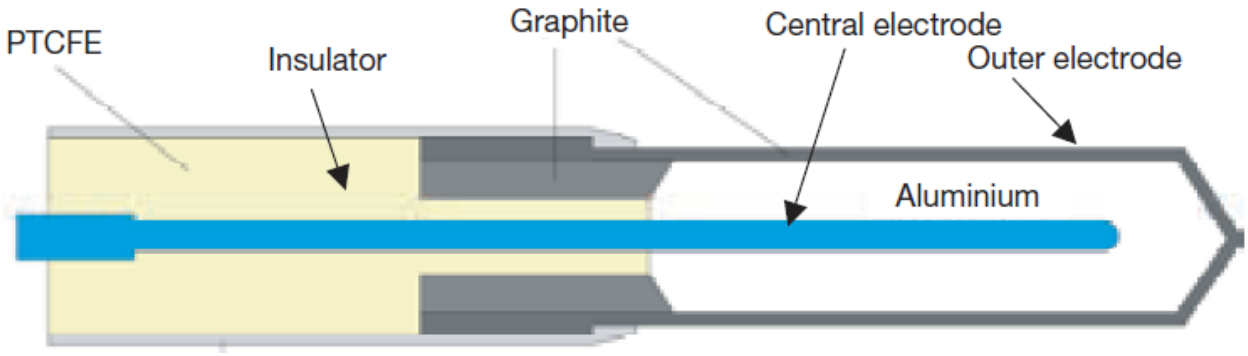


Figure 3.2: A schematic of a Farmer-type air-filled cylindrical ionization chamber. Reproduced with permission from [6].

by applying a voltage across them. This voltage is generally between +150 V and +300 V. The high voltage is generally applied to the central electrode while the outer electrode is at ground. The electric field lines point radially outwards from the high voltage electrode The charges move towards the high or low voltage electrode, depending on the sign of their charge. The amount of charge measured for a pulsed MV or MeV photon beam is in the nC range. The measured charge is read out by an electrometer which is accurate to within a few pC. The charge measurement is a time integration of the measured current.

3.3.2 Parallel-plate chambers

Parallel-plate chambers are the type of chamber recommended by TG-51 [1, 7] for MeV electron beam reference dosimetry for electron beams below 10 MeV ($R_{50} \leq 4.3$cm). The IAEA TRS-398 [2, 8] also recommends cylindrical chambers in proton SOBP beams if $R_{\text{res}} <$ 0.5 cm and in single energy beams if $R_{\text{res}} < 15$ cm. Parallel-plate chambers are also used in kilovoltage x-ray dosimetry. A schematic of a parallel-plate chamber can be seen in Figure 3.3. The components and the operating principles of a parallel-plate chamber is the same as that of a cylindrical chamber. The electric field lines point from the high voltage plate to the ground plate and are perpendicular to the plates themselves.

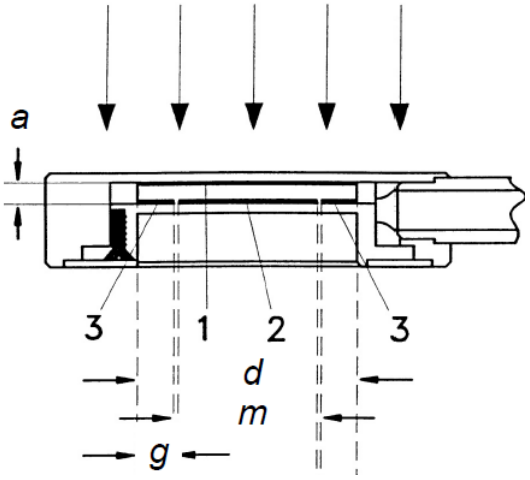


Figure 3.3: A schematic of a parallel-plate ionization chamber with (1) the air volume of diameter d, (2) the collecting electrode of diameter m and (3) the guard electrode of width g. Reproduced with permission from [3].

3.3.3 Ionization chamber characteristics

An ionization chamber should have some desirable characteristics. These are so that the reading of the chamber is reliable and corresponds to a unique absorbed dose value. The response of a chamber is useful in the discussion of these characteristics. The response of a chamber is given by the ratio M/F, where M is the chamber reading and F is the desired quantity, such as kerma or absorbed dose.

Reproducibility

Given the same amount of absorbed dose to the sensitive volume of the chamber, the reading of an ionization chamber should remain the same (to within the Type B uncertainties). Due to these type B uncertainties, ionization chamber measurements should be repeated a few times.

Linearity

The reading of an ionization chamber should be linearly proportional to the absorbed dose deposited in the sensitive volume, given the beam quality is kept constant. In this case, the response should be constant. A user must be aware of any deviation from linearity outside of a certain dose range. In this case, a correction must be applied or another dosimeter used.

Dose rate dependence

Given the same amount of absorbed dose, the reading of the ionization chamber should not depend on the dose rate. Thus, the response should not depend on dose rate. This characteristic also ensures linearity. A chamber may be used outside of the ideal dose rate range but must be calibrated beforehand.

Directional dependence

Ideally, the response of the chamber should not depend on how it is oriented within the radiation beam. For cylindrical chambers, a directional dependence can arise due to construction limitations (asymmetric construction) and air gaps, but these are typically negligible. In this case, they should be consistently used in the same orientation as they have been calibrated (or characterized).

Energy dependence

Ideally, the response of the chamber would be independent of the beam quality of the radiation beam. This is never the case in practice and the response depends on the beam quality. Thus, the chamber should be calibrated at every beam quality needed clinically. This is generally impractical and the chamber is usually calibrated at one reference quality at the PSDL. For MV photons, MeV electrons and protons, this reference beam quality is Cobalt-60. In order for the chamber to be used with other beam qualities, a beam quality correction factor is required. For example, the TG-51 addendum [7] provides a method on determining the value of these beam quality correction factors for several beam qualities and chambers. Monte Carlo simulations can also be used. In the case of protons, the difference in the value of W_{air}, the mean energy expended by a particle to create an ion pair, between Cobalt-60 (33.97 eV) and the protons (34.44 eV) must be taken into account [2]. Energy dependencies of chambers also arise due to the stopping powers having an energy dependence.

3.3.4 Ionization chamber correction factors

As discussed in Section 1.3.3, ionization chambers are calibrated under reference conditions at PSDLs. If in-clinic ionization chamber measurements are not performed at the same reference conditions, then the raw ionization chamber reading, M_{raw}, must be corrected to the value under reference conditions.

Temperature and pressure

In Canada and the US, the reference temperature and pressure are taken to be 22° C and 101.325 kPa, respectively. If the chamber reading is taken at a different temperature and pressure, a temperature and pressure correction factor, P_{TP}, is applied to the raw chamber reading.

Ion recombination

This correction factor has already been discussed in Section 1.4.1 and will be discussed even further in Chapter 4. The ion recombination correction factor is denoted as P_{ion}. According to TG-51, P_{ion} should be no larger than 5%. if using the two-voltage method.

Polarity effects

The raw ionization chamber reading may depend on the polarity of applied voltage. This can be due to a number of causes, such as charges being produced and collected in the collecting electrode (Compton current) and charges produced from radiation interactions outside the collecting volume, such as in the stem or the cable, being collected (extracameral current). Polarity effects are corrected through the polarity correction factor P_{pol}.

In total, the corrected ionization chamber, M, can be calculated as

$$M = M_{\text{raw}} P_{\text{TP}} P_{\text{ion}} P_{\text{pol}} \tag{3.16}$$

References

[1] P. R. Almond *et al.*, "AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon and electron beams," *Medical Physics*, vol. 26, no. 9, pp. 1847–1870, 1999.

[2] *Absorbed Dose Determination in External Beam Radiotherapy* (Technical Reports Series 398 (Rev. 1)). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2024.

[3] P. Andreo, D. T. Burns, A. E. Nahum, J. Suentjens, and F. H. Attix, *Fundamentals of Ionizing Radiation Dosimetry*. Wiley-VCH, 2017.

[4] L. V. Spencer and F. H. Attix, "A theory of cavity ionization," *Radiation Research*, vol. 3(3), no. 3, pp. 239–254, 1955.

[5] A. Nahum, "Water/air mass stopping power ratios for megavoltage photon and electron beams," *Physics in Medicine & Biology*, vol. 23, no. 1, p. 24, Jan. 1978.

[6] I. Izewska and G. Rajan, "Radiation dosimeters," in *Review of Radiation Oncology Physics: A Handbook for Teachers and Students*, E. B. Podgorsak, Ed., Vienna: International Atmojc Energy Agency, 2005, ch. 14, pp. 71–99.

[7] M. McEwen *et al.*, "Addendum to the AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon beams," *Medical Physics*, vol. 41, no. 4, p. 041 501, 2014.

[8] *Absorbed Dose Determination in External Beam Radiotherapy* (Technical Reports Series 398). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2001.

4

Ion Recombination

This chapter reviews previous investigations into ion recombination in air-filled ionization chambers and the development of the theory of ion recombination. Expressions for the chamber collection efficiency (or ion recombination correction factor) that have been developed from these studies are the basis for in-clinic methods of determining the ion recombination correction factor. These methods are the two-voltage method (TVM) and Jaffé plot extrapolation. The issues in determining the ion recombination correction factor in UHDR beams is discussed.

4.1 Specification of dose rate in pulsed beams

The time structure of a pulsed beam is shown in Figure 4.1. A single *macro-pulse* consists of several *micro-pulses*, usually on the order of 1-10 ps. A single macro-pulse can be considered as quasi-continuous because the micro-pulses are very closely bunched together, with a repetition rate of around 10 MHz (but depending on the linac). The pulse width of a single macro-pulse is given by t_{p} (on the order of 1-3 μs) and the intra-pulse dose rate is $\dot{D}_{\mathrm{p}}$ (which can be over 1 MGy/s). The *dose-per-pulse* (DPP) is the total dose delivered by a single

pulse. The time between the macro-pulses is given by t_r. A multi-pulse delivery consists of n_{p} pulses delivered in an irradiation time t_i. The dose rate that is reported in UHDR research is $\bar{\dot{D}}$, the *mean dose rate for a multi-pulse delivery.*

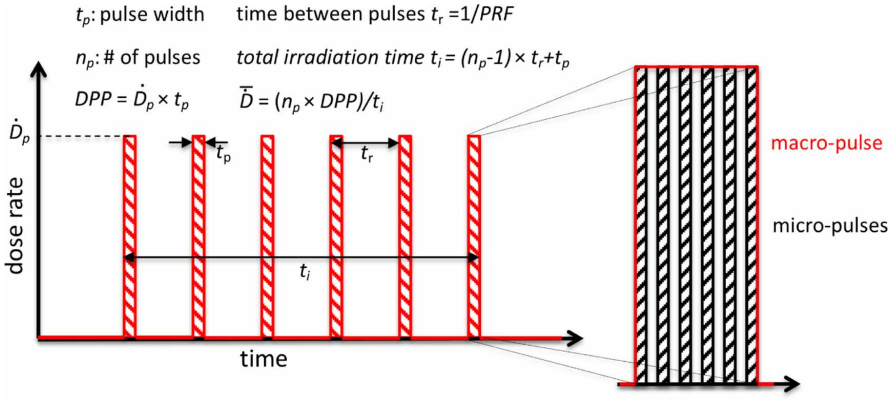


Figure 4.1: A diagram of the time structure of a pulsed beam. Reproduced with permission from [1].

4.2 Previous investigations

4.2.1 Initial recombination

Initial recombination concerns the recombination of positive and negative ions that were initially produced by the same radiation track. Initial recombination was first investigated by Jaffé [2] and Kara-Michailova and Lea [3]. In the Jaffé model, the spatial distribution of ionization in a single track is modelled as a Gaussian with a mean square radius of b. The width of the column widens as time progresses due to diffusion. For a track parallel to the electric field, the collection efficiency $f_{||}$ is given as

$$f_{||} = \frac{y_1}{y_2} e^{-y_1} \left[\mathrm{Ei}(y_1 + \ln(1 + y_2)) - \mathrm{Ei}(y_1) \right], \tag{4.1}$$

where

$$y_1 = \frac{8\pi\mathcal{D}}{\alpha N_0}, \quad y_2 = \frac{2d\mathcal{D}}{\mu b^2 E}, \quad N_0 = \frac{\mathrm{LET}}{W_{\mathrm{air}}}, \tag{4.2}$$

$\mathcal{D}$ is the averaged positive and negative ion diffusion coefficient, α is the positive ion - negative ion recombination coefficient, μ is the averaged positive and negative ion mobility, d is the plate separation, E is the electric field strength and Ei is the exponential integral. The linear charge density along a track, N_0, is equal to the LET of the beam divided by the mean energy required to form an ion pair in air, W_{air}. Additionally, the collection efficiency for a track at an angle θ to the electric field lines, f_θ, is given by

$$f_\theta = \frac{1}{1 + y_1^{-1} S(z)}, \tag{4.3}$$

where

$$S(z) = e^z \left(\frac{i\pi}{2}\right) H_0^{(1)}(iz) \quad \text{and} \quad z = \left(\frac{\mu b E \sin\theta}{2D}\right)^2. \tag{4.4}$$

In Eq. 4.4 $H_0^{(1)}$ is the order zero Hankel function of the first kind. Zanstra [4] experimentally showed that at large electric field strengths and not too high pressures ($E/p > 80$ V cm^{-1} atm^{-1}), the measured charge $1/Q$ followed a linear relationship with $1/E$, i.e.,

$$\frac{1}{Q} = \frac{1}{Q_0} + \frac{c}{V_0}, \tag{4.5}$$

where $E = V_0/d$ when there is no space charge and c is constant for a given chamber. The applied potential to the chamber electrodes is V_0. Thus, $\frac{1}{Q}$ is linear in $\frac{1}{V_0}$ with a y-intercept of $\frac{1}{Q_0}$. Kara-Michailova and Lea also demonstrated this linear relationship theoretically. For conventional MV photon and MeV electron beams, initial recombination is less than 0.2% [5, 6]. For protons, despite being higher LET than electrons, initial recombination is still a small contribution. At higher voltages though, it cannot be ignored. For light ion beams such as helium, carbon and oxygen, initial recombination is dominant at high bias voltages applied to the chamber electrodes [7].

4.2.2 General recombination

General (or volume) recombination concerns the recombination of positive and negative ions formed in different radiation tracks. General recombination is the dominant recombination mechanism for conventional MV photon and MeV electron beams and proton beams. For light ion beams, it is actually dominant at low voltages and high dose rates [7].

Continuous beams

Thomson [8] was one of the earliest investigators into the theory of general ion recombination. He considered a parallel-plate chamber irradiated by a continuous beam and considered only positive and negative ions in a steady-state. He put forward the following partial differential equations (PDEs) for the ion dynamics,

$$\frac{\partial n_+}{\partial t} = 0 = R - \alpha n_+ n_- - \mu_+ \boldsymbol{\nabla} \cdot \left(n_+ \vec{\mathbf{E}} \right), \tag{4.6}$$

$$\frac{\partial n_-}{\partial t} = 0 = R - \alpha n_+ n_- + \mu_- \boldsymbol{\nabla} \cdot \left(n_- \vec{\mathbf{E}} \right), \tag{4.7}$$

$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = \frac{e}{\epsilon} \left(n_+ - n_- \right), \tag{4.8}$$

$$\vec{\mathbf{E}} = -\boldsymbol{\nabla} V, \tag{4.9}$$

where n_+ and n_- are the positive and negative ion concentrations, respectively, μ_+ and μ_- are the positive and negative ion mobilities, respectively, and R is the constant ionization rate. The electric field, $\vec{\mathbf{E}}$, due to space charge is modelled through Gauss' law (Eq. 4.8) and V is the electric potential. Due to the parallel-plate geometry, $\boldsymbol{\nabla} = \frac{\partial}{\partial z}$. The Thomson boundary conditions for a parallel-plate chamber are prescribed as

$$n_+(0,t) = 0 \tag{4.10}$$

$$n_-(d,t) = 0 \tag{4.11}$$

$$V(0,t) = V_0 \tag{4.12}$$

$$V(d,t) = 0 \tag{4.13}$$

where V_0 is the positive applied voltage to the electrode at position $z = 0$. The condition of Eq. 4.10 implies no positive ions are collected at the positive electrode and condition Eq. 4.11 implies no negative ions are collected at the ground electrode at $z = d$. Thomson was able to develop a closed-form solution to Eqs. 4.6-4.8 only when $\mu_+ = \mu_-$. Also, an approximate current-voltage (I-V) relationship was obtained. Mie [9] was able to develop a more accurate series solution to Eqs. 4.6-4.8. Townsend [10], along with Rosen and George [11] expanding on his work, obtained a closed form solution for the ion concentrations, neglecting space charge. Greening [12] showed that Mie's theory agreed with experimental results if the collection efficiency $f \geq 0.7$. Armstrong and Tate [13] demonstrated the same conclusion through numerical solutions. They also demonstrated that the approximate theory of Boag and Wilson [14] is accurate when $f \geq 0.95$. The Mie expression for the collection efficiency can be written as

$$\frac{1}{f} = \frac{Q_0}{Q} = 1 + \frac{\alpha d^4 R}{6\mu_+\mu_- V_0^2}. \tag{4.14}$$

Notice this is an analytic expression for P_{ion}. Equation 4.14 can be re-written as

$$\frac{1}{Q} = \frac{1}{Q_0} + \frac{1}{V_0^2}\left(\frac{\alpha d^4 R}{6\mu_+\mu_- Q_0}\right). \tag{4.15}$$

Notice that in Eq. 4.15, the quantity in parentheses is constant for a given chamber. Thus, $\frac{1}{Q}$ is linear in $\frac{1}{V_0^2}$ with a y-intercept of $\frac{1}{Q_0}$.

Pulsed beams

A theory of general ion recombination in pulsed beams was first put forward by Boag [15]. Boag modelled the beam as a delta function that creates all the initial charge instantly. The positive ions travel towards the ground electrode and the negative ions travel towards the high voltage electrode. Thus, a sheath devoid of positive ions is found adjacent to the high voltage electrode and a sheath devoid of negative ions adjacent to the ground electrode. In-between the sheaths is an overlap region where the positive and negative ion concentrations are equal and undergo recombination until the boundaries of the sheaths meet. Space charge is neglected. From this model, Boag derived the following analytical expression for the collection efficiency

$$f = \frac{1}{u}\ln(1+u), \tag{4.16}$$

where

$$u = \frac{\alpha d^2 n_0}{(\mu_+ + \mu_-)V_0}. \tag{4.17}$$

In Eq. 4.17, n_0 is the concentration of ions of one sign initially produced by the pulsed radiation beam. Near saturation ($f \to 1 \implies u \to 0$), Eq. 4.16 can be expanded as

$$\frac{1}{f} = \frac{Q_0}{Q} = 1 + \frac{u}{2}, \tag{4.18}$$

or

$$\frac{1}{Q} = \frac{1}{Q_0} + \frac{1}{V_0}\left(\frac{\alpha d^2}{(\mu_+ + \mu_-)e\nu}\right), \tag{4.19}$$

where the relation $Q_0 = e\nu n_0$ was used, where e is the electron charge and ν is the volume of the collecting volume of the chamber. Notice that in Eq. 4.19, the quantity in parentheses

is constant for a given chamber. Thus, $\frac{1}{Q}$ is linear in $\frac{1}{V_0}$ with a y-intercept of $\frac{1}{Q_0}$. This is in contrast to the $1/V_0^2$ relation found for continuous beams.

4.3 Jaffé plot extrapolation

A Jaffé plot is one where measured values of $\frac{1}{Q}$ at various voltages V_0 are plotted against $\frac{1}{V_0}$ in the pulsed beam case (including initial recombination) or $\frac{1}{V_0^2}$ in the continuous beam case. According to Eq. 4.15 for continuous beams and Eq. 4.19 for pulsed beams, a linear fit to these data points, yields $1/Q_0$ as the y-intercept. The collection efficiency (at a given voltage) can then be computed as Q/Q_0.

4.4 Two-voltage method

The two-voltage method (TVM) is used by the TG-51 [16] and TRS-398 [6] to determine the ion recombination correction factor in a clinical environment. It is useful because only two, quick chamber measurements are required. For now, let us consider a pulsed beam and a parallel-plate chamber, though this method also works with cylindrical chambers. Consider a measurement at voltages V_1 and V_2 resulting in chamber readings Q_1 and Q_2. In Eq. 4.19, denote the quantity in parentheses as a. Because Q_0 is the same for both measurements, the following equality holds

$$\frac{1}{Q_0} = \left(\frac{1}{Q_1} - \frac{a}{V_1}\right) = \left(\frac{1}{Q_2} - \frac{a}{V_2}\right). \tag{4.20}$$

Solving for a in Eq. 4.20 and then plugging it back into Eq. 4.19 at (V_1, Q_1), the following expression is obtained for the collection efficiency

$$f_1 = \frac{Q_1}{Q_0} = \frac{\frac{V_1}{V_2} - \frac{Q_1}{Q_2}}{\frac{V_1}{V_2} - 1}. \tag{4.21}$$

Note how Eq 4.21 also works for initial recombination. A similar expression can be developed for continuous beams,

$$f_1 = \frac{Q_1}{Q_0} = \frac{\left(\frac{V_1}{V_2}\right)^2 - \frac{Q_1}{Q_2}}{\left(\frac{V_1}{V_2}\right)^2 - 1}. \tag{4.22}$$

In the region where $f > 0.97$, the TVM is accurate to 0.1%.

4.5 Ion recombination in UHDR beams

4.5.1 Free electrons and space charge

It has been known since Boag and Currant [17] that at high V_0, the fraction of free electrons collected is quite large. Therefore, the original Boag formula (Eq 4.16) would underestimate the collection efficiency at higher voltages. Hochhäuser and Balk [18, 19] were able to measure the free electron component. Bielajew [20] considered the free electron fraction for continuous beams. Boag, Hochhäuser and Balk proposed three new analytical expressions for the charge collection efficiency in pulsed beams that incorporated the free electron fraction p [21] (their derivations can be found in the same article)

$$f' = \frac{1}{u} \ln\left(1 + \frac{e^{pu} - 1}{p}\right), \tag{4.23}$$

$$f'' = p + \frac{1}{u} \ln\left(1 + (1-p)u\right), \tag{4.24}$$

$$f''' = \lambda + \frac{1}{u} \ln\left(1 + \frac{e^{\lambda(1-\lambda)u}}{\lambda}\right), \tag{4.25}$$

where $\lambda = 1 - \sqrt{1-p}$. Several attempts have been made to use these expressions to obtain a recombination correction factor in the intermediate DPP range (10-100 mGy/pulse) [22, 23]. Gotz et. al. [24] showed that the modified Boag formulae only work up to intermediate DPPs

and low voltages (< 100 V). The modified Boag formulae, however, do not incorporate the space charge that results from the rapid free electron collection. This space charge distorts the electric field within the ionization chamber. This space charge effect becomes significant at high DPPs [25, 26]. Figure 4.2 shows how the electric field within a parallel-plate chamber distorts with increasing DPP. As a consequence of the free electron fraction and space charge, the TVM and Jaffé plot extrapolation only give accurate results up to 10 mGy/pulse [25, 27].

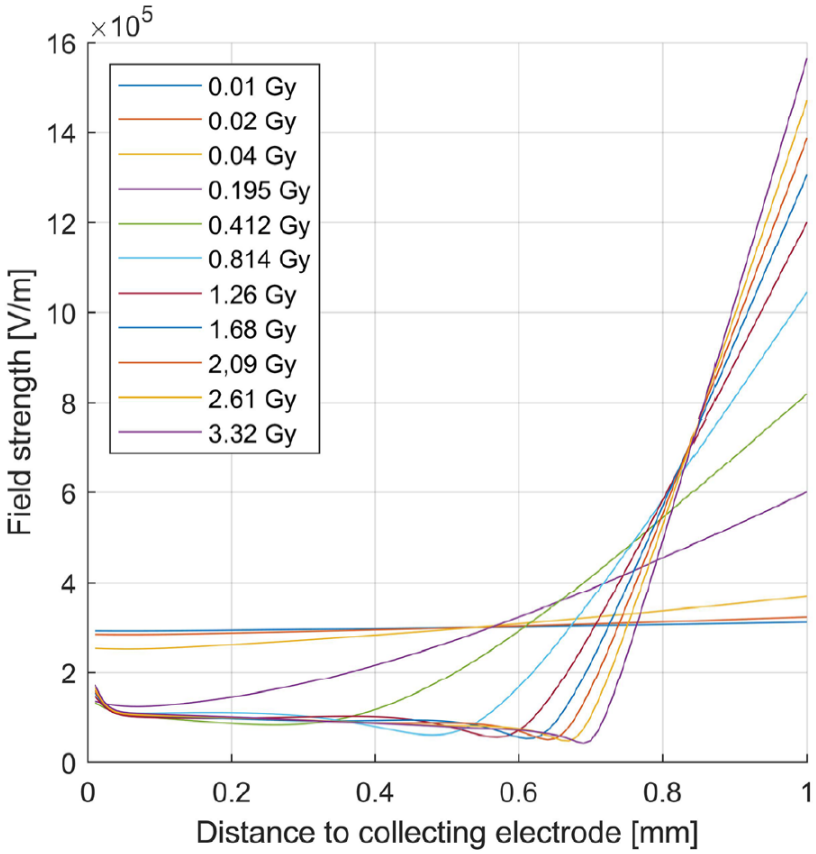


Figure 4.2: The electric field inside an Advanced Markus parallel-plate ionization chamber with 1 mm plate separation. Various dose-per-pulses (DPP) are shown in the legend. As the DPP increases the electric field becomes increasingly more distorted from the initial field (applied voltage of 300 V). The absence of the free electrons at 1 mm leaves a large positive space charge. The field near the collecting electrode at 0 mm (where the free electrons are collected) collapses. Reproduced under the terms of the Creative Commons CC BY license from [25].

4.5.2 Potential solutions

The breakdown in analytic expressions for the ion recombination correction factor in high DPP beams carries over to the TVM and Jaffé plot extrapolation. Therefore, other solution methods must be used to study ion recombination and its respective correction factor. Simply increasing the applied voltage is not an ideal solution because charge multiplication begins to occur and thus the total charge produced by the beam is not proportional to the absorbed dose [26, 28, 29].

Metrological solutions

Metrological solutions use novel measurement equipment or techniques to maximize the charge collection efficiency. Gomez et. al. [30] have developed ultra-thin parallel-plate chambers with plate separations around 0.25 mm. A small plate separation increases the free electron fraction and maximizes the collection efficiency for increasing DPP and decreasing pulse widths. Di Martino et. al. [31] have presented the conceptual design for the so-called ALLS chamber, shown in Figure 4.3. It is a parallel-plate chamber filled with a noble gas such as argon. A noble gas prevents the attachment of electrons to the gas molecules and thus the creation of negative ions. Therefore, ion recombination is eliminated. Electron - positive ion recombination is possible but very minimal (less than 0.1%) due to the much larger electron velocity compared to ion velocity [32]. The space charge due to the positive ions can be minimized by decreasing the pressure of the gas, resulting in a reading accurate to within 1% of the initial charge created by the beam. The ALLS chamber has not been used experimentally yet. Another potential solution is the Razor Nano ionization chamber by IBA. It has a collecting volume of 0.003 cm^3. Cavallone et. al. [33] found that the Razor Nano chamber collection efficiency was 85% and 55% at 1 and 10 Gy DPP in comparison to 60% and 25% for the PTW Advanced Markus parallel-plate chamber (plate separation of 1 mm). Due to the very small electrode separation of the Razor Nano chamber, it is more sensitive to the pulse duration. This is due to the ion collection time being comparable to

Figure 4.3: Left: A conceptualization of the ALLS chamber design. Right: A prototype ALLS chamber. Reproduced under the Creative Commons CC-BY-NC-ND license from [31].

the pulse duration.

Numerical solutions

Numerical solutions aim to numerically solve the PDEs describing the ion and electron dynamics along with space charge. These PDEs are generally written as

$$\frac{\partial n_+}{\partial t} = R(x,y,z,t) - \alpha n_+ n_- - \beta n_+ n_e - \mu_+ \boldsymbol{\nabla} \cdot \left(n_+ \vec{\mathbf{E}}\right) \tag{4.26}$$

$$\frac{\partial n_-}{\partial t} = \gamma n_e - \alpha n_+ n_- + \mu_- \boldsymbol{\nabla} \cdot \left(n_- \vec{\mathbf{E}}\right) \tag{4.27}$$

$$\frac{\partial n_e}{\partial t} = R(x,y,z,t) - \gamma n_e - \beta n_+ n_e + \boldsymbol{\nabla} \cdot (n_e v_e) \tag{4.28}$$

$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = \frac{e}{\epsilon}\left(n_+ - n_- - n_e\right) \tag{4.29}$$

$$\vec{\mathbf{E}} = -\boldsymbol{\nabla} V \tag{4.30}$$

with the Thomson boundary conditions

$$n_+(0,t) = 0 \tag{4.31}$$

$$n_-(d,t) = 0 \tag{4.32}$$

$$n_e(d,t) = 0 \tag{4.33}$$

$$V(0,t) = V_0 \tag{4.34}$$

$$V(d,t) = 0. \tag{4.35}$$

In Eqs. 4.26-4.29, n_e is the electron concentration, γ is the electron attachment to neutral molecules, v_e is the electron velocity and β is the electron - positive ion recombination coefficient. As before, β is commonly taken to be equal to zero. Numerical schemes to solve these PDEs have been developed by Gotz et. al. [24], Kranzer et. al. [25, 26] and Paz-Martín et. al. [34]. In the work of Paz-Martín et. al., their proposed numerical scheme solves the same PDEs as shown in Eqs. 4.26-4.29 on a spatially-discretized mesh with time discretization. They also consider the diffusion of the free electrons and positive and negative ions but ultimately neglect these terms. This is due to the numerical error or "diffusion", caused by spatial discretization, being larger than the physical diffusion, for the implemented number of grid bins, 3000. They also investigate the influence of the pulse time structure on the collection efficiency. The effect is small, within 3%, for pulse widths typically seen in medical linacs, less than 5 μs. For longer pulse widths, over 100 μs, the variation of the collection efficiency over various time structures is around 20%. Christensen et. al. [35] have also developed a numerical scheme beginning with ionization tracks and consider both general and initial recombination. They do not consider Eq. 4.28 and the electron components in Eqs. 4.26, 4.27 and 4.29 but instead use the electron distribution found in Boag et. al. [21]. In comparison to the work that will be presented in Chapter 8, numerical solution schemes yield more accurate results for the charge collection efficiency and recombination correction factor at higher DPP and larger plate separations for parallel-

plate chambers. These numerical solutions however, have to be calculated for every chamber type, voltage and DPP combination which is time-consuming and not suitable for the clinical environment, where it is desirable to determine the recombination correction factor quickly.

References

[1] N. Esplen, M. S. Mendonca, and M. Bazalova-Carter, "Physics and biology of ultra-high dose-rate (flash) radiotherapy: A topical review," *Physics in Medicine & Biology*, vol. 65, no. 23, 23TR03, Dec. 2020.

[2] G. Jaffé, "Zur Theorie der Ionisation in Kolonnen," *Annalen der Physik*, vol. 347, no. 12, pp. 303–344, 1913.

[3] E. Kara-Michailova and D. E. Lea, "The interpretation of ionization measurements in gases at high pressures," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 36, no. 1, pp. 101–126, 1940.

[4] von H. Zanstra, "Ein kurzes verfahren zur bestimmung des sättigungsstromes nach der jaffé'schen theorie der kolonnenionisation," *Physica*, vol. 2, no. 1, pp. 817–824, 1935.

[5] P. B. Scott and J. R. Greening, "The Determination of Saturation Currents in Free-air Ionization Chambers by Extrapolation Methods," *Physics in Medicine & Biology*, vol. 8, no. 1, p. 51, Apr. 1963.

[6] *Absorbed Dose Determination in External Beam Radiotherapy* (Technical Reports Series 398). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2001.

[7] S. Rossomme *et al.*, "Ion recombination correction factor in scanned light-ion beams for absolute dose measurement using plane-parallel ionisation chambers," *Physics in Medicine & Biology*, vol. 62, no. 13, p. 5365, Jun. 2017.

[8] J. J. T. M. F.R.S., "XIX. On the theory of the conduction of electricity through gases by charged ions," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 47, no. 286, pp. 253–268, 1899.

[9] G. Mie, "Der elektrische Strom in ionisierter Luft in einem ebenen Kondensator," *Annalen der Physik*, vol. 318, no. 5, pp. 857–889, 1904.

[10] J. S. Townsend, *Electricity in gases*. Clarendon Pr, 1915.

[11] R. Rosen and E. P. George, "Ion distributions in plane and cylindrical chambers," *Physics in Medicine & Biology*, vol. 20, no. 6, p. 990, Nov. 1975.

[12] J. R. Greening, "Saturation Characteristics of Parallel-Plate Ionization Chambers," *Physics in Medicine & Biology*, vol. 9, no. 2, p. 143, Apr. 1964.

[13] W. Armstrong and P. Tate, "Accuracy of approximate solutions for currents in a plane parallel ion chamber," *Physics in Medicine & Biology*, vol. 10, no. 2, p. 229, Apr. 1965.

[14] J. W. Boag and T. Wilson, "The saturation curve at high ionization intensity," *British Journal of Applied Physics*, vol. 3, no. 7, p. 222, Jul. 1952.

[15] J. W. Boag, "Ionization measurements at very high intensities—Part I," *The British Journal of Radiology*, vol. 23, no. 274, pp. 601–611, 1950.

[16] P. R. Almond *et al.*, "AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon and electron beams," *Medical Physics*, vol. 26, no. 9, pp. 1847–1870, 1999.

[17] J. W. Boag and J. Currant, "Current collection and ionic recombination in small cylindrical ionization chambers exposed to pulsed radiation," *British Journal of Radiology*, vol. 53, no. 629, pp. 471–478, Jan. 2014.

[18] E. Hochhauser and O. A. Balk, "The influence of unattached electrons on the collection efficiency of ionisation chambers for the measurement of radiation pulses of high dose rate," *Physics in Medicine & Biology*, vol. 31, no. 3, p. 223, Mar. 1986.

[19] E. Hochhauser, O. A. Balk, H. Schneider, and W. Arnold, "The significance of the lifetime and collection time of free electrons for the recombination correction in the

ionometric dosimetry of pulsed radiation," *Journal of Physics D: Applied Physics*, vol. 27, no. 3, p. 431, 1994.

[20] A. F. Bielajew, "The effect of free electrons on ionization chamber saturation curves," *Medical Physics*, vol. 12, no. 2, pp. 197–200, 1985.

[21] J. W. Boag, E. Hochhäuser, and O. A. Balk, "The effect of free-electron collection on the recombination correction to ionization measurements of pulsed radiation," *Physics in Medicine & Biology*, vol. 41, no. 5, p. 885, May 1996.

[22] F. Di Martino, M. Giannelli, A. C. Traino, and M. Lazzeri, "Ion recombination correction for very high dose-per-pulse high-energy electron beams," *Medical Physics*, vol. 32, no. 7Part1, pp. 2204–2210, 2005.

[23] R. F. Laitano, A. S. Guerra, M. Pimpinella, C. Caporali, and A. Petrucci, "Charge collection efficiency in ionization chambers exposed to electron beams with high dose per pulse," *Physics in Medicine & Biology*, vol. 51, no. 24, p. 6419, Nov. 2006.

[24] M. Gotz, L. Karsch, and J. Pawelke, "A new model for volume recombination in plane-parallel chambers in pulsed fields of high dose-per-pulse," *Physics in Medicine & Biology*, vol. 62, no. 22, p. 8634, Nov. 2017.

[25] R. Kranzer *et al.*, "Ion collection efficiency of ionization chambers in ultra-high dose-per-pulse electron beams," *Medical Physics*, vol. 48, no. 2, pp. 819–830, 2021.

[26] R. Kranzer *et al.*, "Charge collection efficiency, underlying recombination mechanisms, and the role of electrode distance of vented ionization chambers under ultra-high dose-per-pulse conditions," *Physica Medica*, vol. 104, pp. 10–17, 2022.

[27] K. Petersson *et al.*, "High dose-per-pulse electron beam dosimetry — A model to correct for the ion recombination in the Advanced Markus ionization chamber," *Medical Physics*, vol. 44, no. 3, pp. 1157–1167, 2017.

[28] F. DeBlois, C. Zankowski, and E. B. Podgorsak, "Saturation current and collection efficiency for ionization chambers in pulsed beams," *Medical Physics*, vol. 27, no. 5, pp. 1146–1155, 2000.

[29] S. Rossomme *et al.*, "Correction of the measured current of a small-gap plane-parallel ionization chamber in proton beams in the presence of charge multiplication," *Zeitschrift für Medizinische Physik*, vol. 31, no. 2, pp. 192–202, 2021, Special Issue: Ion Beam Therapy, Part I.

[30] F. Gómez *et al.*, "Development of an ultra-thin parallel plate ionization chamber for dosimetry in FLASH radiotherapy," *Medical Physics*, vol. 49, no. 7, pp. 4705–4714, 2022.

[31] F. Di Martino *et al.*, "A new solution for UHDP and UHDR (Flash) measurements: Theory and conceptual design of ALLS chamber," *Physica Medica*, vol. 102, pp. 9–18, Oct. 2022.

[32] G. Boissonnat, J.-M. Fontbonne, J. Colin, A. Remadi, and S. Salvador, "Measurement of ion and electron drift velocity and electronic attachment in air for ionization chambers," *arXiv e-prints*, Sep. 2016.

[33] M. Cavallone *et al.*, "Determination of the ion collection efficiency of the razor nano chamber for ultra-high dose-rate electron beams," *Medical Physics*, vol. 49, no. 7, pp. 4731–4742, 2022.

[34] J. Paz-Martín *et al.*, "Numerical modeling of air-vented parallel plate ionization chambers for ultra-high dose rate applications," *Physica Medica*, vol. 103, pp. 147–156, 2022.

[35] J. B. Christensen, H. Tölli, and N. Bassler, "A general algorithm for calculation of recombination losses in ionization chambers exposed to ion beams," *Medical Physics*, vol. 43, no. 10, pp. 5484–5492, 2016.

5

Particle acceleration methods

In this chapter, various particle acceleration methods used in radiation therapy are presented. These are the radiofrequency (RF) linear accelerator, cyclotron, synchrotron and linear induction accelerator (LIA). They are introduced so as to compare them to the dielectric wall accelerator (DWA).

5.1 Radiofrequency accelerators

Clinical RF linear accelerators use cylindrical disk-loaded wave guides to establish a travelling or standing wave RF field. The description of a uniform cylindrical waveguide will be presented first and then the disk-loaded variant.

5.1.1 Infinite, uniform cylindrical waveguide

A uniform cylindrical waveguide consists of a hollow cylindrical tube with the walls made of a perfect conductor material, such as copper. The tube extends infinitely in the z direction. The inside of the waveguide is under vacuum. Following Appendix A.1, the dispersion relation of the waveguide is

$$k = \frac{1}{c}\sqrt{\omega^2 - c^2\left(\frac{x_{mn}}{a}\right)^2} = \frac{1}{c}\sqrt{\omega^2 - \omega_c^2}. \tag{5.1}$$

For field propagation to occur, $k > 0$ which implies $\omega > \omega_c$. ω_c can therefore be considered a cutoff frequency for the mn mode, below which the field does not propagate but decay exponentially over the length of the waveguide. The dispersion relation from Eq. 5.1 is shown in Figure 5.1. The equation defines a parabola with a vertex at the cutoff frequency. The asymptote is defined by $\omega = ck$. The phase velocity v_{ph} of the wave is given as

$$v_{ph} = \frac{\omega}{k} = \frac{c}{\sqrt{1 - (\frac{\omega_c}{\omega})^2}} = \tan\alpha_{ph}, \tag{5.2}$$

where α_{ph} is the angle made by the line joining the origin to the point P in Figure 5.1. The phase velocity is always greater than c. The group velocity v_{gr} of the wave is given as

$$v_{gr} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = c\sqrt{1 - \left(\frac{\omega_c}{\omega}\right)^2} = \tan\alpha_{gr}, \tag{5.3}$$

where α_{gr} is the angle of the tangent line to point P in Figure 5.1. The group velocity is always less than c. In order to accelerate particles in the z-direction, a non-zero longitudinal electric field component (E_z) is required. This requirement therefore dictates the use of transverse magnetic or TM_{mn} modes. The simplest TM mode is the TM_{01} mode, where $m = 0$ and $n = 1$. A figure of the TM_{01} mode is shown below in Figure 5.2.

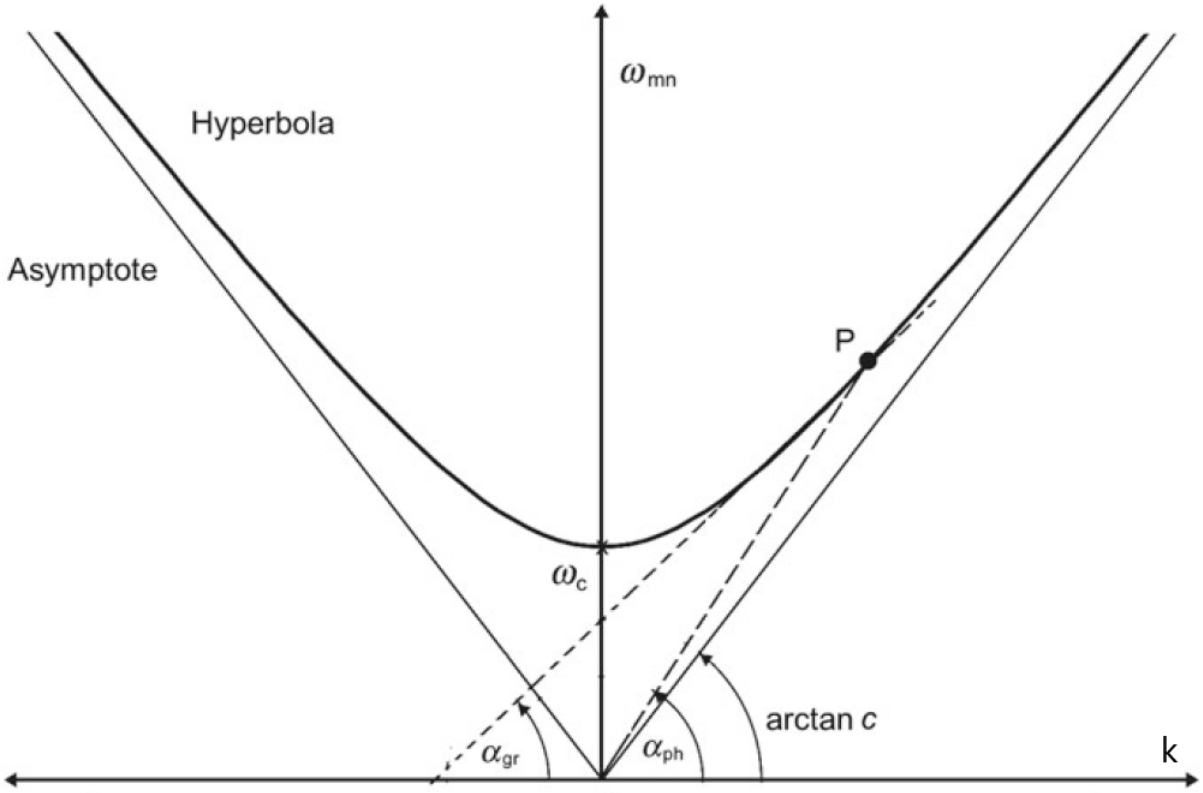


Figure 5.1: The dispersion relation for a hollow cylindrical waveguide. ω_{mn} is the frequency of the mn mode and ω_c is the cutoff frequency for the mn mode. The tangent of the angles α_{ph} and α_{gr} give the phase and group velocity, respectively. The asymptote is defined by $\omega = ck$. Reproduced with permission from [1].

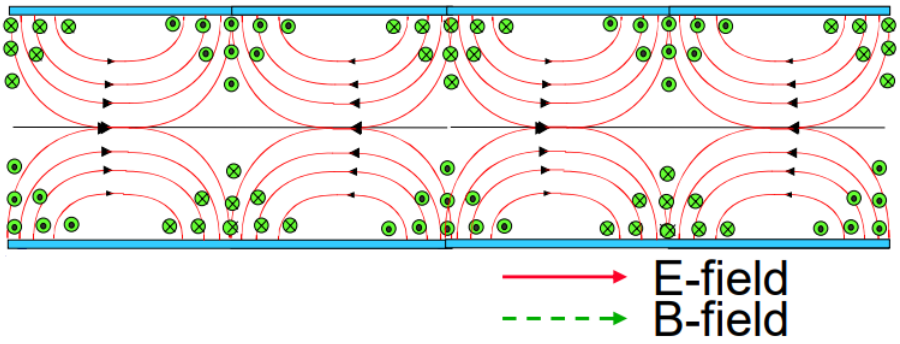


Figure 5.2: Electric and magnetic fields of the TM_{01} in a hollow cylindrical waveguide. Reproduced under the Creative Commons CC-BY 3.0 license from [2].

5.1.2 Issues with the infinite, hollow waveguide

What is the relation between the particle velocity v_{pr} and the wave phase velocity v_{ph}? It is useful to first picture a particle travelling at a velocity $v_{pr} \leq c$ inside the waveguide. If the

phase velocity of the wave is greater than the particle velocity, the accelerating part of the wave will eventually overtake the particle and the particle will begin to be decelerated. This is not optimal in terms of achieving maximum particle acceleration. This concept can be made more concrete by considering the frequency of the wave $\omega^{'}$, as seen from the particle's frame of reference (the derivation can be found in [1])

$$\omega^{'} = \gamma_{pr}\omega \left[\frac{v_{pr}}{v_{ph}} - 1\right], \tag{5.4}$$

where γ_{pr} is the Lorentz factor of the particle. In order for the particle to "see" an electric field with constant phase ($\omega^{'}$=0), i.e., always ever seeing the accelerating phase of the wave, $\omega^{'}$ must equal 0. This implies

$$v_{ph} = v_{pr}. \tag{5.5}$$

This is often known as the synchronous condition for RF acceleration. We have seen above in Eq. 5.2 that the phase velocity of an infinite, hollow cylindrical waveguide is always greater than c. As the particle can never travel faster than c, it is concluded that infinite, hollow cylindrical wave guides are not suitable for particle acceleration and must be modified in some way.

5.1.3 Accelerator parameters

In this section, some parameters that are widely used in accelerator physics are introduced.

Q

As input power fills the cavity, the accelerating gradient (E_0) and the stored energy U increase ($U \propto E_0^2$). The walls of the cavity, which in reality have a finite conductivity, will dissipate power P_w through ohmic losses ($P_w \propto E_0^2$). When the power dissipation in the walls equals the input power P_0, the maximum gradient is achieved. The Q of a cavity is a

measure of its ability to store energy. Mathematically, it is written as

$$Q = \frac{\omega U}{P}, \tag{5.6}$$

and is defined as the ratio of stored energy to power dissipation in 2π RF cycles.

Shunt impedance, R

As seen in the previous section, an input power yields a certain gradient (and therefore accelerating voltage V). Thus, we can define a shunt impedance of the cavity

$$R_{cav} = \frac{V^2}{P}. \tag{5.7}$$

From this definition, we can see that the shunt impedance is a measure of the cavity's ability to generate an accelerating gradient. Sometimes it is easier to work with the shunt impedance per unit length r_{cav} which is defined as

$$r_{cav} = \frac{E_0^2}{\frac{\mathrm{d}P}{\mathrm{d}z}}, \tag{5.8}$$

where E_0 is the longitudinal accelerating gradient and $\frac{\mathrm{d}P}{\mathrm{d}z}$ is the power lost per unit length.

R/Q

Taking the ratio of the shunt impedance and Q, we get

$$\frac{R}{Q} = \frac{V^2}{\omega U}. \tag{5.9}$$

From the above equation, we see that R/Q is independent of the quality of the cavity walls and is purely a geometric quantity. A similar quantity can be formed with the shunt impedance per unit length.

5.1.4 Disk-loaded cylindrical waveguide

As discussed in Section 5.1.2, a method must be found to decrease the phase velocity of the wave below c. One method of achieving this is by adding disk or irises at regular intervals within the waveguide. This method is shown in Figure 5.3. The inner radius of the disks is labelled as b and is less than a. Each pair of disks define a cavity of width d. The disks act

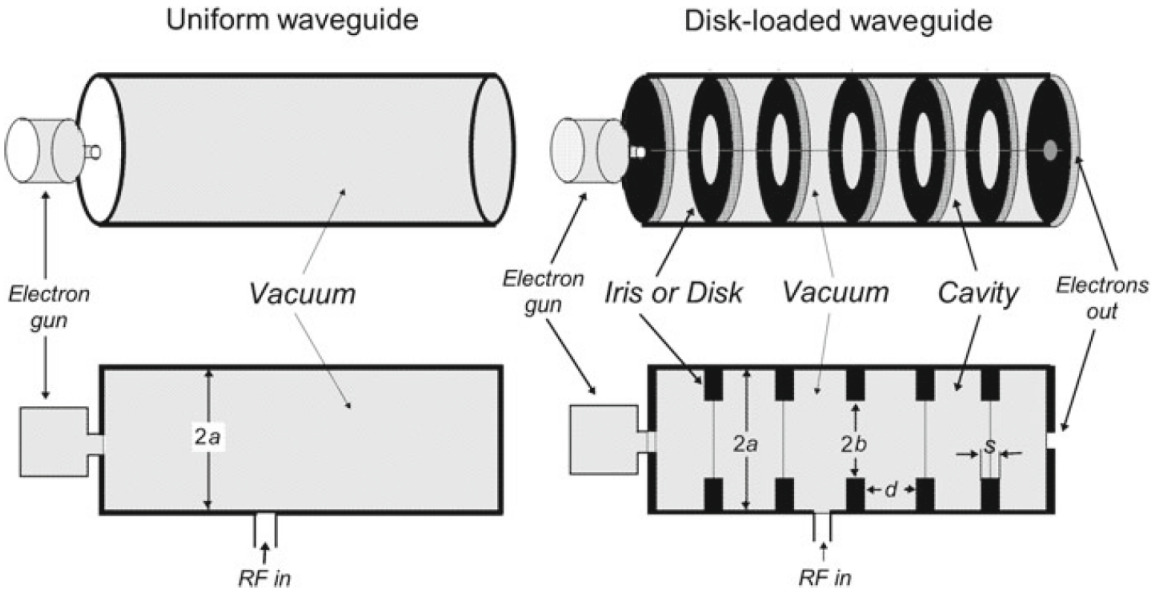


Figure 5.3: A comparison of a hollow and disk loaded cylindrical waveguide. The inner radius of the disks is labelled as b and is less than a. Each pair of disks define a cavity of width d. Reproduced with permission from [1].

as perturbations to the travelling wave in the waveguide. At each disk, the wave is partially reflected. The fraction of reflected wave to transmitted wave is dependent on the relative magnitude of the wavelength λ to the difference $a - b$. The electric field of the disk-loaded waveguide is no longer a single mode, but an infinite collection of so called "space harmonics"

$$E_z(r, \theta, z, t) = \sum_{n=-\infty}^{\infty} E_n(r, \theta) e^{i(k_{zn} z - \omega t)}, \tag{5.10}$$

where $k_{zn} = k_{z0} + 2\pi n/d$. The dispersion relation of the disk-loaded waveguide is shown in Figure 5.4.

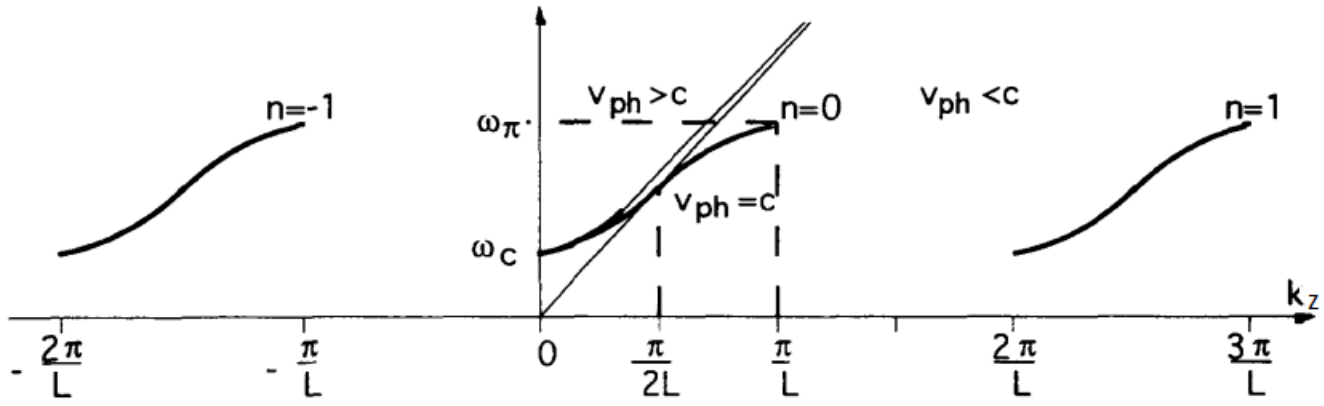


Figure 5.4: Dispersion relation of a disk loaded waveguide where n denotes a specific mode. The hollow waveguide is also shown with the thin line. Modes with $n < 0$ correspond to backwards travelling waves (negative phase velocity) where energy still travels in the positive direction (positive group velocity). Reproduced under the Creative Commons CC-BY 3.0 license from [3].

As can be seen in Figure 5.4, for a given RF frequency, there are an infinite set of space harmonics present in the waveguide. Each harmonic has the same group velocity but different phase velocities. The input RF power must populate the modes $n > 0$ and therefore the R/Q and shunt impedance of the $n = 0$ are reduced by a factor. With the help of Figure 5.5, we look at three regimes.

1. When $a \sim b$, i.e., $a - b << \lambda$ or $k << \frac{1}{a-b}$, the perturbation caused by the disks is small. Since in this regime k is small, the frequency ω is close to $(\omega_c)_1$. The dispersion relation of the disk loaded waveguide is close to that of the hollow waveguide in this regime.

2. As k increases, the reflected fraction increases and the dispersion relation of the disk loaded waveguide deviates from the hollow waveguide. Past point 3, the phase velocity is less than c.

3. When k approaches point 2, $\frac{\pi}{d}$ ($\lambda = 2d$), a standing wave is set-up in each cavity. In this regime, ω is independent of k and there is no energy propagation. Therefore, v_{gr}=0 and α_{gr}=0. This occurs at a second cutoff frequency $(\omega_c)_2$.

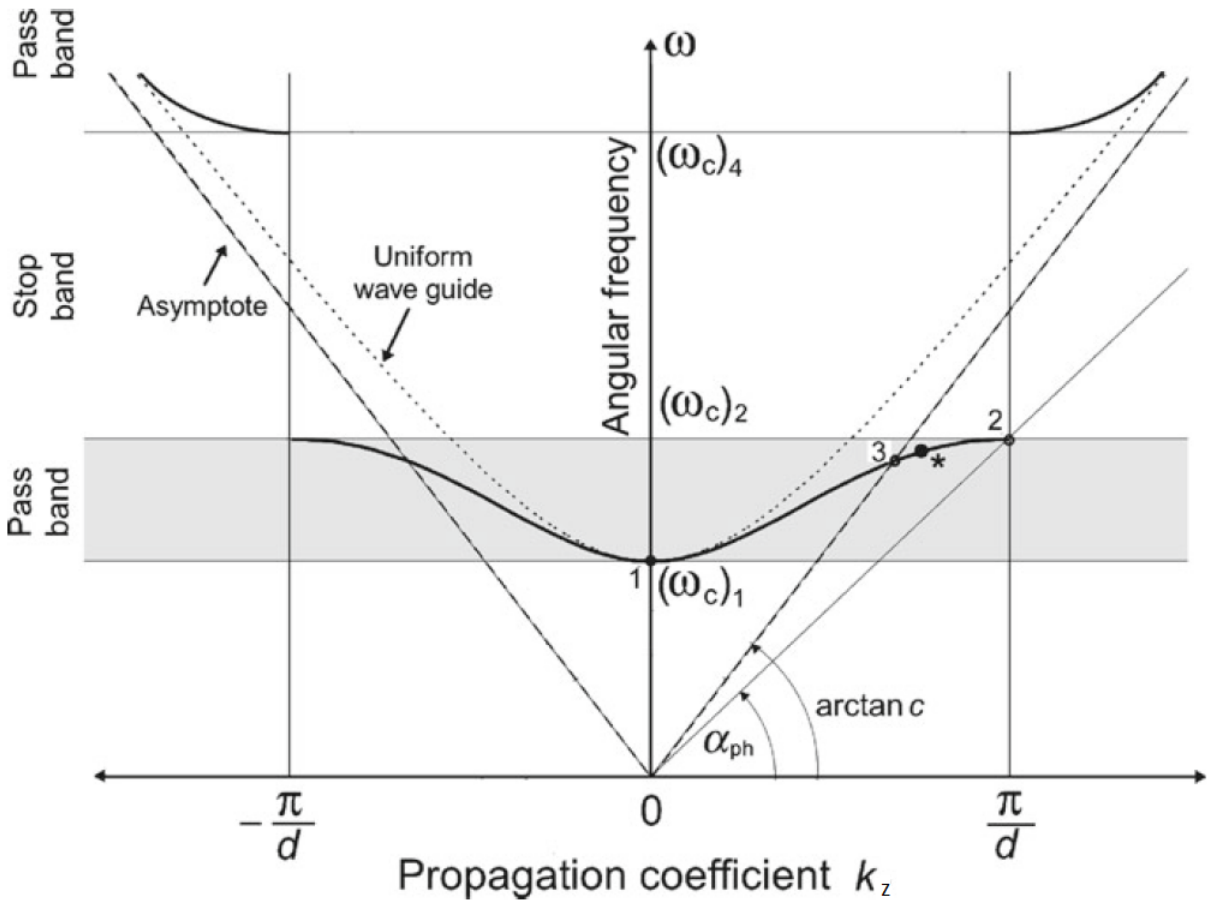


Figure 5.5: Dispersion relation of a disk loaded waveguide. Reproduced with permission from [1].

4. An ideal phase velocity is given by the point denoted with $*$, between point 2 and 3. This is because the phase velocity is less than c at this point but energy propagation is still non-zero.

5. There are bands of frequencies, known as stop bands, where a propagating wave cannot be established.

The above dispersion analysis works only for an infinite disk-loaded waveguide. In reality, there are only a finite number of cells N in the structure. The finite structure can then be though of as N coupled oscillators with N resonant frequencies. The modes are spaced uniformly between $k_z d = 0$ and $k_z d = \pi$. The result of this is that the dispersion relation, as shown in Figures 5.4 and 5.5, is not a continuous curve but a discrete set of points.

5.1.5 Travelling wave structure

The disk-loaded travelling waveguide is one of the two most common disk-loaded wave guides. An example of one is shown below in Figure 5.6. The RF power is fed into the waveguide by an input coupler, travels through the length of the waveguide, and exits through an output coupler. The impedance of the couplers are matched to the wave guide impedance in order to prevent reflections. The stored energy per unit length U' in the structure is related to

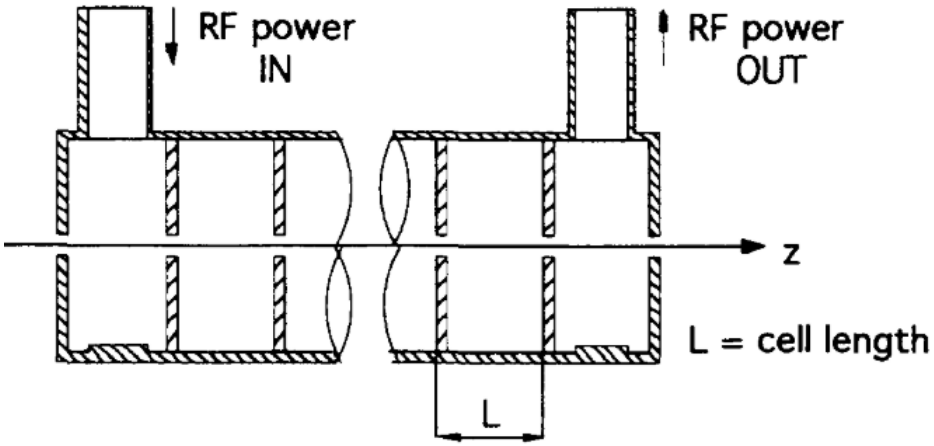


Figure 5.6: A diagram of a travelling disk-loaded wave guide. The input coupler feeds RF power into the wave guide and exits through the exit coupler. The couplers are impedance matched to the wave guide in order to prevent reflections. Reproduced under the Creative Commons CC-BY 3.0 license from [3].

the power loss in the walls per unit length p_w and the power flow $P(z)$ by the following differential equation

$$\frac{\mathrm{d}U'}{\mathrm{d}t} = -p_w - \frac{\mathrm{d}P}{\mathrm{d}z}. \tag{5.11}$$

In the steady state situation, the stored energy is constant and the power loss in the walls is equal in magnitude to the power flow at any point. With the use of the definition of Q (Eq. 5.6),

$$\frac{\mathrm{d}P}{\mathrm{d}z} = -p_w = \frac{-\omega U'}{Q} = \frac{-\omega}{Qv_{gr}}P = -2\alpha_0 P \tag{5.12}$$

where $P = v_{gr}U'$ and $\alpha_0 = \omega/2Qv_{gr}$ is an attenuation coefficient. The solution for the differential equation Eq. 5.12 is

$$P(z) = P_0 e^{-2\alpha_0 z} \tag{5.13}$$

where P_0 is the input RF power. From the definition of the shunt impedance per unit length (Eq. 5.8), the accelerating gradient is

$$E_0 = \sqrt{r_{cav}p_w} = \sqrt{r_{cav}\frac{\mathrm{d}P}{\mathrm{d}z}} = \sqrt{2r_{cav}\alpha_0 P_0 e^{-2\alpha_0 z}} \tag{5.14}$$

There are two common types of travelling wave structures; constant impedance and constant gradient. In the constant impedance approach, all the cells of the waveguide are identical, forcing a constant impedance and attenuation coefficient. In this case, the gradient decreases along the length of the waveguide, as given by Eq. 5.14. The second and far more common structure is the constant gradient structure.

Constant gradient structure

In order to keep the gradient constant, the product $r_{cav}p_w$ must be kept constant over the entire waveguide. The shunt impedance per unit length is typically held constant and therefore the power dissipation must also be held constant. Through Eq. 5.12, the ratio $\omega P/Qv_{gr}$ must be held constant. The Q is also typically held constant. If the power P decreases from front to back, v_{gr} must be slowed down from front to back. This is achieved by reducing the waveguide radius and the inner radius of the irises near the end of the waveguide. The gradient of a constant gradient structure is (derivation can be found in [3])

$$E_0 = \sqrt{r_{cav}\frac{P_0}{d}(1 - e^{-2\tau})}, \tag{5.15}$$

where $\tau = \alpha_0 d$ and the potential is

$$V = \sqrt{r_{cav} P_0 d(1 - e^{-2\tau})}. \tag{5.16}$$

5.1.6 Standing wave structure

Instead of adding an output coupler to the end of the waveguide, the end of the waveguide can be shorted through the addition of another wall. In this case, forward waves are reflected at the back end of the waveguide and travel backwards. The dispersion relation of the waveguide is modified, as can be seen in Figure 5.7. The advantage of using this standing

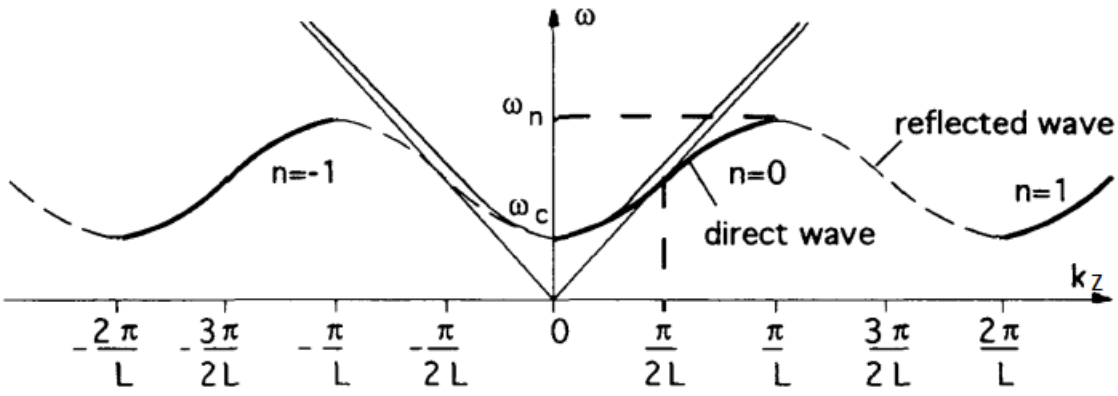


Figure 5.7: Dispersion relation of a disk loaded standing wave structure. Note that both forward and backward waves are now considered. Reproduced under the Creative Commons CC-BY 3.0 license from [3].

wave structure is that the reflected wave can also accelerate a particle. Unlike the travelling wave structure, the standing wave structure can accelerate particles in both wave directions. For a constant input power source, beginning at $t = 0$, the voltage across the structure is (derivation found in [4])

$$V = \sqrt{\frac{R}{Q}\omega t_c P_0 \frac{2\beta_c}{1+\beta_c}}\left(1 - e^{-t/t_c}\right), \tag{5.17}$$

where β_c is the the coupling coefficient which is ratio of the power emitted from the cavity to the power lost in the walls. The characteristic time is $t_c = 2Q_L/\omega$, where $Q_L = Q/(1+\beta_c)$ is the loaded (by the coupler) Q of the cavity.

5.1.7 Beam loading

When a beam is introduced into the waveguide, the effective gradient is the superposition of that from the input power (described in the previous sections) and that induced by the beam itself. Since the beam uses up power for acceleration, Eq. 5.12 is modified in the following manner

$$\frac{\mathrm{d}U'}{\mathrm{d}t} = -p_w - \frac{\mathrm{d}P}{\mathrm{d}z} - I_B E(z), \tag{5.18}$$

where I_B is the beam current. The constant gradient structure voltage then becomes (derivation can be found in [4])

$$V = \sqrt{r_{cav} P_0 d(1 - e^{-2\tau})} - \frac{r_{cav} d I_B}{2}\left(1 - \frac{2\tau e^{-2\tau}}{1 - e^{-2\tau}}\right) \tag{5.19}$$

and the standing wave structure voltage becomes (derivation can be found in [4])

$$V = \sqrt{\frac{R}{Q}\omega t_c P_0 \frac{2\beta_c}{1+\beta_c}}\left(1 - e^{-t/t_c}\right) - \frac{r_{cav} d I_B}{1+\beta_c}\left(1 - e^{-\frac{t-t_B}{t_c}}\right) \tag{5.20}$$

where t_B is the time the beam is introduced into the cavity. It can be seen that beam loading decreases the effective gradient. A power transfer efficiency can be defined as

$$\eta = \frac{V I_B}{P_0}. \tag{5.21}$$

5.1.8 Capture condition

In the previous sections, it was shown that disk loaded wave guides reduce the phase velocity below c so that particles can be accelerated. Despite this, particles enter the waveguide from the gun side with an initial velocity v_0 that is much smaller than c. To achieve $v_{pr} = v_{ph}$, there are two solutions:

1. Lower v_{ph} so that $v_{ph} \sim v_0$ and then gradually increase v_{ph} as the particle gains energy. This is called velocity modulation of the RF wave.

2. Provide a sufficiently large $(E_z)_0$ for the wave to capture the electron at the entrance.

Of the two solutions, the second is much simpler. It is the solution described in this section. In this section, we focus on electrons and make the simplifying assumption that the RF wave propagates through the waveguide with $v_{ph} = c$. The final result (derivation can be found in [1]) shows that the electric field amplitude must satisfy the following condition in order to capture the entering electron

$$(E_z)_0 \geq \frac{\pi m_e c^2}{\lambda e} \sqrt{\frac{1 - v_0/c}{1 + v_0/c}} \tag{5.22}$$

where λ is the wavelength of the RF field. The constant $\frac{\pi m_e c^2}{e}$ is about 1.605 MV for electrons. In an S band (2856 MHz) accelerator and initial electron energy of 20 keV (v_0/c=0.27), the minimum $(E_z)_0$ is 11.6 MV/m. This is achievable with current technology that is able to achieve gradients of up to 20 MV/m.

5.1.9 Limitations

Proton acceleration

It is suitable for this section to look back on the standing wave waveguide. Let us assume we are operating in the π-mode (E-field alternates sign in each cavity) for simplicity and the RF

has a period of time T and a phase velocity of $v_{ph} = c$. In order for the particle to undergo maximum energy gain from the electric field, it must travel from one cavity to another in a time $T/2$. The time for the particle to cross the one cavity is d/v_{pr}. Therefore, equating the two times, we have

$$\frac{T}{2} = \frac{d}{v_{pr}} \implies d = \frac{v_{pr}T}{2} = \frac{\beta cT}{2} = \frac{\beta\lambda}{2} \tag{5.23}$$

where β is v_{pr}/c. This relation is often called the synchronism condition. With respect to electrons, which travel near c (β=1) (except near the initial sections of the waveguide), the synchronism condition gives $d = \lambda/2$. The cavities are of constant length. Note this is equivalent to Eq. 5.5. Concerning protons, they travel at speeds lower than c ($\beta < 1$) and their speeds increase more slowly, compared to electrons. Therefore, according to the synchronism condition, each cavity must be increased in length as the proton gains more energy. These differences find their origins in the differences between the very large proton mass and very small electron mass. As such, a travelling waveguide structure is not suitable for linear proton acceleration. It is therefore necessary to use standing wave wave guides with cavity lengths satisfying Eq. 5.23.

Gradient

A higher electric field gradient within the waveguide allows for larger energy gain in each cavity. This in turn reduces the waveguide length and increases the maximum achievable particle energies. Higher gradients can be achieved by using high RF frequencies, as high as 30 GHz. At such large frequencies the waveguide structure becomes prone to RF breakdown. RF breakdown occurs when a large amount of energy is absorbed by the structure in a very short time. This causes melting and evaporation along with electron emission, X-ray radiation, acoustic waves and gas desorption (the release of trapped gas molecules in the structure) [5]. The physical mechanism of RF breakdown is still not fully understood but

the most likely starting point of the break down is field-induced electron emission or cracks in the micro structure of the wall.

5.2 Cyclotrons

Cyclotrons were first developed in the 1930's in order to accelerate protons and ions to a few MeV. A cyclotron consists of two, semi-circular electrodes that are under vacuum, called dees. The dees are under a uniform magnetic field of about 1 T that is established by large coils. A schematic of a cyclotron is shown in Figure 5.8. The particles in the dees travel with a constant speed and in a spiral orbit due to the magnetic field. Each time the particles cross the gap, they gain some kinetic energy because a RF potential around 150 keV is applied across the gap. The RF potential has a frequency between 10 and 30 MHz and is set to the gyro-frequency of the (non-relativistic) particles, qB/m, where q is the particle charge, B is the magnetic field strength and m is the particle mass. This gain in kinetic energy results in a larger radius of the orbit. Once the particles reach the desired kinetic energy, they are deflected out of the dee. Micro-bunches of particles exit the cyclotron, separated in time by the RF frequency. These micro-bunches are so close together that the beam can be considered as quasi-continuous. Due to this continuous beam, cyclotrons are able to achieve very high beam currents and therefore FLASH dose rates for proton therapy. Cyclotrons are compact because the acceleration always takes place in the same region. Cyclotrons are not compatible with relativistic speeds. In the relativistic regime, the particle gyro-frequency is modified by a factor of $1/\gamma$, where γ is the Lorentz factor from special relativity. Thus, the gyro-frequency depends on the particle velocity and eventually falls out of phase with the RF potential. This is why electrons can not be accelerated by cyclotrons, due to their small rest mass. Cyclotrons are unable to vary energy of the protons, which must be modified by external passive energy modulators.

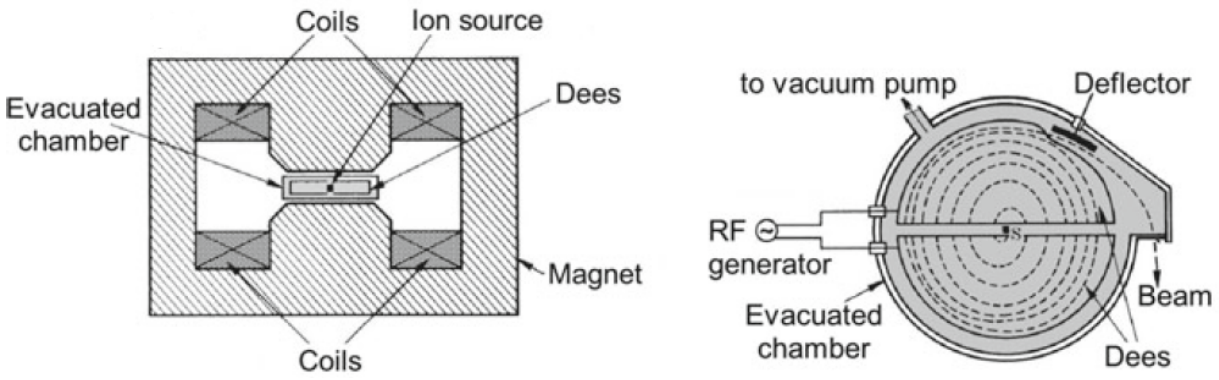


Figure 5.8: A schematic of a cyclotron. Reproduced with permission from [1].

5.3 Synchrotrons

A synchrotron consists of a circular beam pipe under vacuum. Coils cover a majority of the structure and establish a time-varying magnetic field. This is done so as to keep the particles travelling in a circular orbit, even with increasing kinetic energy. A schematic of a synchrotron is shown in Figure 5.9. The particles are first pre-accelerated and then injected into the main structure. One section of the synchrotron contains a RF cavity that provides the RF potential to accelerate the particles. The frequency of the RF potential is also time-varying to account for the decrease in the period of the particle path at relativistic speeds. Synchrotrons can be used to accelerate both electrons and protons. The RF frequency is on the order of 100 MHz for electrons and 1 MHz for protons. Synchrotrons are now even being used for light ion therapy. Synchrotrons produce pulsed beams and are able to vary the energy of the beam by varying the number of revolutions the particle takes through the synchrotron.

5.4 Induction accelerators

Induction accelerators are a class of accelerators that do not use RF power. The betatron, first developed in 1940, is a circular induction accelerator. A magnet that is fed RF current induces an electric field that accelerates electrons in a toroidal chamber, which is placed in

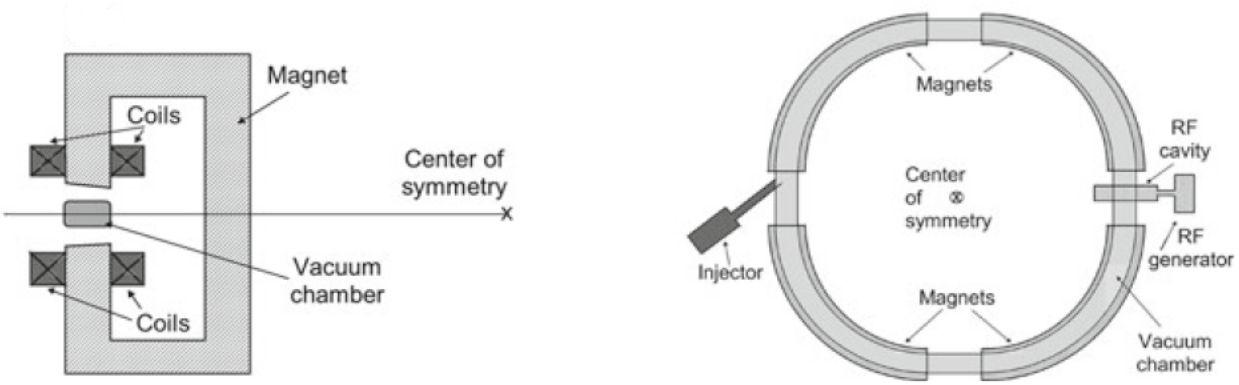


Figure 5.9: A schematic of a synchrotron. Reproduced with permission from [1].

the magnet cavity. The electrons are kept in the circular orbit by the magnetic field. With the introduction of linacs, betatrons ceased to be used for radiation therapy. The linear induction accelerator (LIA) came out of the need for much higher beam currents and charge per pulse than an RF accelerator could provide at the time, during the 1960's [6].

5.4.1 Design and method of acceleration

The induction accelerator is a non-resonant, low impedance structure. The idea behind the induction accelerator is to utilize a series of isolated induction modules. Each module contains an induction cell, which contains a magnetic core, and a pulsed power system, which drives the cell. The cell is composed of two concentric metallic (conductor) toroids, where the smaller one contains the toroidal magnetic core. The pulsed power system is connected to the cell through a transmission line. A standard design of an induction linac is shown below in Figures 5.10 and 5.11. There is a path for the voltage pulse to reach the accelerating gap from the pulsed power system. At the gap, the beam acts as a load and a current travels up the interior of the walls. Here, we see the maximum voltage is limited by what comes from the pulsed power system. There is also another short circuit path from the pulsed power system, through the central conductor of the transmission line, around the core and back through the exterior of the transmission line. Therefore, there is a "magnetizing current"

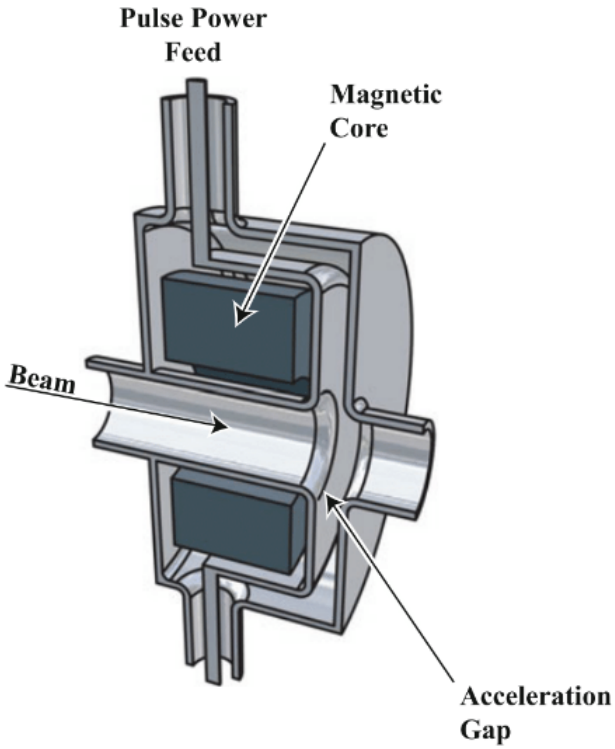


Figure 5.10: A 3D model of an induction cell. Reproduced with permission from [7]

that travels through this path. If the magnetic core presents a large enough impedance, the magnetizing current is small and most of the pulsed power is delivered to the accelerating gap. The purpose of the magnet is twofold. The first purpose is to act as a 1:1 transformer between the pulsed power system and the beam current. The voltage from the pulsed power system induces a magnetic flux through the core which induces an accelerating electric field gradient to be set up in the accelerating gap. This can be seen with Faraday's law, taking the red dotted line in Figure 5.11 as the integration contour

$$V_{pulse} = \oint \mathbf{E} \cdot \mathbf{dl} = -\frac{\mathrm{d}}{\mathrm{d}t} \int \mathbf{B} \cdot \mathbf{dS}. \tag{5.24}$$

The tangential E field along the metallic walls is 0 and therefore the contribution to the loop integral is 0. The contribution across the transmission line is 0 because the electric field flips polarity on either side of the central conductor. Therefore, the only remaining contribution is the accelerating gap region. The second purpose, as discussed in the previous paragraph,

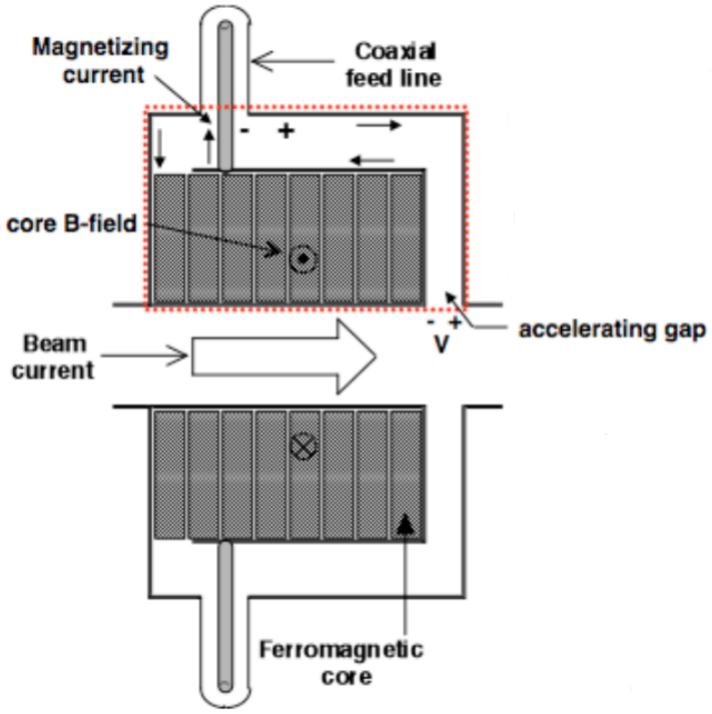


Figure 5.11: Cross section of an induction cell. The red, dotted line is a closed integration contour. Reproduced with permission from [8].

is to present a large impedance to the short circuit path around the core. This ensures most of the pulsed power is delivered to the accelerating gap. For this reason, it is crucial that the magnetic core remains unsaturated. Were the core to become saturated, its impedance would collapse and the magnetizing current would increase rapidly. As with the RF linac, beam loading causes the gradient in the gap to decrease, as some of the input power is used to accelerate the beam.

5.4.2 Magnetic core

The process of reaching saturation can be understood with the hysteresis loop presented in Figure 5.12. The core, prior to the voltage pulse, is reset to a remnant magnetic field value of $-B_r$. The voltage pulse increases the B-field to some value, with a max value being the saturation field at $+B_s$. The max "swing" is defined as $\Delta B = B_r + B_s$. Through Eq. 5.24, we see that

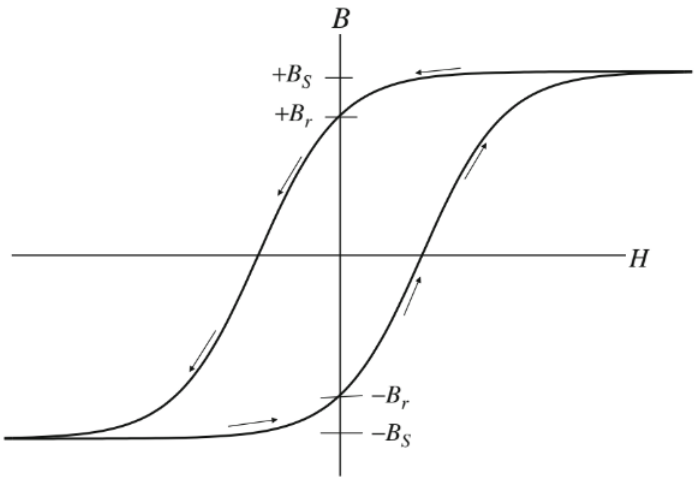


Figure 5.12: Hysteresis loop of a magnet. The saturation field is denoted as B_s. The magnetic core in an induction accelerator is initially reset to $-B_r$. As voltage pulse induces induction, the magnetic field swings upwards, towards a positive B field. Reproduced with permission from [7].

$$V_{pulse} = -\frac{\mathrm{d}}{\mathrm{d}t}\int \mathbf{B}\cdot\mathbf{dS} \implies V_{pulse}t_p \leq (\Delta B)S_c, \tag{5.25}$$

where t_p is the pulse time length and S_c is the core cross sectional area. When this "volt-second" product is reached, the core saturates. The core material and cross sectional area are therefore of crucial importance. To save on space, it seems an induction accelerator that is short longitudinally but long radially is ideal. The downside to this is that energy loss due to magnetization is greater (in comparison to a cell that is long longitudinally but short radially) since energy loss is proportional to magnet volume.

5.4.3 Limitations

The induction linac achieves lower gradients than an RF linac. Over the entire cell, the gradient is on the order of 1 MV/m. This is due to the large space occupied by the magnetic core. One must also be sure, for a large enough voltage, to use a magnetic material with a large enough "swing" to satisfy the volt-second constraint of Eq. 5.25.

5.4.4 "Coreless" or "Line-type" induction accelerator

Beginning in the late 1960's and into the 1970's, a form of induction linacs that did not use a magnetic core was being developed [9]. The removal of the magnetic core resulted in a superior cell form factor (long radially and short longitudinally), less energy loss (no magnetization) and higher accelerating gradients (5 MV/m) [10]. The magnetic core was removed so as to use shorter pulse widths, which enabled higher accelerating gradients. A single line contained a closing switch and is filled with a dielectric material. A vacuum insulator separates the line from the beam tube. These switches are generally solid state switches. Photo conductive semiconductor switches (PCSSs) have been a popular choice. An example of the line is shown below in Figure 5.13. The line and switch SW2 combination is used as a Blumlein pulsed power configuration. The central conductor is initially charged to a specific voltage (through SW1) with the solid-state switch (SW2) open. The gradient across one cell is net zero. When the switch SW2 closes, a voltage (and hence gradient) is eventually sustained along the cell gap for twice the transit time of the line. The electric field between the lines connected by SW2 flips polarity and there is a net accelerating gradient. There were several limitations of this design [11]. First, the switches SW2 and the lines are stressed with high voltage for a considerable amount of time, increasing the likelihood of dielectric breakdown. This limits the initial charging voltage. Additionally, at high voltages, there is electric field enhancement at the edges of the switch contacts which also causes breakdown [12].

5.5 Dielectric wall accelerator

The dielectric wall accelerator (DWA) is a non-resonant accelerator that builds upon the coreless induction accelerator. The operation of the DWA was already presented in Section 1.7. A diagram of the current conception of the DWA is shown in Figure 5.14. In this section, several advancements in semiconductor technology, radial waveguides and vacuum

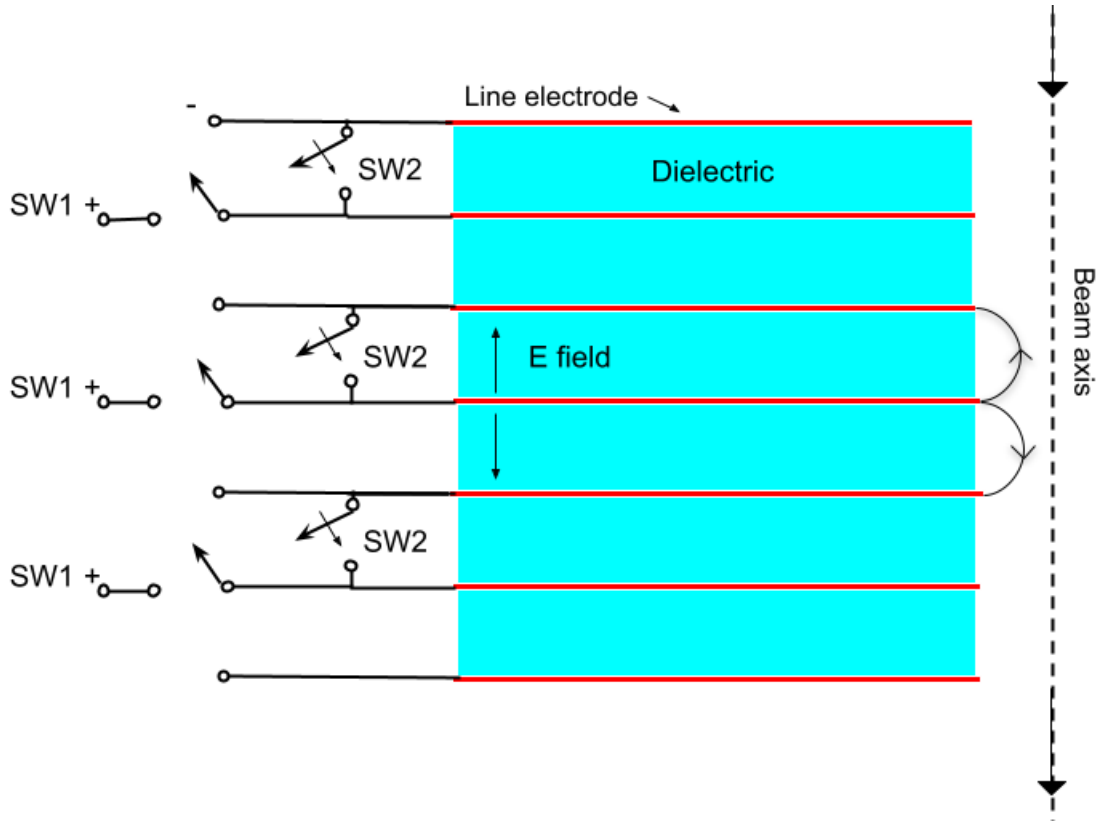


Figure 5.13: Design of a "coreless" induction accelerator. The vacuum insulator is not shown. The central conductor is initially charged to a specific voltage (through SW1) with the solid-state switch (SW2) open. The gradient across one cell is net zero. When the switch SW2 closes, a voltage (and hence gradient) is eventually sustained along cell gap for twice the transit time of the line. The electric field between the lines connected by SW2 flips polarity and there is a net accelerating gradient.

insulator technology that the DWA makes use of are highlighted. Through these technologies, the DWA is theoretically able to achieve very high gradients, approaching 100 MV/m. This high gradient enables a compact design, which, along with lightweight and relatively cheap materials (in relation to current clinical proton accelerators), makes the DWA an ideal candidate for a novel clinical proton accelerator.

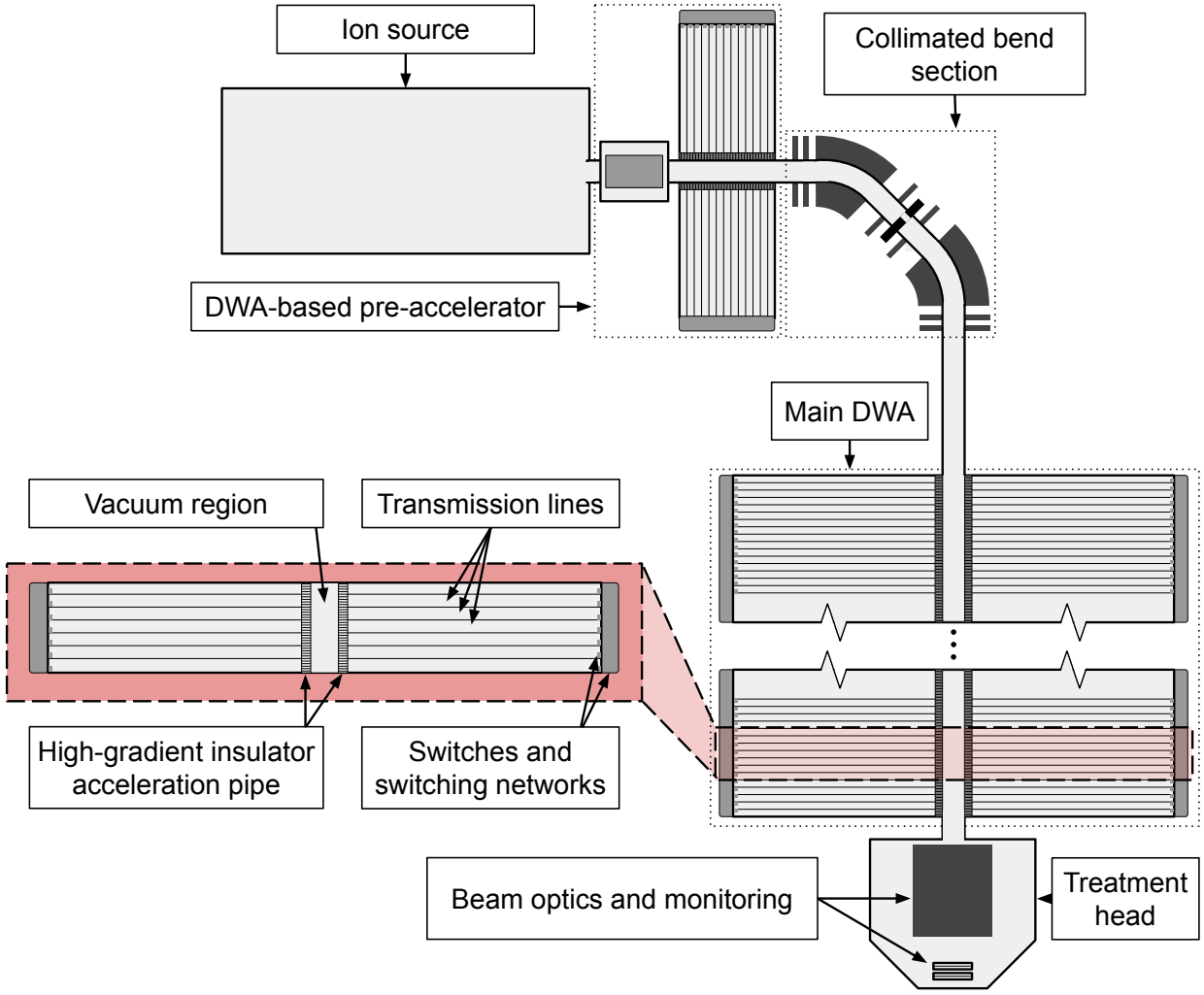


Figure 5.14: A diagram of the current DWA concept. Courtesy of Christopher Lund [13].

5.5.1 Enabling technologies

High gradient insulator

Standard vacuum insulators are prone to dielectric breakdown at high gradients. The dielectric breakdown occurs in the form of surface flashover. The underlying physics of surface flashover are still not well understood, but the current understanding is that it occurs when field-emitted electrons from the insulator continually bombard the surface of the insulator. Leopold et. al. [14] have shown that a periodic structure of layered insulators and dielectrics serves to deflect electrons from the surface through surface fields. This layered structure

serves as the basis of the high gradient insulator (HGI). Studies have also shown that for HGIs, the breakdown threshold is inversely proportional to the pulse timescale [8]. A gradient of 100 MV/m can be sustained without breakdown for a pulse width of 1 ns.

Radial waveguides

Radial waveguides are formed by two electrically conducting annuli with a dielectric in between them. The advantage of radial waveguides is that there are no magnetic field lines that close outside the structure. It is completely contained within the dielectric. Due to this, there is no magnetic coupling between adjacent lines and no induced current. The gradient over each line is therefore only dependent on the respective line and pulser. This is in contrast with the previously used strip line waveguides [8]. Radial waveguides have a low impedance that is dependent on the radius of the line. The equation for the impedance is

$$Z_{\text{radial}} = \sqrt{\frac{\mu}{\epsilon}} \frac{d}{2\pi r} \tag{5.26}$$

where d is the separation of the plates, ϵ is the dielectric constant, μ is the magnetic permeability and r is radial distance from the plate centre. Due to the small inner radius of the line, the acceleration gradient formed is larger with radial waveguides than strip line waveguides. The differences between the radial and strip line waveguides is shown in Figure 5.15. However, the radial dependence of the impedance does pose some issues. The radial dependence causes reflections and the pulse is distorted as it travels from the pulser end to the beam end. Despite this, preliminary work from Maher et. al. [15] has shown that an input pulse is amplified as it travels through a line. The radial dependence also implies that at the outer radius, where it connects to the HV pulser, the load presented to the HV pulser is quite small, limiting the magnitude of the output HV pulse. There is potentially an impedance mismatch between the waveguide at the inner radius and the beam, which causes reflections and distorts the gradient. In terms of the dielectric, there is a wide variety

of dielectrics with high breakdown thresholds, solid and able to be formed into the desired shape.

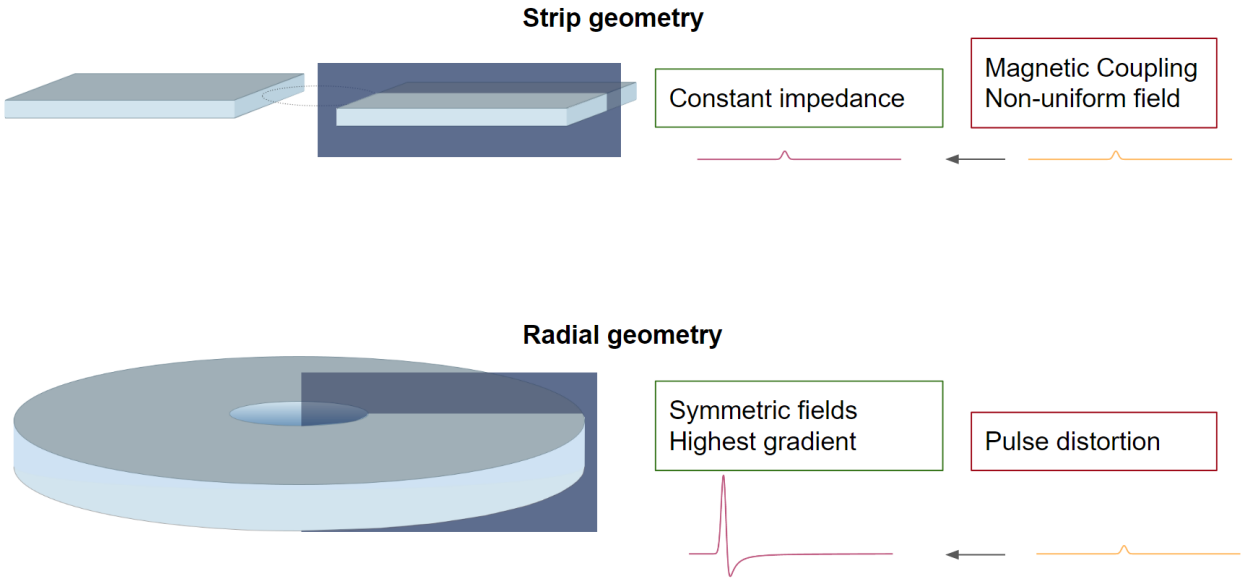


Figure 5.15: The differences between strip line waveguides and radial waveguides. Courtesy of Morgan Maher [16].

Drift step recovery diodes

Drift step recovery diodes (DSRDs) are Silicon (Si)-based opening switches, developed during the 1980's [17]. Their operation is discussed in detail in Section 7.4.1. Their potential application to coreless induction accelerators was investigated early on as well [18]. As discussed in Section 1.7, the HV pulse from the pulser is directly injected into the radial waveguide upon opening of the DSRD. This was also the case for the coreless induction accelerator incorporating DSRDs. As a consequence of it's opening nature, the DSRD itself and the waveguide are stressed with HV only for a short duration, the duration of the pulse width. This is an improvement upon the previous Blumlein design, where the lines are pre-charged and the switch and lines are stressed for a long duration. These design considerations were then carried over to the DWA during the 2000's [11].

References

[1] E. B. Podgorsak, *Radiation Physics for Medical Physicists*. Springer International Publishing, 2016.

[2] M. Vretenar, "Differences between electron and ion linacs," *CERN*, 2006, `http://cds.cern.ch/record/1005048`.

[3] M. Weiss, "Introduction to RF linear accelerators," 1994, `https://cds.cern.ch/record/261732`.

[4] P. Tenenbaum, "Fields in waveguides – a guide for pedestrians," Jun. 2003, `https://www.desy.de/~njwalker/uspas/coursemat/notes/unit_2_notes.pdf`.

[5] M. Ferrario and R. Assmann, "Advanced Accelerator Concepts," *arXiv e-prints*, arXiv:2103.10843, Mar. 2021.

[6] N. C. Christofilos, R. E. Hester, W. A. S. Lamb, D. D. Reagan, W. A. Sherwood, and R. E. Wright, "High Current Linear Induction Accelerator for Electrons," *Review of Scientific Instruments*, vol. 35, no. 7, pp. 886–890, Jul. 1964.

[7] R. J. Briggs, "Principles of induction accelerators," in *Induction Accelerators*, K. Takayama and R. J. Briggs, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 23–37.

[8] G. J. Caporaso, Y.-J. Chen, and S. E. Sampayan, "The Dielectric Wall Accelerator," *Reviews of Accelerator Science and Technology*, vol. 2, no. 1, 2009.

[9] A. Pavlovskii, A. Gerasimov, D. Zenkov, V. Bosamykin, A. Klement'ev, and V. Tananakin, "An iron-free linear induction accelerator," *Soviet Atomic Energy*, vol. 28, no. 5, Jan. 1970.

[10] A. Krasnykh, "Analysis of modulator schemes for nanosecond induction linacs," in *Proc. of the Meeting on Problems at Collective Method Acceleration*, 1982, p. 136.

[11] F. Arntz, A. Kardo-Sysoev, and A. Krasnykh, "Slim, short-pulse technology for high gradient induction accelerators," *Particle Accelerators*, Dec. 2008.

[12] G. J. Caporaso *et al.*, "Status Of The Dielectric Wall Accelerator For Proton Therapy," *AIP Conference Proceedings*, vol. 1336, no. 1, pp. 369–373, Jun. 2011.

[13] C. Lund *et al.* "Towards accessible proton radiotherapy: Beam optics design for a compact medical proton accelerator," ePoster presentation, 64th Annual Meeting & Exhibition of the AAPM, Washington, DC, USA. (Jul. 2022).

[14] J. G. Leopold, U. Dai, Y. Finkelstein, and E. Weissman, "Optimizing the performance of flat-surface, high-gradient vacuum insulators," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 12, pp. 530–536, 2005.

[15] M. Maher, C. Lund, J. Bancheri, D. Cooke, and J. Seuntjens, "An exploration of annular parallel plate waveguides for electric field generation in a dielectric wall accelerator," *Radiotherapy and Oncology*, vol. 186, S117, Sep. 2023.

[16] Morgan Maher. "Waveguides for a dielectric wall accelerator," Unit presentation, Medical Physics Unit, McGill University. (May 5, 2023).

[17] I. Grekhov, V. Efanov, A. Kardo-Sysoev, and S. Shenderey, "Power drift step recovery diodes (DSRD)," *Solid-State Electronics*, vol. 28, no. 6, pp. 597–599, 1985.

[18] A. Krasnykh, "High gradient linear collider," in *Proc. of the 11th All Union Meeting on Charge Particle Accelerators*, vol. 2, 1988, p. 103.

6

Semiconductor Modelling

The drift-step recovery diodes (DSRDs) are crucial to the DWA operation and one of the main objects under investigation in this thesis. As they are semiconductor devices, some semiconductor physics is presented to further understand their operation. The semiconductor simulation software, Sentaurus TCAD by Synopsys®, used to simulate the DSRDs in Chapter 7 utilizes these concepts.

6.1 Semiconductor physics

Semiconductors are a class of solids that fall between conductors and insulators. They have an electrical conductivity that falls between an insulator and a conductor (such as a metal). The unique organization of the electronic energy states gives rise to this behaviour. In semiconductors, the electronic energy states are organized in such a way that two distinct energy

bands form, the valence band and the conduction band. The formation of these band gaps is due to the small inter-atomic distances. The overlap of individual atomic electronic orbitals raise some electronic energy levels and lower others until the valence and conduction bands form with a band gap between them. For common semiconductors such as Silicon (Si), the band gap energy is 1.12 eV [1]. The energy states in the conduction band are mostly empty. For a semiconductor, the band gap is small enough that thermal excitations or an external perturbation, such as an applied electric field, can excite electrons near the top of the valence band to cross the band gap and enter the conduction band, where they behave as free particles with an effective mass m_n. The absence of an electron at the top of the valence band can be modelled as a positively charged free particle, known as a hole, with an effective mass m_p. In an intrinsic semiconductor, under thermal equilibrium, the hole concentration in the valence band equals the electron concentration in the conduction band. In an intrinsic semiconductor, the dominant mechanism of conduction is thermal excitations. A semiconductor with impurities is known as an extrinsic semiconductor. N-dopants or donors, such as phosphorus, introduce additional energy states just below the conduction band from which electrons can easily be excited to the conduction band. P-dopants or acceptors, such as Boron, introduce additional energy states just above the conduction band from which holes can easily be excited to the valence band. There are a wide variety of doping profiles used in semiconductors that enable various different behaviours under applied electric fields (voltages). Such examples are PN junctions, PIN diodes and MOSFETS.

The transport of the electrons and the holes in a semiconductor at constant temperature T is modelled through the drift-diffusion equations, the continuity equation and Gauss' law.

The drift-diffusion equations describe the electron and hole current density, $\vec{\mathbf{J_n}}$ and $\vec{\mathbf{J_p}}$, as

$$\vec{\mathbf{J_n}} = \mu_n \left(n \nabla E_C - \frac{3}{2} n k T \nabla \ln m_n \right) + D_n \left(\nabla n - n \nabla \ln \gamma_n \right), \tag{6.1}$$

$$\vec{\mathbf{J_p}} = \mu_p \left(p \nabla E_V + \frac{3}{2} p k T \nabla \ln m_p \right) - D_p \left(\nabla p - p \nabla \ln \gamma_p \right), \tag{6.2}$$

where n is the electron concentration, p is the hole concentration, μ_n and μ_p are the electron and hole mobilities, E_C is the energy of the bottom of the conduction band, E_V is the energy of the top of the valence band, D_n and D_p are the electron and hole diffusion coefficient, γ_n and γ_p are factors that account for Fermi statistics, k_B is the Boltzmann constant and ∇ is the spatial gradient operator. The continuity equation Eq. 6.3 for electrons and Eq. 6.4 for holes ensures the conservation of charge of each species

$$\nabla \cdot \vec{\mathbf{J_n}} = -e \left(R_{net,n} - G_{net,n} \right) - e \frac{\partial n}{\partial t}, \tag{6.3}$$

$$-\nabla \cdot \vec{\mathbf{J_p}} = -e \left(R_{net,p} - G_{net,p} \right) - e \frac{\partial p}{\partial t}, \tag{6.4}$$

where $e > 0$ is the electron charge, R_{net} is the net carrier recombination rate and G_{net} is the net carrier generation rate. Gauss' law is also used to solve for the electric field $\vec{\mathbf{E}}$ in the semiconductor

$$\nabla \cdot \vec{\mathbf{E}} = e \left(p - n + N_D - N_A \right), \tag{6.5}$$

where N_D is the ionized donor concentration and N_A is the ionized acceptor concentration. Additionally, boundary conditions must be specified at the semiconductor boundaries. In the 2D models used in Chapter 7, the left and right boundaries were specified to be Ohmic contacts.

6.1.1 Generation and recombination models

The models used for the generation and recombination terms in Eqs. 6.3 and 6.4 for the simulations carried out in Chapter 7 are described in this section. The respective equations can be found in [2].

Shockley-Read-Hall recombination

Shockley-Read-Hall (SRH) recombination occurs due to the presence of energy levels deep in the band gap. These levels arise due to defects in the semiconductor crystal structure. Thus, these energy levels are also known as trap or defect levels. It is the most dominant recombination mechanism in Si. In a two-step process, an electron can fall into a defect level from the conduction band and then into the valence band. This is as if a hole leaves the valence band to the same defect level and recombines with the trapped electron. The released energy excites a vibrational mode of the crystal. The SRH recombination rate depends on the electron and hole lifetimes. These lifetimes depend on the local dopant concentration. At high dopant concentrations, the lifetimes are drastically reduced due to the increasing amount of defects.

Auger recombination

An electron can also fall from the conduction band directly to the valence band. The energy released by the recombination of the electron and hole is transferred to a third electron or hole, exciting it to a higher energy state without crossing the band gap. This third particle losses its energy through collisions with the crystal lattice.

Avalanche generation

At high enough kinetic energies, an electron or hole is able to ionize an atom through a collision, creating an electron-hole pair. This is known as impact ionization. If the electric field is sufficiently strong, these secondary electrons and holes are accelerated to high kinetic energies and undergo their own collisions, creating additional electron-hole pairs. If this chain reaction is able to be sustained, the number of pairs increases exponentially. This chain reaction is known as avalanche generation. In the simulations performed in Chapter 7, the avalanche generation was modelled with the Okuto-Crowell model [3]. In some cases, the presence of avalanche generation results in the breakdown (rapid rise in conductivity) of the device.

6.1.2 Mobility models

The mobilities found in Eqs. 6.1 and 6.2 are not constant under various conditions such as doping, high electric fields and carrier-carrier scattering.

Doping dependence

As electrons and holes traverse the semiconductor, they are scattered by ionized dopant atoms. This leads to the reduction of the mobilities at high doping concentration. In the simulations performed in Chapter 7, the Masetti model is used [4].

Carrier-carrier scattering

At large electron and hole concentrations, there is the possibility for them to scatter off of one another. Thus, their mobilities will be reduced. The Conwell–Weisskopf model describes the effect of carrier-carrier scattering on the mobility [5].

High electric field saturation

Under low electric field strengths, the (drift) velocities of the carriers are equal to μE. As the electric field strength increases, the velocities cease to be proportional to the field strength and saturate to a constant value. The Canali model is used in the simulations in Chapter 7 [6].

6.1.3 Wide-band gap semiconductors

Over the past 2 decades, a class of semiconductors known as wide-band gap semiconductors have gained significant interest. Common wide-band gap semiconductors are Silicon Carbide (4H-SiC) and Gallium Nitiride (GaN). The band gap energy of these semiconductors is between 3-4 eV. As a result of the large band gap, wide-band gap semiconductors have a very high breakdown threshold and electron saturation velocity. For example, the breakdown threshold for 4H-SiC is ten times that of Si and the saturation velocity of electrons is two times that in Si. Wide-band gap semiconductors can therefore withstand very strong electric fields and temperatures without suffering breakdown. This enables the development of high power and high frequency devices. The size of the devices can also be reduced due to the large breakdown threshold. Due to these features, wide-band gap semiconductors are of interest to the DWA HV pulsers. In order to increase the accelerating gradient, a larger HV pulse is required to be generated by the pulser and therefore requires a switch that can withstand higher voltages.

One effect that pertains to wide-band gap semiconductors that is not found in Si at conventional temperatures (such as room temperature) is incomplete ionization. As discussed in Section 6.1, doping introduces impurity states in the band gap region. For Si, these impurity states are so close to the conduction or valence band (depending on the dopant type), that at room temperature, the thermal energy is sufficient enough for all of the dopants to

be ionized and the carrier concentration is equal to that of the doping concentration. In wide-band gap semiconductors, this is not the case because the energy difference between the impurity levels and the conduction or valence band is large in comparison to the thermal energy. In this case, not all the dopants are ionized. This is known as incomplete ionization. The degree of dopant ionization decreases with decreasing temperature and increasing doping concentration.

6.2 Sentaurus TCAD

The simulations performed in Chapter 7 were done with Sentaurus Technology Computer Aided Design (TCAD), by Synopsys®. In the field of electronics, TCAD refers to computer design software that can numerically model semiconductor device fabrication processes and the electrical, thermal, mechanical and optical properties of the device. Numerical modelling is extremely beneficial because physical modelling of semiconductor device behaviour under various conditions is often impossible in the context of modern semiconductor devices which are highly complex. Sentaurus TCAD provides several simulation suites to achieve the device modelling. The Sentaurus structure editor (`sde`) enables the definition of device geometries, doping profiles and electrical contacts. In this step, the mesh is also defined. The mesh divides the device regions into a finite number of discrete elements. The Sentaurus device (`sdevice`) suite defines the physical models (drift-diffusion, Gauss' law and continuity equation in our case), the materials of each region of the device and any parameters that pertain to the numerical solver. The type of simulation, such as steady-state or time transient is also defined in this step. Here, desired solution variables, such as electron density, are defined so that they are saved and output at the end of the simulation. The actual device simulation is performed by `sdevice`.

As mentioned above, the mesh divides the device region into a finite number of discrete

elements. `Sdevice` then solves the equations representing the physical models in each of these elements. This scheme for numerically solving partial differential equations is known as the Finite Element Method (FEM). A detailed theoretical treatment of the FEM can be found in [7].

References

[1] S. Sze and K. K. Ng, "Appendix G Properties of Si and GaAs," in *Physics of Semiconductor Devices*. John Wiley & Sons, Ltd, 2006, pp. 790–790.

[2] *Sentaurus tm device user guide*, English, version Version 2023.09, Synopsys, Sep. 2023, 1779 pp., September 2023.

[3] Y. Okuto and C. Crowell, "Threshold energy effect on avalanche breakdown voltage in semiconductor junctions," *Solid-State Electronics*, vol. 18, no. 2, pp. 161–168, 1975.

[4] G. Masetti, M. Severi, and S. Solmi, "Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon," *IEEE Transactions on Electron Devices*, vol. 30, no. 7, pp. 764–769, 1983.

[5] E. Conwell and V. F. Weisskopf, "Theory of impurity scattering in semiconductors," *Phys. Rev.*, vol. 77, pp. 388–390, 3 Feb. 1950.

[6] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," *IEEE Transactions on Electron Devices*, vol. 22, no. 11, pp. 1045–1047, 1975.

[7] O. Zienkiewicz, R. Taylor, and J. Zhu, *The Finite Element Method: Its Basis and Fundamentals*. Elsevier Science, 2005.

7

Evaluation of DSRD-based pulsers for a Dielectric Wall Accelerator

Julien Bancheri, Andrew Currell, Anatoly Krasnykh, Morgan Maher, Christopher Lund, Alaina Giang Bui and Jan Seuntjens

Article published in: Bancheri J, et. al. Evaluation of DSRD-based pulsers for a Dielectric Wall Accelerator. *NIM A*. 2024; 169935. https://doi.org/10.1016/j.nima.2024.169935

7.1 Preface

The DWA and its historical development have been discussed in the previous chapter. The DWA consists of several different components. Each component must be studied in order to understand how it functions and determine the design and mode of operation that yields the desired accelerating gradients. The accelerating gradients are formed from HV pulses that are initially generated by the pulsers that include a semiconductor switch. The DSRD has been identified as a suitable switch due to its nanosecond scale operation and high breakdown threshold. Previous studies have all used DSRDs with different doping profiles and materials. This makes it difficult to establish which pulser and DSRD combination yields optimal pulses. This necessitates (1) the direct comparison of pulsers while keeping the DSRD constant and (2) the study of how various DSRD profiles and materials influence the output pulse, given a constant pulser. This chapter presents the work undertaken to address these points. A magnetic saturation transformer (MST)-DSRD-based pulser and a MOSFET-DSRD-based multi-module pulser were built and tested. The magnitude, rise time and pulse width of the output pulses generated by the pulsers are observed. For the MOSFET-DSRD-based pulser, the output of a multi-module setup has not yet been observed. The simulation of the two pulsers with silicon carbide (SiC) DSRDs are also performed, as SiC DSRDs have yet to be incorporated with the MOSFET-DSRD-based pulser.

7.2 Abstract

Silicon (Si) drift step recovery diodes (DSRD) are opening switches that form the basis for nanosecond-scale high voltage (HV) pulsers. These pulsers are crucial to the operation of a dielectric wall accelerator (DWA), a compact particle accelerator design for proton beam radiation therapy. For the DWA to operate properly, the pulser must generate HV pulses with fast rise times and widths on the order of 1 ns at an operating frequency above 1 kHz. Several pulser designs are available that can optimally drive DSRDs. The purpose

of this study is to investigate the capabilities of two DSRD-based pulsers in the context of the DWA and the required pulse parameters. The first pulser is based on a magnetic saturation transformer (MST) and a DRSD. The second pulser considered is a multi-module MOSFET-DSRD-based pulser. Additionally, Sentaurus TCAD simulations are used to study the influence of the Si DSRD doping profile on the output pulse shape. The MST-DSRD-based pulser is able to generate higher amplitude pulses than the MOSFET-DSRD-based pulser with a single module (9645 V and 1807 V, respectively). However, the multi-module MOSFET-DSRD-based pulser achieves the maximum pulse amplitude (10.9 kV) compared to MST-DSRD-based pulser (9645 V). The MOSFET-DSRD-based pulser also generates pulses with shorter pulse widths compared to the MST-DSRD-based pulser (minimum of 3 ns vs. minimum of 8.36 ns, respectively). With no compression stage, the MOSFET-DSRD-based pulser (with more than one module) generates pulses with consistently faster rise times ($\leq$ 3.06 ns) compared to the MST-DSRD-based pulser ($\leq$ 4.31 ns). With a compression stage, the rise times were comparable between the two pulsers (MOSFET-DSRD: 1.66 ns and MST-DSRD: 1.48 ns). Both pulsers are able to operate at 1 kHz and the MOSFET-DSRD-based pulser up to 10 kHz. Parasitic capacitance and coupling between modules of the MOSFET-DSRD-based pulser affect the pulse through their influence on the forward and reverse pumping duration. Simulations show that the p region of the DSRD should be larger than the n doped region. Simulations including a parasitic capacitance, modelling the electrical connections to the DSRD, reduce the pulse amplitude. Simulations of the pulsers incorporating SiC DSRDs show a reduction in the pulse amplitude and a widening of the pulse width in comparison to the Si DSRDs.

7.3 Introduction

Applications of nanosecond, high voltage (HV) pulsed power technology are found in laser pulse formation, water treatment, food processing, microorganism sterilization [1] and particle acceleration [2]. An enabling technology of these applications has been the development

of fast, HV semiconductor switches. The operation of these switches is based on nanosecond-scale ionization and plasma processes. Through the integration of these switches into HV pulsers, stored inductive energy is able to be quickly commuted to a load, resulting in a HV pulse [3]. One such semiconductor switch is the drift step recovery diode (DSRD). The high power DSRD was designed in the 1980s in an effort to overcome the low blocking voltages and slow rise times of standard step recovery diodes (SRDs) [4]. Since then, extensive research has been conducted into understanding the dynamics of the DSRD and determining optimal driving conditions [5–10]. A DSRD is an opening (or turn-off) switch that generally has a p^+-p-n-n^+ V-type doping profile and is most commonly made of silicon (Si). The high breakdown threshold of Si DSRDs can be attributed to their doping profile. The high field saturation velocity of electrons in Si (2×10^7 cm/s) enable their fast turn-off capabilities. Individual DSRDs can also be stacked in series to increase the breakdown threshold. Silicon carbide (4H-SiC) DSRDs have also been investigated [11–14]. Due to 4H-SiC having a breakdown threshold 10 times that of Si, smaller DSRDs can conceivably be built which do not experience breakdown under very high voltage. Additionally, the electron saturation drift velocity in 4H-SiC is two times that in Si, enabling a faster DSRD operation.

The application of interest of pulsed power technology to this work is in the design and operation of particle accelerators. The most common example is the linear induction accelerator (LIA), which was first developed in order to deal with very high pulsed beam currents that radiofrequency (RF) accelerators (e.g. cyclotrons) could not handle without instabilities [15]. Since the 1950s, particle accelerators have been used for radiotherapy, which is the treatment of cancer with radiation. Conventional medical linear accelerators are RF linear accelerators that accelerate electrons to produce electron or photon beams. Over the past three decades, proton radiotherapy has gained significant interest. Proton beams are able to be delivered more conformally and precisely to tumour volumes compared to photon and electron beams [16]. Additionally, ultra-high dose rate radiotherapy ($>$ 40 Gy/s), also known as FLASH radiotherapy, has recently gained interest due to evidence indicating superior normal tissue

sparing [17, 18]. Due to the technology behind clinical proton accelerators, proton therapy is well suited for FLASH RT. Modern medical proton accelerators, which are cyclotron or synchrotron-based, and the facilities required to house them are both expensive and large in size [19]. These factors limit the global availability and accessibility of proton radiotherapy [20]. The Dielectric Wall Accelerator (DWA) is a compact accelerator concept that aims to address these issues and provide affordable high dose rate proton radiotherapy. The DWA design emerged from the "coreless" or "line-type" induction accelerator design, which itself emerged from the LIA design [21–23]. The DWA consists of individual modules stacked on top of each other in the direction parallel to the beam axis which generate pulsed electric fields [24]. These electric fields coincide with passing proton bunches and accelerate them. In this work's approach to the DWA, each module consists of a pulser containing a fast, HV switch and a radial waveguide. The pulser generates a HV pulse that is injected into the radial waveguide. The radial waveguide provides a path by which the pulse can travel to the beam pipe and form the accelerating gradient. Preliminary simulation studies show that the radial waveguides also amplify the pulse [25]. A high gradient insulator (HGI) beam pipe is employed to limit surface-flashover dielectric breakdown [26] and keep the beam pipe under vacuum. The high accelerating gradient of the DWA along with the compact design and relatively low cost of materials make it an attractive alternative acceleration method for FLASH proton radiotherapy.

In this work's approach to the DWA, a HV pulse of sufficient amplitude and fast rise time is required to be injected into the radial waveguide. A pulser incorporating an opening switch is practically simpler to operate in comparison to one incorporating a closing switch [27], and, as argued by Rukin [3], yields higher power pulses. Pulse widths on the order of 1 ns have been shown to withstand a 100 MV/m gradient at the HGI beam pipe without inducing dielectric breakdown [26]. Particle-in-cell simulations show that longitudinal stability of the particle bunch can be achieved with a fast pulse rise time and a pulse width on the order of 1 ns [28]. The operating frequency should be 1 kHz or higher to maintain a sufficient dose rate

for applications in conventional or FLASH RT. In order to meet these pulse requirements, an appropriate switch and pulser must be incorporated into the DWA concept.

The purpose of this study is to determine the suitability of a DSRD-based pulser in the context of the DWA and the above pulse parameters. Several different pulser designs incorporating DSRDs have been presented in the literature [27, 29–32]. Two common pulser designs found in previous studies [27, 29, 33] are considered here; (1) a magnetic saturation transformer (MST)-DSRD-based design and (2) a MOSFET-DSRD-based multi-module design. Though many results have been published about these pulser designs, there are currently no comparative studies between the two pulser designs that keep the DSRD doping and dimensions constant. A consistent comparative study is especially crucial in order to determine a suitable pulser for the DWA, based on it meeting the above listed pulse requirements. Both pulsers are constructed and tested under various conditions. Simulations are also performed with Synopsys Sentaurus TCAD in order to investigate the effect of the DSRD doping profile and the inclusion of 4H-SiC DSRDs on the output pulse.

7.4 Experimental designs and simulations

7.4.1 DSRD operation

A general overview of the DSRD operation is given in this section. Specific values of the forward and reverse pumping times and currents for each pulser will be given in Sections 7.4.3 and 7.4.4. These parameters are fully determined by the pulser components and operation. The DSRD's fast turn-off operation is based on subsequent forward and reverse current pumping stages [5, 6]. During the forward pumping stage, which lasts for a duration of t_+ = 100-300 ns, a forward current I_+ through the DSRD ionizes the dopants and generates an electron-hole plasma in the diode. During the reverse pumping stage, which lasts for a duration of t_- = 50-100 ns, a reverse current I_- extracts the plasma from the diode. The peak reverse current is typically a factor of 1.5-2 larger than the peak forward current [14, 34, 35]. As minority carriers are extracted, the plasma region shrinks and space charge

regions form on either side of the plasma region, which persists in the region about the *pn* junction. The plasma is fully extracted when the amount of charge extracted during the reverse pumping stage is equivalent to that stored during the forward pumping stage, i.e.,

$$\int I_+ \, dt = \int I_- \, dt. \tag{7.1}$$

Ideally, the boundaries of the plasma region meet near the *pn* junction [9]. At this point, the DSRD "switches off" or "opens" and the current is interrupted as equilibrium majority carriers are extracted and a large space charge region forms. This is known as the current breaking (CB) stage. The reverse current is commuted from the DSRD to the load at this stage and a HV pulse is produced with a rise time on the order of nanoseconds due to the high saturation drift velocity of electrons in Si (2×10^7 cm/s). A short t_+ and small I_+ (relative to I_-) causes the electron-hole plasma to be non-uniform and not extend too far into the DSRD. This ensures a fast t_- and that the plasma boundaries meet near the *pn* junction [6].

7.4.2 VMI DSRDs

The DSRDs (part #: H140S) used in this study were purchased from Voltage Multiplier Inc (VMI, Visalia, California). Each VMI DSRD consists of a stack of 16 individual DSRDs. They have a hexagonal face with a surface area of 1.3 cm^2. The height of each stack is 2.46 mm, so the thickness of an individual DSRD is 154 μm. While the exact doping profile is proprietary information and not available to the authors, it is known that the DSRD doping is done through high temperature diffusion of boron and phosphorus dopants in the silicon base. Gold and silver ohmic contacts are deposited on each end of the stack. The DSRD fabrication technology was initially transferred to VMI from Russia through a Phase II Department of Energy SBIR grant DE-FG02-06ER84459 and further developed with the Department of Energy grant DE-AC02-76SF00515.

7.4.3 Magnetic saturation transformer-DSRD-based pulser

A circuit diagram of the MST-DSRD-based pulser is shown in Fig. 7.1 and a photo of the experimental setup is shown in Fig. 7.2. A single turn is used for each side of the 1:1 MST in order to reduce the leakage inductances L_1 and L_2. Additional turns would increase L_2 and therefore increase the forward and reverse pumping times as well as the rise time of the pulse [14, 36]. Initially, the switch S_1 is open. The HV power supply V_0 charges the capacitor C_1 through the charging resistor R_1. Once C_1 is fully charged, switch S_1 is closed and C_1 begins to discharge through L_1 and L_2. Simultaneously, C_2 is charged and the DSRD is forward pumped. Referenced to the primary side, $C_2 = C_1$ because a 1:1 MST is used. The current I_+ in each side of the MST and the voltages over C_1 and C_2 are:

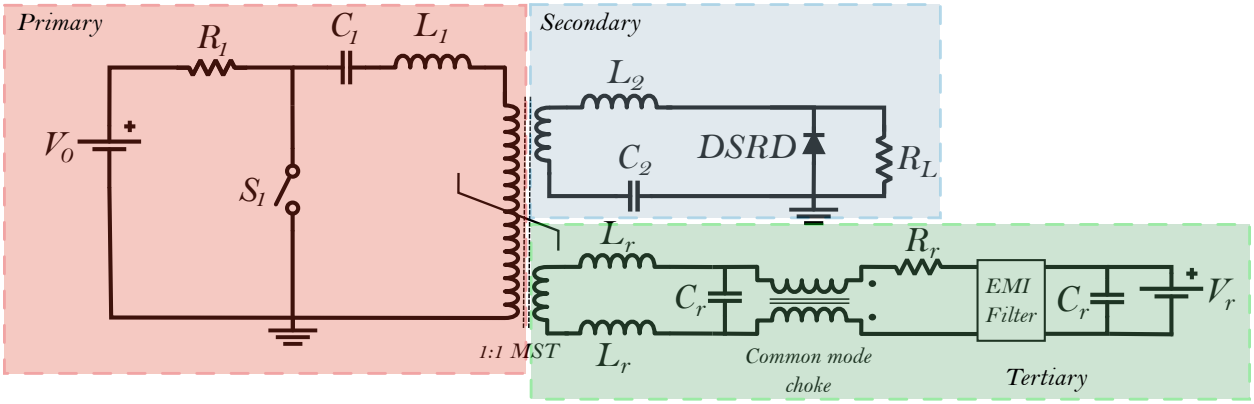


Figure 7.1: Circuit diagram of the 1:1 MST-DSRD-based pulser.

$$I_+(t) = V_0\sqrt{\frac{C_1/2}{L_1 + L_2}}\sin(\omega_+ t), \tag{7.2}$$

$$V_{C_1}(t) = \frac{V_0}{2}(1 - \cos(\omega_+ t)), \tag{7.3}$$

$$V_{C_2}(t) = \frac{V_0}{2}(1 + \cos(\omega_+ t)), \tag{7.4}$$

where

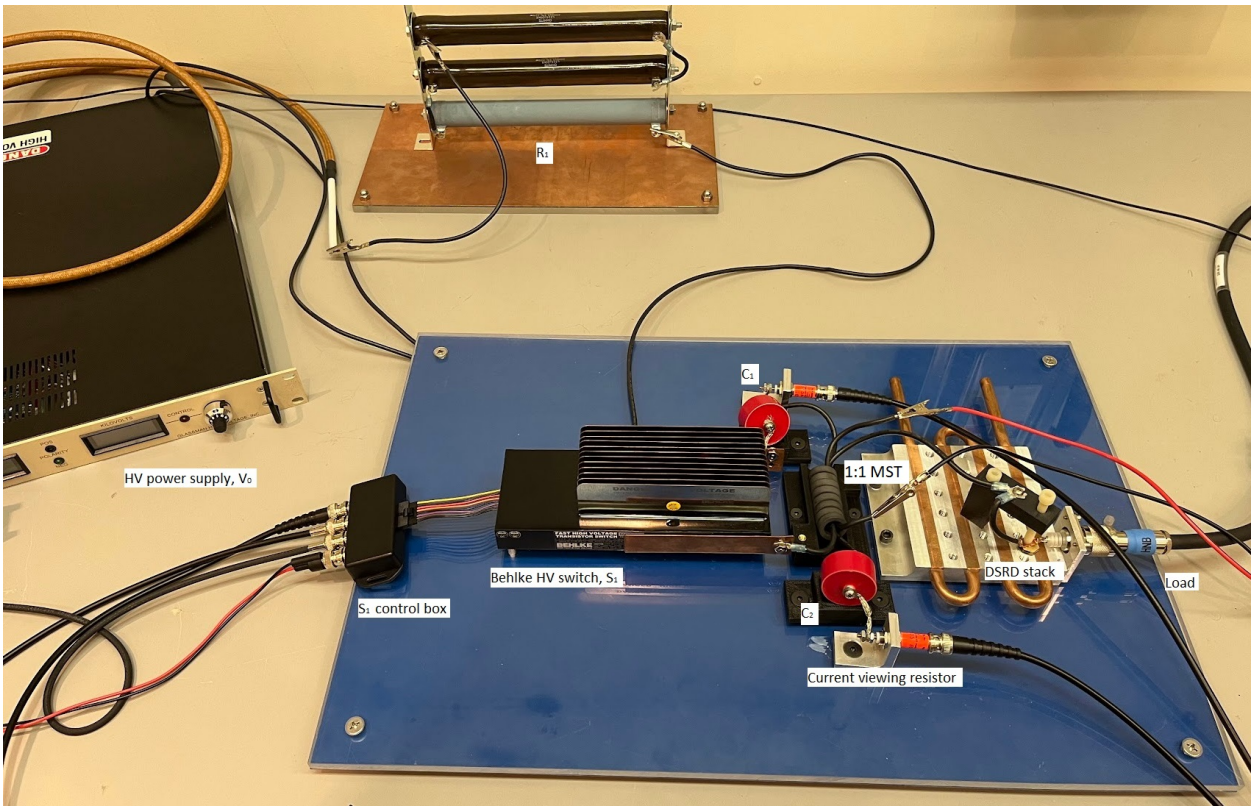


Figure 7.2: Experimental setup of the 1:1 MST-DSRD-based pulser. The DC reset circuit is not shown.

$$\omega_+ = \sqrt{\frac{1}{(L_1 + L_2)\frac{C_1}{2}}}. \tag{7.5}$$

The DSRD is forward pumped until the MST saturates. The time of saturation of the MST, t_{sat}, is given by the volt-second product:

$$\int_0^{t_{sat}} V_{C_1}(t)\, dt = N \Delta B A_c, \tag{7.6}$$

where ΔB is the available magnetic flux density swing of the MST, A_c is the cross-sectional area of the magnetic material of the MST, and $N = 1$ is the number of turns. In order for the MST to saturate prior to the current I_+ (7.2) changing sign (which would decrease the charge stored in the DSRD), the following condition must be met;

$$\frac{V_0}{2}\frac{T}{2} \geq N\Delta BA_c, \tag{7.7}$$

where $T = {}^{2\pi}/_{\omega_+}$ is the period of (7.2) and $T/2 = t_+$ represents the half period corresponding to $I_+ \geq 0$. Once the MST saturates (when the magnetic flux through the MST cannot change and a voltage can no longer be sustained, see Fig. 7.3), the primary side decouples from the secondary side and C_2 begins to discharge in the reverse direction through L_2 and the DSRD is reverse pumped. The reverse current I_- is given by:

$$I_-(t) = -V_0\sqrt{\frac{C_1/2}{L_2}}\sin(\omega_- t), \tag{7.8}$$

where $\omega_- = {}^{1}/_{\sqrt{L_2C_2}}$ and $t_- = {}^{\pi}/_{2\omega_-}$ The energy stored in C_2 is now stored in L_2. Once all of the excess charge carriers are removed, the DSRD switches off and the CB stage begins. The energy stored in L_2 is commuted to the load R_L and the HV pulse is formed. The peak reverse current was 1.3-1.5 times larger than the peak forward reverse current.

The tertiary winding of the MST connects to a DC reset circuit. This circuit resets the magnetic flux density of the MST between pulses to some negative value past $-B_r$. This is done by flowing a reverse current $I_r = {}^{V_r}/_{R_r}$ through the MST. The hysteresis loop of the MST is shown in Fig. 7.3. The initial operating point of the MST with the reset circuit is at point 3 with $H = {}^{-I_-}/_{l_e}$, where l_e is the effective magnetic path length of the magnetic material of the MST. The available magnetic flux density swing is then ΔB_2. The MST can then operate between points 1 and 3. Without this reset circuit, the available magnetic flux density swing would be limited to $\Delta B_1 = B_s - B_r$ and the MST would operate between points 1 and 2. This would limit the voltage transferred from C_1 to C_2, as the MST would saturate prior to full voltage transfer. The reset inductors L_r, reset capacitors C_r, common mode choke, and EMI filter are implemented to isolate the power supply V_r from the high frequency HV pulse. Therefore, the impedance of the reset inductors should be much larger than that of the pulser, i.e., $\omega L_r \gg Z_{pulser}$. Air core inductors were used in the reset circuit

to avoid magnetic saturation.

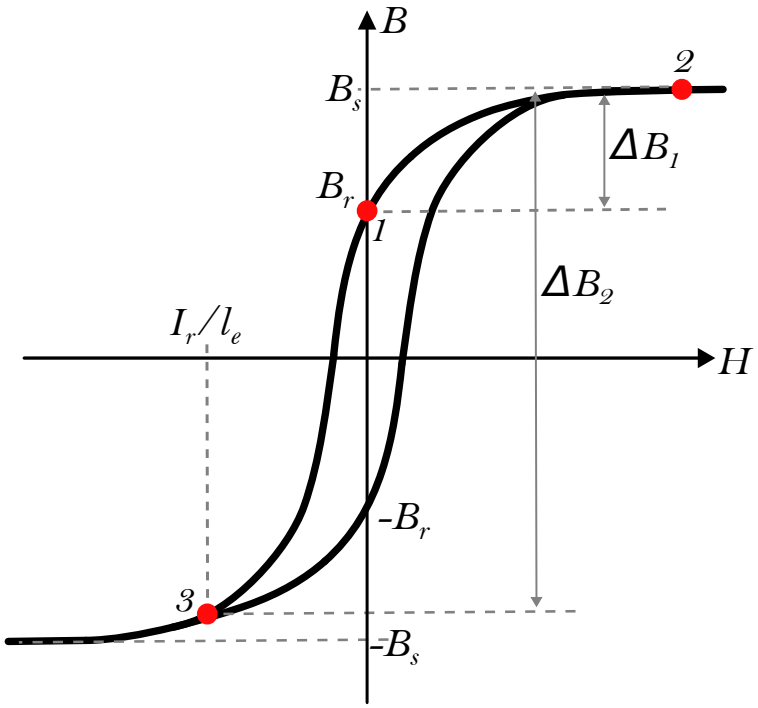


Figure 7.3: Hysteresis loop of the 1:1 MST. The reset circuit resets the initial magnetic flux density to some negative value past $-B_r$ by flowing a reverse current I_- through the MST. Without the reset circuit, the MST operates between points 1 and 2. With the reset circuit, the MST operates between points 1 and 3.

The addition of a compression stage and its influence on the output pulse was investigated in a series of experiments. The compression stage is formed by placing an inductor L_3 in between two stacks (in series) of DSRDs (DSRD1) and three stacks (in series) of DSRDs (DSRD2). In this configuration, the forward pumping stage pumps both DSRD1 and DSRD2. Upon saturation of the MST, the discharge of C_2 reverse pumps DSRD1. When DSRD1 turns off, the current flows through L_3 and is commuted to DSRD2. Once DSRD2 turns off, the current is commuted to the load. Compression stages such as this are used to decrease the rise time and increase the peak voltage of the pulse [37]. Due to the small amount of charge stored in DSRD2 during the forward pumping phase, the rise time is very fast (see Table 7.3). The numerical values for the circuit parameters used in this work are listed in Table 7.1. The MST is formed of 8 cores of CMD5005 NiZn ferrite toroids, provided by National

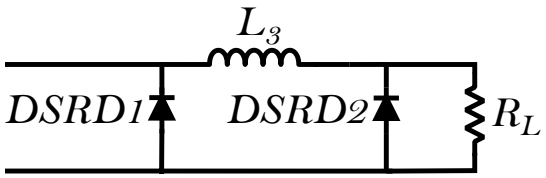


Figure 7.4: A compression stage with two DSRD stacks. This stage can be used to decrease the pulse rise time and increase the peak voltage.

Table 7.1: Numerical values of the parameters used for the MST-DSRD-based pulser.

Parameter	Value
V_0	2000 - 6000 V
R_1	140 kΩ
C_1	2 nF
C_2	2 nF
R_L	50 Ω
L_r	10 μH
C_r	1 nF
R_r	10 Ω
V_r	20 V
MST	CMD5005 NiZn ferrite [a]

[a] $B_s = 0.33$ T, $B_r = 0.13$ T, $H_c = 0.12$ Oe, $A_c = 0.26$ cm^2, $l_e = 5.41$ cm

Magnetics Group Inc (Bethlehem, Pennsylvania). The switch S_1 is a Behlke HTS 61-40, pulsed by an HP 81101A pulse generator. The common mode choke is formed of 6 NiZn ferrite toroids (5943001001) by Fair-Rite Products Corp. The EMI filter is a 10DKAG5 by Delta Electronics. The current in both the primary and secondary side of the MST was measured with 0.05 Ω SSDN-414-05 current viewing resistors from T&M Research products. The load $R_L = 50\,\Omega$ consisted of a 20 dB Barth Model 2536, 2 × 26 dB Barth Model 202 and a 20 dB Tektronix 011-0059-02 attenuator, for a total attenuation of 92 dB. These attenuators were connected to a Tektronix 694C 3 GHz oscilloscope.

7.4.4 MOSFET-DSRD-based multi-module pulser

A circuit diagram of the MOSFET-DSRD-based multi-module pulser is shown in Fig. 7.5 and is based on the work of Fang et. al. [33]. The topology of this type of PFN was originally investigated for applications to the DWA in the work of Arntz et al. [27]. Initially,

the MOSFET S_1 is open and the capacitor C_2 is charged to the voltage $V_1 - V_2$. When S_1 is closed, C_2 discharges and forward pumps the DSRD. The current through the DSRD during this stage is given by

$$I_+(t) = (V_1 - V_2)\sqrt{\frac{C_2}{L_2}}\sin(\omega_+ t), \tag{7.9}$$

where $\omega_+ = 1/\sqrt{L_2 C_2}$. The forward pumping time is determined by the gate driver input pulse width $t_+ = \Delta T$. When S_1 is opened, the reverse pumping stage begins. The frequency of this stage is given by $\omega_- = 1/\sqrt{L_\parallel C_1}$, where $L_\parallel$ is the equivalent parallel inductance of L_1 and L_2. Since $C_1 \ll C_2$, the reverse pumping time is much shorter than the forward pumping time. The reverse pumping time is given by $t_- = \pi/2\omega_-$. The reverse current $I_-(t)$ is a complicated function of the pulser components, the voltages V_1 and V_2 and ΔT. However, its amplitude is mostly determined by the product $V_1\,\Delta T$[33]. Due to this large value, the peak reverse current was 7-7.3 times larger than the peak forward reverse current. When the DSRD switches off, the energy stored in L_2 is commuted to the load R_L. The capacitor C_3 is present to block the load from the DC voltage source V_2 and reverse bias the DSRD. Additional modules (indicated by the red dashed line) can be incorporated into the pulser by parallel connections to the gate driver input, V_1, V_2, and the DSRD. A photo of a 6 module experimental setup is shown in Fig. 7.6. The pulse capacitors seen in Fig. 7.6 (not shown in Fig. 7.5) supply the forward current through L_1 because the power supply V_1 itself cannot supply such a current.

The numerical values for the circuit parameters used in this work are listed in Table 7.2. The gate driver is a 1ED3121MC12H by Infineon Technology. The MOSFET is a SiC C3M0040120K by Wolfspeed.

7.4.5 Sentaurus TCAD simulations

Simulations were performed in order to study the effect of the Si DSRD doping profiles on the breakdown voltage and the output HV pulses, given constant circuit parameters.

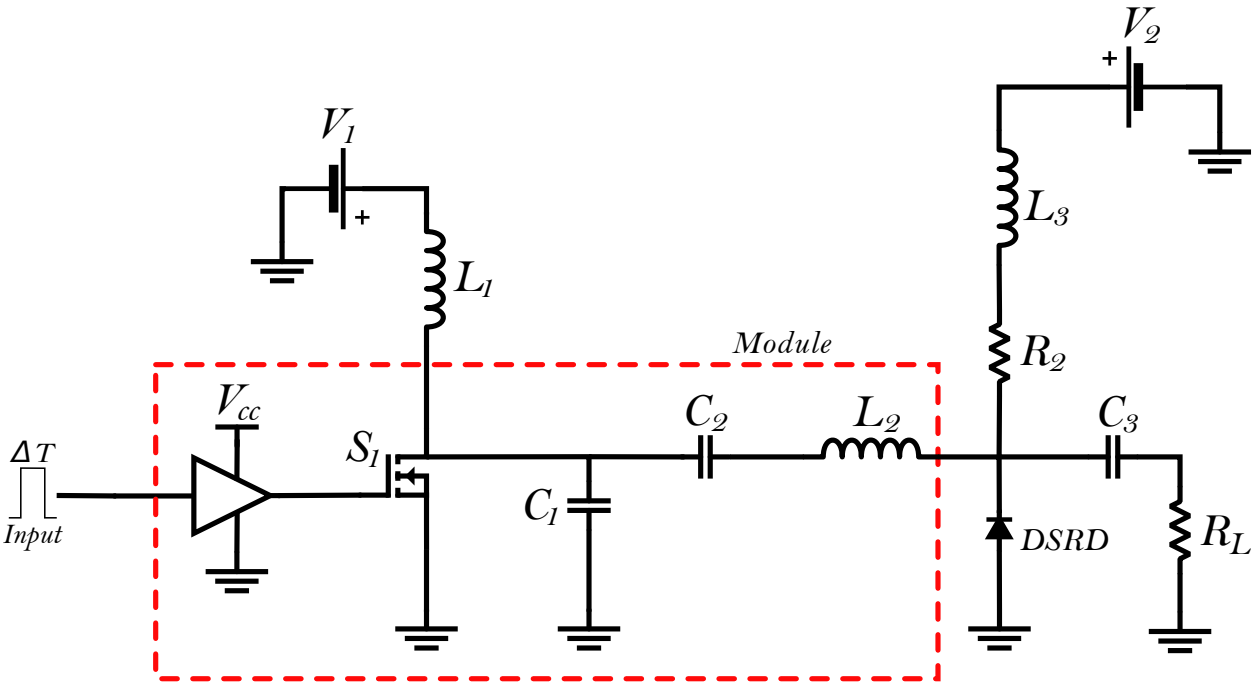


Figure 7.5: Circuit diagram of a single module of the MOSFET-DSRD-based pulser. Additional modules, enclosed by the red dashed line, are incorporated by parallel connections to the gate driver input, V_1, V_2, and the DSRD.

Sentaurus TCAD by Synopsys® was used to perform the simulations. An individual DSRD was modelled as a 2D structure with a length l. The width was kept constant at 50 μm. The parameter AREAFACTOR was set to 2535200 in order to model the correct surface area of the VMI DSRD samples. This parameter specifies the extension of the model (in μm) in the unsimulated dimension. By scaling down the device width and using a large AREAFCTOR, the computation time is reduced. In 2D simulations, the simulated current is given in units of $A/\mu m$ and therefore must be scaled correctly in circuit simulations where different size devices interact. The acceptor and donor concentrations, $N_a(x)$ and $N_d(x)$ respectively, were modelled as

$$N_a(x) = N_{sa} e^{-\frac{x^2}{\lambda_p^2}}, \tag{7.10}$$

$$N_d(x) = N_{sd} e^{-\frac{(l-x)^2}{\lambda_n^2}}, \tag{7.11}$$

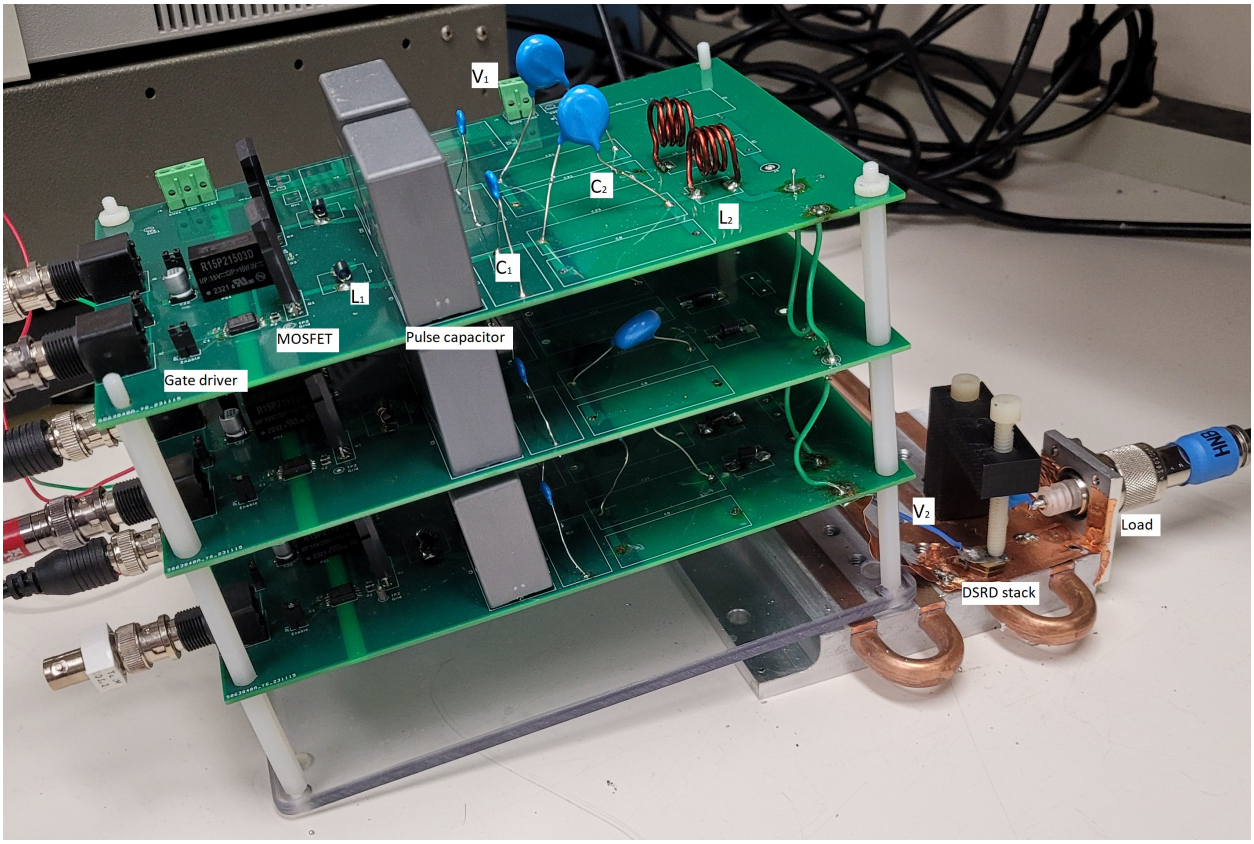


Figure 7.6: Experimental setup of a 6 module MOSFET-DSRD-based pulser. The pulse capacitors are connected in series to V_1 and supply the current through L_1 when the MOSFET S_1 is closed.

where

$$\lambda_{p,n} = \frac{l_{p,n}}{\sqrt{\ln\left(\frac{N_{sa,sd}}{N_0}\right)}}. \tag{7.12}$$

For (7.10), the surface was defined at the $x = 0$ plane. For (7.11), the surface was defined at the $x = l$ plane. Here, N_{sa} and N_{sd} are the acceptor and donor surface concentrations, respectively, and N_0 is the doping concentration at the pn junction which was kept constant at $1{\times}10^{14}$ cm^{-3}. The parameters $l_{p,n}$ are the distances from the surface of the respective dopant profile to the pn junction and $l = l_p + l_n$. Examples of (7.10) and (7.11) are shown in Fig. 7.7. The simulation physics used the drift-diffusion transport model and included Fermi statistics, Shockley-Read-Hall recombination with temperature and doping enhancement,

Table 7.2: Numerical values of the parameters used for the MOSFET-DSRD-based pulser.

Parameter	Value
V_1	300 V
V_2	90 V
L_1	120 nH
L_2	100 nH
L_3	1 μH
C_1	220 pF
C_2	2 nF
C_3	1 nF
R_L	50 Ω
ΔT	53.6 ns

Auger recombination, and avalanche generation (Okuto-Crowell model). The electron and hole mobility models included doping dependence, high electric field saturation and carrier-carrier scattering.

The SiC DSRD was modelled as a $p^+pn_0n^+$ epitaxial structure and is based on the work of Sun et. al. [38]. The acceptor concentration of the 1 μm p^+ region is 2×10^{19} cm^{-3}, the acceptor concentration of the 4 μm p region is 1×10^{17} cm^{-3}, the donor concentration of the 85 μm n_0 region is 7×10^{14} cm^{-3}, and the donor concentration of the 180 μm n^+ substrate is 5×10^{18} cm^{-3}. For the SiC DSRD, the p-dopant was aluminum and the n-dopant was nitrogen. Along with the simulation physics used for the Si DSRD simulations, the 4H-SiC DSRD model also included incomplete ionization and anisotropic mobility and avalanche generation.

To obtain the breakdown voltage values, the IV curves of the DSRD with various doping profiles were computed in quasistationary mode. First, a positive voltage was applied to the cathode of the DSRD and the current computed. This was done for various voltages until the simulation no longer converged. At this point, the IV curve is parallel to the current axis. To continue the simulation, the boundary conditions of the cathode are changed so that a current is applied to the cathode and the voltage computed. It is at this point that the breakdown voltage is extracted.

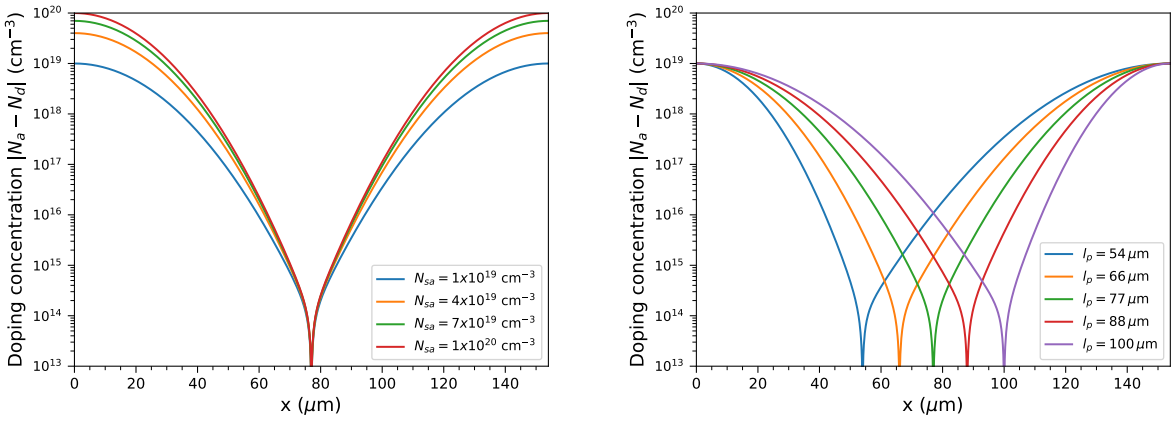


Figure 7.7: The doping profile of an individual DSRD used in the TCAD simulations with (a) varying surface concentration and (b) varying junction depth.

For the circuit simulations of the MST with a Si DSRD, the current measured in the secondary side of the experimental MST was used as a current source. This was done so as to remain as close as possible to the experimental setup. It is justifiable to use the measured current (taken from experiments using the VMI DSRDs and therefore a constant doping profile) in the simulations (where the doping profile is varied) because the carrier dynamics near the *pn* junction at the DSRD turn-off stage depend primarily on the local doping concentration gradient and that largely does not change between doping profiles. For the SiC DSRD simulations, an alternative current source was required, leading to a different simulation set up. This was done because the current measured in the experimental MST includes effects from the Si DSRD turn-off and the dynamics of this process are different in the SiC DSRD due to the different material and the doping profile. For the forward pumping stage, the measured current was used as a current source with an ideal switch S open in parallel. This charged the capacitor C_2 to the exact voltage as in the experiments. We considered the use of the measured current during this stage to be acceptable because the forward-biased DSRD does not have a significant effect on the waveform. For the reverse pumping stage, the switch S is closed and C_2 discharges through the leakage inductor L_2, forming the reverse current. The inductance value was determined by fitting the measured current to (7.8) and

extracting the frequency ω_-, from which L_2 can be determined through $\omega_- = {}^1/\sqrt{L_2 C_2}$. It was found to be 482 nH.

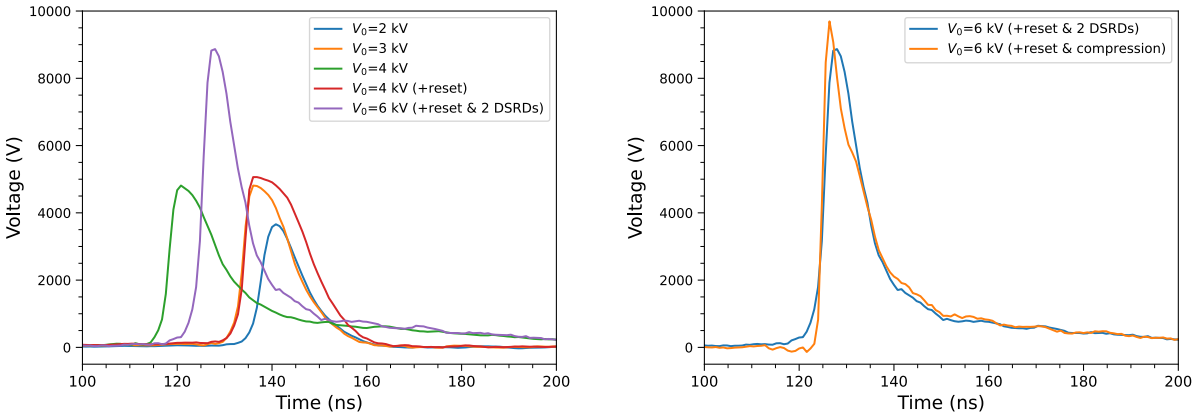
7.5 Results and discussion

7.5.1 MST-DSRD-based pulser experiments

Fig. 7.8a presents the output of the MST-DSRD-based pulser for various input voltages V_0. Unless otherwise specified, only one DSRD stack is present. It can be seen at $V_0 = 4$ kV (no reset) the pulse forms earlier than with the reset circuit. This is due to limited available magnetic flux swing with no reset circuit. Of note is the widening of the pulse and the formation of a flat and decaying pulse top at $V_0 = 3$ kV and 4 kV (with reset). These pulse features are characteristic of avalanche breakdown [34, 39, 40]. Avalanche breakdown occurs when there is a rapid increase in carrier generation. This is due to strong electric fields accelerating carriers to high enough energies that they are able to generate electron-hole pairs through collisions. The rapid increase in conductivity of the DSRD shunts the load, widening the pulse. From these experimental observations, it can be concluded therefore that the breakdown threshold of the DSRD stack is between 4.5-5 kV. It can be seen that at $V_0 = 6$ kV with the use of 2 DSRD stacks, this effect is mitigated due to the increased breakdown threshold of 2 DSRD stacks in series. Fig. 7.8b presents the output pulse with the load section of the MST modified to include a compression stage (see Fig. 7.4).

7.5.2 MOSFET-DSRD-based pulser experiments

Fig. 7.9 presents the experimental results for the MOSFET-DSRD-based pulser. In Fig 7.9a, the output pulse for 1-4 modules with a single DSRD stack is shown. In Fig 7.9b, the output pulse for 4-6 modules with two DSRD stacks in series is shown. Two DSRD stacks are used to accommodate the higher voltage at these number of modules. The gate driver input pulse width ΔT had to be increased as the number of modules increased to achieve the maximum output pulse amplitude. A maximum pulse amplitude (10.9 kV) for 6 modules and two



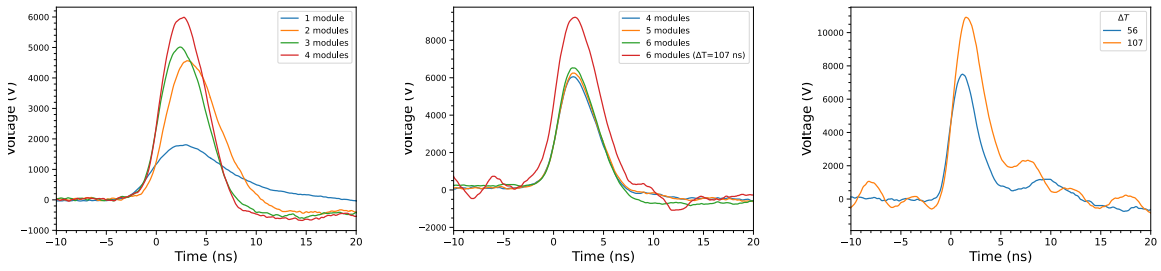
(a) (b)

Figure 7.8: Experimental output pulses from the MST-DSRD-based pulser. (a) Pulses with various input voltages V_0 and inclusion of the the reset circuit and additional DSRD stacks. (b) Pulses at $V_0 = 6$ kV with and without the inclusion of a compression stage.

DSRD stacks with compression was obtained with a gate driver input pulse width of ΔT = 107 ns, in comparison to ΔT = 53.6 ns with 1 module. In Fig 7.9c, the output pulse for a 6 module setup with compression is shown. A potential explanation as to why the gate driver input pulse width that yielded the maximum output pulse amplitude increased as additional modules were added is the presence of parasitic capacitance and parasitic coupling. In the experimental setup used for this work (see Fig. 7.6), two modules were present per PCB. At six modules, three PCBs were stacked on top of each other and could be inductively coupled together due to any stray inductance from leads or traces. Any parasitic capacitance between the modules and coupling between the PCBs would have influenced both the forward pumping stage (through the frequency ω_+, see Eq. 7.9) and the duration of the reverse pumping stage. This in turn would determine how much charge is stored during the forward pumping stage and how quickly it is extracted. These parasitics would also store some energy which is not transferred to the load, further limiting the pulse amplitude. Efforts are on going to reduce these parasitics by fabricating new PCBs with smaller traces and removing leads from through-hole components and stray leads. Additional experiments were performed with two pulsers on separate PCBs fed by separate gate driver inputs to

determine whether the inherent variations of the MOSFETs were the cause of the need to increase the gate driver input pulse width. Minor improvements were made by varying the phase between the inputs as well as their widths, but not enough to make up the large difference observed above. In order to ensure the MOSFETs were not failing, a lower voltage was used for V_1. Again, only minor changes were observed.

In Fig. 7.9a, the output pulse obtained with 4 modules and 1 DSRD stack shows no indication of pulse widening (pulse width of 5.04 ns), despite the peak voltage being 6 kV. This is above the breakdown threshold of a single DSRD stack established in section 7.5.1, between 4.5-5 kV. For the MST and MOSFET-DSRD-based pulser, the pulse width is proportional to L_2/R_L (9.64 ns) and $L_{\parallel}/R_L$ (1.1 ns), respectively. Since this value is 10 times smaller for the MOSFET-DSRD-based pulser, the pulse is too short lived for any considerable charge multiplication (through avalanche generation) to occur.



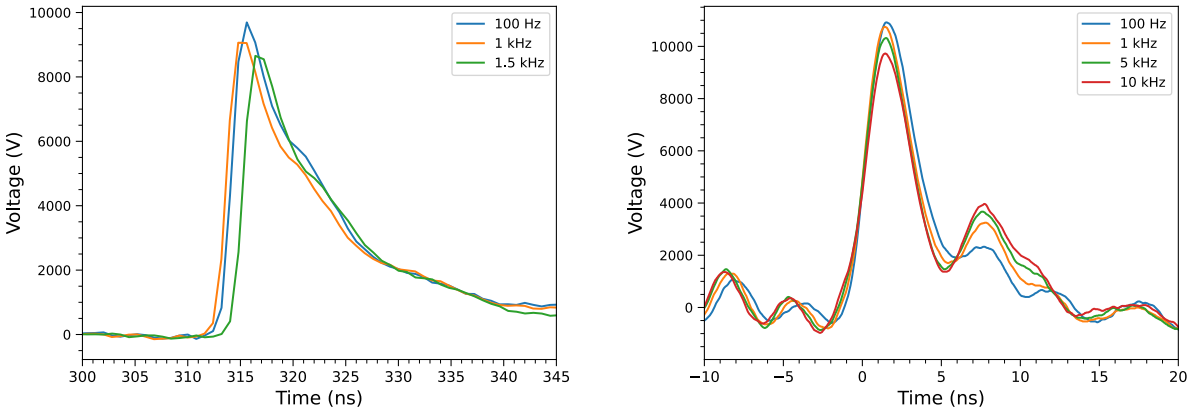
(a) (b) (c)

Figure 7.9: Experimental output pulses from the MOSFET-DSRD-based pulser. (a) Pulses with 1-4 modules and 1 DSRD stack. (b) Pulses with 4-6 modules and 2 DSRD stacks. (c) Pulses with 6 modules and the inclusion of a compression stage with a gate driver input pulse width of ΔT= 56 and 107 ns.

7.5.3 Frequency experiments

Fig. 7.10a and 7.10b present the average of 50 pulses for varying operating frequencies with the MST-DSRD-based pulser and the MOSFET-DSRD-based pulser, respectively. Both pulsers can be reliably run at an operating frequency of 1 kHz and the MOSFET-DSRD-based pulser up to 10 kHz, though at a reduced peak voltage. The operating frequency of the

MST-DSRD-based pulser is limited by the long charging time constant, that is $\tau_1 = R_1 C_1$. Considering the values in Table 7.1, $\tau_1 = 280$ μs. In a single cycle at an operating frequency of 1 kHz, 3.6 time constants elapse, limiting the amount of energy initially stored in C_1 and available to be commuted to the load. A higher voltage power supply would enable the use of a smaller capacitance value and therefore higher operating frequency. The large charging resistance is necessary so as to not exceed the power and voltage rating of any one resistor. The MOSFET-DSRD-based pulser operating frequency is limited by the heating of the MOSFETs. This heating will be mitigated by mounting the MOSFETs on a heat sink or on a water-cooled cold plate. The stability of the pulse amplitude is calculated as $100\% \times \sigma/\bar{V}$ where σ is the standard deviation of the pulse amplitudes and $\bar{V}$ is the mean pulse amplitude. For the MST-DSRD-based pulser, the pulse amplitude stability was on average 1.2%. For the MOSFET-DSRD-based pulser, the pulse amplitude stability was on average 0.67%.



(a) (b)

Figure 7.10: An average of 50 pulses from the (a) MST-DSRD-based pulser and (b) MOSFET-DSRD-based pulser at high frequencies.

7.5.4 Comparison between pulsers

The pulse features of the two pulsers are shown in Tables 7.3 and 7.4. The rise time is defined as the time required for the pulse to rise from 10% to 90% of its peak value, V_{max}.

Table 7.3: Features of the output pulse generated by the MST-DSRD-based pulser with various configurations.

$V_0(V)$	$V_{max}(V)$	Rise time (ns)	Pulse width (ns)
2000	3649	4.31	9.86
3000	4843	3.68	11.29
4000	4812	3.53	11.79
4000 (+reset)	5063	3.54	14.59
6000 (+reset & 2 DRSDs)	8943	3.61	8.84
6000 (+reset & compression)	9645	1.48	8.36

The pulse width is defined as the full width half maximum of the pulse. It can be seen that a single MST-DSRD-based pulser can generate higher peak voltages than a single module MOSFET-DSRD-based pulser (3649 V and 1807 V, respectively). This is due to the high voltage power supply. With a MOSFET-DSRD-based pulser comprised of 6 modules along with a gate driver input width of $\Delta T = 107$ ns and a compression stage, the maximum observed pulse amplitude for any configuration (10.9 kV) is achieved, in comparison to a maximum amplitude of 9.65 kV with the MST-DSRD-based pulser ($V_0 = 6000$ V with reset and compression). In comparing these two pulses, a fast rise time is observed for both (1.66 ns and 1.48 ns, respectively) but the pulse width achieved by the MOSFET-DSRD-based pulser (3.5 ns) was much closer to the desired 1 ns width than the MST-DSRD-based pulser (8.4 ns). Overall, the MOSFET-DSRD-based pulser is able to generate shorter pulses with faster rise times, even without compression. This can be attributed to the much shorter forward and reverse pumping times. Additionally, one must consider the size of each of the pulsers, especially if they are to be integrated in a compact DWA structure. The MOSFET-DSRD-based pulser is significantly smaller in comparison to the MST-DSRD-based pulser, due to the presence of the large HV switch and pulse transformer in the latter.

7.5.5 Si DSRD simulations

The simulated breakdown voltage was found to be between 500-560 V, decreasing slightly with increasing surface doping concentration, as seen in Fig. 7.11a. The junction depth

Table 7.4: Features of the output pulse generated by the MOSFET-DSRD-based pulser with various configurations.

# modules	$V_{max}(V)$	Rise time (ns)	Pulse width (ns)
1	1807	3.72	8.10
2	4562	3.06	5.89
3	5013	2.74	5.16
4	5989	2.66	5.04
4 (2 DSRDs)	6052	2.42	4.32
5 (2 DSRDs)	6247	2.50	4.32
6 (2 DSRDs)	6530	2.37	4.23
6 (ΔT = 107 ns)	9231	3.05	4.87
6 (+compression)	7498	1.53	3.03
6 (ΔT = 107 ns & compression)	10919	1.66	3.48

has a negligible effect, as seen in Fig. 7.11b. The simulated breakdown voltage of one VMI DSRD stack (16 DSRDs) is therefore between 8-9 kV. Referring back to Fig. 7.8a for V_0 = 3 kV and 4 kV (with reset), it can be seen that the actual breakdown voltage for a DSRD stack is between 4.5-5 kV. The discrepancy between the simulated and experimental value can be attributed to defects in the DSRD crystals due to fabrication processes [41] and surface breakdown effects [42, 43] that are not accounted for in the simulations. It is known that the surface breakdown field of doped Si is an order magnitude lower than the bulk breakdown field. This reduction is caused by electric field enhancement near the surface of the semiconductor which induces avalanche generation.

Fig. 7.12a presents the simulation of the MST-DSRD-based pulser with V_0 = 2 kV and varying surface doping concentrations with $N_{sd} = N_{sa}$. The parameter l_p is kept constant at 77 μm. Evidently there is a negligible influence on the pulse amplitude and rise time. Fig. 7.12b presents the simulation of the MST-DSRD-based pulser with V_0 = 2 kV and varying pn junction positions. The surface concentrations were kept constant at $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}. A maximum peak voltage was achieved for larger values of l_p (77-100 μm). Simulations of the MOSFET-DSRD-based pulser also demonstrated this trend. The pedestal effect, which is the portion of the pulse prior to DSRD turn-off, increases as

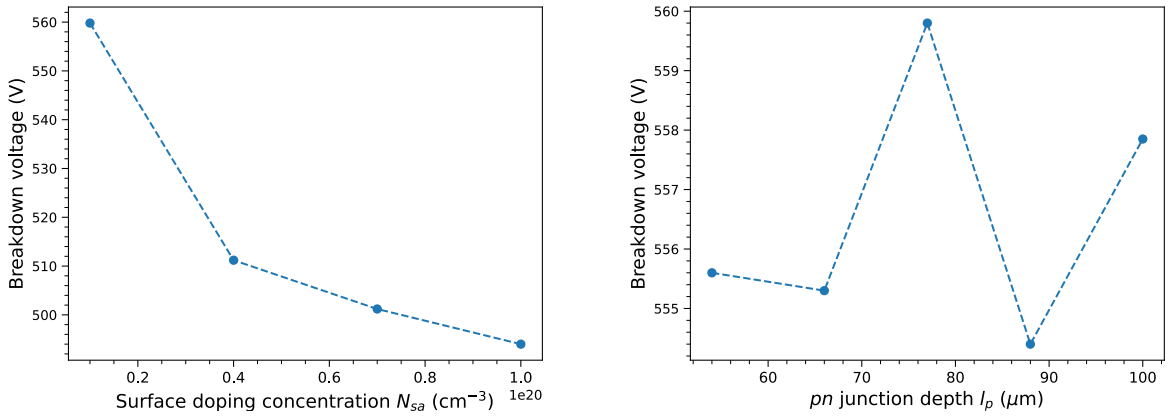
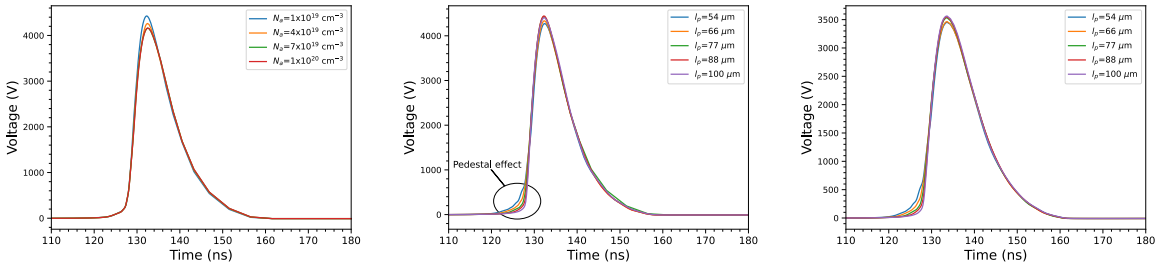


Figure 7.11: Simulation results of the breakdown voltage depending on (a) the surface doping concentration N_{sa} ($l_p = 77$ μm) and (b) the pn junction depth l_p ($N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}).

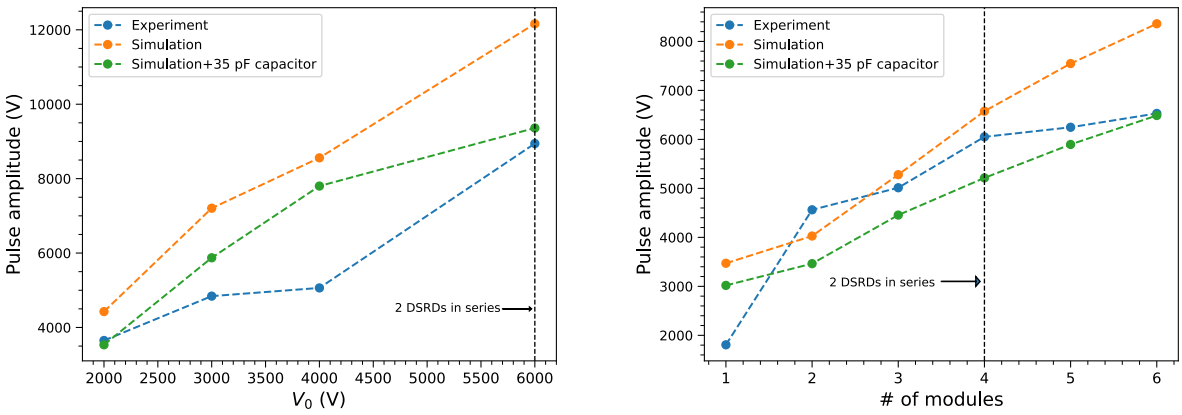
l_p decreases. This pedestal effect can be explained by the boundaries of the electron-hole plasma meeting away from the pn junction, in the n-doped region, at the end of the reverse pumping stage when the DSRD begins to turns off. To the left of the left plasma region boundary and in the junction region, a space charge region consisting of uncompensated holes is present in order to sustain the reverse current [44]. Since the left boundary of the plasma region forms first and moves with a greater velocity (due to the higher electron mobility compared to holes) [45], the meeting of the plasma boundaries occurs further in the n-doped region as l_p decreases, increasing the left space charge and the pedestal effect. It can be seen in Fig. 7.13a that the simulated ($l_p = 77$ μm and $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}) pulse amplitude is larger than that observed in the experiments. Additional simulations were performed that included a 35 pF parasitic capacitance in parallel with the DSRD, shown in Fig. 7.12c. The value of 35 pF was chosen This parasitic capacitance is used to model that which is present from the electrical connection between the DSRD electrode, leads and board. The presence of this parasitic capacitance brings the simulated amplitudes closer in agreement with experiment at $V_0 = 2$ kV and 6 kV. A disagreement remains at $V_0 = 3$ kV and 4 kV, which can be attributed to the lower DSRD breakdown threshold voltage in the

experiments compared to the simulations. Comparing Figs. 7.12a and 7.12c, the presence of this parasitic capacitance reduces the peak voltage and slightly increases the pulse width.



(a) (b) (c)

Figure 7.12: Simulation results of the MST-DSRD-based pulser with the DSRD doping profile varying through (a) N_{sa} ($N_{sa} = N_{sd}$ and $l_p = 77$ μm), (b) l_p ($N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}) and (c) l_p and an additional 35 pF parasitic capacitance in parallel with the load ($N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}). This parasitic capacitance models that which is present between the DSRD electrodes and leads.



(a) (b)

Figure 7.13: Experimental and simulation ($l_p = 77$ μm and $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}) results of the pulse amplitude of the (a) MST-DSRD-based pulser at different input voltage V_0 (with reset at 4 kV) and (b) MOSFET-DSRD-based pulser with varying number of modules. Note that all data points to the right of the black line used 2 DSRD stacks in series.

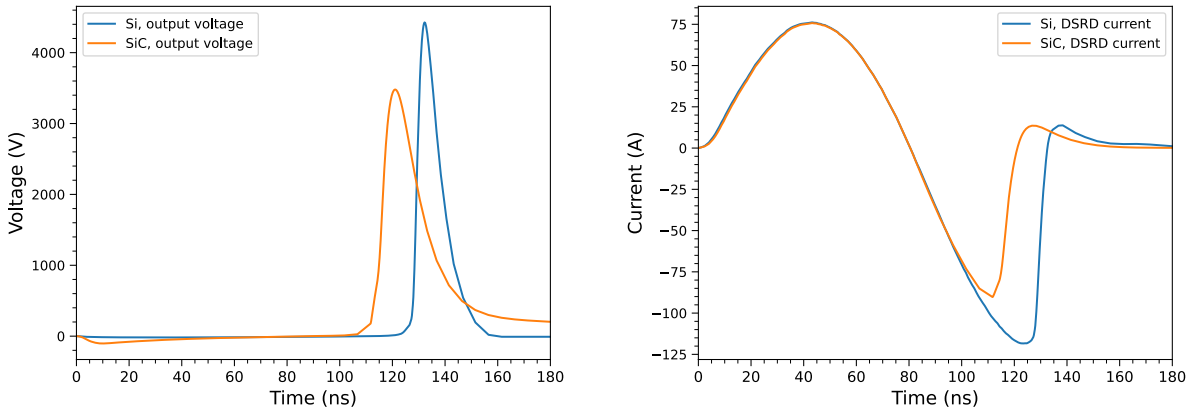
The results of the MOSFET-DSRD-based pulser simulations ($l_p = 77$ μm and $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}) are shown in Fig. 7.13b. The percent difference between the simulations and experiments is less than 12%, for 2-4 modules. The presence of a 35 pF parasitic capacitance lowers the amplitude, as in the MST-DSRD-based pulser case. The inclusion

of the parasitic capacitance brings the simulated pulse amplitudes in agreement at 5 and 6 modules. In addition, Fig. 7.12 demonstrates the relative insensitivity of the pulse shape to doping profile parameters. Thus parasitic coupling appears to be a significant reason for the discrepancy between simulations and experiments, as expected from the discussion in section 7.5.2.

7.5.6 SiC DSRD simulations

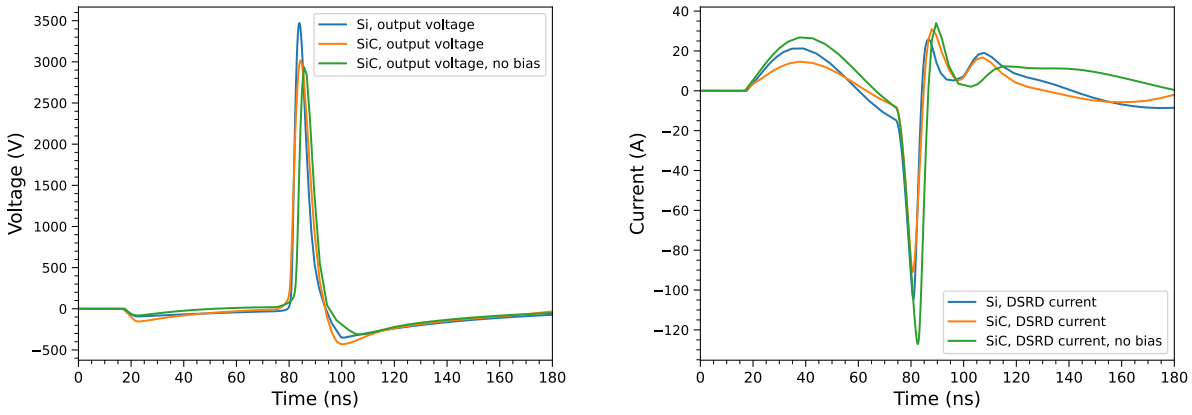
Figs. 7.14 and 7.15 present the simulation results for the MST ($V_0 = 2$ kV) and MOSFET-DSRD-based pulser (1 module) respectively, incorporating Si ($l_p = 77$ μm and $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}) and SiC DSRDs. It can be seen that despite the higher breakdown threshold of SiC in comparison to Si, the peak voltage of the pulse is less than that obtained with Si DSRDs. This is due to the incomplete ionization of the dopants during the forward pumping stage. Minority carrier recombination, e.g., electrons recombining in a p-doped region, is not a concern because the lifetime of minority carriers in SiC is 150-250 ns [13] and the forward pumping times used in this study are much shorter, < 80 ns. Fewer electrons and holes are generated during the forward pumping stage due to incomplete ionization and therefore, during the reverse pumping stage, there is less charge extracted. The SiC DSRD turns off earlier and less current is commuted to the load. Smirnov and Shevchenko (2018) showed that a significant gain in charge extraction by using short forward pumping stages (< 40 ns) [46]. With the MOSFET-DSRD-based pulser and those similar to it, such as the SLIM design [27], such a small forward pumping time is achievable with appropriate component values, such as those used in Table 7.2. A short forward pumping time is difficult to achieve with the MST-DSRD-based pulser due to the large leakage inductances of the MST. Furthermore, Yang et al. (2023) showed that a MST-DSRD-based pulser with a SiC DSRD is much more sensitive to the leakage inductance of the MST than with Si DSRDs [14]. This makes control and knowledge of the leakage inductance crucial in the development of a MST-DSRD-based pulser with SiC DSRDs. In this scenario, it is therefore recommended to use a 1:1 MST.

Figs. 7.14 and 7.15 also demonstrate that a SiC-based pulser generates a pulse with a larger pulse width than that obtained with Si DSRDs. This is due to the stored energy in the capacitor (C_2 in the MST and MOSFET-DSRD-based pulser case) being directly discharged into the load upon the turn-off of the SiC DSRD. As the SiC DSRD turns off earlier (in comparison to Si), not all of the initial energy stored in C_2 is transferred to the inductive storage during the reverse pumping stage. Optimization of the value for C_2 would aid in generating higher quality pulses. Fig. 7.15 also demonstrates that the removal of the bias voltage (V_2 in Fig. 7.5) does not increase the output pulse amplitude. Removal of the bias results in a larger peak forward and reverse current through the SiC DSRD. This is due to a non-zero bias voltage enhancing the current blocking properties of the DSRD during the forward pumping stage. As a greater amount of charge is stored with no bias, it must be compensated for with a larger peak reverse current (see Eq. 7.1). Despite the increased amount of charge stored, incomplete ionization still limits the pulse amplitude.



(a) (b)

Figure 7.14: Simulated (a) output pulse and (b) DSRD current of the MST-DSRD-based pulser ($V_0 = 2$ kV) with Si and SiC DSRDs. The Si parameters are $l_p = 77$ μm and $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}.



(a) (b)

Figure 7.15: Simulated (a) output pulse and (b) DSRD current of the MOSFET-DSRD-based pulser (1 module) with Si and SiC DSRDs. The Si parameters are $l_p = 77$ μm and $N_{sd} = N_{sa} = 1 \times 10^{19}$ cm^{-3}. "No bias" indicates $V_2 = 0$ in Fig. 7.5.

7.6 Conclusions

The purpose of this study was to determine the suitability of two DSRD-based pulsers for the purposes of a dielectric wall accelerator for proton therapy. This was achieved through the testing and comparison of a MST-DSRD-based pulser and a MOSFET-DSRD-based-pulser. The criteria for which the pulsers are considered suitable are (1) large pulse amplitude, (2) fast pulse rise time (3) nanosecond-scale pulse width and (4) an operating frequency of 1 kHz or higher. Experimental setups of both pulsers were built to study their performance with the same DSRDs. While a single stage MST-DSRD-based pulser is able to generate higher pulse amplitudes than a single module MOSFET-DSRD-based pulser, the MOSFET-DSRD-based pulser consistently achieves faster rise times and pulse widths. At 6 modules and a compression stage, the MOSFET-DSRD-based pulser with a MOSFET gate driver input pulse width of 53.6 ns and 107 ns achieves the shortest pulse widths, 3 ns and 3.5 ns respectively. The maximum pulse amplitude, 10.9 kV, is obtained at 6 modules with a compression stage and a gate driver input pulse width of 107 ns. Both pulsers can operate at a frequency of 1 kHz with no pulse degradation and the MOSFET-DSRD-based pulser can

operate up to 10 kHz. The MOSFET-DSRD-based pulser also has a smaller form factor which makes it more compatible with a compact DWA structure. Considering these results, the MOSFET-DSRD-based pulser appears more suitable for the DWA and may be relevant for future designs. However, one must ensure that parasitic coupling present between the PCBs of the MOSFET-DSRD-based pulser is properly characterized or minimized. Sentaurus TCAD simulations demonstrate that the p doped region of the DSRD should be larger than the n doped region to reduce the pedestal effect. Additionally, the electrical connections to the DSRD should be made as close and tight as possible to reduce parasitic capacitance which reduces the pulse amplitude and increases the pulse width. Simulations with SiC DSRDs show that the pulse amplitude is reduced due to incomplete ionization and the pulse width is increased. Considering the criteria above, these issues should be resolved prior to the use of SiC DSRDs in a pulser for use in a DWA.

CRediT authorship contribution statement

Julien Bancheri: Conceptualization, Methodology, Investigation, Visualization, Writing - Original draft **Andrew Currell:** Methodology, Investigation, Resources, Writing - Review & editing **Anatoly Krasnykh:** Conceptualization, Methodology, Supervision, Writing - Review & editing **Morgan Maher:** Conceptualization, Writing - Review & editing **Christopher Lund:** Conceptualization, Writing - Review & editing **Alaina Giang Bui:** Writing - Review & editing **Jan Seuntjens:** Conceptualization, Writing - Review & editing, Supervision, Funding acquisition

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available upon request.

Acknowledgements

Julien Bancheri is a recipient of the Natural Sciences and Engineering Research Council Canada (NSERC) CGSD award (NSERC PGSD3-559436-2021). This work was supported in part by Canadian Institutes of Health Research (CIHR) Project Grant CIHR PJT 183753.

References

[1] H. Akiyama, T. Sakugawa, T. Namihira, K. Takaki, Y. Minamitani, and N. Shimomura, "Industrial applications of pulsed power technology," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 14, no. 5, pp. 1051–1064, 2007.

[2] I. D. Smith, "Induction voltage adders and the induction accelerator family," *Phys. Rev. ST Accel. Beams*, vol. 7, p. 064 801, 6 Jun. 2004.

[3] S. N. Rukin, "Pulsed power technology based on semiconductor opening switches: A review," *Review of Scientific Instruments*, vol. 91, no. 1, p. 011 501, Jan. 2020.

[4] I. Grekhov, V. Efanov, A. Kardo-Sysoev, and S. Shenderey, "Power drift step recovery diodes (DSRD)," *Solid-State Electronics*, vol. 28, no. 6, pp. 597–599, 1985.

[5] I. V. Grekhov and G. A. Mesyats, "Nanosecond semiconductor diodes for pulsed power switching," *Physics-Uspekhi*, vol. 48, no. 7, p. 703, Jul. 2005.

[6] I. Grekhov and G. Mesyats, "Physical basis for high-power semiconductor nanosecond opening switches," *IEEE Transactions on Plasma Science*, vol. 28, no. 5, pp. 1540–1544, 2000.

[7] X. Yang, Y. Li, H. Wang, Z. Li, and Z. Ding, "Numerical investigation of the nanosecond opening-mechanism of drift step recovery diodes," *Journal of Applied Physics*, vol. 109, no. 1, p. 014 917, Jan. 2011.

[8] A. S. Kyuregyan, "Theory of drift step-recovery diodes," *Technical Physics*, vol. 49, no. 6, pp. 720–727, Jun. 2004.

[9] A. Kyuregyan, "High-voltage diffused step recovery diodes: I. Numerical simulations," *Semiconductors*, vol. 53, no. 7, pp. 962–968, Jul. 2019.

[10] A. S. Kyuregyan, "High-voltage diffused step recovery diodes: II. Theory," *Semiconductors*, vol. 53, no. 7, pp. 969–974, Jul. 2019.

[11] I. V. Grekhov, P. A. Ivanov, D. V. Khristyuk, A. O. Konstantinov, S. V. Korotkov, and T. P. Samsonova, "Sub-nanosecond semiconductor opening switches based on 4H–SiC p+pon+-diodes," *Solid-State Electronics*, vol. 47, no. 10, pp. 1769–1774, 2003, Contains papers selected from the 12th Workshop on Dielectrics in Microelectronics.

[12] P. A. Ivanov, O. I. Kon'kov, T. P. Samsonova, A. S. Potapov, and I. V. Grekhov, "Dynamic characteristics of 4H-SiC drift step recovery diodes," *Semiconductors*, vol. 49, no. 11, pp. 1511–1515, Nov. 2015.

[13] A. V. Afanasyev *et al.*, "Temperature dependence of minority carrier lifetime in epitaxially grown p+-p—n+ 4H-SiC drift step recovery diodes," in *Silicon Carbide and Related Materials 2014*, ser. Materials Science Forum, vol. 821, Trans Tech Publications Ltd, Jul. 2015, pp. 632–635.

[14] Z. Yang, L. Liang, and X. Yan, "Characteristics comparison of SiC and Si drift step recovery diode," *Power Electronic Devices and Components*, vol. 6, p. 100 042, 2023.

[15] R. J. Briggs, "Principles of induction accelerators," in *Induction Accelerators*, K. Takayama and R. J. Briggs, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 23–37.

[16] M. Durante and H. Paganetti, "Nuclear physics in particle therapy: A review," *Reports on Progress in Physics*, vol. 79, no. 9, p. 096 702, Aug. 2016.

[17] N. Esplen, M. S. Mendonca, and M. Bazalova-Carter, "Physics and biology of ultra-high dose-rate (flash) radiotherapy: A topical review," *Physics in Medicine & Biology*, vol. 65, no. 23, 23TR03, Dec. 2020.

[18] E. C. Daugherty *et al.*, "FLASH radiotherapy for the treatment of symptomatic bone metastases (FAST-01): Protocol for the first prospective feasibility study," *JMIR Res Protoc*, vol. 12, e41812, Jan. 2023.

[19] R. Mohan, "A review of proton therapy – current status and future directions," *Precision Radiation Oncology*, vol. 6, no. 2, pp. 164–176, 2022.

[20] S. Yan, T. A. Ngoma, W. Ngwa, and T. R. Bortfeld, "Global democratisation of proton radiotherapy," en, *The Lancet Oncology*, vol. 24, no. 6, e245–e254, Jun. 2023.

[21] A. Pavlovskii, A. Gerasimov, D. Zenkov, V. Bosamykin, A. Klement'ev, and V. Tananakin, "An iron-free linear induction accelerator," *Soviet Atomic Energy*, vol. 28, no. 5, Jan. 1970.

[22] A. Krasnykh, "Analysis of modulator schemes for nanosecond induction linacs," in *Proc. of the Meeting on Problems at Collective Method Acceleration*, 1982, p. 136.

[23] A. Krasnykh, "High gradient linear collider," in *Proc. of the 11th All Union Meeting on Charge Particle Accelerators*, vol. 2, 1988, p. 103.

[24] G. J. Caporaso, Y.-J. Chen, and S. E. Sampayan, "The dielectric wall accelerator," *Reviews of Accelerator Science and Technology*, vol. 02, no. 01, Jan. 2009.

[25] M. Maher, C. Lund, J. Bancheri, D. Cooke, and J. Seuntjens, "An exploration of annular parallel plate waveguides for electric field generation in a dielectric wall accelerator," *Radiotherapy and Oncology*, vol. 186, S117, Sep. 2023.

[26] S. Sampayan, P. Vitello, M. Krogh, and J. Elizondo, "Multilayer high gradient insulator technology," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 7, no. 3, pp. 334–339, Jun. 2000.

[27] F. Arntz, A. Kardo-Sysoev, and A. Krasnykh, "Slim, short-pulse technology for high gradient induction accelerators," *Particle Accelerators*, Dec. 2008.

[28] C. Lund *et al.*, "A novel approach to longitudinal beam stability in a compact dielectric wall accelerator for proton therapy," *Radiotherapy and Oncology*, vol. 186, S15–S16, Sep. 2023.

[29] A. Krasnykh, A. Benwell, T. Beukers, and D. Ratner, "R&D at SLAC on Nanosecond-Range Multi-MW Systems for Advanced FEL Facilities," in *Proc. of International Free Electron Laser Conference (FEL'17), Santa Fe, NM, USA, August 20-25, 2017*, (Santa Fe, NM, USA), ser. International Free Electron Laser Conference, Geneva, Switzerland: JACoW, Feb. 2018, pp. 404–406.

[30] A. G. Lyublinsky, S. V. Korotkov, Y. V. Aristov, and D. A. Korotkov, "Pulse power nanosecond-range DSRD-based generators for electric discharge technologies," *IEEE Transactions on Plasma Science*, vol. 41, no. 10, pp. 2625–2629, 2013.

[31] L. M. Merensky, A. S. Kesar, and A. F. Kardo-Sysoev, "Efficiency study of a 2.2-kV, 1-ns, 1-MHz pulsed power generator based on a drift-step-recovery diode," in *2013 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS 2013)*, 2013, pp. 1–4.

[32] V. S. Belkin and G. I. Shulzchenko, "Forming of powerful nanosecond and sub-nanosecond pulses using standard power rectifying diodes," Budker Institute Nuclear Physics, Novosibirsk, Russia, 1991.

[33] X. Fang *et al.*, "Driving mechanism of drift-step-recovery diodes," *Review of Scientific Instruments*, vol. 92, no. 8, p. 084 702, Aug. 2021.

[34] V. A. Kozlov, A. F. Kardo-Sysoev, and V. I. Brylevskii, "Impact ionization wave breakdown of drift step recovery diodes," *Semiconductors*, vol. 35, no. 5, pp. 608–611, May 2001.

[35] A. Kardo-Sysoev, M. Cherenev, A. Lyublinsky, and M. Vexler, "New concept of two-cascade energy compression systems based on drift step recovery diodes," *Microelectronic Engineering*, vol. 284-285, p. 112 126, 2024.

[36] A. Krasnykh, "Performances of induction system for nanosecond mode operation," May 2006.

[37] A. S. Kesar, "A compact, 10-kV, 2-ns risetime pulsed-power circuit based on off-the-shelf components," *IEEE Transactions on Plasma Science*, vol. 46, no. 3, pp. 594–597, 2018.

[38] R. Sun *et al.*, "10-kV 4H-SiC drift step recovery diodes (DSRDs) for compact high-repetition rate nanosecond hv pulse generator," in *2020 32nd International Symposium on Power Semiconductor Devices and ICs (ISPSD)*, Sep. 2020, pp. 62–65.

[39] K. Mayaram, C. Hu, and D. Pederson, "Oscillations during inductive turn-off in rectifiers," *Solid-State Electronics*, vol. 43, no. 3, pp. 677–681, 1999.

[40] S. A. Darznek, S. K. Lyubutin, S. N. Rukin, and B. G. Slovikovskii, "Generation of microwave oscillations in a no-base diode," *Semiconductors*, vol. 36, no. 5, pp. 599–604, May 2002.

[41] H.-J. Schulze and B. Kolbesen, "Influence of silicon crystal defects and contamination on the electrical behavior of power devices," *Solid-State Electronics*, vol. 42, no. 12, pp. 2187–2197, 1998.

[42] C. G. B. Garrett and W. H. Brattain, "Some Experiments on, and a Theory of, Surface Breakdown," *Journal of Applied Physics*, vol. 27, no. 3, pp. 299–306, Mar. 1956.

[43] R. J. Feuerstein and B. Senitzky, "Surface breakdown of silicon," *Journal of Applied Physics*, vol. 70, no. 1, pp. 288–298, Jul. 1991.

[44] Y. Sharabani, Y. Rosenwaks, and D. Eger, "Mechanism of fast current interruption in $p-\pi-n$ diodes for nanosecond opening switches in high-voltage-pulse applications," *Phys. Rev. Appl.*, vol. 4, p. 014 015, 1 Jul. 2015.

[45] H. Benda and E. Spenke, "Reverse recovery processes in silicon power rectifiers," *Proceedings of the IEEE*, vol. 55, no. 8, pp. 1331–1354, 1967.

[46] A. A. Smirnov and S. A. Shevchenko, "A study of the influence of forward bias pulse duration on the switching process of a 4H-SiC drift step recovery diode," *Journal of Physics: Conference Series*, vol. 993, no. 1, p. 012 033, Mar. 2018.

8

A semi-analytical procedure to determine the ion recombination correction factor in high dose-per-pulse beams

Julien Bancheri and Jan Seuntjens

Article published in: Bancheri J, Seuntjens J. A semi-analytical procedure to determine the ion recombination correction factor in high dose-per-pulse beams. *Med Phys.* 2024; 51: 4458–4471. https://doi.org/10.1002/mp.17005

8.1 Preface

In Chapter 4, the previous investigations into the theory of ion recombination were discussed. Concerning pulsed beams, the pulsed Boag theories appear to fail in the intermediate DPP range (10-100 mGy DPP). This is because they do not account for the space charge due to the free electron collection. Furthermore, the TVM and the Jaffé plot extrapolation fail above 10 mGy DPP, which are based on theory. Currently, there is no existent analytical formula for the ion recombination correction factor in UHDR beams. This carries over to evaluation methods such as the TVM. As a consequence, the current methods to determine the correction factor in UHDR beams, numerical solutions or metrological, are not suitable for the clinical environment. In this chapter, we use a semi-analytical solution method to solve the charge carrier and electric field PDEs to derive an expression for the ion recombination correction factor and use it to develop an extrapolation procedure.

8.2 Abstract

Background: The conventional theories and methods of determining the ion recombination correction factor, such as Boag theory and the related two voltage method and Jaffé plot extrapolation, do not seem to yield accurate results in FLASH/high dose per pulse (DPP) beams ($>$ 10 mGy DPP). This is due to the presence of a large free electron fraction that distorts the electric field inside the chamber sensitive volume. To understand the influence of these effects on the ion recombination correction factor and to develop new expressions for it, it is necessary to re-visit the underlying physics.

Purpose: To present a mathematical procedure to develop an analytical expression for the ion recombination correction factor. The expression is the basis for an extrapolation method

so the correction factor can be determined in a clinical setting.

Methods: A semi-analytical solution method, the homotopy perturbation method (HPM), is used to solve the partial differential equations (PDEs) describing the charge carrier physics, including space charge and free electrons. The electron velocity and attachment rate are modelled as functions of the electric field strength. An expression for the charge collection efficiency and ion recombination correction factor are developed. A fit procedure based on this expression is used to compare it to measured data from previously published articles. Another fit procedure using a general equation is also proposed and compared to the data.

Results: The series obtained for the charge collection efficiency and the ion recombination correction factor are determined to be asymptotic series and the optimal truncation established. The ion recombination correction factor exhibits a $1/V^2$ dependency due to the free electron presence. The fit using this expression agrees well with measured data as long as (1) the DPP is below 1 Gy for chambers with a 1 mm plate separation and (2) when the DPP is below 3 Gy for chambers with a 0.5 mm plate separation. In these DPP ranges, the deviation between measured and fit value did not exceed 6%. In both chamber cases the voltage range where the fit applies decreases as DPP increases. The general equation yielded comparable results.

Conclusions: The HPM was shown to be applicable to a complex system of PDEs and generate meaningful and novel solutions, as they include both space charge and free electrons. The HPM also lends itself to other chamber geometries. The fit procedure was also shown to yield accurate results for the ion recombination correction up to the 1 Gy DPP level.

8.3 Introduction

The clinical reference dosimetry of clinical photon and electron radiotherapy beams involves the determination of the absorbed dose to water, D_w. Protocols such as the AAPM TG-51 [1] and the IAEA TRS-398 [2], which outline the procedure for measuring D_w, are based on ionization chamber measurements. In this case, the raw ionization chamber reading must be corrected through correction factors. These correction factors correct for a variety of chamber non-idealities. One such correction factor is the ion recombination correction factor, often denoted as P_{ion} or k_s, which corrects for the incomplete collection of charge initially released by the radiation beam. When radiation is incident upon an ionization chamber, the radiation interacts with the molecules of the chamber sensitive volume medium. Bound electrons are freed and eventually re-attach to other molecules (due to their electro-negativity). The final products of the interaction are positive and negative ions. In an ideal case, 100% of these charge carriers are collected and the charge reading requires no correction. In practice, opposite signed charge carriers can recombine, decreasing the amount of charge collected and resulting in an underestimated signal.

Recombination can be split into two categories; initial recombination and volume recombination. Initial recombination concerns the recombination of charge carriers generated by the same radiation track. Therefore, it is considered to be dose-rate independent but dependent on linear energy transfer (LET). The theoretical study of initial recombination has been undertaken by Jaffé [3] and Kara-Michailova and Lea [4]. For clinically relevant photon and electron beams, the initial recombination correction factor has been shown to be small and therefore need not be considered [5]. We assume that the same conclusion holds for electron FLASH dose rates and initial recombination will not be considered henceforth [6, 7].

Volume recombination concerns the recombination of charge carriers originating from different tracks. It is considered dose-rate dependent and the main concern in clinical photon and electron beams. The theory of volume recombination has been studied from the standpoint

of charge carrier transport within an irradiated ionization chamber [8–15]. Particularly, Boag studied the recombination of positive and negative ions in continuous and pulsed beams and developed expressions for the ion recombination correction factor [11, 12]. From these expressions, the two voltage method (TVM) [16] was developed and is now widely used in many dosimetry protocols such as TG-51 and TRS-398. The Jaffé plot extrapolation method, where $1/Q$ is plotted against $1/V$ and extrapolated to $1/V = 0$, is based on these expressions [17]. It was also understood at the time that free electrons contribute to the measured charge. These free electrons, released through the molecule-radiation interaction, do not re-attach to the medium molecules and are directly collected, decreasing the negative ion concentration, decreasing ion recombination and increasing the charge collection efficiency. Formulae for the recombination correction factor in pulsed beams including the effect of free electron collection were proposed by Boag and Hochhäuser [18]. Several approaches that incorporate these formulae have been published, aiming to go beyond the TVM [19–21]. Despite these attempts, the modified Boag formulae have been shown to be inaccurate at high chamber voltages (> 100 V) and in high dose per pulse (DPP) beams, above 100 mGy DPP [22, 23]. Additionally, they neglect the distortion of the electron field by the presence of free electron induced space charge at high DPP [24, 25]. As a consequence of the electric field distortion and free electron collection, the TVM and Jaffé plot extrapolation do not give accurate results above 10 mGy DPP [22, 24]. This is a critical issue due to the increasing proliferation of FLASH radiotherapy and the dosimetry that necessarily follows [26].

Many studies have been published on determining the ion recombination correction factor in high DPP beams. They can be broadly categorized into three categories; (1) Numerical solutions of the ion transport equations [23–25], (2) simulations of the charge carrier transport physics [6, 27] and (3) metrological-based, such as the use of ultra-thin parallel plate ionization chambers [28] or using an independent modality (from chambers) to determine dose to water, such as alanine. While the first two categories offer quite accurate results, the methods employed are not suitable for quick evaluation in a clinical setting. Ultra-thin

parallel plate chambers are currently not widely available. In contrast to the above methods, a fourth method can be conceived; studies based on underlying physics resulting in analytical expressions. Analytical expressions offer several advantages; they can be evaluated much more quickly, can be used to develop methods such as the TVM and Jaffé plot extrapolation and all the parameters involved gain physical meaning, since physics assumptions underlie the theory behind the analytical expression. An eventual extrapolation technique is thus much better guided by this approach. However, this way of approaching recombination at high DPP remains quite difficult, mostly due to the complexity of the transport equations when considering free electrons and space charge (non-linear, coupled system of partial differential equations). Different theoretical analyses have been used to approach this problem [29, 30]. In particular, Chabod utilizes perturbation theory in order to solve the charge carrier transport partial differential equations. He considers only positive and negative ions and three scenarios; no space charge and no time dependence [31], space charge and no time dependence [32] and no space charge and time dependence [33].

The purpose of this study is to present a mathematical procedure to develop an analytical expression for the ion recombination correction factor. A semi-analytical solution method is used to solve the system of charge carrier transport partial differential equations (PDEs) in a parallel plate ionization chamber in the presence of a high DPP radiotherapy beam. Crucial effects such as the beam time profile, free electrons and space charge are included. From the solutions of the system of PDEs, an analytical expression for the recombination correction factor can be obtained. The chamber voltage dependence of this expression is established so an extrapolation method can be developed and compared to previously published data. An experimental clinical procedure is also proposed.

8.4 Methods

8.4.1 Charge carrier physics

The physics of ion recombination are modelled by a system of non-linear, coupled partial differential equations. The system of PDEs describing the time evolution of the positive ion concentration n_+, negative ion concentration n_-, free electron concentration n_e, electric field strength E and voltage V are

$$\frac{\partial n_+}{\partial t} = R(x,y,z,t) - \alpha n_+ n_- - \beta n_+ n_e - \mu_+ \boldsymbol{\nabla} \cdot (n_+ E) \tag{8.1}$$

$$\frac{\partial n_-}{\partial t} = \gamma n_e - \alpha n_+ n_- + \mu_- \boldsymbol{\nabla} \cdot (n_- E) \tag{8.2}$$

$$\frac{\partial n_e}{\partial t} = R(x,y,z,t) - \gamma n_e - \beta n_+ n_e + \boldsymbol{\nabla} \cdot (n_e v_e) \tag{8.3}$$

$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = \frac{e}{\epsilon} (n_+ - n_- - n_e) \tag{8.4}$$

$$\vec{\mathbf{E}} = -\boldsymbol{\nabla} V \tag{8.5}$$

where R is the ionization rate of the radiation beam, α is the positive ion - negative ion recombination rate, β is the positive ion-electron recombination rate, γ is the electron attachment rate, μ_i $(i = +, -)$ are the positive and negative ion mobilities, v_e is the electron velocity, e is the electron charge and ϵ is the medium permittivity. The numerical values for α, μ_+, μ_- are taken from Gotz et al. [23]. The values for μ_+ and μ_- were initially taken from Boissonnat et al. [34]. The space charge is modelled through Gauss' law, Eq. 8.4. Some simplifications are made in this work. First, β is set to 0. This is done because the recombination of electrons with positive ions is much lower than that between positive and negative ions [25, 35]. This is due to the electron velocity, which is over 100 times greater than that of ions. Therefore, electrons spend less time around positive ions and are less likely to recombine. Additionally, a parallel plane geometry is considered, with a plane spacing d.

Therefore, $\boldsymbol{\nabla} = \frac{\partial}{\partial z}$. This choice is not a real limitation and one could express $\boldsymbol{\nabla}$ in cylindrical or other coordinates, should the geometry of the problem require this. Lastly, R is considered only as a function of time with no spatial dependence, i.e., $R = R(t)$. In addition to Eqs. 8.1-8.5, Thomson boundary conditions for a parallel plate chamber are prescribed

$$n_+(0,t) = 0 \tag{8.6}$$

$$n_-(d,t) = 0 \tag{8.7}$$

$$n_e(d,t) = 0 \tag{8.8}$$

$$V(0,t) = V_0 \tag{8.9}$$

$$V(d,t) = 0 \tag{8.10}$$

where V_0 is the positive operating voltage of the chamber applied at electrode position $z = 0$. Condition 8.6 implies no positive ions are collected at the positive electrode and conditions 8.7-8.8 imply no negative ions and electrons are collected at the ground electrode at $z = d$.

Electron velocity attachment rate

The electron velocity v_e is in fact a function of the electric field strength E within the ionization chamber. Therefore, a constant mobility cannot be used to model the electron velocity. In similar studies, it is customary to express the electron velocity in terms of a saturation equation with exponentials. In this work, the electron velocity in air model is taken directly from [36], given by

$$v_e(E) = v_1(1 - e^{-E/E_1}) + v_2(1 - e^{-E/E_2}) \tag{8.11}$$

where

$$v_1 = 1.717 \times 10^5 \text{ m/s} \tag{8.12}$$

$$v_2 = 6.499 \times 10^3 \text{ m/s} \tag{8.13}$$

$$E_1 = 28.19 \text{ kV/cm} \tag{8.14}$$

$$E_2 = 0.1272 \text{ kV/cm} \tag{8.15}$$

The parameters given in Eqs. 8.12-8.15 are fit parameters to data from Biagi-v8.9 Boltzamnn calculations in dry air. The goodness of fit is $\chi^2/\nu = 5.3$. Equation 8.11 is plotted in Figure 8.1a.

The electron attachment rate γ is also a function of the electric field strength. Based on measurement data in humid air from Hochhäuser et al. (60% humidity) [37] and Boissonnat et al. (2% humidity) [34], the following model is proposed;

$$\gamma(E) = a\left(1 + \frac{E}{b}\right)e^{-E/c} + \gamma_0 \tag{8.16}$$

where

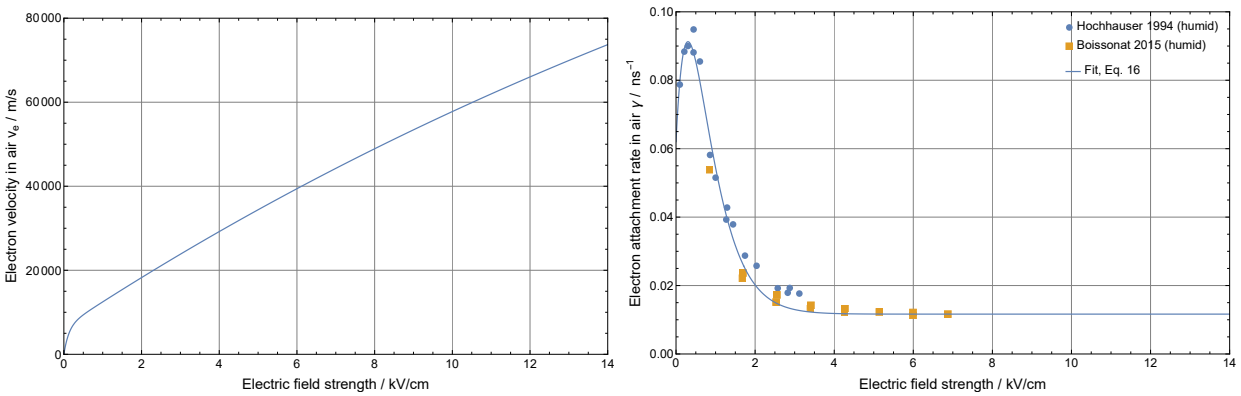
$$a = 0.05041 \text{ ns}^{-1} \tag{8.17}$$

$$\gamma_0 = 0.01165 \text{ ns}^{-1} \tag{8.18}$$

$$b = 0.1433 \text{ kV/cm} \tag{8.19}$$

$$c = 0.4463 \text{ kV/cm.} \tag{8.20}$$

The parameters a, b and c are best fit parameters and the fit was forced to take the value of γ_0 at high field strengths. The form of the model was chosen as a comprise between compactness and goodness of fit to the data. The experimental data and Eq. 8.16 are plotted in Figure 8.1b. The goodness of fit is $\chi^2/\nu = 0.00045$.



(a) (b)

Figure 8.1: (a) Electron velocity in air as a function of the electric field strength, based on data from Boissonnat [36]. (b) Measured electron attachment rate in (humid) air from Hochhäuser et al. [37] and Boissonnat et al. [34]. The fit is modelled according to Eq. 8.16.

8.4.2 Homotopy Perturbation Method

Semi-analytical solution methods are solution techniques that result in power series or even sometimes closed form solutions, They differ from numerical methods as they require no discretization or linearization. The Homotopy Perturbation Method (HPM) is one such semi-analytical method that can be used to solve non-linear ordinary or partial differential equations, including systems of differential equations. The HPM was first proposed by J. He in 1999 [38]. A description of the HPM is now given, following [39]: A general differential equation is written as

$$A(u) - f(z) = 0, \quad z \in \Omega, \tag{8.21}$$

where A is a general differential operator, $f(z)$ is a known analytic function and $u(z)$ is the solution sought on the domain Ω whose boundary is Γ. The boundary conditions are expressed as

$$B\left(u, \frac{\partial u}{\partial n}\right) = 0, \quad z \in \Gamma, \tag{8.22}$$

where B is a boundary operator. Here, $\frac{\partial}{\partial n}$ denotes the derivative in the direction of the

normal to the boundary Γ.

The operator A is separated into a linear part L and a non-linear part N. Therefore, Eq. 8.21 is re-written as

$$L(u) + N(u) - f(z) = 0. \tag{8.23}$$

Using an artificial parameter $p \in [0, 1]$, the homotopy $H(w, p)$ is then constructed as

$$H(w, p) = 0 = (1 - p)[L(w)] + p[L(w) + N(w) - f(z)] \tag{8.24}$$

where $w(z, p) : \Omega \times [0, 1] \to \mathbb{R}$. Note that the new function w is dependent on *both* z and p. It can be seen that by setting $p = 0$ in Eq. 8.24

$$H(w, 0) = L(w) = 0 \tag{8.25}$$

which is a linear approximation to the full differential equation, still satisfying the boundary conditions Eq. 8.22. With $p = 1$,

$$H(w, 1) = L(w) + N(w) - f(z) = 0 \tag{8.26}$$

the full differential equation is returned, which is just Eq. 8.23. Therefore, varying p from 0 to 1 results in deforming the linear approximation to the full solution, that is, varying from $w(z, 0)$ to $w(z, 1) = u(z)$. In mathematical terms, $L(u)$ and $A(u) - f(z)$ are homotopic to one another. The parameter p is used to construct the solution of Eq. 8.24 as

$$w(z, p) = w_0(z) + p^1 w_1(z) + p^2 w_2(z) + ... = \sum_{n=0}^{\infty} p^n w_n(z). \tag{8.27}$$

With Eq. 8.27, Eq. 8.24 is expanded in powers of p. The n^{th} order term (p^n), once solved, yields w_n. The full solution to Eq. 8.21 is then

$$u(z) = \lim_{p\to 1} w(z,p) = w_0(z) + w_1(z) + w_2(z) + ... = \sum_{n=0}^{\infty} w_n(z). \tag{8.28}$$

In applying the HPM to Eqs. 8.1-8.5, four homotopies are constructed, one for each charge carrier PDE and one for Gauss' law

$$H_V(w_+, w_-, w_e, \tilde{V}, p) = 0 = (1-p)[-\frac{\partial^2 \tilde{V}}{\partial z^2}] + p[-\frac{\partial^2 \tilde{V}}{\partial z^2} - \frac{e}{\epsilon}(w_+ - w_- - w_e)] \tag{8.29}$$

$$H_+(w_+, w_-, w_e, \tilde{E}, p) = 0 = (1-p)[\mu_+ \frac{\partial(w_+\tilde{E})}{\partial z}] + p[\frac{\partial w_+}{\partial t} - R(t) + \mu_+ \frac{\partial(w_+\tilde{E})}{\partial z} + \alpha w_+ w_-] \tag{8.30}$$

$$H_-(w_+, w_-, w_e, \tilde{E}, p) = 0 = (1-p)[-\mu_- \frac{\partial(w_-\tilde{E})}{\partial z}] + p[\frac{\partial w_-}{\partial t} - \mu_- \frac{\partial(w_-\tilde{E})}{\partial z} + \alpha w_+ w_- - \gamma w_e] \tag{8.31}$$

$$H_e(w_+, w_-, w_e, \tilde{E}, p) = 0 = (1-p)[-\frac{\partial(w_e v_e)}{\partial z}] + p[\frac{\partial w_e}{\partial t} - R(t) - \frac{\partial(w_e v_e)}{\partial z} + \gamma w_e] \tag{8.32}$$

where

$$V(x,t) = \lim_{p\to 1} \tilde{V}(x,t,p) \tag{8.33}$$

$$E(x,t) = \lim_{p\to 1} \tilde{E}(x,t,p) \tag{8.34}$$

$$n_+(x,t) = \lim_{p\to 1} w_+(x,t,p) \tag{8.35}$$

$$n_-(x,t) = \lim_{p\to 1} w_-(x,t,p) \tag{8.36}$$

$$n_e(x,t) = \lim_{p\to 1} w_e(x,t,p). \tag{8.37}$$

As described in Section 8.4.1, the electron velocity v_e and electron attachment rate γ are functions of the electric field strength. Therefore, they must also be expanded in powers of p in order to be used in the above homotopies. As seen in Eqs. 8.11 and 8.16, the electric

field strength appears in an exponential term. These exponential terms can be expanded in powers of p with exponential Bell polynomials $B_{n,k}\left(x_1, ..., x_{n-k+1}\right)$. For example,

$$e^{-\tilde{E}/c} = e^{-\frac{1}{c}\sum_{n=0}^{\infty} p^n \tilde{E}_n} \tag{8.38}$$

$$= e^{-\frac{\tilde{E}_0}{c}} e^{-\frac{1}{c}\sum_{n=1}^{\infty} p^n \tilde{E}_n} \tag{8.39}$$

$$= e^{-\frac{\tilde{E}_0}{c}} e^{-\sum_{n=1}^{\infty} \frac{p^n}{n!}\frac{(n!\tilde{E}_n)}{c}} \tag{8.40}$$

$$= e^{-\frac{\tilde{E}_0}{c}} \sum_{n=0}^{\infty} \frac{p^n}{n!} \sum_{k=0}^{n} \left(-\frac{1}{c}\right)^k B_{n,k}\left(1!\ \tilde{E}_1, ..., (n-k+1)!\ \tilde{E}_{n-k+1}\right) \tag{8.41}$$

where

$$\tilde{E}_0 = \frac{V_0}{d} \equiv E_0 \tag{8.42}$$

A Wolfram Mathematica (V13.0) script was written to solve the homotopies 8.29-8.32, up to order p^8. As will be seen in Section 8.5, calculating higher order terms does not improve the agreement between the calculated expression for the recombination correction factor and experimental data. The boundary conditions 8.6-8.8 and 8.10 are applied at every order of p. The boundary condition 8.9 is applied at the zeroth order (p^0) and then set to zero for every other order. The solutions for the charge concentrations, voltage and electric field strength are found through Eqs. 8.33-8.37. The ionization rate function $R(t)$ is a user input, the only requirement being that it is analytic. A benefit of using the HPM is that other geometries, such as spherical or cylindrical geometries, are easily implemented since only the spatial gradient has to be re-written in the appropriate coordinate system. The computations were performed on a desktop computer with an AMD Ryzen 3900X 12 core processor, 3.8 GHz and 128 GB RAM and parallelization in Mathematica is utilized. The HPM computation part, at order p^8, takes less than 6 minutes with an ionization rate function $R(t)$ and 2 minutes with a constant ionization rate n_0. The evaluation of Eq. 8.43 and 8.44 depends on the complexity

of the ionization rate function and the number of terms, so an average of 30 minutes is given. With a constant ionization rate function, the evaluation of the integrals takes less than 2 minutes. Note that once the analytical form of the charge collection efficiency is obtained, the integrals do not have to be re-evaluated. The numerical values of the parameters of the beam under study have to be simply inserted into the formula. In comparison, Christensen et al. [27], who use a physics simulation method, reports that the simulation of pulsed or continuous beams requires several minutes. With this method though, the simulation has to be re-run for a change in any single beam parameter.

8.4.3 Charge collection efficiencies

The negative charge collection efficiency f_- and the electron charge collection efficiency f_e are first defined. They are computed at the positive electrode position as

$$f_- = \frac{\int_{pulse} \mu_- n_-(0,t) E(t)\, dt}{d \int_{pulse} R(t)\, dt} = \frac{Q_e}{Q_0} \tag{8.43}$$

and

$$f_e = \frac{\int_{pulse} v_e(0,t) n_e(0,t)\, dt}{d \int_{pulse} R(t)\, dt} = \frac{Q_-}{Q_0}. \tag{8.44}$$

The integrals in Eqs. 8.43 and 8.44 are meant to be taken over the full pulse width. Each quantity represents the ratio of the amount of the respective charge species collected to the total amount of charge of one sign released by the beam, denoted as Q_0. Note that f_e is by definition the *free electron fraction*, often denoted as p in the literature. The symbol f_e will be kept throughout the rest of the article to avoid confusion with the HPM artificial parameter mentioned above.

The total charge collection efficiency f, which is the inverse of the ion recombination correction factor k_s, is calculated as

$$f = \frac{1}{k_s} = f_e + f_-. \tag{8.45}$$

8.4.4 Experimental data

In order to compare the calculated collection efficiencies and ion recombination correction factors to experimental data, data is culled from various, previously published articles. In this work, the data from Gotz et al. [23] and Kranzer et al. [24] is used. Gotz et al. performed measurements with a PTW Advanced Markus ionization chamber (d = 1 mm) and the Electron Linac of high Brilliance and low Emittance (ELBE) superconducting linear electron accelerator. The 20 MeV electron beam consisted of pulses of width of 3.77 μs, delivering up to 1 Gy DPP. Figure 9 from Gotz et al. presents the ion recombination correction as a function of (inverse) voltage for 1 mGy, 30 mGy, 100 mGy and 300 mGy DPP. Kranzer et al. performed measurements with a PTW Advanced Markus ionization chamber and custom made parallel plate chambers EWC2, EWC1, EWC05 with plate separation separation d = 2 mm, 1 mm and 0.5 mm, respectively. A 24 MeV electron beam was used with a pulse width of 2.5 μs and a 5 Hz repetition rate, delivering up to 3 Gy DPP. Figure 5a and 5b from Kranzer et al. present $Q(300V)/Q(V_0)$ as a function of (inverse) voltage for various DPP levels for the Advanced Markus and the EWC05 chamber, respectively. Here, $Q(V_0)$ is the collected charge at a chamber voltage V_0. Note that this ratio is equivalent to $f(300V)k_s(V_0)$ and approaches $f(300V)$ as $V_0 \to \infty$, since $k_s \to 1$. Figures 6a and 6d present the charge collection efficiency $f(300V)$ vs. DPP for the Advanced Markus and EWC05 chamber, respectively. An uncertainty of 0.5% is assigned to all of the data points used in this study, as a conservative estimate, except for those from Figures 6a and 6d from Kranzer et al. where an uncertainty of 3% is assigned. In practice, uncertainties for the collection efficiency and ion recombination correction factor can be as high as 3-4% [20–22].

8.4.5 Selection of ionization rate function

To solve the homotopies 8.29-8.32, a selection for explicit form of $R(t)$ has to be made. In the studies by Boag et al. [11, 12], the beam is modelled as either a constant dose rate or a delta function pulse, where the dose is delivered instantaneously. For the constant dose rate approximation to be considered valid, the ion collection time (the time for the slowest ion to cross the plate gap) t_{ion} has to be much smaller than the pulse width t_p, $t_{ion} \ll t_p$. For the delta function pulse approximation to be considered valid, the opposite must hold, $t_p \ll t_{ion}$. Considering the values of the pulse widths of the above described beams (on the order of μs) and the usual ion collection times, about 20 μs for a 1 mm plate separation, the delta function pulse appears to be an appropriate choice. In fact, this reasoning motivates the continued use of pulsed Boag theory. However, in high DPP or FLASH beams, due to the large free electron fraction, their presence must also be taken into account. Considering a nominal voltage of 300 V and a plate separation of 1 mm, the electron velocity is about 20000 m/s and the time it takes for an electron to cross the gap, t_e, is 50 ns. This is two orders of magnitude less than t_p. Therefore, the delta function pulse approximation is not suitable with respect to the free electrons. The situation in a high DPP beam is then

$$t_e \ll t_p \ll t_{ion} \tag{8.46}$$

which appears to contradict both the constant dose rate and the delta function pulse approximation. Ideally then, a square pulse would be used for $R(t)$, where the instantaneous dose rate in the pulse is determined by the DPP divided by the pulse width and has a width defined by the pulse width. Unfortunately, an explicit square pulse function cannot be used with the HPM because it is not an analytic function. Approximations to a square pulse, such as sums of error functions or tanh functions, yield integrals (Eqs. 8.43 and 8.44) that cannot be computed symbolically. A *pseudo square pulse approximation* is proposed through the following procedure:

1. Set $R(t) = n_0$ where n_0 is a constant ionization rate.

2. Compute solutions of HPM and the collection efficiency f.

3. In the final expressions, perform the substitution, $n_0 \to c \cdot \frac{DPP}{s}$

where c is a conversion factor from DPP to carrier concentration, determined through the following relation

$$c = \left(\frac{1}{\text{Vol.}}\right)\frac{Q_0/e}{\text{DPP}} = \frac{\rho}{e}\frac{1}{\left(\frac{\bar{W}}{e}\right)_{air}} = 2.21 \times 10^{17}\ \text{Gy}^{-1}\ \text{m}^{-3} \tag{8.47}$$

where Vol. is the volume of the collecting region of the chamber, ρ is the density of air and $(\bar{W}/e)_{air}$ is the average energy deposited per coulomb of charge of one sign released by one electron in air.

The parameter s is necessary to convert from a constant dose rate in step 1 to an instantaneous dose rate of a square pulse in step 3. It encapsulates the discrepancies between carrying out the HPM with a constant dose rate but fitting to data taken with a pulsed beam. It is a phenomenological parameter that depends on the DPP and chamber geometry and must be determined by fitting to experimental data. Since it has units of time, it can be considered as an *effective* pulse width.

8.5 Results

8.5.1 Verification

In order to verify that the HPM provides sensible expressions for the charge collection efficiency, base cases were studied. First, considering a constant dose rate and only positive and negative ions with space charge, the HPM reproduced Eq. 3.1 in the article by Chabod [32]. Second, considering beams with different time profiles and only positive and negative ions with no space charge, the HPM reproduced Eqs. 3.19, 3.21 and 3.22 in the article by Chabod [33].

8.5.2 Charge collection efficiency and free electron fraction expressions

The charge carrier concentrations, electric field strength and the collection efficiency were calculated up to order p^8. The expressions become very large, so for demonstration purposes the 4$^{\text{th}}$ order calculation of f, $f^{(4)}$, is given below in Eq. 8.48.

$$f^{(4)} = 1 - \frac{\alpha d^4 n_0 \, a e^{-E_0/c}}{24 b \mu_- \mu_+ (v_1(1-e^{-E_0/E_1}) + v_2(1-e^{-E_0/E_2}))V_0} - \frac{\alpha d^5 n_0 (a e^{-E_0/c} + \gamma_0)}{24 \mu_- \mu_+ (v_1(1-e^{-E_0/E_1}) + v_2(1-e^{-E_0/E_2}))V_0^2} + \ldots \tag{8.48}$$

and the 2$^{\text{nd}}$ order calculation of f_e, $f_e^{(2)}$, is given below in Eq. 8.49.

$$f_e^{(2)} = \frac{(2v_1(1-e^{-E_0/E_1}) + 2v_2(1-e^{-E_0/E_2})) - a e^{-E_0/c} d - d\gamma_0}{2(v_1(1-e^{-E_0/E_1}) + v_2(1-e^{-E_0/E_2}))} - \frac{a e^{-E_0/c} V_0}{2b(v_1(1-e^{-E_0/E_1}) + v_2(1-e^{-E_0/E_2}))} + \ldots \tag{8.49}$$

There are several interesting features to note; (1) both even and odd powers of V_0 appear in f and f_e, (2) there exist terms in both f and f_e where V_0 appears in the numerator. This does not cause an issue as $V_0 \to \infty$ because they are always accompanied by an exponential factor such as $e^{-E_0/c}$. This is a consequence of the second term in the parentheses of Eq. 8.16, (3) the series solution for f is no longer a simple power series in powers of $1/V_0$ due to the exponential factors that also contain a V_0, (4) though not shown in Eq. 8.49, higher order terms of f_e depend on n_0, which goes against conventional thought, (5) it can be seen that in taking the limit $V_0 \to \infty$

$$f_e \xrightarrow{V_0 \to \infty} 1 - \frac{\gamma_0 d}{2(v_1+v_2)} + \frac{\gamma_0^2 d^2}{6(v_1+v_2)^2} - \frac{\gamma_0^3 d^3}{24(v_1+v_2)^3} + \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n \gamma_0^n d^n}{(n+1)!(v_1+v_2)^n} = \frac{1 - e^{-\gamma_0 d/(v_1+v_2)}}{\frac{\gamma_0 d}{(v_1+v_2)}} \tag{8.50}$$

which is exactly the conventional definition for the free electron fraction given by Boag [18]. Thus, Boag's expression for the free electron fraction can be thought of as a high voltage limit. Another crucial observation is that the $1/V_0$ term of Eq. 8.48 is suppressed by the

exponential factor $e^{-E_0/c}$, whereas the γ_0 term in the parentheses of the second term is not. Thus, it appears as though the collection efficiency follows a $1/V_0^2$ relationship as $V_0 \to \infty$, in contrast to the pulsed Boag theory.

In mathematical terms, the series solutions for f and f_e are *asymptotic series* with the limit point $V_0 = \infty$. A defining feature of an asymptotic series is that for a fixed number of terms in the series, the series converges as the argument approaches the limit point. A consequence of this feature is that as additional terms are added to the series, past an optimal number of terms (optimal truncation), the series diverges very quickly, even for arguments very close to the limit point. The optimal truncation also decreases as the series coefficients increase, which in the current case is when n_0 increases, (when the dose or dose rate increases) and when d increases.

8.5.3 Comparison of analytic expressions to experimental data

Due to the asymptotic nature of the series solution for f, the optimal truncation of the series is determined. The truncated ion recombination correction factors, $k_s^{(n)}$, were calculated by computing the inverses of the n^{th} order collection efficiency $f^{(n)}$ up to order 8, for example,

$$k_s^{(5)} = \frac{1}{f^{(5)}}. \tag{8.51}$$

Fitting the truncated ion recombination correction factor expressions to the data from Figure 9 from Gotz et al. [23] over a limited voltage range yielded better agreement than fitting over the entire voltage range. The expressions were fit down to 200 V for the 30 mGy and 100 mGy data sets and down to 300 V for the 300 mGy data set, shown in Figure 8.2. The data sets from Gotz et al. [23] were chosen to determine the optimal truncation limit because k_s is reported as a function of voltage, whereas the Kranzer et al. [24] article does not.

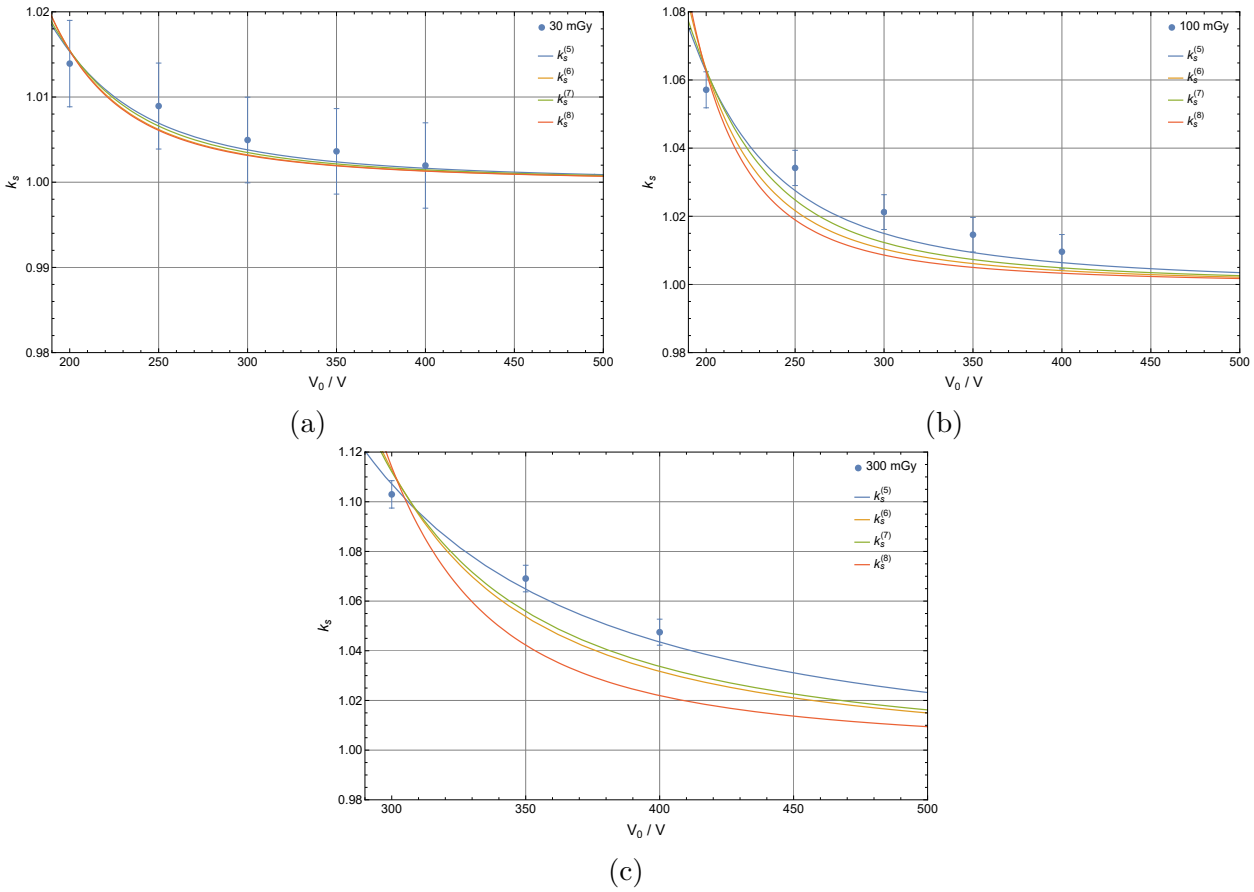


Figure 8.2: Fit of truncated ion recombination correction factor of various orders to the data from Figure 9 from Gotz et al. [23]; down to 200 V for (a) 30 mGy and (b) 100 mGy and down to 300 V for (c) 300 mGy. Note the change in the y-axis scale from left to right.

It is clear from Figure 8.2 that the 5^{th} order truncation yields the best approximation. For the remainder of the study, only this expression is used when considering the expressions derived from the HPM.

As discussed in Section 8.5.2, the collection efficiency appears to follow a $1/V_0^2$ relation as $V_0 \to \infty$. The series expansion of Eq. 8.51 is calculated (without expanding the exponential terms) and only the $1/V_0^0$ and $1/V_0^2$ terms kept. This new expression denoted as $k^{(5)}_{s,1/V^2}$. In order to observe if the $1/V_0^2$ relation holds true, $k^{(5)}_{s,1/V^2}$ is fit down to 200 V for the 30 mGy and 100 mGy data sets and down to 250 V for the 300 mGy data set, shown in Figure 8.3. These voltage cut offs were chosen because they yielded the best fits, in terms of root mean square error (RMSE), compared to other cut off choices. Table 8.1 presents the value of the

fit parameter s, the RMSE and R^2 value of the fit.

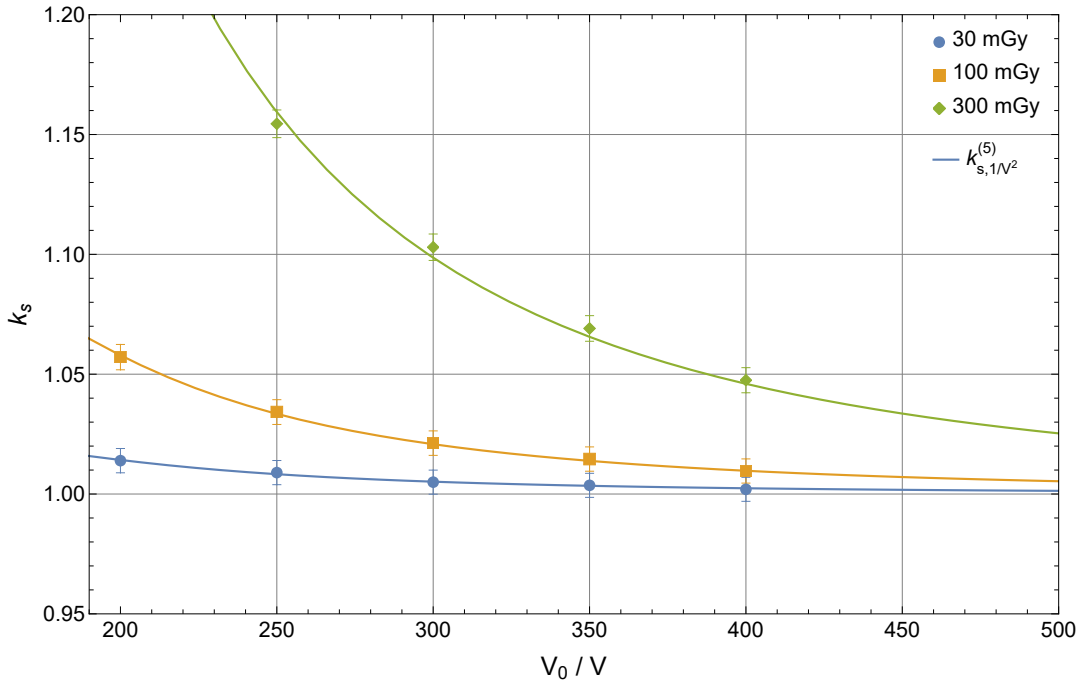


Figure 8.3: Fit of $k^{(5)}_{s,1/V^2}$ expression to data subsets; down to 200 V for (a) 30 mGy and (b) 100 mGy and down to 250 V for (c) 300 mGy. Data is taken from Figure 9 from Gotz et al. [23].

Table 8.1: Value of the fit parameter s, RMSE and R^2 value for the $k^{(5)}_{s,1/V^2}$ fit to the data from Figure 9 from Gotz et al. [23].

DPP (mGy)	s (μs)	RMSE	R^2
30	8.74 ± 0.220	0.0004	1
100	7.21 ± 0.070	0.0006	1
300	4.56 ± 0.098	0.0038	0.9999

Due to the $1/V_0^2$ relationship of k_s, a general expression of the form

$$g + q\left(\frac{1}{V_0}\right)^2 \tag{8.52}$$

was fit to the same data sets. Extrapolating the fit to $1/V_0 = 0$ yields g, which is an estimate of the value of k_s at $V_0 = \infty$, which theoretically equals 1. The fits are shown in Figure 8.4. The deviation of the fit at $1/V_0 = 0$ from 1 is an indication of the accuracy of the fit. The values of the parameter g, the RMSE and R^2 value of the fit are shown in Table 8.2.

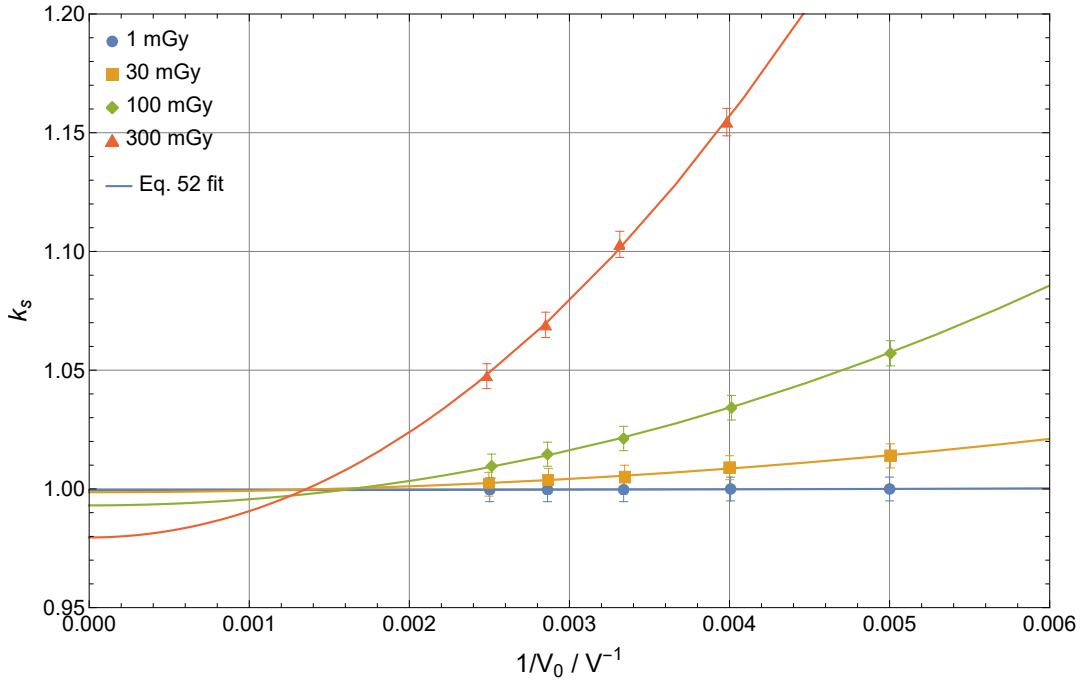


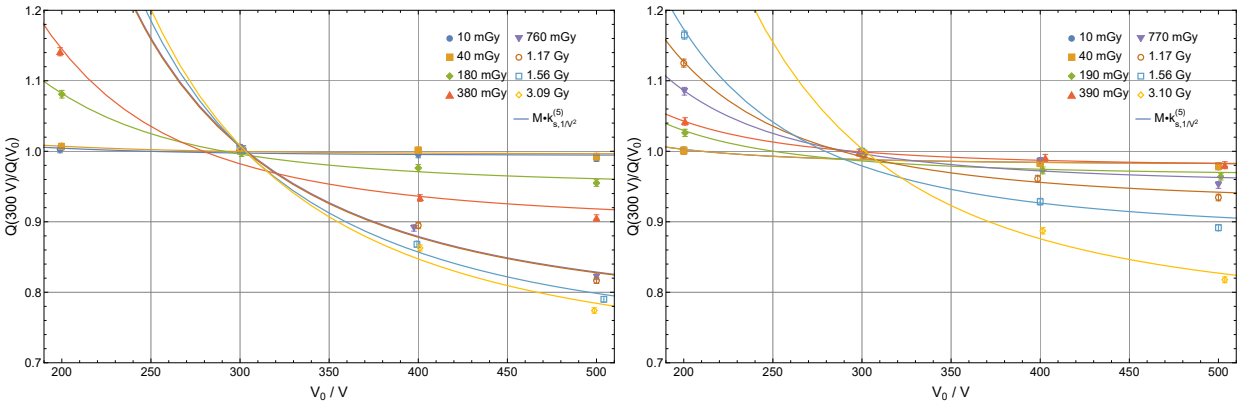
Figure 8.4: Eq. 8.52 fit to the data from Figure 9 from Gotz et al. [23]. The deviation of the fit at $1/V_0 = 0$ from 1 indicates how much the calculated correction factor is expected to deviate from the measured value.

Table 8.2: Extrapolation of the fit in the form of Eq. 8.52 at $1/V_0 = 0$, giving the parameter g. The deviation of g from 1 is an indication of the accuracy of the fit. The RMSE and R^2 value of the fit are given.

DPP (mGy)	g	RMSE	R^2
1	0.99947 ± 0.00005	0.00008	1
30	0.9986 ± 0.0005	0.00085	1
100	0.9931 ± 0.0003	0.00036	1
300	0.980 ± 0.002	0.00100	1

The ratio $Q(300V)/Q(V_0)$ is presented in Figure 5a and 5b of Kranzer et al. [24] for the Advanced Markus and EWC05 chamber, respectively. This ratio is equivalent to $f(300V)k_s(V_0)$. The expression $M \cdot k^{(5)}_{s,1/V^2}$ is fit to the data, where M is a fit parameter. Extrapolation of the fit to $V_0 \to \infty$ yields the parameter M, which is an estimate for $f(300V)$ since $k_s \to 1$ as $V_0 \to \infty$. The fit of $M \cdot k^{(5)}_{s,1/V^2}$ to the data are shown in Figures 8.5a and 8.5b for the Advanced Markus and EWC05 chamber, respectively. In Figure 8.5a, the DPP values up to 380 mGy are fit down to 200 V while the DPP values from 760 mGy to 3.09 Gy are fit down

to 300 V. In Figure 8.5b, the DPP values up to 1.56 Gy are fit down to 200 V while the DPP value of 3.10 Gy is fit down to 300 V. Table 8.3 presents the value of the fit parameter s, the RMSE and R^2 value of the fit. The reasoning for the voltage cut offs is that they yield values for $f(300V)$ that are closest to the experimental values, as will be discussed in the following section.



(a) (b)

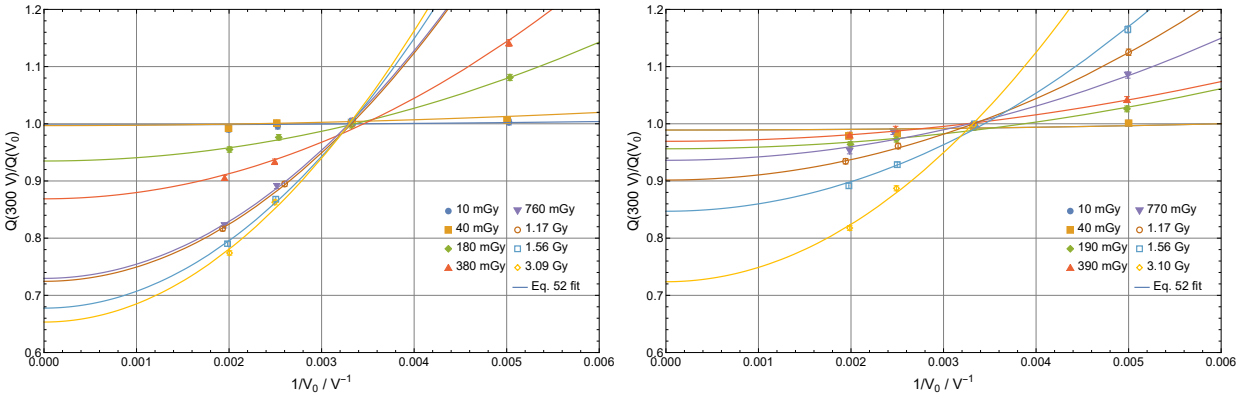
Figure 8.5: $M \cdot k^{(5)}_{s,1/V^2}$ fit to data of $Q(300V)/Q(V_0)$ from Figure 5a and 5b of Kranzer et al.[24], where M is a fit parameter, for (a) the Advanced Markus chamber (d = 1 mm) and (b) the EWC05 (d = 0.5 mm) chamber. This ratio is equivalent to $f(300V)k_s(V_0)$, so M is an estimate of $f(300V)$. In (a), the DPP values up to 380 mGy are fit down to 200 V while the DPP values from 760 mGy to 3.09 Gy are fit down to 300 V. In (b), the DPP values up to 1.56 Gy are fit down to 200 V while the DPP value of 3.10 Gy is fit down to 300 V.

Table 8.3: Value of the fit parameter s, root mean square error (RMSE) and R^2 value for the $M \cdot k^{(5)}_{s,1/V^2}$ fit to the data from Figure 5a and 5b of Kranzer et al. [24]

	Advanced Markus (d = 1 mm)			EWC05 (d = 0.5 mm)		
DPP (mGy)	s (μs)	RMSE	R^2	s (μs)	RMSE	R^2
10	3.83 ± 2.47	0.0035	1	0.0411 ± 0.0153	0.0041	1
40	14.1 ± 7.10	0.0029	1	0.164 ± 0.0611	0.0041	1
180/190	5.45 ± 0.360	0.0042	1	0.258 ± 0.0510	0.0063	1
380/390	5.94 ± 0.679	0.013	0.9998	0.531 ± 0.0373	0.0022	1
760/770	3.63 ± 0.426	0.0073	0.9999	0.497 ± 0.0706	0.0089	0.9999
1170	5.61 ± 1.08	0.012	0.9998	0.490 ± 0.0277	0.0051	1
1560	6.10 ± 0.654	0.0075	0.9999	0.432 ± 0.0452	0.013	0.9998
3099 / 3100	10.9 ± 1.75	0.012	0.9998	0.309 ± 0.0445	0.0088	0.9999

Expressions of the form 8.52 were also fit to the same data, shown in Figure 8.6. Extrapo-

lating the fit to $1/V_0 = 0$ yields g, which is an estimate for $f(300V)$. The fits and data are shown in Figures 8.6a and 8.6b for the Advanced Markus and EWC05 chamber, respectively. The same voltage cut offs were used as in Figure 8.5 except for the 10 and 40 mGy data set, where the whole voltage range was used. Table 8.4, the root mean square error (RMSE) and R^2 value of the fit.

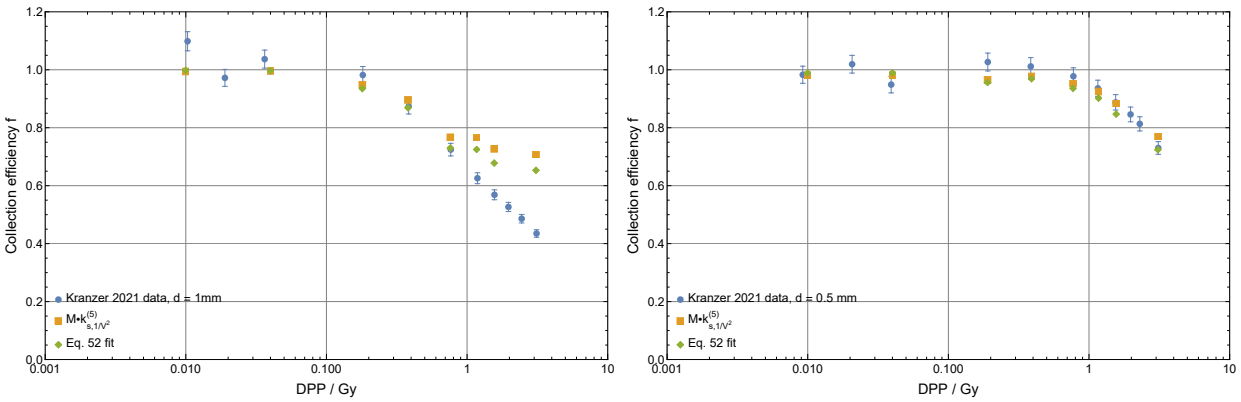


(a) (b)

Figure 8.6: Eq. 8.52 fit to data of $Q(300V)/Q(V_0)$ from Figure 5a and 5b of Kranzer et al.[24] for (a) the Advanced Markus chamber (d = 1 mm) and (b) the EWC05 (d = 0.5 mm) chamber. This ratio is equivalent to $f(300V)k_s(V_0)$, so g is an estimate of $f(300V)$. In (a), the DPP values up to 40 mGy are fit to the whole voltage range, the DPP values of 180 and 380 mGy are fit down to 200 V while the DPP values from 760 mGy to 3.09 Gy are fit down to 300 V. In (b), the DPP values up to 40 mGy are fit to the whole voltage range, the DPP values from 190 to 1.56 Gy are fit down to 200 V while the DPP value of 3.10 Gy is fit down to 300 V.

8.5.4 Determination of collection efficiencies

From Figures 8.5a and 8.5b, the results for the fit parameter M are plotted against the measured collection efficiencies from Figures 6a and 6d from Kranzer et al. [24] in Figures 8.7a and 8.7b for the Advanced Markus chamber (d = 1 mm) and (b) the EWC05 (d = 0.5 mm) chamber, respectively. The results of the Eq. 8.52 fit, Figures 8.6a and 8.6b, are also plotted in Figure 8.7. The percent difference between the measured data and the values obtained with the fit procedures is an indication of whether the fit procedure is accurate at the given DPP value. These percent differences (if measured values above 1 are brought down to a value of 1) are presented in Table 8.5.



(a) (b)

Figure 8.7: The parameter M from the fit $M \cdot k^{(5)}_{s,1/V^2}$ and the parameter g of Eq. 8.52 plotted against the data from Figures 6a and 6d from Kranzer et al. [24] for (a) the Advanced Markus chamber (d = 1 mm) and (b) the EWC05 (d = 0.5 mm) chamber.

Table 8.4: Root mean square error (RMSE) and R^2 value for the Eq. 8.52 fit to the data from Figure 5a and 5b of Kranzer et al. [24]

	Advanced Markus (d = 1 mm)		EWC05 (d = 0.5 mm)	
DPP (mGy)	RMSE	R^2	RMSE	R^2
10	0.0050	1	0.0087	0.9999
40	0.0086	0.9999	0.0087	0.9999
180/190	0.0027	1	0.0048	1
380/390	0.0080	0.9999	0.0027	1
760/770	0.0029	1	0.0090	0.9999
1170	0.0011	1	0.0025	1
1560	0.0040	1	0.0057	1
3090/3100	0.0082	0.9999	0.0056	1

Table 8.5: Percent differences between the charge collection efficiencies determined with the fit procedure using $M \cdot k^{(5)}_{s,1/V^2}$ or Eq. 52 and those from Figures 6a and 6d from Kranzer et al. [24] Measured values above 1 are brought down to a value of 1. Negative values indicate the fit procedure yields a value higher than the measured data point.

	Advanced Markus (d = 1 mm)		EWC05 (d = 0.5 mm)	
DPP (mGy)	$M \cdot k^{(5)}_{s,1/V^2}$	Eq. 52	$M \cdot k^{(5)}_{s,1/V^2}$	Eq. 52
10	0.60%	0.20%	0.19%	-0.63%
40	0.40%	0.30%	-3.4%	-4.2%
180/190	3.3%	4.8%	3.4%	4.3%
380/390	-2.6%	0.53%	2.3%	3.1%
760/770	-5.9%	-0.76%	2.6%	4.3%
1170	-22%	-16%	1.2%	3.6%
1560	-28%	-19%	0.51%	4.6%
3090/3100	-63%	-50%	-5.3%	0.84%

8.6 Discussion

The asymptotic series nature of the series expression of f can be seen in Figure 8.2, where the 5^{th} order truncation provides the best fit to the data. That the series expression for f is an asymptotic series is not an exact statement because the series coefficients are non-constant with respect to the argument V_0. This is due to the electron velocity function v_e and electron attachment function γ being non-polynomial in form. A polynomial expression for these functions would not be suitable because it would be of a large order and over fit the data. The interesting behaviour of this series is perhaps an indication of why it has been difficult thus far to derive an analytic expression for the charge collection efficiency including free electrons based on first principles (excluding those from Boag et al. [18] but all electron dynamics are condensed into a single term p.).

Figures 8.3, 8.4, 8.5 and 8.6 indicate that from a DPP value of 380 mGy and below, the

expression $k^{(5)}_{s,1/V^2}$ fits the Gotz and Kranzer data set well above 200 V. For DPP values above 760 mGy, the expression fits both data sets well above 300 V. From Tables 8.1 and Table 8.3, the R^2 values are all practically 1 in both DPP ranges. The same observation holds for the Eq. 8.52 fit. In contrasting the RMSE between the $k^{(5)}_{s,1/V^2}$ fit and Eq. 8.52 at a given DPP for the Advanced Markus for both data sets, Eq. 8.52 generally yields a lower RMSE. This is because the fit parameters in Eq. 8.52 are not constrained by other factors in the expression, as seen in 8.48. Despite this, referring to Figure 8.7a, using both expressions does not yield accurate results above 760 mGy DPP and therefore should not be used above 1 Gy DPP. For the EWC05 chamber, the RMSE values are similar between the two different fits, except at high DPP, where the Eq. 8.52 fit yields a lower RMSE. The extrapolation method seems to yield accurate results up to 3 Gy DPP. When looking at Tables 8.3 and Table 8.4, the RMSE decreases in going from 380 mGy to 760 mGy DPP. This occurs when the voltage range used in the fitting process is cut from 200 V to 300 V. The different cut off voltages, concerning $k^{(5)}_{s,1/V^2}$, can be understood through it being a series. At higher DPP values the expression diverges at larger voltage values than at lower DPP values (see Eq. 8.48). Using a higher order expansion does not increase the voltage range where the series converges, as demonstrated in Figure 8.2, due it being an asymptotic series. The higher order expansions go beyond the optimal truncation. Concerning Eq. 8.52, higher order terms are required to cover a larger voltage range. The voltage range where both expressions fit the data expands as the plate separation d decreases, as shown in Figures 8.5 and 8.6. This feature can be attributed d appearing in the numerator of the series. Due to the large gap between 380 mGy and 760 mGy, a future study can determine at what, exact DPP the transition from 200 V to 300 V as the voltage cut off occurs.

The fit parameter s, the effective pulse width, appears to decrease with increasing DPP in the 30 - 400 mGy (low) DPP range, as shown in Tables 8.1 and 8.3. In the low DPP range, the value of s is consistent between the Gotz data set and the Kranzer data set. In the 760 mGy - 3.10 Gy (high) DPP range, the values of s are above that of the low DPP range

and appear to increase with DPP. Due to s being a fit parameter, it is dependent on DPP and chamber geometry, which explains why its value does not follow a single trend. Due to this behaviour, it is not currently clear whether this parameter contains any meaningful information.

In Table 8.2, the largest deviation from 1 is 2%, at 300 mGy DPP. This is certainly within realistic type B uncertainties, between 3-4%. Additional data points at higher voltages around 500 V would potentially reduce this deviation. In Figure 8.7a, for the Advanced Markus and below 1 Gy DPP, the largest percent difference between the measured value and that using the $k^{(5)}_{s,1/V^2}$ fit is -5.9% at 760 mGy DPP (if measured values above 1 are brought down to a value of 1). Using Eq. 8.52 yields slightly more accurate results with the largest percent difference between the measured value and fit value being 4.8% at 180 mGy DPP. Above 1 Gy DPP, both fit procedures do not give accurate results. The use of higher order truncations would not improve the accuracy of the calculated k_s at DPP values above 1 Gy due to the nature of asymptotic series. As the series coefficients increase (in this case, DPP) the optimal truncation limit decreases. Referring back to Figure 8.2c, the higher order truncations do not fit the data as well as the fifth order truncation. As the DPP increases, these higher order truncations will fit the data even more poorly. The fourth order truncation results in fits similar to the fifth order truncation and below that $f^{(n)} = 1$. In Figure 8.7b, for the ECW05 chamber, the largest percent difference between the measured value and that using the $k^{(5)}_{s,1/V^2}$ fit is -5.3% at 3.1 Gy DPP, and 4.6% at 1.56 Gy DPP using Eq. 8.52. These percent differences are within realistic type B uncertainties with a coverage factor of $k = 2$. Additionally, both fits give reasonable results up to a high DPP, ≤ 1 Gy for $d = 1$ mm and ≤ 3 Gy for $d = 0.5$ mm. In the case of both chambers, the extrapolation procedure presented in this study gives much more accurate results than that by the usual $1/V$ Jaffé plot extrapolation method, which is accurate only up to 10 mGy DPP [22, 24]. It is currently unclear whether this procedure works for DPP values above 3 Gy for very thin parallel plate chamber ($d \leq 0.5$ mm).

Based on the reported results, an experimental clinical procedure is presented to measure the ion recombination correction factor

1. At a specific DPP, take chamber readings $Q(V_0)$ for various chamber voltages V_0.
2. Plot $1/Q(V_0)$ vs V_0.
3. (a) Fit $M \cdot k^{(5)}_{s,1/V^2}$ to the measured data. The parameter M is an estimate for $1/Q_0$.

   (b) Fit Eq. 8.52 to the measured data. The parameter g is an estimate $1/Q_0$.
4. Calculate $k_s(V_0) = Q_0/Q(V_0)$ for the given DPP and voltage V_0.
5. Keep V_0 constant for subsequent measurements.

Step 3 is presented with either expression but given the reported results and the simplicity of Eq. 8.52 vs. $k^{(5)}_{s,1/V^2}$, it may be simpler to implement this procedure with Step 3b. In order to use the expression for $k^{(5)}_{s,1/V^2}$, the DPP must be determined experimentally. Doing so requires a chamber-independent method such as alanine or a Faraday cup. These methods are generally time consuming in a clinical setting. Additionally, using an incorrect DPP can result in greatly under or over-estimating the recombination correction factor. Of course, this discrepancy depends the chamber voltage, plate separation and DPP regime (high or low). At a low voltage alone, < 300 V, an incorrect DPP results in quite large discrepancies. Considering a usual plate separation of 1 mm, such as that of the Advanced Markus, at low DPP (< 100 mGy) and high voltage, an incorrect DPP will have little impact on the corrected signal. As DPP increases, the recombination correction factor can be off by over $10\% - 20\%$. The limit between high and low DPP also decreases as the plate separation increases. In any case, using Eq. 8.52 eliminates these DPP-related issues. A user should be made aware of the limitations of this procedure; (1) for a chamber with $d = 1$ mm, the fit should be done with voltages above 200 V when the DPP is below 400 mGy and above 300 V when the DPP is between 800 mGy - 1 Gy and (2) for a chamber with $d = 0.5$ mm, the fit should be done with voltages above 200 V when the DPP is below 2 Gy and above

300 V when the DPP is above 3 Gy. Note that the proposed procedure requires chamber measurements at several different voltages. Some clinics have access to electrometers with only two voltage settings. In these scenarios, the proposed fit procedure cannot be performed accurately due the sparsity of the measured data.

A modification to the proposed method presented in this article that would potentially enable the calculated series solution for the charge collection efficiency to be more accurate at high DPP would be to simplify the main problem presented in Eqs 8.1 - 8.10. As currently stated, the proposed HPM generates a global solution in terms of DPP. A simplification, only valid at high DPP, would generate a local solution instead. This assumption would assume that the electrons only experience the electric field near the positive electrode. This is because the electric field is large near the negative electrode, due to space charge, and electrons move away from it quickly due to their large velocity. The resulting solutions would then only be valid in the high DPP regime.

8.7 Conclusion

The aim of this study is to use a semi-analytic solution method to obtain an expression for the ion recombination correction factor in high ($\geq$ 100 mGy) DPP beams and develop a procedure to measure the correction factor in a clinical setting. The homotopy perturbation method was used to solve the partial differential equations describing the charge carrier transport with space charge and the presence of free electrons. A series solution for the charge collection efficiency was obtained and subsequently an expression for the recombination correction factor. It was shown that in high DPP beams, the recombination correction factor exhibits a $1/V_0^2$ dependency, more so than a $1/V_0$ dependency, due to the free electron presence. It was shown that the analytical expression agrees well with measured data from Gotz et al. [23] and Kranzer et al. [24] when (1) the DPP is below 1 Gy for chambers with a 1 mm plate separation and (2) when the DPP is below 3 Gy for chambers with a 0.5 mm plate separation. A general expression containing a $1/V_0^2$ term was also proposed which yielded

similar results as the analytic expression. In both cases, the percent differences between expression and the measured data were below 6%, which is within clinical uncertainties with a coverage factor of $k = 2$. A clinical procedure to measure the ion recombination correction factor was also proposed. A subsequent study applying this methodology to cylindrical chambers would be beneficial, as these types of chambers are the most common in clinical environments.

8.8 Acknowledgements

Thanks to Morgan Maher, Christopher Lund and Alaina Giang Bui for their insightful comments concerning the manuscript. Julien Bancheri is a recipient of the NSERC CGSD award (NSERC PGSD3-559436-2021).

8.9 Conflict of Interest Statement

The authors have no relevant conflicts of interest to disclose.

References

[1] P. R. Almond *et al.*, "AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon and electron beams," *Medical Physics*, vol. 26, no. 9, pp. 1847–1870, 1999.

[2] *Absorbed Dose Determination in External Beam Radiotherapy* (Technical Reports Series 398). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2001.

[3] G. Jaffé, "Zur Theorie der Ionisation in Kolonnen," *Annalen der Physik*, vol. 347, no. 12, pp. 303–344, 1913.

[4] E. Kara-Michailova and D. E. Lea, "The interpretation of ionization measurements in gases at high pressures," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 36, no. 1, pp. 101–126, 1940.

[5] P. B. Scott and J. R. Greening, "The Determination of Saturation Currents in Free-air Ionization Chambers by Extrapolation Methods," *Physics in Medicine & Biology*, vol. 8, no. 1, p. 51, Apr. 1963.

[6] J. B. Christensen *et al.*, "Mapping initial and general recombination in scanning proton pencil beams," *Physics in Medicine & Biology*, vol. 65, no. 11, p. 115 003, Jun. 2020.

[7] S. Rossomme *et al.*, "Ion recombination correction factor in scanned light-ion beams for absolute dose measurement using plane-parallel ionisation chambers," *Physics in Medicine & Biology*, vol. 62, no. 13, p. 5365, Jun. 2017.

[8] J. J. T. M. F.R.S., "XIX. On the theory of the conduction of electricity through gases by charged ions," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 47, no. 286, pp. 253–268, 1899.

[9] G. Mie, "Der elektrische Strom in ionisierter Luft in einem ebenen Kondensator," *Annalen der Physik*, vol. 318, no. 5, pp. 857–889, 1904.

[10] R. Seeliger, "Beitrag zur Theorie der Elektrizitätsleitung in dichten Gasen," *Annalen der Physik*, vol. 338, no. 12, pp. 319–380, 1910.

[11] J. W. Boag, "Ionization measurements at very high intensities—Part I," *The British Journal of Radiology*, vol. 23, no. 274, pp. 601–611, 1950.

[12] J. W. Boag and T. Wilson, "The saturation curve at high ionization intensity," *British Journal of Applied Physics*, vol. 3, no. 7, p. 222, Jul. 1952.

[13] J. R. Greening, "Saturation Characteristics of Parallel-Plate Ionization Chambers," *Physics in Medicine & Biology*, vol. 9, no. 2, p. 143, Apr. 1964.

[14] E. Sprinkle and P. A. Tate, "The saturation curve in cylindrical and spherical ionization chambers," *Physics in Medicine & Biology*, vol. 11, no. 1, pp. 31–46, 1966.

[15] R. Rosen and E. P. George, "Ion distributions in plane and cylindrical chambers," *Physics in Medicine & Biology*, vol. 20, no. 6, p. 990, Nov. 1975.

[16] J. W. Boag and J. Currant, "Current collection and ionic recombination in small cylindrical ionization chambers exposed to pulsed radiation," *British Journal of Radiology*, vol. 53, no. 629, pp. 471–478, Jan. 2014.

[17] D. Burns and M. McEwen, "Ion recombination corrections for the NACP parallel-plate chamber in a pulsed electron beam," *Physics in Medicine & Biology*, vol. 43, no. 8, p. 2033, 1998.

[18] J. W. Boag, E. Hochhäuser, and O. A. Balk, "The effect of free-electron collection on the recombination correction to ionization measurements of pulsed radiation," *Physics in Medicine & Biology*, vol. 41, no. 5, p. 885, May 1996.

[19] A. Piermattei *et al.*, "The saturation loss for plane parallel ionization chambers at high dose per pulse values," *Physics in Medicine & Biology*, vol. 45, no. 7, p. 1869, Jul. 2000.

[20] F. Di Martino, M. Giannelli, A. C. Traino, and M. Lazzeri, "Ion recombination correction for very high dose-per-pulse high-energy electron beams," *Medical Physics*, vol. 32, no. 7Part1, pp. 2204–2210, 2005.

[21] R. F. Laitano, A. S. Guerra, M. Pimpinella, C. Caporali, and A. Petrucci, "Charge collection efficiency in ionization chambers exposed to electron beams with high dose per pulse," *Physics in Medicine & Biology*, vol. 51, no. 24, p. 6419, Nov. 2006.

[22] K. Petersson *et al.*, "High dose-per-pulse electron beam dosimetry — A model to correct for the ion recombination in the Advanced Markus ionization chamber," *Medical Physics*, vol. 44, no. 3, pp. 1157–1167, 2017.

[23] M. Gotz, L. Karsch, and J. Pawelke, "A new model for volume recombination in plane-parallel chambers in pulsed fields of high dose-per-pulse," *Physics in Medicine & Biology*, vol. 62, no. 22, p. 8634, Nov. 2017.

[24] R. Kranzer *et al.*, "Ion collection efficiency of ionization chambers in ultra-high dose-per-pulse electron beams," *Medical Physics*, vol. 48, no. 2, pp. 819–830, 2021.

[25] R. Kranzer *et al.*, "Charge collection efficiency, underlying recombination mechanisms, and the role of electrode distance of vented ionization chambers under ultra-high dose-per-pulse conditions," *Physica Medica*, vol. 104, pp. 10–17, 2022.

[26] N. Esplen, M. S. Mendonca, and M. Bazalova-Carter, "Physics and biology of ultra-high dose-rate (flash) radiotherapy: A topical review," *Physics in Medicine & Biology*, vol. 65, no. 23, 23TR03, Dec. 2020.

[27] J. B. Christensen, H. Tölli, and N. Bassler, "A general algorithm for calculation of recombination losses in ionization chambers exposed to ion beams," *Medical Physics*, vol. 43, no. 10, pp. 5484–5492, 2016.

[28] F. Gómez *et al.*, "Development of an ultra-thin parallel plate ionization chamber for dosimetry in FLASH radiotherapy," *Medical Physics*, vol. 49, no. 7, pp. 4705–4714, 2022.

[29] F. Di Martino *et al.*, "A new solution for UHDP and UHDR (Flash) measurements: Theory and conceptual design of ALLS chamber," *Physica Medica*, vol. 102, pp. 9–18, Oct. 2022.

[30] J. D. Fenwick and S. Kumar, "Collection efficiencies of ionization chambers in pulsed radiation beams: An exact solution of an ion recombination model including free electron effects," *Physics in Medicine & Biology*, vol. 68, no. 1, p. 015 016, Dec. 2022.

[31] S. P. Chabod, "A perturbation method to examine the steady-state charge transport in the recombination and saturation regimes of ionization chambers," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 595, no. 2, pp. 419–425, 2008.

[32] S. P. Chabod, "Impact of space charges on the saturation curves of ionization chambers," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 602, no. 2, pp. 574–580, 2009.

[33] S. P. Chabod, "Charge collection efficiency in ionization chambers operating in the recombination and saturation regimes," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 604, no. 3, pp. 632–639, 2009.

[34] G. Boissonnat, J.-M. Fontbonne, J. Colin, A. Remadi, and S. Salvador, "Measurement of ion and electron drift velocity and electronic attachment in air for ionization chambers," *arXiv e-prints*, Sep. 2016.

[35] A. F. Bielajew, "The effect of free electrons on ionization chamber saturation curves," *Medical Physics*, vol. 12, no. 2, pp. 197–200, 1985.

[36] G. Boissonnat, "Chambres d'ionisation en protonthérapie et hadronthérapie," Ph.D. dissertation, Université Caen Normandie, 2015.

[37] E. Hochhauser, O. A. Balk, H. Schneider, and W. Arnold, "The significance of the lifetime and collection time of free electrons for the recombination correction in the ionometric dosimetry of pulsed radiation," *Journal of Physics D: Applied Physics*, vol. 27, no. 3, p. 431, 1994.

[38] J.-H. He, "Homotopy perturbation technique," *Computer Methods in Applied Mechanics and Engineering*, vol. 178, no. 3, pp. 257–262, 1999.

[39] S. Chakraverty, N. R. Mahato, P. Karunakar, and T. D. Rao, *Homotopy Perturbation Method, Advanced Numerical and Semi-Analytical Methods for Differential Equations.* John Wiley & Sons, Ltd, 2019.

[40] D. K. Davies and P. J. Chantry, *Air chemistry measurements 2*, Final Report, Jan. 1983 - Aug. 1984 Dikewood Corp., Albuquerque, NM. May 1985.

8.10 Outline of modified theory

An outline of a modified theory is presented here that considers the pulsed nature of the beam. Since it is more accurate to the actual physics of a parallel-plate chamber under

UHDR radiation, a more accurate expression for the ion recombination correction factor may be obtained. This expression will be valid over a greater range of plate separations and DPP values. Consider a pulse $R(t)$ of duration T which generates an initial ion concentration of N_0. The applied voltage is V_0 with an associated electric field of V_0/d. The associated electron velocity is $v_e^{(0)}$ and electron attachment coefficient $\gamma^{(0)}$ (use Eqs. 8.11 and 8.16). Since T, on the order of 2 μs, is much smaller than the ion collection time, around 20 μs, it can be assumed that the ions are stationary for the duration of the pulse, after which they begin to move. The problem would have to be separated into four time intervals, each with a respective system of PDEs:

1. $0 < t < d/v_e^{(0)}$

$$\frac{\partial n_+}{\partial t} = R(t) \tag{8.53}$$

$$\frac{\partial n_-}{\partial t} = \gamma n_e \tag{8.54}$$

$$\frac{\partial n_e}{\partial t} = R(t) - \gamma n_e + \boldsymbol{\nabla} \cdot (n_e v_e) \tag{8.55}$$

$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = \frac{e}{\epsilon}(n_+ - n_- - n_e) \tag{8.56}$$

$$\tag{8.57}$$

Because $d/v_e^{(0)}$ is very fast, between 10-100 ns (depending on the applied voltage and plate separation), the positive ion-negative ion recombination is negligible.

2. $d/v_e^{(0)} < t < T$

$$\frac{\partial n_+}{\partial t} = R(t) - \alpha n_+ n_- \tag{8.58}$$
$$\frac{\partial n_-}{\partial t} = \gamma n_e - \alpha n_+ n_- \tag{8.59}$$
$$\frac{\partial n_e}{\partial t} = R(t) - \gamma n_e + \boldsymbol{\nabla} \cdot (n_e v_e) \tag{8.60}$$
$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = \frac{e}{\epsilon} (n_+ - n_- - n_e) \tag{8.61}$$
$$\tag{8.62}$$

3. $T < t < T + d/v_e^{(0)}$

$$\frac{\partial n_+}{\partial t} = 0 \tag{8.63}$$
$$\frac{\partial n_-}{\partial t} = \gamma n_e \tag{8.64}$$
$$\frac{\partial n_e}{\partial t} = -\gamma n_e + \boldsymbol{\nabla} \cdot (n_e v_e) \tag{8.65}$$
$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = \frac{e}{\epsilon} (n_+ - n_- - n_e) \tag{8.66}$$
$$\tag{8.67}$$

$R(t)$ is set to 0 because the pulse has finished. Electrons cease to be produced by the beam. Again, recombination is ignored because the duration of this interval is very short, $d/v_e^{(0)}$. Due to the short interval, the ions are still considered stationary.

4. $T + d/v_e^{(0)} < t$

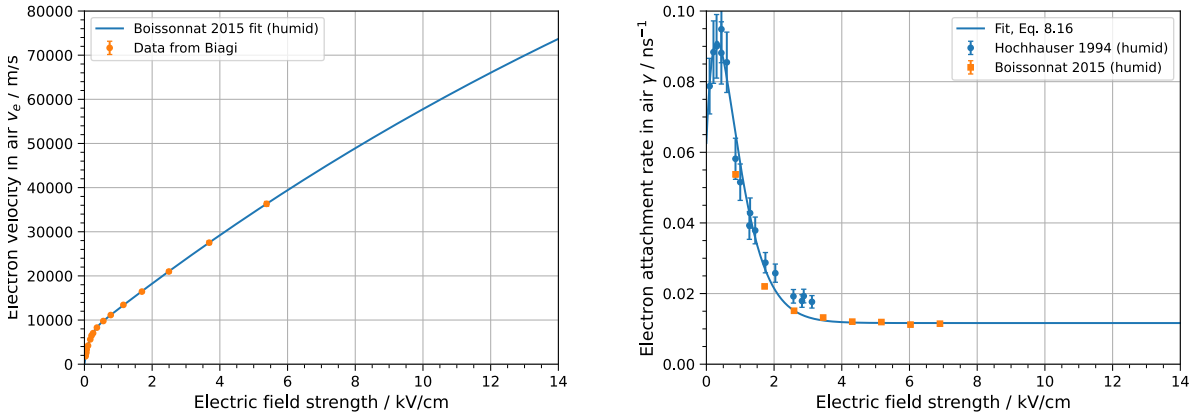
$$\frac{\partial n_+}{\partial t} = -\mu_+ \boldsymbol{\nabla} \cdot (n_+ E) - \alpha n_+ n_- \tag{8.68}$$
$$\frac{\partial n_-}{\partial t} = \mu_- \boldsymbol{\nabla} \cdot (n_- E) - \alpha n_+ n_- \tag{8.69}$$
$$\tag{8.70}$$

Here, the electron concentration is 0 because they all have been collected or have attached to molecules. Space charge is also ignored, as it is small for just positive and negative ions [11, 32].

A perturbation approach is used. In time interval 1, the PDEs are first solved with the constant $v_e^{(0)}$ and $\gamma^{(0)}$. Higher order corrections to the electron and ion concretions and electric field are then iteratively computed. These expressions are then used as initial conditions for time interval 2. The process is then repeated for the other time intervals. Another benefit of this theory is that the PDEs for each time interval are simpler than the full system of PDEs found in Eqs.8.1 - 8.5. This allows for easier computation of the concentrations and electric field.

8.11 Additional comments

Updated versions of Figures 8.1a and 8.1b are shown below in Figures 8.8a and 8.8b. The error bars in Figure 8.8a represent an uncertainty of 1.5% on the plotted data. This uncertainty value is taken from Davies [40]. The data from Davies was demonstrated to be in agreement with the data from Biagi-v8.9 Boltzamnn calculations in dry air, seen in Boissonnat [34], which in turn were used to extract the Eq. 8.11 fit parameters. The updated χ^2/ν value is 10.6 with ν =12 degrees of freedom. To reduce the goodness of fit statistic, more high E-field data is required. The error bars in Figure 8.8b represent an uncertainty of 10%, taken from Hochhäuser [37]. To obtain the fit parameters in Eq. 8.16, only the Hochhäuser data set was used. This is because there are too few data points in the Boissonnat data set. The updated χ^2/ν value is 2.34 with ν =14 degrees of freedom. This goodness of fit value is slightly high because γ_0 is fixed in the fit, according to the Boissonnat value. This leads to larger residuals for the last few data points of the Hochhäuser data set with respect to the fit.



(a) (b)

Figure 8.8: (a) Electron velocity in air as a function of the electric field strength, based on data from Boissonnat [36]. Error bars represent an uncertainty of 1.5%. The χ^2/ν value is 10.6 with ν =12 degrees of freedom. (b) Measured electron attachment rate in (humid) air from Hochhäuser et al. [37] and Boissonnat et al. [34]. The fit is modelled according to Eq. 8.16 and the fit parameters are obtained by fitting only to the Hochhäuser data set. The χ^2/ν value is 2.34 with ν =14 degrees of freedom.

The χ^2/ν statistic is used for the models of electric drift velocity and attachment because there are a relatively large number of data points, and therefore degrees of freedom, available from the literature. Additionally, the literature offers experimental uncertainties that can be assigned to this data. This is not the case with the other figures presented in this chapter, where the degrees of freedom are small, ranging from 2-4, and the experimental uncertainties are either unknown or not directly quoted and difficult to determine. These reasons motivate the use of the RMSE and R^2 coefficient. The validity of the analytical expression for the ion recombination correction factor (within a certain DPP range depending on the applied voltage and plate separation) is justified post-hoc by Figure 8.7. In comparing Figures 8.5 and 8.6 to Figure 8.7, the DPP values where the fit models for the recombination correction factor do not agree with the measured data (above 1 Gy DPP in Figure 8.7a and above 3 Gy DPP in Figure 8.7b) is when only three data points are used in the fit procedure. Along with the observation that using lower voltage data points in the fit procedure does not improve the agreements, the validity of these models at high DPP can be questioned and it is advised

that the models presented in this chapter should not be used for higher DPP values.

9

Conclusions and Future Work

9.1 Summary

The advent of UHDR radiation therapy brings about two needs. The first is the need for low-cost and compact proton accelerators for more accessible and affordable proton therapy. This thesis aims to address this need by contributing to the development of the DWA system. More specifically, this work describes the HV pulsers and constituent DSRDs that generate the nano-second scale HV pulses that go on to form the large accelerating gradient of the DWA. For the purposes of the DWA, the pulses need to have a large magnitude, be of nanosecond scale, have a fast rise time and a repetition rate of at least 1 kHz.

A magnetic saturation (MST)-DSRD-based pulser and a MOSFET-DSRD-based multi-module pulser, were built and comparatively tested with the DSRD being kept constant.

This is was done so as to ascertain which pulser would perform better in the context of the DWA and the pulse requirements. It was found that the most optimal pulse with a single MST-DSRD-based pulser (9645 V pulse amplitude, 1.48 ns rise time, 8.36 ns pulse width) outperformed a single module of the MOSFET-DSRD-based pulser (1807 V pulse amplitude, 3.72 rise time, 8.10 pulse width). However, the most optimal pulse with a six-module MOSFET-DSRD-based pulser, presented for the first time in this work, was superior (10.9 kV pulse amplitude, 1.66 ns rise time and 3.48 ns pulse width). Based on this comparative study of the two pulsers, and considering the smaller form factor of the MOSFET-DSRD-based pulser and its modular design, which adheres to the desired compactness of the DWA, the MOSFET-DSRD-based pulser seems to be more suitable for the purposes of the DWA. The comparisons presented in this thesis can also aid in other HV applications that require certain pulse and pulser specifications. The six module MOSFET-DSRD-based pulser with compression also yields higher magnitude pulses and shorter pulse widths in comparison to the previous study by Fang et. al. [1]. One issue with the MOSFET-DSRD-based pulser is the inductive coupling between the pulser modules. This must be properly characterized in order for the most optimal pulse to be generated. Simulations of the DSRD with both pulsers were carried out with Sentaurus TCAD by Synopses. This was done to observe the effect of the DSRD doping profile and material on the output pulse, which will aid in future DSRD design for the DWA. Silicon DSRD simulations indicate the surface doping concentrations are negligible with respect to the pulse. The depth of the *pn* junction is far more crucial, with a deeper *pn* junction increasing the pulse amplitude and decreasing the pedestal effect. The simulations demonstrated the insensitivity of the pulse shape (except in the pedestal region) with respect to the doping profile. Simulations with an additional 35 pF parasitic capacitance in parallel with the DSRD, modelling the electrical connections between the DSRD, leads and ground plane, reduced the pulse amplitude. The implementation of 4H-SiC DSRDs in simulations with both pulsers yielded lower amplitude pulses, due to incomplete ionization. The SiC simulations also indicate that SiC-based DSRDs and other diode types should

be carefully studied within the context of the specific application and not always be immediately considered as superior to Si, solely by virtue of being a wide-bang gap semiconductor.

The second is the need for analytical expressions for the ion recombination correction factor, critical in ionization-based dosimetry reference standards in UHDR beams. Analytical expressions are quick to evaluate in comparison to metrological-based solutions or numerical solutions. Analytical expressions can also be used to generate quick evaluation methods such as the TVM, which are able to be used in a clinical environment. This need is addressed by developing a mathematical procedure to solve the charge carrier transport equations.

The PDEs describing the positive and negative ion and electron transport dynamics as well as the electric field were solved with the homotopy perturbation method (HPM), a semi-analytical solution method. The free electron velocity and attachment rate were modelled as functions of the electric field according to experimental data. The analytical expressions for the charge collection efficiency and ion recombination correction factor were found to exhibit a $1/V_0^2$ relation, in contrast with the pulsed Boag theory. This expression was fit to previously published experimental data from Gotz et. al. [2] and Kranzer et. al. [3]. The collection efficiency or ion recombination correction factor obtained from these fits agreed with the experimentally-determined values to within 6% (k=2) below 1 Gy DPP for 1 mm plate separation and below 3 Gy DPP for a 0.5 mm plate separation chamber. It was demonstrated that a generic expression with just a constant and $1/V_0^2$ term fit the data just as well. One must take care to use a appropriate voltage range for the fit. At higher DPP, a more limited voltage range (above 300 V) seems to work well. An extrapolation procedure based on the derived expressions that can be used in the clinic is also proposed.

9.2 Future work

9.2.1 Dielectric wall accelerator

The next steps for the HV pulsers is to reduce the pulse widths, closer to 1 ns. Looking back at Table 7.4, the shortest pulse generated has a width of 3.48 ns. For a radial waveguide of inner radius 1.5 cm, outer radius of 30 cm and a dielectric constant of 3.4 (PCB), the transit time is 1.75 ns. Two transit times is 3.5 ns. Therefore, the reflection will interfere with the remainder of the pulse and a pile up will occur, impacting the accelerating gradient. This is one reason for decreasing the pulse width; a smaller outer radius radial waveguide can be used, decreasing the overall size of the DWA. The other reason, already mentioned, is the inverse relationship of the HGI breakdown threshold with pulse width. One method of decreasing the pulse width is by implementing additional pulse compression circuity at the DSRD end of the pulser. This can be be done with a magnetic switch [4]. Another possible solution is the use of smaller surface area DSRDs. Previous studies have shown that smaller area DSRDs generate pulses with smaller widths and faster rise times [5]. A modified pulser design that incorporates a second DSRD stage with a smaller area DSRD, after the first DSRD stage with a larger DSRD, can readily be built from our existing MOSFET-DSRD-based pulser.

As research on the radial waveguides also progresses, the two areas of research increasingly converge and must be done in tandem. There are several questions that arise. The first is how to maximize the load presented to the pulser by the radial waveguide in order to maximize the amplitude of the pulse. A preliminary idea is to have the radial waveguide have strip line protrusions at the outer radius. The width and height of the strip line section can be produced to have a specific impedance. Another question that arises is how many pulsers are required per line to achieve a certain accelerating gradient. To first order, this is determined by space considerations and the input voltage. The stacking of waveguides may

also increase the gradient. The development of the combined pulsers+waveguide technology will have to be done in an iterative process. The performance (in terms of accelerating gradient pulse width, magnitude and shape) of a given prototype will inform how the following prototype should be altered in order to improve the performance.

9.2.2 UHDR ion recombination

Parallel plate chambers are the chamber of choice for reference dosimetry in UHDR beams. Nonetheless, in the clinic, cylindrical chambers are much more common. To further extend the clinical relevance of the semi-analytical method presented in Chapter 8, it can be used to solve the charge carrier PDEs in cylindrical coordinates. For an average cylindrical chamber, however, the gap between the central electrode and wall is about 3 mm. This is quite large in comparison to parallel plate chambers. Since the DPP level at which the method presented in Chapter 8 yields accurate results decreases with increasing electrode separation, it will most likely only be accurate to the order of a few 100 mGy DPP. Recall in this work, the HPM method is first applied to a constant dose rate problem and then a pseudo square pulse approximation is applied. Using a more accurate theory, one that takes into account the pulsed nature of the beam and the resulting carrier and electric field evolution, would result in an expression for the ion recombination correction factor that would be more accurate and therefore applicable for a larger range of plate separations (and voltages and DPP values). The outline of a modified theory was presented in Section 8.10 that could address these issues. The computations are still required to be done in order to see if the resulting expression for the recombination correction factor is valid over a larger range of plate separations and DPP values.

References

[1] X. Fang *et al.*, "Driving mechanism of drift-step-recovery diodes," *Review of Scientific Instruments*, vol. 92, no. 8, p. 084 702, Aug. 2021.

[2] M. Gotz, L. Karsch, and J. Pawelke, "A new model for volume recombination in plane-parallel chambers in pulsed fields of high dose-per-pulse," *Physics in Medicine & Biology*, vol. 62, no. 22, p. 8634, Nov. 2017.

[3] R. Kranzer *et al.*, "Ion collection efficiency of ionization chambers in ultra-high dose-per-pulse electron beams," *Medical Physics*, vol. 48, no. 2, pp. 819–830, 2021.

[4] J.-G. Choi, "Introduction of the magnetic pulse compressor (MPC) - fundamental review and practical application," *Journal of Electrical Engineering and Technology*, vol. 5, no. 3, pp. 484–492, Sep. 2010.

[5] A. Kyuregyan, "High-voltage diffused step recovery diodes: I. Numerical simulations," *Semiconductors*, vol. 53, no. 7, pp. 962–968, Jul. 2019.

Appendices

A

Theoretical treatment of particle accelerators

A.1 Infinite, uniform cylindrical waveguide

Consider an infinite, uniform cylindrical waveguide with walls made of a perfect conductor material. The inner radius of the waveguide is $r = a$. Under these conditions, the Maxwell equations are

$$\boldsymbol{\nabla} \cdot \vec{\mathbf{E}} = 0, \tag{A.1}$$

$$\boldsymbol{\nabla} \cdot \vec{\mathbf{B}} = 0, \tag{A.2}$$

$$\boldsymbol{\nabla} \times \vec{\mathbf{E}} = -\frac{\partial \vec{\mathbf{B}}}{\partial t}, \tag{A.3}$$

$$\boldsymbol{\nabla} \times \vec{\mathbf{B}} = \frac{1}{c^2}\frac{\partial \vec{\mathbf{E}}}{\partial t}, \tag{A.4}$$

or in wave equation form

$$\nabla^2 \vec{\mathbf{E}} = \frac{1}{c^2}\frac{\partial^2 \vec{\mathbf{E}}}{\partial t^2}, \tag{A.5}$$

$$\nabla^2 \vec{\mathbf{B}} = \frac{1}{c^2}\frac{\partial^2 \vec{\mathbf{B}}}{\partial t^2}, \tag{A.6}$$

where $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are the electric and magnetic field vectors, respectively. Furthermore, the boundary conditions are

$$\vec{\mathbf{n}}_{12} \times (\vec{\mathbf{E}}_2 - \vec{\mathbf{E}}_1) = 0, \tag{A.7}$$

$$(\vec{\mathbf{B}}_2 - \vec{\mathbf{B}}_1) \cdot \vec{\mathbf{n}}_{12} = 0, \tag{A.8}$$

where subscript 1 denotes the field in the vacuum and 2 denotes the field in the conductor wall. These boundary conditions are evaluated at $r = a$ (interface of the two media). The vector $\vec{\mathbf{n}}_{12}$ is the vector from media 1 to 2 in the normal (radial) direction of the waveguide. As there are no fields present in the conductor wall ($\vec{\mathbf{E}}_2$ and $\vec{\mathbf{B}}_2$ are zero). We therefore have,

$$\vec{\mathbf{E}}_1 \times \vec{\mathbf{n}}_{12} = 0 \text{ or } \vec{\mathbf{E}}_{tang}|_{r=a} = 0, \tag{A.9}$$

$$\vec{\mathbf{B}}_1 \cdot \vec{\mathbf{n}}_{12} = 0 \text{ or } \vec{\mathbf{B}}_{norm}|_{r=a} = 0. \tag{A.10}$$

Eq. A.9 shows that the E-field in the z (longitudinal) and θ (azimuthal) direction at the wall are zero. Eq. A.10 shows the B-field in the radial direction at the wall is zero. Using the separation of variables solution method with Eqs. A.5 and A.6, along with boundary conditions A.9 and A.10, the following solutions are obtained for the z-component of the E and B-field, E_z and B_z,

$$E_z(r,\theta,z,t) = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} J_m(\gamma_{mn}r)[a_{mn}\cos m\theta + b_{mn}\sin m\theta]e^{i(kz-\omega t)}, \tag{A.11}$$

$$B_z(r,\theta,z,t) = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} J_m(\gamma_{mn}r)[c_{mn}\cos m\theta + d_{mn}\sin m\theta]e^{i(kz-\omega t)}, \tag{A.12}$$

$$\tag{A.13}$$

where J_m is the mth order Bessel function of the first kind, $a_{mn}, b_{mn}, c_{mn}, d_{mn}$ are constants determined from initial and boundary conditions and

$$k^2 = \left(\frac{\omega}{c}\right)^2 - \gamma_{mn}^2. \tag{A.14}$$

Through Eqs. A.1-A.4, the remaining field components are

$$E_r = \frac{1}{\gamma_{mn}^2}\left[ik\frac{\partial E_z}{\partial r} + \frac{i\omega}{r}\frac{\partial B_z}{\partial \theta}\right], \tag{A.15}$$

$$E_\theta = \frac{1}{\gamma_{mn}^2}\left[\frac{ik}{r}\frac{\partial E_z}{\partial \theta} - i\omega\frac{\partial B_z}{\partial r}\right], \tag{A.16}$$

$$B_r = \frac{1}{\gamma_{mn}^2}\left[-\frac{i\omega}{c^2}\frac{1}{r}\frac{\partial E_z}{\partial \theta} + ik\frac{\partial B_z}{\partial r}\right], \tag{A.17}$$

$$B_\theta = \frac{1}{\gamma_{mn}^2}\left[\frac{i\omega}{c^2}\frac{\partial E_z}{\partial r} + \frac{ik}{r}\frac{\partial B_z}{\partial \theta}\right]. \tag{A.18}$$

From boundary condition A.9 with Eq. A.11, one finds

$$E_z(r=a) = 0 \implies J_m(\gamma_{mn}a) = 0 \implies \gamma_{mn}a = x_{mn}, \tag{A.19}$$

where x_{mn} is the n^{th} root of the m^{th} order Bessel function of the first kind. From Eq. A.14, we see that

$$k = \frac{1}{c}\sqrt{\omega^2 - c^2\left(\frac{x_{mn}}{a}\right)^2} = \frac{1}{c}\sqrt{\omega^2 - \omega_c^2}. \tag{A.20}$$

For the TM_{01} mode, the field components are

$$E_z = (E_z)_0 J_0\left(\gamma_{01}r\right) e^{i(kz-\omega t)}, \tag{A.21}$$

$$E_r = \frac{-ika}{x_{01}}(E_z)_0 J_1\left(\gamma_{01}r\right) e^{i(kz-\omega t)}, \tag{A.22}$$

$$E_\theta = 0, \tag{A.23}$$

$$B_z = 0, \tag{A.24}$$

$$B_r = 0, \tag{A.25}$$

$$B_\theta = \frac{-i\omega a}{c^2 x_{01}}(E_z)_0 J_1\left(\gamma_{01}r\right) e^{i(kz-\omega t)}, \tag{A.26}$$

where $\gamma_{01} = x_{01}/a$ and $x_{01} = 2.405$.