5-1-2015

New Optimization Technique for Radiation Therapy Planning

Daniel Riofrio

Recommended Citation

Riofrio, Daniel. "New Optimization Technique for Radiation Therapy Planning." (2015). https://digitalrepository.unm.edu/cs_etds/51

Daniel A. Riofrio A.

Candidate

Computer Science

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Shuang Luan, Chairperson

Trilce Estrada Piedra

Adam Hecht

Michael Holzscheiter

New Optimization Technique for Radiation Therapy Planning

by

Daniel Andrés Riofrío Almeida

Ingeniero en Sistemas Informáticos y de Computación, Escuela
Politécnica Nacional, Quito, Ecuador, 2006.

M.S., Computer Science, University of New Mexico, Albuquerque,
New Mexico, 2012.

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy in Computer Science

The University of New Mexico

Albuquerque, New Mexico

May, 2015

Dedication

To Mom, Dad, Carlos, María Cris, and Daniela.
Thank you so much for all your support.
It's been a long ride.

Acknowledgments

I would like to thank all the people that contributed to the development of this dissertation, specially, to my advisor, Prof. Shuang Luan, for his constant support, trust, and all the time and dedication he invested in this work; to Jun Zhou, from the William Beaumont Hospital, Radiation Oncology, for suggesting the minimization of interstitial implants for HDR prostate brachytherapy that started this research; to Lijun Ma, from the University of California in San Francisco, Department of Radiation Oncology, for his suggestions and valuable feedback in both our Gamma Knife® and HDR brachytherapy results. In addition, I would like to thank the members of my Ph.D. dissertation committee: Prof. Trilce Estrada, Prof. Michael Holzscheiter, and Prof. Adam Hecht for your valuable feedback and help to make this dissertation better.

This dissertation was partially supported by NSF grant CBET-0853157 as well as the Ecuadorian Senescyt scholarship. Part of this research was presented in the 2014 AAPM Anual Meeting in Austin, Texas which was made possible by the Ecuadorian Senescyt travel grant and to the travel and research grants from the UNM Graduate and Professional Student Association (GPSA) and UNM Office of Graduate Studies (OGS).

Last but not least, I would like to thank my family and friends, and specially to my wife Daniela.

New Optimization Technique for Radiation Therapy Planning

by

Daniel Andrés Riofrío Almeida

Ingeniero en Sistemas Informáticos y de Computación, Escuela
Politécnica Nacional, Quito, Ecuador, 2006.
M.S., Computer Science, University of New Mexico, Albuquerque,
New Mexico, 2012.

Ph.D., Computer Science, University of New Mexico, 2015

Abstract

The main goal of radiation therapy is to deliver a lethal dose of radiation to the targeted tumor while minimizing the radiation dose to the surrounding normal tissues and critical organs. Modern cancer therapy has benefited enormously from computer controlled treatment devices with increased precision and capability. However, this increased sophistication also creates new challenges for treatment planning. As the number of parameters in a treatment plan increases, the traditional computational approaches are no longer adequate to fully exploit the potential of the latest treatment devices. This is because at the heart of treatment planning is often a set of substantially non-trivial constrained geometric optimization problems.

In this dissertation, we present a new optimization framework combining Particle Swarm Optimization (PSO) with deterministic optimization (e.g., least distance

programming, non-negative least-square optimization, etc.). For our new PSO framework, we moved away from the classical view of a particle representing a potential solution of the optimization function; instead, we use the whole particle distribution as the solution. We modeled tumors, critical organs and other tissues as geometric volumes, whose surfaces have an associated potential function. The radiation source is modeled as kinetic particles subject to the forces from the potential functions from both the particles and the various geometric objects. The final configuration of the swarm of particles including their trajectories is the treatment plan.

To demonstrate the potential of our new optimization paradigm, we have applied it to Gamma Knife® radiosurgery and High-Dose Rate Brachytherapy (HDR) for prostate cancer. Mathematically, Gamma Knife® radiosurgery is a ball-packing process whose goal is to "pack" some spherical high dose volumes into a tumor volume, while brachytherapy is to find the trajectories of some spherical high dose volumes. Both problems are computationally intractable. Our new framework models the spherical high dose volume as kinetic particles and simulates the "swarm" of these particles through a potential field created based on medical constraints and prescriptions. The resulting stable swarm, further refined by a deterministic optimization algorithm (e.g., non-negative least squares and least distance programming), is the final treatment plan.

In the medical field, the adoption of new technologies take several years (sometimes even decades) due to the long trials they undergo, their high complexity, the lack of specific additional characteristics physicians demand or their high cost of infrastructure. As a consequence, Gamma Knife® radiosurgery and HDR brachytherapy treatment planning are mainly performed as manual processes by physicians, despite the existence of algorithms that attempt to make them fully or partially automatic. Our experiments with challenging simulated and real clinical data have shown that our new framework significantly outperforms current clinical systems. In particular for Gamma Knife® radiosurgery, our algorithm is able to produce high quality treatment plans that meet clinical standards. For HDR brachytherapy planning, our method can generate optimal (i.e., minimal and error tolerant) implant trajectories, which is a feature that no known algorithm has attempted to solve.

Finally, we expect that due to the evidence shown in this dissertation, the simplicity of implementing our framework, and the ease of understanding the concepts of our approach, we will be able to widely impact the technologies currently being used not only in Gamma Knife® radiosurgery and HDR brachytherapy, but also in other radiation therapy modalities.

Contents

List of Figures

List of Tables

Glossary

Anti-proton	is a charged particle with the same mass and opposite exact charge as a proton. It is the anti particle of proton and is represented by the symbol $\bar{p}$ [8].
Brachytherapy	comes from the greek word *brachus* which means short and refers to short-distance therapy; i.e., the treatment of tumors at short-range with radioactive sources [9, 4, 10].
Chemotherapy	is the (invasive) use of drugs (chemicals) to target cells of rapid reproduction, such as cancer cells.
Collimator	is a device that usually consists of interleaved bars of a high density material that is used to shape, trim and define the field size of a beam [4].
Dose	is defined as the amount of energy deposited per unit mass (for more details, see section 2.1.1).
DVH	Dose volume histogram is a tool used to understand the quality of a treatment plan in terms of its dose distributions. It provides the amount of volume per biological structure that received at least a certain percentage of the maximum dose.

Heuristic	refers to the use of experience to do something. In this dissertation's scope, it relates to an empirical or practical technique that encodes knowledge of a certain problem in order to guide the search process towards a better solution.
Ionization	is the effect of physically adding or removing orbital electrons from an atom, converting it into an ion.
Photon	is a wave-particle that moves at the speed of light, has zero mass and rest energy, carries energy and momentum, and can collide with other particles such as electrons [11].
Proton	is a fundamental particle that constitutes matter. It was named by Niels Bohr in 1914 when he referred to the nucleus of the lightest atom (hydrogen) [8]. This atom is characterized by possessing one electron (e^-) orbiting a nucleus of one proton (p^+).
Stereotactic Radiosurgery (SRS)	refers to the use of small-field, high localized dose fields to treat intracraneal or extracraneal tumors [4].
IMRT	Intensity-modulated radiation therapy is an irradiation technique that uses non-uniform beam intensities for delivering highly conformal dose to a localized tumor while sparing healthy tissues.

Chapter 1

Introduction

> “The proof that the little prince existed is that he was charming, that he laughed,
> and that he was looking for a sheep.
> If anybody wants a sheep, that is a proof that he exists.”
> —Antoine de Saint-Exupéry, The Little Prince.

Cancer is a wide spread disease that expresses itself through the errant growth of abnormal cells. If the uncontrolled growth of these cells is not stopped, it can cause death [12, 13]. As the worldwide cancer’s fatal trend increases, in 2015 alone, 589,430 deaths are expected due to cancer in America, which is almost 1,620 people per day [13]. This along with the overall estimated cost of $216.6 billion USD for cancer management [14] has led to an increased demand from the general public to develop more effective tools and technologies for treating and curing the disease which include radiation therapy, brachytherapy, particle therapy, and chemotherapy, among others. With the rapid advancement of medical imaging, tumors are being diagnosed in early stages when they are still local or regional. External beam therapy, radiosurgery and brachytherapy are effective means to treat local-regional tumors [12, 15, 16, 17, 18, 19]. In this dissertation, we focus on two radiation therapy modalities: Gamma Knife® Radiosurgery and High-Dose Rate Brachytherapy.

Radiation therapy is a modality of cancer treatment that uses ionizing radiation. This type of ionizing radiation (e.g., high energy X-rays) damages the DNA and causes cell death within the region being irradiated [4, 15, 9, 20, 21, 18, 22]. Hence, the goal of radiation therapy is to deliver a radiation dose high enough to kill all the targeted tumor cells while simultaneously minimizing the damage to surrounding normal structures. The quality of a radiotherapy plan is usually judged by its *dose conformity* and *treatment time*. The dose conformity describes how well the high radiation dose region conforms to the targeted tumor and spares the surrounding normal tissues, while treatment time describes how long treatment takes and the efficiency of the treatment machines used. Any improvement in the dose conformity in radiation therapy will likely improve tumor control and reduce the likelihood of complications, and any improvement in treatment time will likely lower the treatment cost and improve patient throughput and comfort.

Many types of ionizing radiation have been experimented through the history of radiotherapy. These include high energy photons (e.g., γ-rays and high energy X-rays), electrons, and charged particles heavier than electrons such as protons, pions, alpha particles, carbon ions, and even antiprotons [15, 23, 12, 9, 18, 4]. In terms of the energy deposited on matter along the path of this ionizing radiation, they have different benefits and drawbacks. Figure 1.1 shows a depth-dose diagram and compares the energy deposited (dose in relative dose units) in matter (water) along the straight path (depth) followed by these ionizing radiations. The energy deposited by carbon ions from radiation beams of 200 and 270 MeV is shown in red. The energy deposited by a ^{60}Co (cobalt-60) radiation source is shown in red, and the energy deposited by photons of 25 MeV produced by a linear accelerator is shown in

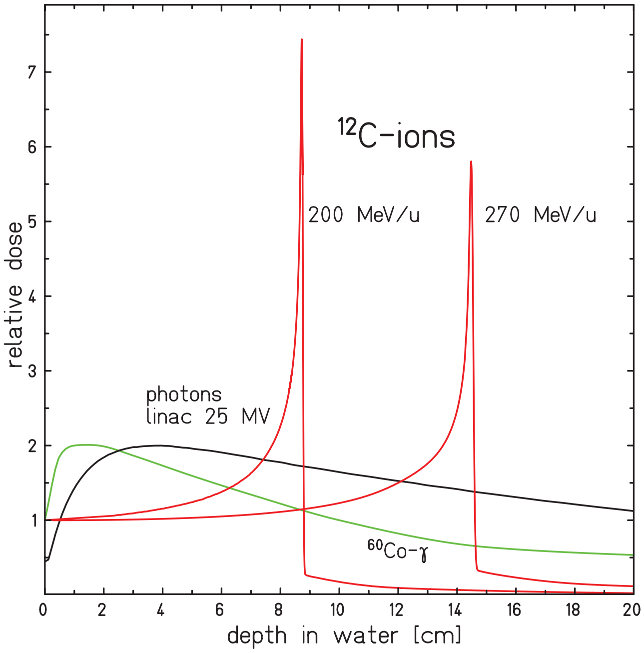


Figure 1.1: Depth-dose diagram: Comparison high of energy photons vs. charged particles energy beams [1]. Red: energy deposited by carbon ions from radiation beams of 200 and 270 MeV. Green: energy deposited by a ^{60}Co (cobalt-60) radiation source. Black: energy deposited by photons of 25 MeV produced by a linear accelerator. Dose units are relative and arbitrary.

black.

In order to provide dose conformity and efficient treatment time, radiotherapy relies on specialized optimization algorithms, for instance simulated annealing, deterministic optimization models such as linear programming, non-negative least squares, linear programming, among others; genetic algorithms, neural networks, mixed integer linear programming, and graph algorithms, etc. Usually these algorithms try to model all competing treatment goals and radiation source configurations as a unique optimization problem. In this dissertation, we present a new optimization technique that can be used for radiation therapy and radiosurgery. In particular, the algorithm showcased in this dissertation combines the maturity and fast convergence of deterministic optimization methods (e.g., non-negative least squares, least-distance programming, least squares optimization, etc.) with the flexibility of a stochastic optimization method, i.e., particle swarm optimization (PSO).

In a nutshell, the classical PSO algorithm consists of an iterative algorithm that uses the notion of individuals carrying potential solutions to the problem one is trying to solve. A set of individuals comprise the swarm, and each individual, i.e., a potential solution, in the swarm is represented by a multidimensional location vector and a velocity vector. At each iteration, the location and the velocity vectors are updated based on the current velocity and both the best location each individual has explored and the global best location the swarm has explored. This process repeats for a fixed number of iterations or until a desired minimum error is achieved or a desired fitness level is achieved by one individual. Finally, the individual that has visited the best location is the one carrying the optimal solution.

In our novel approach for radiation therapy planning, we do not use the classical PSO modeling, where a particle or individual represents a potential solution of the optimization problem; instead, the whole swarm is the solution we are looking for. Tumors, critical organs, and other tissues are modeled as geometric volumes, whose surfaces have an associated potential function. The radiation source is modeled as kinetic particles subject to the forces from the potential functions from both the particles and the various geometric objects. The final configuration of the swarm of particles including their trajectories is the treatment plan. To demonstrate the effectiveness of our model, we applied this new framework to two challenging radiation therapy problems: Gamma Knife® radiosurgery inverse planning and high-dose rate brachytherapy inverse planning for prostate cancer.

Gamma Knife®radiosurgery inverse planning: Radiosurgery is a radiation therapy technique based on the principle that a single high dose of radiation delivered precisely to a small area will arrest or kill the tumor. Hence, radiosurgeries usually use a number of precisely aimed highly focused external beams of radiation to target a specific area [24]. Gamma Knife® radiosurgery uses the γ-ray radiations from cobalt-60 (^{60}Co) sources to eradicate tumors. The sources (typically about 200) are arranged either in a torus or hemisphere (depending on the machine model), and their emitted γ-rays are focused at a single point (called the iso-center or focus), which creates a spherical high dose volume [25]. A Gamma Knife® procedure is essentially a "ball-packing" process, where the goal is to pack the spherical high dose volumes into a target tumor volume and create a radiation dose deposition that conforms to the target tumor volume and spares the surrounding normal tissues and structures. Even though there are some attempts to develop automatic planning systems for Gamma Knife® the procedure is still planned by physicians manually using a "trial-and-error" approach.

In our solution, the Gamma Knife® dose kernels are mapped to kinetic particles inside the tumor with zero initial speed. After these particles stabilize, a non-negative least squares algorithm is used to filter-out redundant radiation sources and efficiently shape the final dose distribution. The prescribed dose is an input parameter provided by the physician.

HDR brachytherapy inverse planning: Brachytherapy is a radiation therapy modality in which the tumor is eradicated by placing small radioactive isotope sources in short range of a tumor. We have applied our optimization technique to the interstitial high-dose rate (HDR) brachytherapy treatment for prostate cancers, which is the most common disease among men in the United States and the second leading cause of death [26].

Generally speaking, HDR brachytherapy treatment involves the following key steps. First, the physician inserts the interstitial catheters or implants (i.e., hollow needles with bevel tips) inside the prostate under ultrasound guidance. After enough needles (typically about 20) are placed, they are hooked up to an afterloader unit. The afterloader (containing an iridium-192 (^{192}Ir) source) then sends the radioactive source through the hollow needles one by one to deliver the radiation dose to the target. A computer program, typically based on linear programming or simulated annealing, is used to calculate the dwell-time of the source at each location along the needles [27].

Not surprisingly, the biggest challenge for HDR brachytherapy for prostate cancer is the needle placement. There are no computer algorithms to calculate the needle positions. As a result, the needles are placed in real time and the physician tends to use many more needles than necessary to ensure proper coverage of the entire prostate. The drawbacks of using large number of needles are: (1) longer procedure times, (2) longer patient recovery times, and (3) longer treatment times.

In our solution, we used nested particle swarming strategies to solve the HDR inverse planning problem. First, kinetic particles with zero momentum are allowed to stabilize in a cross-section of the tumor to establish the initial locations. Then, an initial speed is assigned to each kinetic particle, which pushes these particles to traverse the tumor along a desired direction. Finally, we map the path followed by each particle to an implant trajectory and each discrete position in this path as a potential position for a radioactive source. From these potential positions, the best ones are selected through our filtering process that uses least distance programming. The prescribed dose is an input parameter provided by the physician.

We have experimented our approach against a challenging simulated case of a C-shaped tumor surrounding a critical spherical organ, which resembles brain tumors usually treated with Gamma Knife® radiosurgery, and find that our method produces high quality plans with respect to the clinical standard. In addition, we show our experimental results with real patient data from prostate cancer cases. Our method is able to determine optimal and reliable implant trajectories using a minimal number of implants while meeting the prescription. It also allows flexible configuration in order to obtain plans with different levels of dose homogeneity. To the best of our knowledge, there are no known algorithms that can minimize the number of implants and optimize implant trajectories. The current commercial system can only calculate dwell times once the implants are in place. Hence, we expect the most important impact of our work will be to significantly improve HDR systems by providing the following key aspects: (1) optimal implant trajectories with minimum number of implants; (2) real-time imaging to guide physicians for implant placement; and, (3) real-time assessment for manual errors while placing needles.

The rest of this document is organized in the following manner: *chapter 2* includes the background information needed to better understand this document, i.e., basic concepts and detailed information about radiation energies and radiation therapy modalities, treatment planning systems and their components, and mathematical optimization. *Chapter 3* details our novel optimization technique. Then, *chapters 4 and 5* show how our optimization technique was applied to Gamma Knife® radiosurgery and high-dose rate brachytherapy for prostate cancer, respectively. In each chapter, we provide details of each phase in our optimization and discuss the results obtained. Finally, *chapter 6* includes our conclusions and lightens the path for future research.

Chapter 2

Background

> “Only the children know what they’re looking for,” said the little prince.
> “They spend their time on a rag doll and it becomes very important, and if it’s
> taken away from them, they cry...”
> —Antoine de Saint-Exupéry, The Little Prince.

2.1 Radiation Therapy

Radiological physics is the science that studies ionizing radiation and its interaction with matter. The field began in the 1890s immediately following the discovery of X-rays by Wilhelm Röntgen [28], radioactivity by Henri Becquerel [29], and radium by the Curies [30]. The therapeutic use of radiation started as early as the discovery of X-rays and radium, where X-rays were used to treat a patient with breast cancer [15]. It can be argued that both external-beam radiation therapy as well as brachytherapy started as early as the first applications of radiation in the medical field [31, 32].

The most important types of ionizing radiation for therapeutic usage are photons (e.g., γ-rays and X-rays), electrons, hadrons (e.g., protons, alpha particles, carbon ions, etc.), and neutrons [23, 33, 2, 34]. In this disseration, we focus on two γ-ray based radiation therapy modalities: Gamma Knife® radiosurgery (using γ-rays from the radioactive decay of ^{60}Co), and HDR Brachytherapy (using γ-rays from the radioactive decay of ^{192}Ir).

Photons deposit energy along their path through matter in such a way that higher doses are deposited at short distances (e.g., the highest dose is deposited at 1.5 cm in water). Figure 2.1 shows this behavior in a depth-dose diagram in water for different types of photons and energies: 25 MeV, 10 MeV, 4 MeV, 1.33 MeV (i.e., ^{60}Co), and 511 keV (i.e., 3.0 mm Cu, copper, Half Value Layer) [23, 15, 2, 35]. Depth-dose diagrams in water are especially useful in the context of radiation therapy since we can assume that similar dose deposition will take place in soft tissues, which are usually the composition of tumors and nearby organs. The photon energies produced by ^{60}Co sources are 1.17 and 1.33 MeV, and by ^{192}Ir sources are mainly 0.296, 0.309, 0.317, and 0.468 MeV [32].

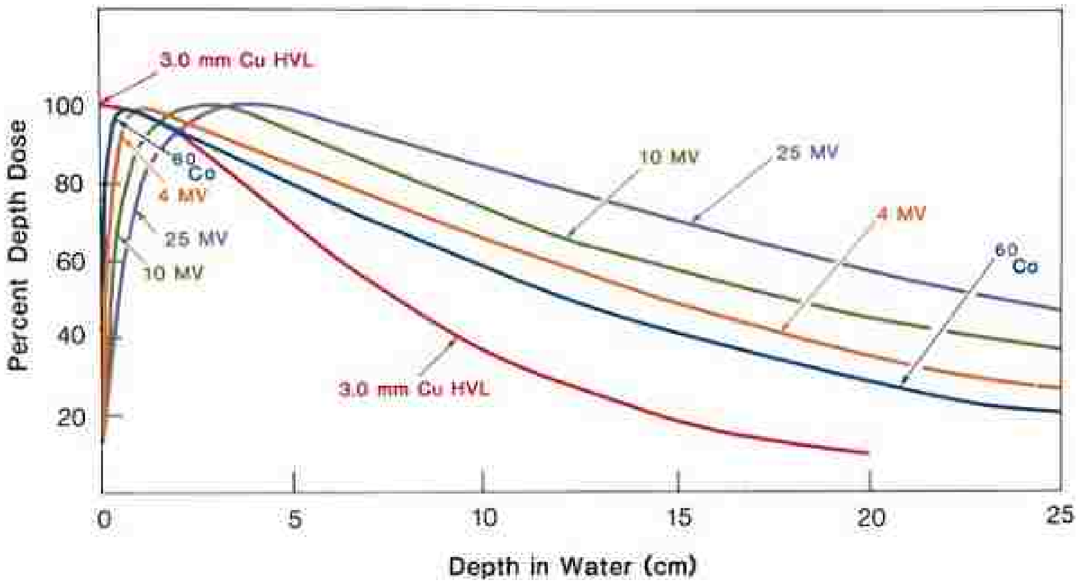


Figure 2.1: Photon beams - depth-dose diagram [2]. Dose deposition along the photon path in water are shown for photon energies of 25 MeV, 10 MeV, 4 MeV, 1.33 MeV (i.e., ^{60}Co), and 511 keV (i.e., 3.0 mm Cu Half Value Layer).

2.1.1 Dose

One of the most important concepts in radiological physics is **absorbed dose** (or simply **dose**). Generally speaking, the absorbed dose D is the amount of energy deposited by the incident radiation in a medium per unit mass. The mathematical definition of dose is

$$D = \frac{\mathrm{d}E}{\mathrm{d}m}, \tag{2.1}$$

where $\mathrm{d}E$ is the differential of energy deposited in a mass differential $\mathrm{d}m$. The SI unit for absorbed energy is $\left[{}^{J}\!/_{kg}\right]$ or gray $[Gy]$[2]. In radiation therapy, the unit centigray ($[cGy]$), which equals to 0.01 Gy, is also commonly used.

2.1.2 Gamma Knife® radiosurgery

Radiosurgery is a radiation therapy modality originally developed by Professor Lars Leskell of the Karolinska Institute in Sweden in the 1950s. It is a minimally invasive technique aiming to perform surgery-like radiation treatment by delivering an extremely high dose of radiation precisely to a selected area or region [24].

One radiosurgery modality is the Gamma Knife® radiosurgery. Each year, about sixty thousand patients receive this treatment from over hundreds of Gamma Knife® centers worldwide [36, 37]. In Gamma Knife® radiosurgery, γ-rays from the radioactive decay of ^{60}Co are used to eradicate tumors. The ^{60}Co sources, which range from 30 to 200, are arranged either in a torus or hemisphere (depending on the machine model). The γ-rays emitted from these sources are collimated and focused at a single point called the iso-center or focus. Figures 2.2, 2.3, and 2.4 show the dose distributions and dose iso-contours of the spherical dose kernels generated at the focus of a Gamma Knife® unit (for radii 4, 8, and 16mm respectively) when all collimators are open [4, 36, 38, 25].

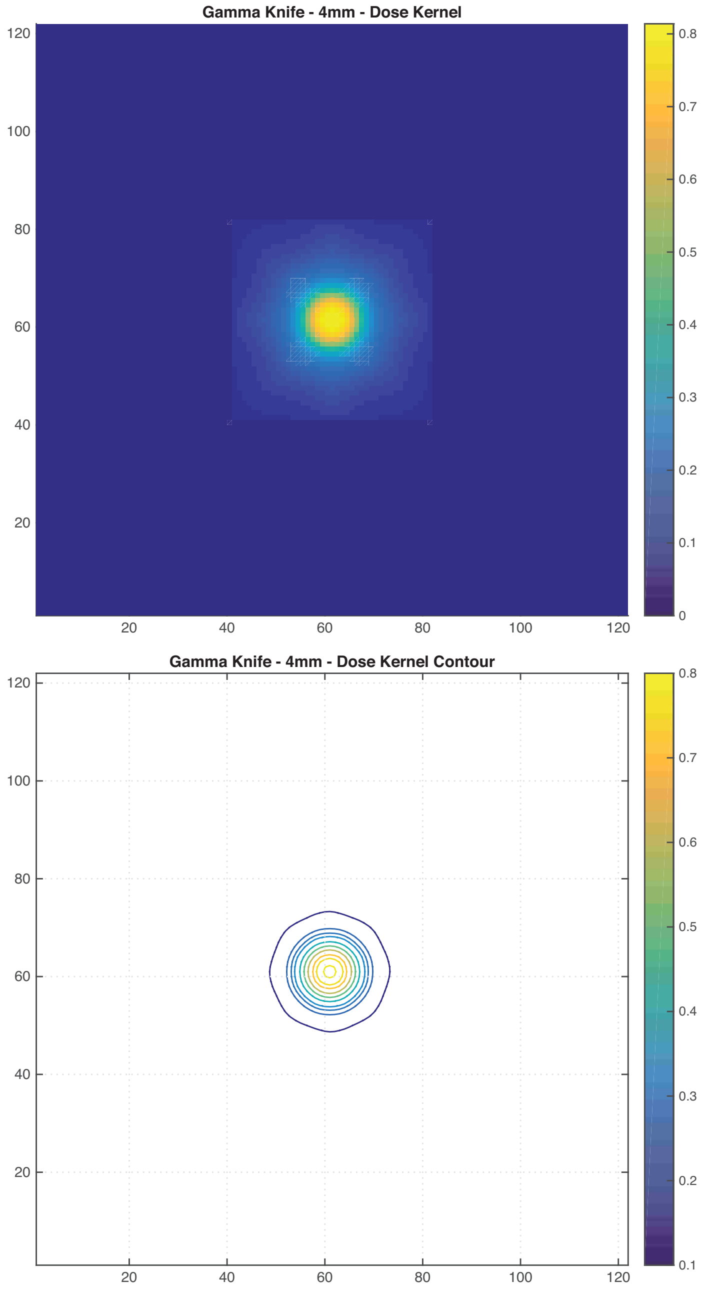


Figure 2.2: Top: Elekta Gamma Knife® 4mm dose kernel. Bottom: Elekta Gamma Knife® 4mm dose kernel contours. Grid resolution is 0.5 x 0.5 mm^2. The kernel has been normalized to the highest dose from the 16mm dose kernel.

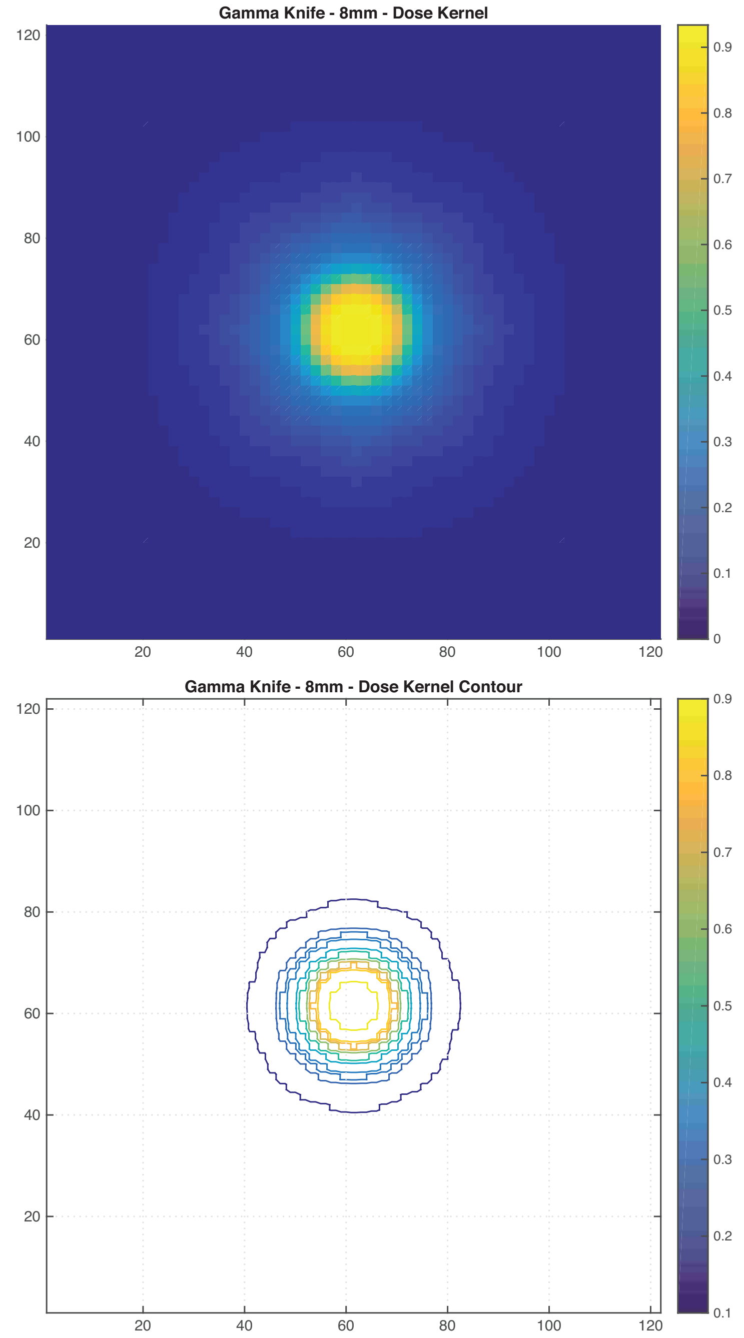
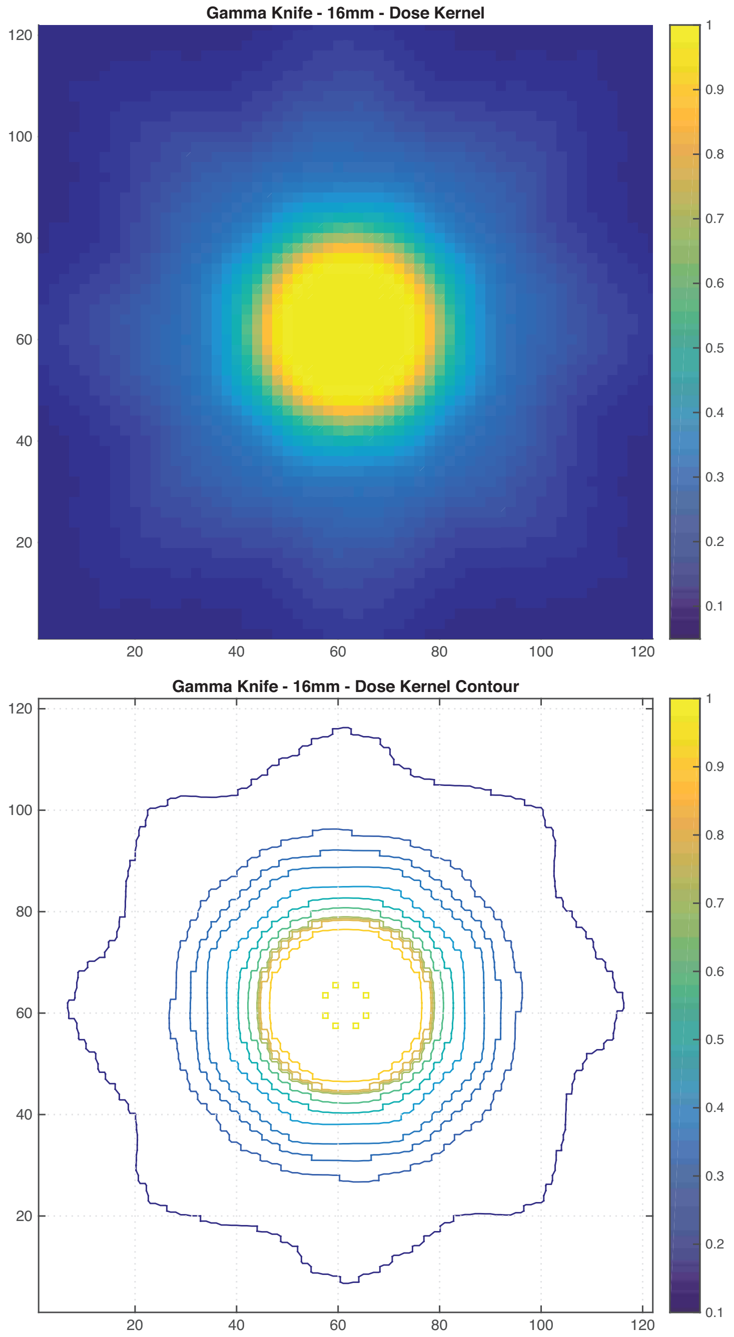


Figure 2.3: Top: Elekta Gamma Knife® 8mm dose kernel. Bottom: Elekta Gamma Knife® 8mm dose kernel contours. Grid resolution is 0.5 x 0.5 mm^2. The kernel has been normalized to the highest dose from the 16mm dose kernel.



Gamma Knife - 16mm - Dose Kernel

Gamma Knife - 16mm - Dose Kernel Contour

Figure 2.4: Top: Elekta Gamma Knife® 16mm dose kernel. Bottom: Elekta Gamma Knife® 16mm dose kernel contours. Grid resolution is 0.5 x 0.5 mm^2. The kernel has been normalized to its highest dose value.

Figures 2.5, 2.6, and 2.7 show an Elekta Gamma Knife® machine. Figure 2.5 shows the schematic of this machine. During treatment, the patient lies on a retractable table (patient's support). The patient's head is fixed to a helmet (depending on the version of the machine the helmet may not be present). The upper and lower doors open and the table carries the patient focusing the region to be treated into the focus or isocenter where a high dose radiation is delivered. Figure 2.6 shows the treatment room with the treatment chamber closed. Figure 2.7 shows a patient attached to the helmet immobilizer lying on the retractable table entering the treatment chamber.

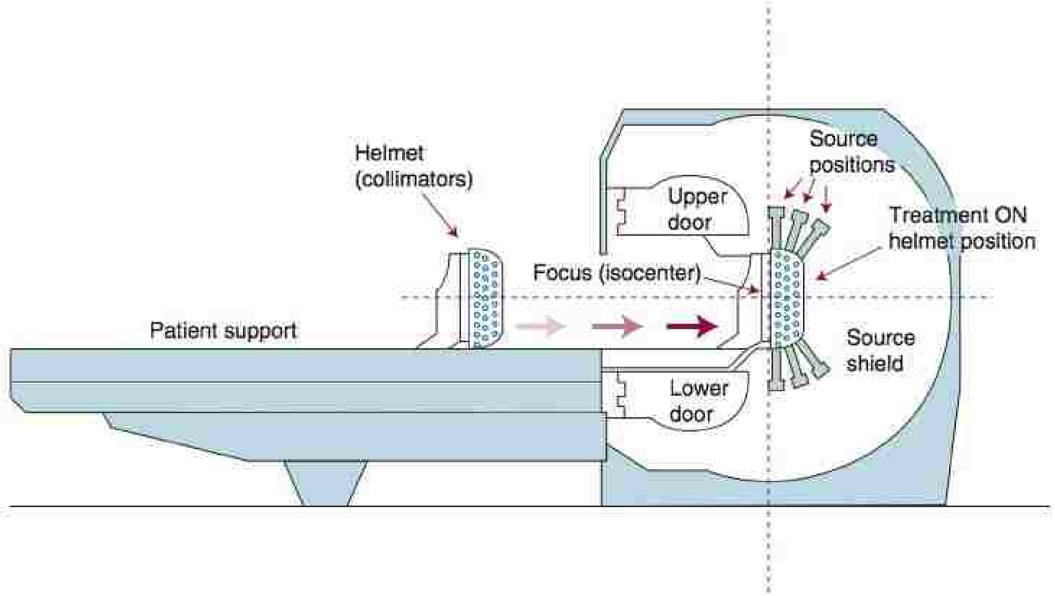


Figure 2.5: Elekta Gamma Knife® Schematic - the patient lies on the patient's support with his/her immobilized head in the helmet, then the patient is moved inside the chamber where treatment takes place [4].

Figure 2.6: Elekta Gamma Knife® machine [3].

Figure 2.7: Elekta Gamma Knife® - patient support and collimator helmet [3].

A Gamma Knife® procedure is essentially a "ball-packing" process, where the goal is to pack the spherical high dose volumes into a target tumor volume and create a radiation dose deposition that conforms to the target tumor volume and spares the surrounding normal tissues and structures. Even though there are some attempts to develop automatic planning systems for Gamma Knife® radiosurgery planning, for instance, the one developed by D. M. Shepard from the University of Maryland School of Medicine [39] which is an algorithm for finding optimal beam configurations, the procedure is still planned by physicians manually using a "trial-and-error" approach guided by a computer system.

The current Elekta treatment planning system, Gammaplan® 10, performs Gamma Knife® radiosurgery planning as a two-step process. First, the target is preprocessed through a packing algorithm that allows the use of templates, i.e., the result of a precomputed collimator set-up of its radioactive sources, also known as kernels, to fill in the tumor. This algorithm focuses on the geometry of this template up to an iso-dose line of tolerance. Usually, plans that target 50% of the iso-dose line target dose to the tumor are accepted clinically. In other words, 50% of the prescribed dose has to cover the tumor [39]. After filling is done, a stochastic optimization algorithm (simulated annealing) is run such that the position and weight of each template is optimized according to the prescribed dose provided by physicians [40].

In this work, we follow the two-step process commonly used for Gamma Knife® radiosurgery treatment planning systems and use pre-computed dose kernels (see figures 2.2, 2.3 and 2.4). These pre-computed dose kernels are calculated in water, i.e., a homogeneous medium. Figure 2.2 shows a heat-map of the center cross-section on top and its iso-dose contour on the bottom for the 4 [mm] dose kernel. Figures 2.3 and 2.4 do the same for 8 and 16 [mm] dose kernels, respectively. The shown kernels are normalized with respect to the highest dose from the 16 [mm] dose kernel.

2.1.3 Brachytherapy

Brachytherapy is the application of ionizing radiation from small radio isotope sources at short range to a tumor. There are three ways to apply this treatment: surface mold, interstitial treatment, and intracavity or intraluminal treatment. In *Surface mold*, an applicator is used to mold the surface to be treated and sources are fixed to it. This technique is used to treat superficial tumors such as ocular melanoma. In *Interstitial treatment*, sources are implanted either directly to tissues or through catheters that are previously inserted into the target. Last, in *Intracavity treatment*, applicator devices carrying radiation sources are inserted into body cavities [31, 4, 10].

Historically, radium and radon sources were used in brachytherapy treatment until the 1950s when artificially produced nuclides such as cesium-134 (^{137}Cs), ^{192}Ir, ^{60}Co, gold-198 (^{198}Au), and iodine-125 (^{125}I) became available [31, 41, 9]. Brachytherapy sources can be placed either permanently using radionuclides with short half-life (e.g., ^{198}Au and palladium-103 (^{103}Pd)) that stay in the tissue until their activity decays, or temporarily using sources of a higher activity (e.g., ^{192}Ir and ^{137}Cs) that are removed after the prescribed dose is delivered [31, 10]. Brachytherapy treatment can also be characterized by the energies used, i.e., low dose rate (LDR) brachytherapy energies range from 0.4 to 2 $\left[{}^{Gy}/_{h}\right]$. Medium dose rate (MDR) brachytherapy energies range from 2 to 12 $\left[{}^{Gy}/_{h}\right]$. Finally, high-dose rate (HDR) brachytherapy energies exceed 12 $\left[{}^{Gy}/_{h}\right]$ [31, 42].

In this study, we focus on interstitial high-dose rate (HDR) brachytherapy treatment using temporary implants and remote-controlled after-loading systems, which is one of the popular treatment modalities of prostate cancer [43, 44]. Prostate cancer is the most common disease among men in the United States aside from non-melanoma skin cancer and it is the second leading cause of death behind lung cancer.

This year alone, about 220,800 new cases will be diagnosed and about 27,540 men will die due to this illness [26, 13].

HDR brachytherapy treatment typically takes place in three phases: catheter placement, imaging and contouring, and dose delivery. *Imaging* usually comes from computer tomographies (CTs) or ultrasounds. The patient is prepared for surgery where the physician is responsible for placing the interstitial catheters (*catheter placement*) inside the prostate, guided by a grid applicator fixed to the patient and ultrasound imaging. Once enough catheters have been placed, a computer tomography (CT) image is taken (*imaging*) to accurately account for the location of the catheters, and *contouring* is made by the physician aided by computer systems. Finally, *dose delivery* takes place following the instructions generated in the treatment planning system used. These instructions describe how the after-loading mechanism should place the radiation sources into each interstitial catheter.

Dose Calculation for brachytherapy radiation sources

The calculation of dose distributions for common radiation source implants is done following The American Association of Physicists in Medicine (AAPM) Task Group No. 43 recommendations, which provides an analytical dosimetry formalism to perform both 2D (cylindrically symmetric line source approximation) and 1D (point source approximation) dose calculations [5, 45, 46, 47, 48, 49, 32].

In this study, we focus our attention on the 2D cylindrically symmetric line source approximation. The source is assumed to be a line segment of length L, and the dose deposited (dose rate) in the 2D plane can be calculated using

$$\dot{D}\left(r,\theta\right) = S_K \cdot \Lambda \cdot \frac{G_L\left(r,\theta\right)}{G_L\left(r_0,\theta_0\right)} \cdot g_L\left(r\right) \cdot F\left(r,\theta\right), \tag{2.2}$$

where r denotes the distance in $[cm]$ from the center of the source to the point of interest, and r_0 is a reference distance specified to be 1 $[cm]$. θ is the polar angle at the point of interest $\dot{D}(r,\theta)$ relative to the source longitudinal axis Z. The angle θ_0 is $^{\pi}/_{2}$ and is a reference angle that defines the source transverse plane. Figure 2.8 shows the schematic of the coordinate system used for brachytherapy dosimetry calculations.

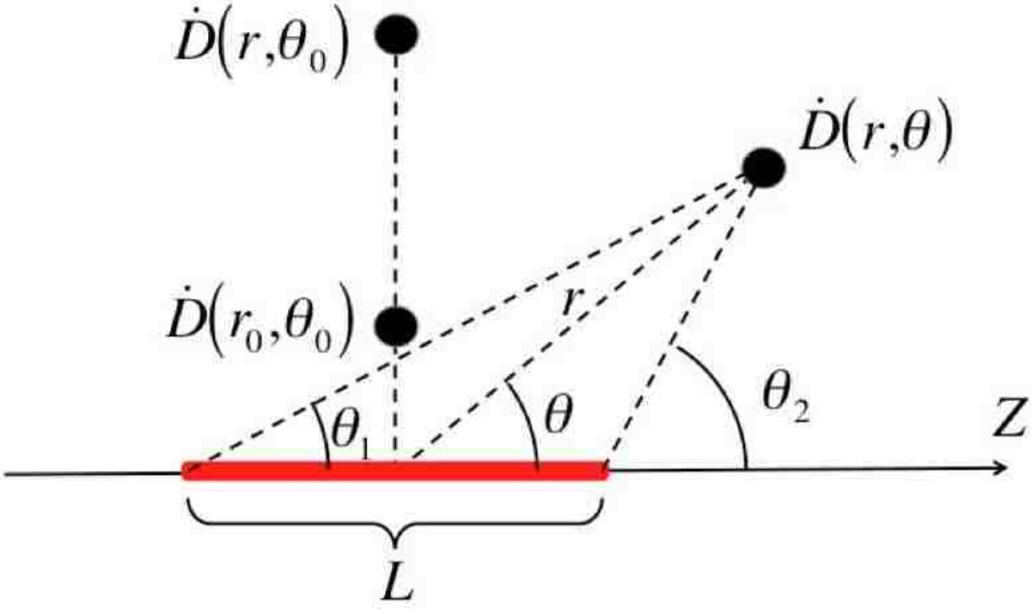


Figure 2.8: Coordinate system used for brachytherapy dosimetry calculations [5]. In red, the implant is shown as a line of length L. The dose rate at any position, $\dot{D}(r,\theta)$, is calculated with respect to the angles θ_1 and θ_2 and the dose rate at a known position $\dot{D}(r_0,\theta_0)$ which is perpendicular to the implant's axis.

Air-kerma strength is defined by

$$S_K = \dot{K}_\delta(d)\, d^2, \tag{2.3}$$

as the air-kerma rate, $\dot{K}_\delta(d)$, due to photons of energy greater than δ in vacuum at distance d. It has unit $1\ U = 1\ ^{\mu Gy \cdot m^2}/_{h}$.

Dose-rate constant is the dose rate in water at a distance of 1 $[cm]$ on the transverse axis of a unit air-kerma strength source. The following equation,

$$\Lambda = \frac{\dot{D}\left(r_0, \theta_0\right)}{S_K}, \tag{2.4}$$

is the mathematical expression for dose-rate constant.

Geometry function is defined by

$$G_L\left(r, \theta\right) = \begin{cases} \frac{\beta}{Lr\sin\theta} & \text{if } \theta \neq 0^\circ,\ \beta = \theta_2 - \theta_1, \\ \left(r^2 - {}^{L^2}\!/_{4}\right)^{-1} & \text{if } \theta = 0^\circ, \end{cases} \tag{2.5}$$

for a line-source approximation. The purpose of this function is to improve accuracy from interpolated data obtained from discrete points.

Radial dose function is defined by

$$g_X(r) = \frac{\dot{D}\left(r, \theta_0\right)}{\dot{D}\left(r_0, \theta_0\right)} \frac{G_X\left(r_0, \theta_0\right)}{G_X\left(r, \theta_0\right)}. \tag{2.6}$$

It accounts for the effects of photon absorption and scatter in the medium along the transverse axis of the source. Table 2.1 shows the data used for the MicroSelectron-HDR ^{192}Ir source (mHDR-v2r) dose calculation.

r, $[cm]$	$g_X(r)$
0.1	1.004
0.2	1
0.3	1.001
0.5	1
1	1
1.5	1.003
2	1.007
2.5	1.008
3	1.008
4	1.004
5	0.995
6	0.981
7	0.964
8	0.94
9	0.913
10	0.882
11	0.844
12	0.799
13	0.747
14	0.681

Table 2.1: Radial dose function, $g_X(r)$, for Nucletron® HDR ^{192}Ir source.

2D anisotropy function is defined by

$$F(r,\theta) = \frac{\dot{D}(r,\theta_0)}{\dot{D}(r_0,\theta_0)} \frac{G_L(r,\theta_0)}{G_L(r,\theta)}. \tag{2.7}$$

It describes the variation in dose as a function of the polar angle relative to the transverse plane.

Air-kerma strength, dose-rate constant, radial dose function, and 2D anisotropy function can be looked up in tables or in the description of the source/implant to be used. Tables 2.2 and 2.3 show the tabulated data for 2D anisotropy function,

$F(r, \theta)$, for the Nucletron® HDR ^{192}Ir source (mHDR-v2r). For all r or θ values within range and not tabulated, linear or bilinear interpolation is used. For values outside the range, nearest neighbor extrapolation is recommended [32].

r [cm] **θ in degrees**	**0.25**	**0.5**	**1**	**2**	**3**	**5**
0	0.729	0.667	0.631	0.645	0.66	0.696
1	0.73	0.662	0.631	0.645	0.661	0.701
2	0.729	0.662	0.632	0.652	0.67	0.709
3	0.73	0.663	0.64	0.662	0.679	0.718
4	0.731	0.664	0.65	0.673	0.69	0.726
5	0.733	0.671	0.661	0.684	0.7	0.735
6	0.735	0.68	0.674	0.696	0.711	0.743
7	0.734	0.691	0.687	0.708	0.723	0.753
8	0.739	0.702	0.7	0.72	0.734	0.763
10	0.756	0.727	0.727	0.745	0.758	0.782
12	0.777	0.751	0.753	0.769	0.781	0.804
14	0.802	0.775	0.778	0.791	0.802	0.822
16	0.82	0.797	0.8	0.812	0.822	0.84
20	0.856	0.836	0.839	0.846	0.854	0.872
24	0.885	0.868	0.869	0.874	0.877	0.888
30	0.92	0.904	0.902	0.907	0.906	0.911
36	0.938	0.93	0.929	0.931	0.934	0.933
42	0.957	0.949	0.949	0.955	0.956	0.954
48	0.967	0.963	0.965	0.965	0.969	0.965
58	0.982	0.982	0.982	0.982	0.983	0.978
73	0.994	0.997	0.997	0.998	0.996	0.985
88	0.997	1.001	1	1	1	1.001
90	1	1	1	1	1	1

Table 2.2: 2D anisotropy function, $F(r, \theta)$, for Nucletron® HDR ^{192}Ir source.

MicroSelectron-HDR ^{192}Ir source (mHDR-v2r)

In our work, we used the microSelectron-HDR ^{192}Ir source (mHDR-v2r) which is built by Nucletron® B.V. in The Netherlands. This radiation source is used worldwide

r [cm] **θ in degrees**	**0.25**	**0.5**	**1**	**2**	**3**	**5**
103	0.995	0.995	1.001	0.999	1	0.995
118	0.987	0.987	0.987	0.989	0.989	0.983
128	0.974	0.972	0.976	0.976	0.98	0.979
133	0.969	0.961	0.966	0.965	0.973	0.973
138	0.957	0.949	0.952	0.952	0.959	0.96
143	0.942	0.933	0.935	0.935	0.944	0.941
148	0.924	0.912	0.914	0.915	0.924	0.926
153	0.899	0.886	0.887	0.889	0.899	0.905
158	0.873	0.85	0.85	0.856	0.863	0.87
165	0.806	0.779	0.778	0.791	0.801	0.816
169	-	0.725	0.723	0.741	0.754	0.785
170	-	0.71	0.707	0.727	0.742	0.774
172	-	0.678	0.675	0.697	0.714	0.748
173	-	0.662	0.657	0.682	0.7	0.733
174	-	0.642	0.64	0.667	0.686	0.72
175	-	0.623	0.624	0.652	0.672	0.707
176	-	0.605	0.608	0.637	0.658	0.695
177	-	0.606	0.594	0.624	0.645	0.686
178	-	0.608	0.586	0.612	0.634	0.675
179	-	0.609	0.585	0.604	0.624	0.665
180	-	0.609	0.585	0.603	0.622	0.662

Table 2.3: 2D anisotropy function, $F(r, \theta)$, for Nucletron® HDR ^{192}Ir source (cont.).

[45, 6] and has 3.5 [mm] of active length, and its dose-rate constant is $\Lambda = (1.109 \pm 0.012)$ $\left[cGy\ h^{-1}\ U^{-1}\right]$. The active source is encapsulated and attached to a stainless steel retractable cable of 0.7 [mm] of diameter. Figure 2.9 shows the design of the source, and figure 2.10 shows the final dose distribution calculated using The AAPM Task Group 43 2D formalism.

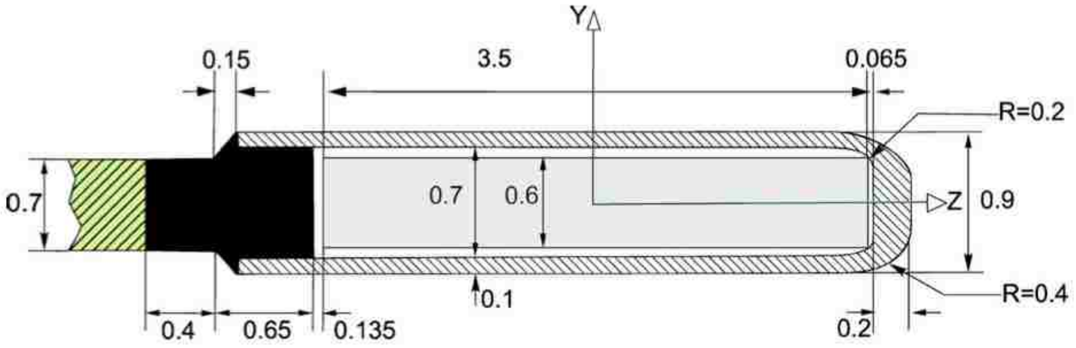


Figure 2.9: Nucletron® HDR ^{192}Ir model mHDR-v2r. The active radiation source is shown in gray, the encapsulation is shown as a dashed surface, and the retractable cable and its support are shown in dashed green and black respectively. All measurements are in mm. The drawing is not to scale [6].

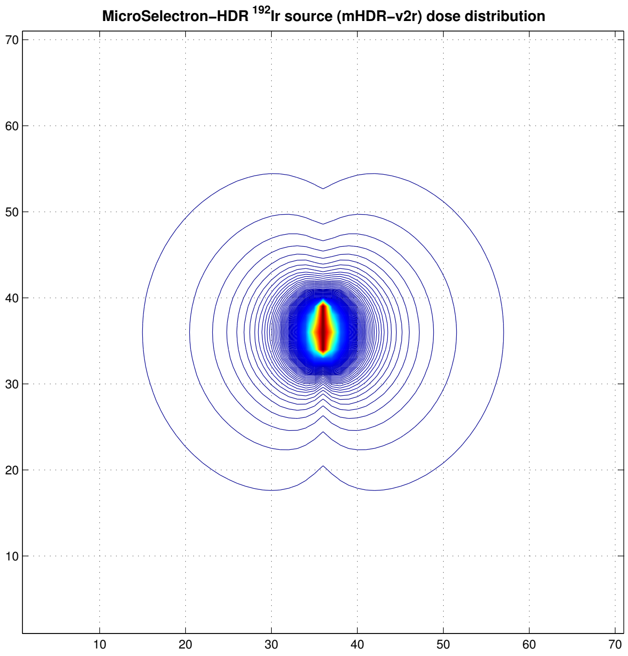


Figure 2.10: Nucletron® HDR ^{192}Ir model mHDR-v2r dose distribution cross-section iso-contour from 1% to 100% every 1%. Dose calculation was performed using The AAPM Task Group 43 2D formalism in a grid of 2 x 2 mm^2 resolution.

2.2 Treatment Planning

Modern radiation therapy treatment planning typically involves the following set of steps: patient imaging, target definition (i.e., structure contouring), dose prescription, beam configuration optimization, plan generation, and quality assurance [50].

Imaging is performed by taking computer tomography scans (CT scans), magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasound or combinations of these depending on the type of cancer. CT scan is the most widely used imaging modality and can provide anatomical information of the patient (see figure 2.11). Once these images are obtained, physicians *contour* the tumor and organs at risk (OARs) as well as *prescribe* the desired dose to treat the tumor.

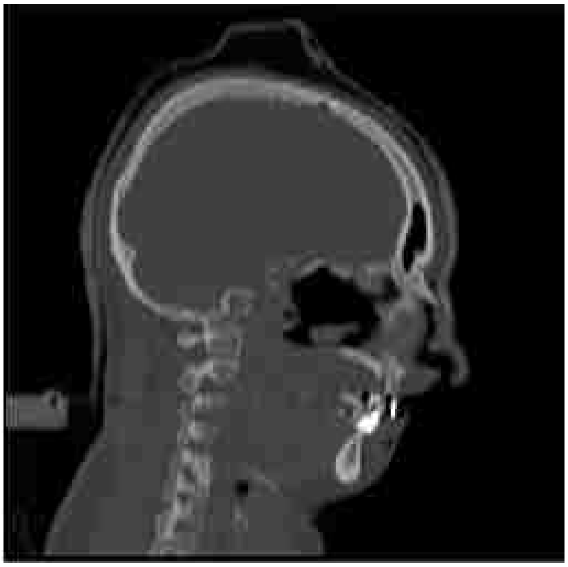


Figure 2.11: Lateral cut (CT scan) of a human head.

Modern radiation therapy relies on computer based optimization algorithms and software to generate the *beam configuration* for delivering the prescribed treatment. Generally speaking, a treatment planning system includes the following functionalities in order to provide an optimal treatment:

- **Patient Representation:** the computational model of a patient is represented as a three dimensional voxel array with resolution inherited from the type of imaging used (e.g., $1mm \times 1mm \times 1mm$). While contouring is performed, each voxel is associated with a particular structure which allows the identification of tumor voxels (target) versus organ-at-risk voxels.

- **Ideal Dose Distribution:** Aided by the prescription obtained from the physicians, a desired dose distribution is generated. A desired dose distribution usually consists of a maximum and minimum dose tolerance per structure. There are many ways to represent this ideal dose distribution. For this study, we shall assume the ideal dose distribution as two three-dimensional array of dose values corresponding to a maximum and minimum tolerance per structure's voxel.

- **Preparation for Dose Calculation:** The goal of treatment planning is to select a subset of beam configurations (e.g., beam energies, locations, angles, size, etc.) from a set of candidate beam configurations that can meet the ideal or prescribed dose distribution. Beam configurations are usually selected using randomized sampling or by applying other optimization algorithms. Once beam configurations are selected, the dose contribution from each candidate beam configuration is calculated before beam-on time optimization. This fact imposes a constraint on the cardinality of the candidate beam configurations set, since it is a process bounded by the amount of memory RAM available in the system.

- **Optimization Problem:** The conceptual optimization problem in radiation therapy can be stated as

$$\underset{t}{\text{minimize}} \quad \left\| \sum_j \dot{D}_j t_j - D^* \right\|_1 \quad \text{subject to } t_j \geq 0, \tag{2.8}$$

or

$$\underset{t}{\text{minimize}} \quad \left\| \sum_j \dot{D}_j t_j - D^* \right\|_2^2 \quad \text{subject to } t_j \geq 0, \tag{2.9}$$

where D^* is the ideal dose distribution or prescribed dose. $\dot{D}_j$ is the dose rate contributed by the $j-$th candidate beam, and t_j is the weighting or beam-on time for the $j-$th beam. The constraint $t_j \geq 0$ reflects that the beam-on time must be non-negative. Thus the goal of the optimization is to find the beam-on times t_j so that the created dose distribution $\sum_j \dot{D}_j \times t_j$ is as close to D^* as possible. Many metrics can be used to model the "closeness" (L^1 and L^2 norms are shown in equations 2.8 and 2.9, respectively), for example the non-negative least square problem [51] as shown in equation 2.9. The L^1 norm is defined by

$$\|x\|_1 = \sum_{m=1}^{n} |x_m|, \tag{2.10}$$

and the L^2 norm, also know as the Euclidean distance, by

$$\|x\|_2 = \sqrt{\sum_{m=1}^{n} |x_m|^2}. \tag{2.11}$$

In certain situations, there may be an additional constraint, such as the total beam-on time, $\sum_j t_j$, must be below a certain threshold, t_0, $\sum_j t_j \leq t_0$. The new optimization problem is a constrained least square problem [51].

- **Plan generation:** after optimization is over, the optimal result is processed in order to generate a set of instructions compatible with the radiation modality and technology used.

After a plan has been generated, physicians and medical physicists work together reviewing the plan. Then, they perform machine calibration and plan testing as part of *quality assurance.*

2.3 Optimization

Optimization is a fundamental tool in science, and it can be defined as the systematic process to achieve a pre-specified purpose. Generally speaking, the purpose is related to two states: maximize a benefit or minimize an effort. In exact sciences, optimization is highly coupled with mathematical models that describe a problem in terms of input or unknown variables and output or desired results. This functional representation allows for the construction of theoretical and numerical methods that explore the search space of possible solutions to find the unknown variables efficiently [52, 53, 54].

The classical mathematical optimization problem can be stated as

$$\underset{x}{\text{minimize or maximize}} \quad f(x) \quad \forall x \in \mathbb{X}, \tag{2.12}$$

where given a function $f(x) : x \in \mathbb{X} \mapsto \mathbb{Y}$; the goal is to find an $x^* \in \mathbb{X}$ such that: $\forall x \in \mathbb{X}$, $f(x^*) \leq f(x)$ when minimizing or $f(x^*) \geq f(x)$ when maximizing. $f(x)$ is called the objective function, x is referred to as the decision variable or unknown variable. The general mathematical optimization function described (see equation 2.12) is also known as an unconstrained optimization problem since no restrictions are imposed on the domain of f, $\mathbb{X}$. Hence, if a restriction is imposed on $\mathbb{X}$, for instance, $x \geq 0$, then the problem is categorized as a constrained optimization problem. Consider the following equation

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & c_j(x) \leq 0, \; j = 1, \ldots, m, \\ & c_j(x) = 0, \; j = m+1, \ldots, m+k, \end{aligned} \tag{2.13}$$

It shows the general form for a constrained optimization problem. The restrictions $c_j(x)$ correspond to a set of constraints as a function of x [55].

Optimization algorithms are classified according to different aspects of their formulation, for instance, according to the nature of their objective function, they can be linear, non-linear, quadratic, convex, concave, differentiable, non differentiable, etc. According to the domain of the decision variables, they can be continuous or discrete. In addition, they can be classified according to the method of operation into two groups: deterministic and probabilistic algorithms.

Deterministic algorithms are used when there exists a clear understanding of the characteristics of the possible solutions and their utility for a given problem. When there is not a clear understanding of a possible solution and its utility or the search space has high dimensionality, these algorithms may not succeed in finding a solution. Then, probabilistic algorithms are used [56, 55, 57]. Probabilistic or stochastic algorithms are those that sample the space of possible solutions for a given problem and through heuristics iteratively improve the solution.

In our work, we use a combination of both optimization classes. First we use a probabilistic algorithm: particle swarm optimization; and then, once we have lowered the complexity of the search space, we use deterministic algorithms: least-distance programming, non-negative least squares, least squares, etc. to refine our result.

2.3.1 Deterministic Optimization

Deterministic optimization algorithms refer to those algorithms that at a certain step of their execution have at most one way to proceed; if there is not one, then the algorithm terminates [56]. We detail non-negative least squares and least-distance programming because they are suitable for the type of optimization problems we discuss in this dissertation.

Non-negative least squares (NNLS)

Non-negative least squares is a constrained optimization algorithm. It is formulated as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \|Ax - b\|_2 \\ \text{subject to} \quad & x \geq 0, \end{aligned} \tag{2.14}$$

where A is a $m \times n$ matrix and both x and b are $n \times 1$ vectors [51]. The goal is to find a vector x such that Ax gets as close as possible to b, minimizing their squared difference. There are several numerical methods that solve a non-negative least squares problem, for instance the one implemented in Matlab®, which is based on the QR-factorization of matrix A and is an implementation of the method proposed by Charles L. Lawson and Richard J. Hanson in their book "Solving Least Squares" [51].

Least-distance programming

Least-distance programming is a constrained optimization algorithm. It is formulated as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \|x\| \\ \text{subject to} \quad & Cx \geq d, \end{aligned} \tag{2.15}$$

where x is a $n \times 1$ vector, C is a $n \times m$ matrix, and d is a $n \times 1$ vector [51]. The goal is to find the smallest (in length, with respect to the norm being used) vector x such that the linear constraints are met. The algorithm proposed on Lawson's and Hanson's book (please refer to [51] for more details) consists of solving a related

NNLS problem. Consider

$$\text{minimize } \|Eu - f\|_2 \text{ subject to } u \geq 0, \tag{2.16}$$

where $E = \begin{bmatrix} C^T \\ d^T \end{bmatrix}$ is $(n+1) \times m$ matrix and consists of an augmented matrix with the transposed values from C and d, and $f = [0, \cdots, 0, 1]^T$ is a $(n+1) \times 1$ augmented vector with the initial n elements zero and the $n+1$ element 1. Now, consider

$$r = Eu^* - f, \tag{2.17}$$

where u^* is the optimal solution of 2.16. Then, the optimal solution to 2.15 is computed as follows: if $\|r\| = 0$ computation is completed with no solution, otherwise x^*, the optimal solution for the initial problem, is updated with the values from r, $x_j^* = \frac{-r_j}{r_{n+1}}$. Notice that r is a $(n+1)$ vector.

2.3.2 Probabilistic Optimization

Probabilistic or stochastic optimization algorithms perform a guided search through the solution space, testing the "fitness" of each candidate solution from an initial random population [55]. These algorithms usually follow a common procedure as shown in table 2.4, where the fitness of each individual of the population is evaluated, and then a new population is created, using heuristics, based on one individual or many individuals with the highest fitness. In particular, particle swarm optimization follows this pattern and is studied in more detail in the following section.

```
Data: Initial Population, $P_0$, Maximum time, $T_{max}$.
Result: $P_t$.
Set $t \leftarrow 0$;
repeat
    Evaluate $P_t$ ;
    Set $t \leftarrow t + 1$;
    Update $P_t$ using $P_{t-1}$;
until A fitness condition is met on $P_t$ or $t$ exceeds $T_{max}$;
Return $P_t$;
```

Table 2.4: Classical Probabilistic Optimization Algorithm. Given both an initial population and a maximum iteration time, the algorithm updates the next population based on the characteristic of the given problem to solve. Once a certain fitness condition is met, the current population is returned.

Particle Swarm Optimization

Particle swarm optimization (PSO) is an stochastic optimization algorithm and is considered a modern method for solving complex problems. It was first proposed by Eberhart and Kennedy in 1995 in the context of solving continuous nonlinear functions. Eberthart and Kennedy's PSO algorithm was the result of realizing that the behavior of social behavior simulators (bird flocking, fish schooling, and swarming theory in general), where subjects have a location, speed or acceleration, and interactions among them, could serve as an optimization algorithm [58, 53, 59, 54, 7, 60, 61, 62, 63, 64, 65, 66].

In a nutshell, PSO consists on a iterative algorithm where each individual in the swarm is represented by a multidimensional vector of its location (i.e., a potential solution of the optimization problem), associated with a velocity vector that will determine its next position. This velocity vector is updated based on the current velocity and both the best location the individual and the global best location the swarm has explored. This process repeats until a fixed number of iterations or until a desired minimum error is achieved [58, 53, 59, 54, 7, 60, 61, 62, 63, 64, 67, 65, 66].

Since the initial formulation of PSO, the variations and applications of PSO have exploded exponentially demonstrating the vast potential of applications that this method possesses. Figure 2.12 shows the number of journal papers related to the words "particle swarm" published from 2000 to 2008 in Elsevier, Springer and IEEE [7].

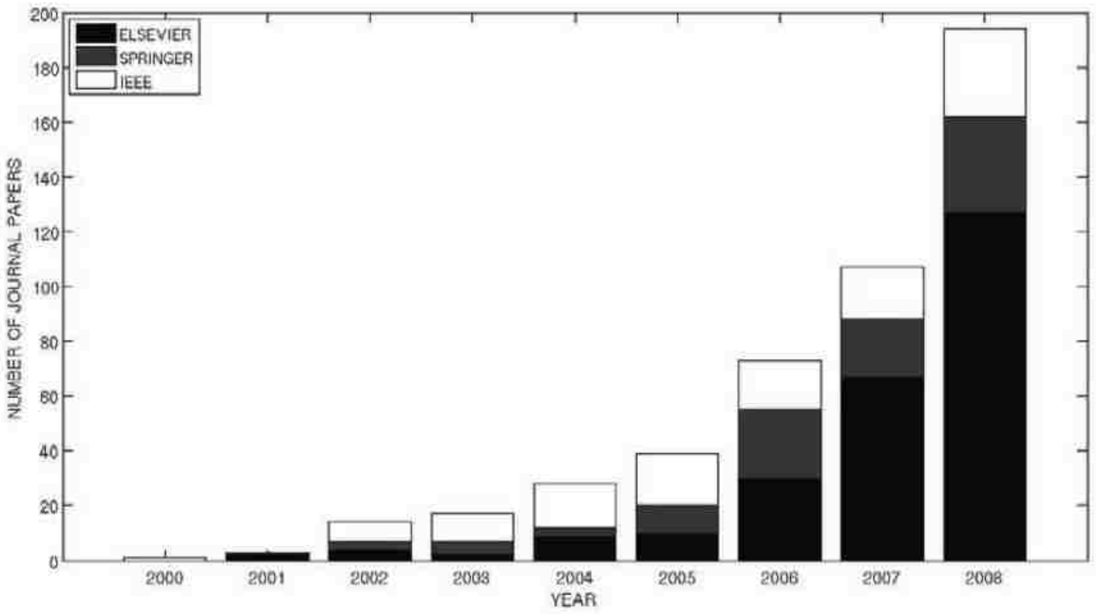


Figure 2.12: Number of journal papers with the term "particle swarm" in their titles, published by Elsevier, Springer, and IEEE, during the years 2000-2008 [7].

Algorithm

Formally, the classical particle swarm optimization is formulated as follows:

- Let $A \subset \mathbb{R}^n$ be the search space, and $f_i = f(x_i) : \quad x_i \in A \to f(x_i) \in Y \subseteq \mathbb{R}$ be available for all x_i.

- Let S be the swarm of N particles (potential solutions) be defined as $S = \{x_1, x_2, \ldots, x_N\}$.
- Let $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T \in A$, $i = 1, 2, \ldots, N$.
- Let each x_i be associated to a v_i, such that $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})^T$, $i = 1, 2, \ldots, N$.
- Let t denote the iteration counter, hence the position and velocity of the i^{th} particle at iteration t is $x_i(t)$ and $v_i(t)$, respectively.
- Let P be the set that stores the current best positions of particles, i.e. $P = \{p_1, p_2, \ldots, p_N\}$ which contains best positions $p_i = (p_{i1}, p_{i2}, \ldots, p_{in})^T \in A$, $i = 1, 2, \ldots, N$. Each $p_i(t)$ is defined as $p_i(t) = \arg\min_t f_i(t)$.
- Let g be the index of the best position with the lowest function value in P at iteration t, i.e. $p_g(t) = \arg\min_i f_i(p_i(t))$.

As shown in the pseudo-code in table 2.5, each particle's location and velocity get updated at iteration $t + 1$ by

$$v_{ij}(t+1) = v_{ij}(t) + \alpha_1 \cdot \varphi_1 \cdot (p_{ij}(t) - x_{ij}(t)) + \alpha_2 \cdot \varphi_2 \cdot (p_{gj}(t) - x_{ij}(t)), \quad (2.18)$$

and

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1), \qquad i = 1, 2, \ldots, N. \;\; j = 1, 2, \ldots, n. \quad (2.19)$$

The parameters φ_1 and φ_2 are random variables uniformly distributed in the interval $[0, 1]$. α_1 and α_2 were initially considered to have the same value, $\alpha = \alpha_1 = \alpha_2$, but latter they were divided and now, they are known as the *cognitive* and the *social* parameters, respectively. The new best positions at iteration $t+1$ are updated following the update rule

$$p_i(t+1) = \begin{cases} x_i(t+1), & \text{if } f(x_i(t+1)) \leq f(p_i(t)) \\ p_i(t), & \text{otherwise.} \end{cases} \tag{2.20}$$

```
Data: Number of particles, $N$; swarm, $S$, best positions, $P$.
Result: $P_g$.
Set $t \leftarrow 0$;
Initialize $S$ and set $P \equiv S$;
Evaluate $S$ and $P$, and define index $g$ of the best position;
while termination criterion not met do
    Update $S$ using equations 2.18 and 2.19;
    Evaluate $S$;
    Update $P$ and redefine index $g$;
    Set $t \leftarrow t + 1$;
end
Return $P_g$;
```

Table 2.5: Classical PSO Pseudocode [7]. Given a initial particle swarm, a new swarm configuration is calculated based on both the best location visited by each particle and the best global location visited by the swarm. Once a certain fitness or termination criterion is achieved the fittest particle is the solution.

Variations

In order to control that a particle does not step outside A, i.e., the solution search space, one simple variation proposed is to clamp the particle at its boundary. Hence, if A is restricted by rectangular boundaries, $A = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$,

$$x_{ij}(t+1) = \begin{cases} a_i, & \text{if } x_{ij}(t+1) < a_j, \\ b_j, & \text{if } x_{ij}(t+1) > b_j, \;\; i = 1, 2, \ldots, N. \;\; j = 1, 2, \ldots, n, \end{cases} \tag{2.21}$$

controls that particles stay within the desired search space.

Since its conception, the classical PSO algorithm has been refined in order to improve convergence, as well as its search capabilities. Due to the great deal of applications found in this topic, we focus on the historically general enhancements of the algorithm that have lead to the modern general PSO algorithm used in most applications today [64, 54, 7, 61].

It is important to mention that there are several variations of the algorithm that specialize on the way particles interact, e.g., charged swarms [66], and multi-swarms [65]; and, others, on the computational performance to solve large scale data problems using parallel and graphics processing units (GPU) [68, 69]. Latter in next chapter, we will discuss an interesting physical formulation made by Mikki and Kishk, when they argue about the close relationship between the state of physical systems and PSO [53].

Swarm Explosion: After the first formulation of the PSO algorithm, researches realized that as the number of iterations increased, $t \leftarrow t + 1$, velocity magnitude kept increasing, resulting in swarm divergence. The solution to this problem included adding a threshold verification to the velocity, after its evaluation (please refer to equation 2.18 in pseudo-code in table 2.5). The verification statement is perform by

$$|v_{ij}(t+1)| \leq v_{max}, \qquad i = 1, 2, \ldots, N. \;\; j = 1, 2, \ldots, n, \tag{2.22}$$

and

$$v_{ij}(t+1) = \begin{cases} v_{max}, & \text{if } v_{ij}(t+1) > v_{max}, \\ -v_{max}, & \text{if } v_{ij}(t+1) < -v_{max}. \end{cases} \tag{2.23}$$

Depending on the problem being solved, one can choose from using a unique maximum velocity,

$$v_{max} = \frac{\min\limits_{i} b_i - a_i}{k}, \tag{2.24}$$

or individual maximum velocities per particle,

$$v_{max,i} = \frac{b_i - a_i}{k} \qquad i = 1, 2, \ldots, n. \tag{2.25}$$

The parameter k is commonly chosen to be 2.

Inertia Weight: PSO algorithm showed additional convergence problems while exploring promising solutions which led to a key observation, “refined search in promising regions requires strong attraction of the particles towards them, and small position changes to control particles from escaping their close vicinity” [7]. For this purpose, a new parameter was introduced, w, or inertia weight. Equation 2.18 is replaced by

$$v_{ij}(t+1) = w(t) \cdot v_{ij}(t) + \alpha_1 \cdot \varphi_1 \cdot (p_{ij}(t) - x_{ij}(t)) + \alpha_2 \cdot \varphi_2 \cdot (p_{gj}(t) - x_{ij}(t)). \tag{2.26}$$

The inertia

$$w(t) = w_{up} - (w_{up} - w_{low}) \cdot \frac{t}{T_{max}} \tag{2.27}$$

is generally decreased as the algorithm approaches the maximum number of iterations, T_{max}. w_{up} and w_{low} are upper and lower bounds of w.

Neighborhood influence: One problem that the previous observation generated was fast convergence, i.e., the algorithm was susceptible to the loss of particle diversity due to the presence of local minima. This issue was an effect due to the influence of the best global solution, p_{gj}, known at iteration t. Hence, in order to reduce the amount of knowledge about the global best solution, particles were only allowed to share its best local solution within a neighbor topology. Equation 2.26 is rewritten as

$$v_{ij}(t+1) = w(t) \cdot v_{ij}(t) + \alpha_1 \cdot \varphi_1 \cdot (p_{ij}(t) - x_{ij}(t)) + \alpha_2 \cdot \varphi_2 \cdot (p_{g_i j}(t) - x_{ij}(t)). \quad (2.28)$$

The neighborhood is defined as

$$NB_i = \{x_{i-r}, x_{i-r+1}, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{i+r-1}, x_{i+r}\}, \quad (2.29)$$

where $\{i-r, i-r+1, \ldots, i-1, i, i+1, \ldots, i+r-1, i+r\} \subseteq \{1, 2, \ldots, N\}$. The best known global position in the neighborhood of particle i is updated by

$$p_{g_i j} = \underset{j \text{ s.t. } x_j \in NB_i}{\arg\min} f(p_j). \quad (2.30)$$

Standard PSO: In 2002, Clerc and Kennedy provided a solid theoretical background to their variation of the algorithm, which has become the contemporary PSO standard [70]. Equation 2.26 is replaced by

$$v_{ij}(t+1) = \chi \cdot \left[v_{ij}(t) + \alpha_1 \cdot \varphi_1 \cdot (p_{ij}(t) - x_{ij}(t)) + \alpha_2 \cdot \varphi_2 \cdot (p_{gj}(t) - x_{ij}(t)) \right] \quad (2.31)$$

and equation 2.28 is replaced by

$$v_{ij}(t+1) = \chi \cdot \left[\cdot v_{ij}(t) + \alpha_1 \cdot \varphi_1 \cdot (p_{ij}(t) - x_{ij}(t)) + \alpha_2 \cdot \varphi_2 \cdot (p_{g_{ij}}(t) - x_{ij}(t)) \right], \tag{2.32}$$

for the global best PSO model, $p_{gj}(t)$, and the local best PSO model (neighborhood) respectively. χ is known as the constriction coefficient and it is computed by

$$\chi = \frac{2}{\left| 2 - \phi - \sqrt{\phi^2 - 4 \cdot \phi} \right|}, \quad \text{where } \phi = \alpha_1 + \alpha_2 \text{ and } \phi > 4. \tag{2.33}$$

It has been shown that convergence is improved for $\chi = 0.729$ and $\alpha = \alpha_1 = \alpha_2 = 2.05$ [70].

Chapter 3

New optimization technique

> “What makes the desert beautiful,’ said the little prince,‘ is that somewhere it hides a well...”
>
> —Antoine de Saint-Exupéry, The Little Prince.

3.1 New optimization technique

In this section, we introduce our novel optimization approach that applies some general concepts from the PSO framework in radiation therapy planning. Particle swarm optimization has been widely used in many different fields (e.g., evolving artificial neural networks, training neural fuzzy networks, color image quantization, computational biology, modeling biomechanical movement, etc.), although there is very little evidence that it has been widely applied in the field of radiation therapy. In fact, the only reference identified in our searches regarding PSO and radiation therapy is the paper written by *Yongjie Li et. al.* titled: *“Adaptive particle swarm optimizer for beam angle selection in radiotherapy planning”*. This paper focused on Intensity-modulated radiation therapy (IMRT) and the ability of PSO to try several discrete sets of angle configurations to find the optimal set that provided a prescribed dose using a reasonable computational time [71].

In addition, as we briefly mentioned in section 2.3.2, an important parallelism between the PSO algorithm and physical systems made by Mikki and Kishk in their book: *"Particle Swarm Optimization: A Physics-Based Approach"*. In this book, the authors present a physical formalism for PSO by formulating the equations of molecular dynamics assuming that particles interact following a certain potential function. Although they mention several times that solving the system of second order diferential equations is by itself a very hard problem, possibly with no closed form solution, they argue that the importance of constructing physical methods can aid to tune the parameters used in the PSO algorithm [53].

As a result, we use the idea that physical models can be used to guide an optimization algorithm consisting of the interactions of a system of stationary and moving particles, such that a desired behavior can be simulated in order to obtain an optimum maximum or minimum state. We introduce a novel algorithm inspired by particle swarm optimization that evolves in time until mapping its state of minimum energy to optimal radiation sources' locations in radiotherapy and radiosurgery. We distance ourselves from the classical PSO algorithm where a particle represents a potential solution of the optimization function by using the whole swarm or particle distribution as the solution.

We model tumors, critical organs and other tissues as geometric volumes, whose surfaces have an associated potential function. The radiation source is modeled as kinetic particles subject to the forces from the potential functions from both the particles and the various geometric objects. The final configuration of the swarm of particles including their trajectories is the treatment plan. The intuition and motivation behind this potential model is to propagate the geometric characteristics of the treatment case throughout all the system so that a charged particle in the interior of the tumor can be aware of near-by critical organs or other tissues. Consider figure 3.1 that depicts the relative awareness of a set of kinetic particles when there is not a critical structure versus when there is one. The tumor is depicted in yellow and is covered by static charges. The critical structure is shown in blue, and it is also covered with static charges. In the interior of the tumor, there are kinetic particles that look for a minimum potential and stabilize. Notice that kinetic charges are displaced to the left once a critical structure is placed on the right; that effect is due to the awareness being influenced in the system due to the potential field associated to the critical structure.

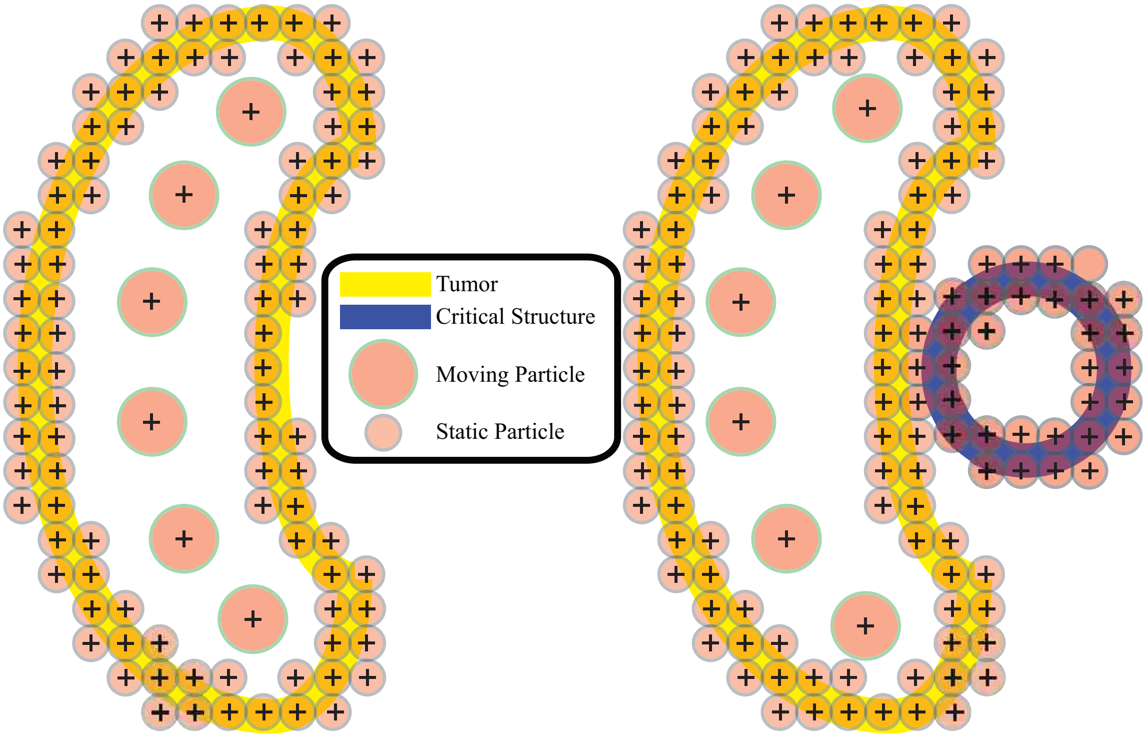


Figure 3.1: Geometrical awareness: the potential field associated to the tumor keeps kinetic particles in a lower potential inside the tumor (left schematic). In the presence of a charged critical structure to the right, kinetic particles find a lower potential inside the tumor that is displaced in comparison to the previous setup (right schematic).

In order to demonstrate the effectiveness of this model, we applied this new PSO framework to two challenging radiation therapy modalities: Gamma Knife® radiosurgery inverse planning and solving for beam location, and high-dose rate brachytherapy inverse planning for prostate cancer and solving for interstitial implant trajectories.

For the general case, consider the scheme depicted in figure 3.2, where charged particles are placed on the boundary of a container. These charged particles on the boundary are static. Several kinetic particles with different potential functions are placed inside the container with a positive charge. These kinetic particles move under the influence of a potential field. The intuition is that particles will move to lower the energy of the system and evenly spread out in the container. In the presence of a positively charged circular structure placed next to the container, the interaction induced by the particles on this structure pushes the kinetic charged particles further away from it. When kinetic particles are given an initial velocity, particles traverse the container in evenly spread trajectories.

We use these ideas to pack dose kernels for Gamma Knife® radiosurgery inverse planning, i.e., kinetic particles are given initial velocity zero and left to stabilize in the container, and to build interstitial implant trajectories in HDR prostate brachytherapy, i.e., kinetic particles are given an initial velocity in the direction used in current clinical surgeries and allowed to traverse the tumor.

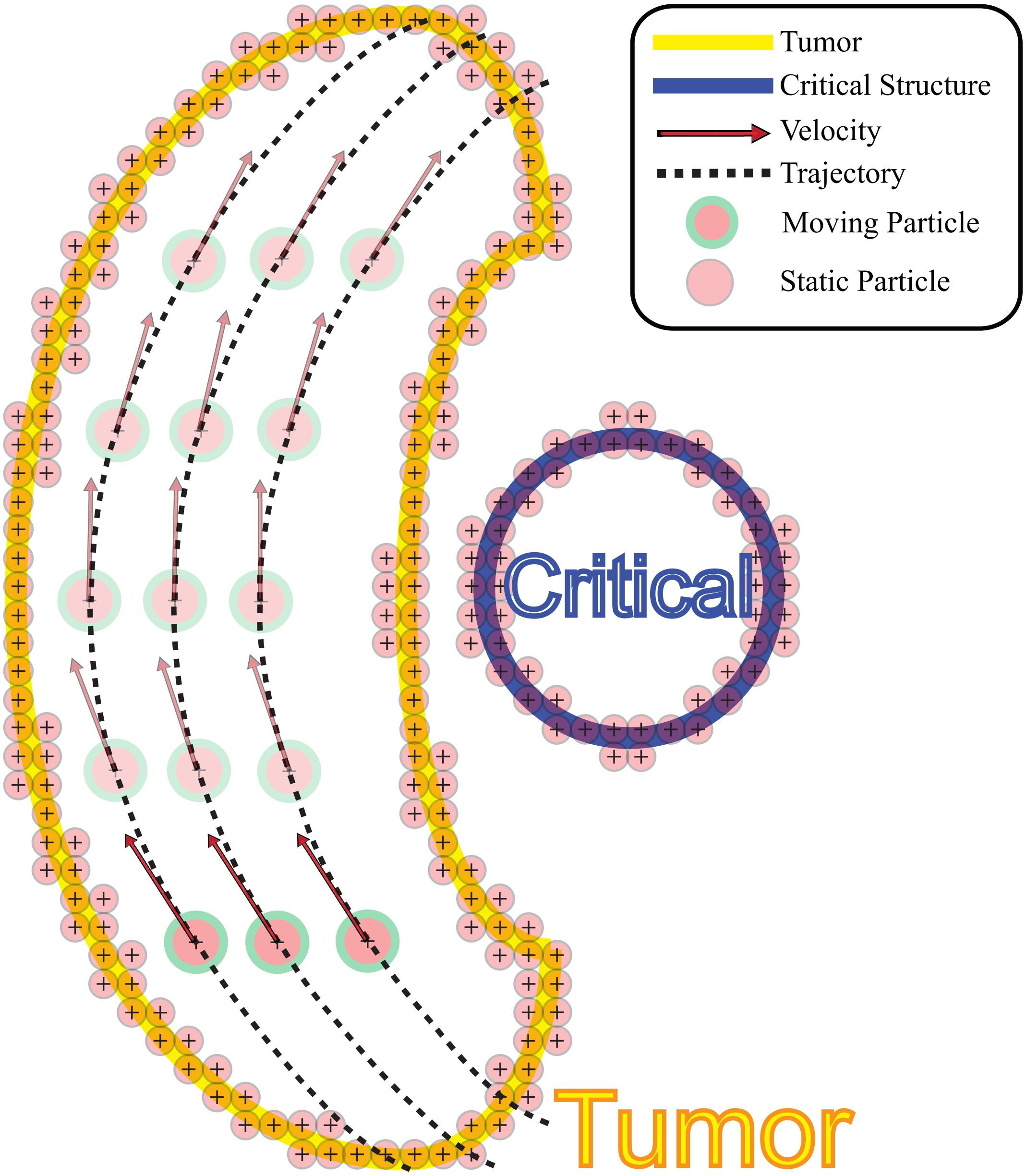


Figure 3.2: New optimization algorithm illustration: Static particles lie on the tumor's surface (in yellow) and the critical organ (in blue). Kinetic particles are placed inside the tumor (big red circles). In the presence of an initial velocity (red arrow), these particles move and produce evenly separated trajectories (dotted black lines).

3.1.1 Algorithm

In section 2.2, we showed a set of common steps followed by any treatment planning system. We use these steps here to formally define the general characteristics of our algorithm.

Patient Representation:

Patient geometrical information is described using set notation:

- Let V be the set of all voxels in a CT or MRI scan of the patient.
- Let T be the set of tumor voxels, and C_i be the set of voxels for the i^{th} critical structure. (Usually, there are multiple critical structures involved). Note that T and the C_i's are pairwise disjoint.
- Let ∂T and ∂C_i denote the voxels on the surface of tumor and the critical structures.

Ideal Dose Distribution:

The objective dose distribution

$$D^* : V \rightarrow [a, b] \in \mathbb{R} \tag{3.1}$$

is a function that maps all voxels in V into a real interval prescribed by the physician, which corresponds to the minimum and maximum dose per voxel. For critical structures, the minimum dose is 0, while the maximum dose is determined by organ/tissue specific radiation tolerance. For the tumor, the minimum dose is determined by tumor control probability, and the gap between minimum and maximum dose, i.e., $b - a$, indicates the required dose uniformity.

Preparation for Dose Calculation:

We define a particle, p_i, as an entity that has:

- Location: $\vec{x_i} = [x_i, y_i, z_i]$,
- Velocity: $\vec{v_i} = [v_{x_i}, v_{y_i}, v_{z_i}]$,
- Type: static or kinetic, and
- A potential function.

Optimization:

Optimization is performed in two phases: PSO and deterministic optimization. For PSO, we divide particles in two disjoint sets according to their type K or S, for kinetic and static particles, respectively. The particle potential functions are created from the corresponding radiation source dose distribution and the prescribed dose distributions. For static particles, we pre-compute the static potential field, U_S. For a kinetic particle, p_i, we compute the total external force incident on it by

$$\sum_{j \neq i} \vec{F_j} = -\vec{\nabla} \sum_{j \neq i} U_j = m \cdot \vec{a_i}, \quad m \text{ stands for the mass of particle } p_i, \tag{3.2}$$

i.e., the negative gradient of the total potential field of all particles. Notice that the total external potential incident on a particle can also be expressed by

$$-\vec{\nabla} U_i = -\vec{\nabla} \left(U_{K-\{p_i\}} + U_S \right). \tag{3.3}$$

Then, we calculate each particle's acceleration, $\vec{a_i}$, and update each particle's velocity

$$\vec{v_i}(t) = \vec{v_i}(t-1) + \vec{a_i}(t-1) \cdot \Delta t \tag{3.4}$$

and each particle's position

$$\vec{x_i}(t) = \vec{x_i}(t-1) + \frac{\vec{v_i}(t) + \vec{v_i}(t-1)}{2} \cdot \Delta t \tag{3.5}$$

in a given time step, Δt.

This process is repeated until a maximum time, T_{max}, is reached or a tolerance error, ξ, is achieved. Once a steady location of particles is achieved, we map particle positions to radiation sources' positions in the treatment plan. In the deterministic optimization phase, we use numerical techniques such as non-negative least squares, least distance programming, etc., to further refine the plan produced by PSO. The pseudo-code of our proposed algorithm is shown in table 3.1.

```
Data: Kinetic particle swarm, $K$; static particle pre-computed
      potential, $U_S$.
Result: $K_t$.
Set $t \leftarrow 0$;
Set $K_{t-1} \equiv Inf$ and $K_t \equiv K$;
while $t < T_{max}$ and $metric\,(K_{t-1}, K_t) \geq \xi$ do
    For each $p_i \in K_t$ calculate $\sum_i \overrightarrow{F_i}$ using equation 3.2;
    Update $K_{t-1} \leftarrow K_t$;
    Update $K_t$ using the motion equations 3.4, and 3.5;
    Set $t \leftarrow t + 1$;
end
Return $K_t$;
```

Table 3.1: Our PSO Algorithm: Given an initial kinetic particle swarm and a pre-computed potential, each particle position gets updated according to the total external forces. This process is repeated until a stable configuration of particles is reached or a minimum error is achieved. The algorithm returns the final stable swarm.

3.2 Implementation

In this section, we present the main strategies used to implement our new optimization technique. We chose Matlab® as our implementation environment. Matlab® is a scientific environment widely used around the world by many research teams and provides a great deal of built-in functions as well as the ability to add third-party libraries. We took advantage of the following Matlab® characteristics: the new object oriented programming extensions added to its scripting language, the direct integration with CERR (a computational environment for radiotherapy research) [72], and the direct integration with CVX (Matlab® software for disciplined convex programming) [73, 57].

CERR is an environment for radiotherapy research developed by Joseph O. Deasy and his collaborators from the Memorial Sloan Kettering Cancer Center, and although the project has been abandoned for a couple of years (no significant updates since 2007), it still provides great tools for loading DICOM (Digital Imaging and Communications in Medicine) files and contouring organs. Additionally, CERR still has an active forum that can aid anyone with questions [74]. A DICOM file is the international standard for medical images and related information (ISO 12052) [75].

CVX is a convex optimization environment developed by Stephen Boyd and his collaborators from Stanford University. CVX extends Matlab®'s scripting language with syntax that allows the formulation of mathematical optimization problems using a more natural mathematical notation. In addition, CVX allows the use of Matlab®'s optimization toolbox, CVX's own built-in optimization solvers, and third party solvers such as MOSEK and Gurobi [73].

In the following sections, we explain our treatment planning system (TPS) prototype accordingly to the common modules a TPS is expected to implement (see section 2.2 for more details): patient representation, ideal dose distribution, preparation for dose calculation, optimization, and plan generation.

3.2.1 Patient Representation

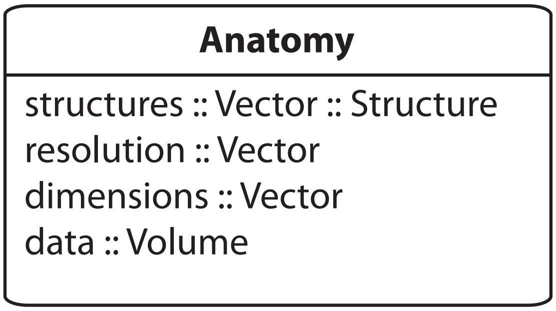
We define two classes, Anatomy and Structure (see figures 3.3 and 3.4), to handle patient representation. An anatomy object has four attributes:

- `structures`: refers to a vector of the subtype structure and stores all the different structures defined in the anatomy.
- `resolution`: refers to a vector that carries the resolution of each voxel in $[cm]$ of the `anatomy↦data` attribute.
- `dimensions`: refers to a vector that carries the $[x, y, z]$-dimensions of the `anatomy↦data` attribute.
- `data`: refers to a three-dimensional voxel array carrying the correspondent `structure↦id` on each voxel.

A structure object has seven attributes:

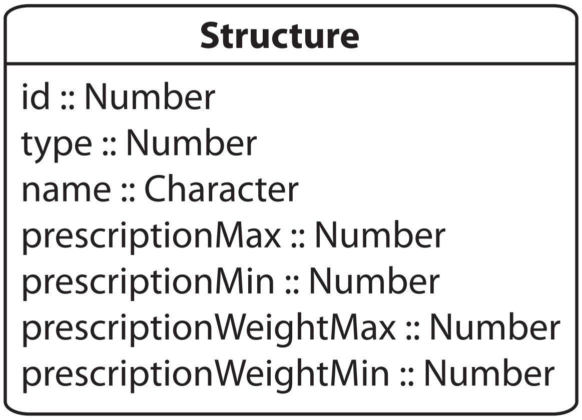
- `id`: refers to a numeric value that is a unique identifier number of a specific structure.
- `type`: refers to a numeric value that is either 0 for air, 1 for tissue, 2 for a dose targeted structure, and 3 for a critical structure.
- `name`: refers to a string that carries the name of the structure.
- `prescriptionMax`: refers to a numeric value that corresponds to the maximum prescribed dose to that structure.
- `prescriptionMin`: refers to a numeric value that corresponds to the minimum tolerance prescribed dose to that structure.

- `prescriptionWeightMax`: refers to a numeric value that corresponds to the relative importance of the maximum dose prescribed for that structure.
- `prescriptionWeightMin`: refers to a numeric value that corresponds to the relative importance of the minimum dose prescribed for that structure.


Anatomy

structures :: Vector :: Structure
resolution :: Vector
dimensions :: Vector
data :: Volume

Figure 3.3: Anatomy class definition: an anatomy object has the following attributes: structures, resolution, dimensions and data. In the figure, the entries follow this structure: <Attribute name>:: <type>:: <subtype>.


Structure

id :: Number
type :: Number
name :: Character
prescriptionMax :: Number
prescriptionMin :: Number
prescriptionWeightMax :: Number
prescriptionWeightMin :: Number

Figure 3.4: Structure class definition: a structure object has the following attributes: id, type, name, prescriptionMax, prescriptionMin, prescriptionWeightMax, and prescriptionWeightMin. In the figure, the entries follow this structure: <Attribute name>:: <type>:: <subtype>.

We support two ways to load an anatomy: CERR's loaded DICOM file representation or a specially formatted text file. In the first scenario, provided a DICOM file, we load it and use CERR for adding or editing contours. Once this representation is available in memory, we run our converter script and translate contours to binary masks and later identify those structure masks with an identifier number. Each structure is loaded providing a unique identifier and a type: 0 for air, 1 for tissue, 2 for a dose targeted structure, and 3 for a critical structure. We inherit the resolution and dimensions of the CT-scan from the DICOM file. In this process, we eliminate the CT-scan image from our patient representation.

In the second scenario, we read an anatomy file which has a header specifying the total number of structures followed by each structure definition (e.g., identifier, type, name, and prescription information) and volumetric data written as 2D slices from the anatomy dimensions.

3.2.2 Ideal Dose Distribution

The ideal dose distribution is calculated as two volumetric matrices, i.e., a three-dimensional matrix, such that one matrix contains the prescribed dose value per voxel for minimum prescribed dose (`Structure`↦`prescriptionMin`, see figure 3.4) and the second contains the prescribed dose value per voxel for maximum prescribed dose (`Structure`↦`prescriptionMax`, see figure 3.4). These matrices are converted to a one-dimensional vector and use later in optimization.

3.2.3 Preparation for Dose Calculation

In our implementation, we can switch from performing real-time dose calculations to using precomputed dose kernels. If kernels are to be used, we load kernels into a vector of kernel objects (see figure 3.5). A kernel object has three attributes:

- `resolution`: refers to a vector that carries the resolution of each kernel voxel in $[cm]$.

- `centerCoordinates`: refers to a vector that carries the coordinates of the center of the kernel dose data.
- `data`: refers to a three-dimensional voxel array carrying the correspondent dose rate information for the kernel.

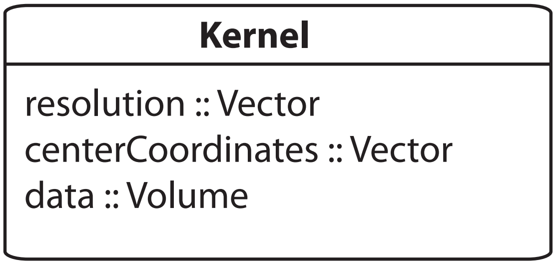


Figure 3.5: Kernel class definition: a kernel object has the following attributes: resolution, centerCoordinates, and data. In the figure, the entries follow this structure: <Attribute name>:: <type>:: <subtype>.

Due to memory constraints, before running any dose calculation or dose optimization, the size of the previously calculated matrices are reduced. We calculated a bounding box which contains all tumor voxels and preferably those voxels from the most critical organs at risk. The margins of the bounding box can be configured by the user depending on the amount of RAM available in the system.

The tolerance margin is an user input, that can be configured in the main script. In this phase, kinetic particles are created and randomly placed inside the tumor T. A particle object has four attributes (see figure 3.6):

- `position`: refers to a vector that carries the $[x, y, z]$ values of the location of a particle.
- `velocity`: refers to a vector that carries the $[vx, vy, vz]$ values of the current velocity of a particle.

- `type`: refers to a numeric value 0 for a static particle and 1 for a kinetic particle.
- `potentialFunction`: refers to a kernel data object or a function representing the potential of that particle.

Only kinetic particles are used in our calculations. The total potential field produced by static particles (U_S) is precomputed using the location of each voxel at ∂T and ∂C. In the event that a kernel's resolution (`Kernel↦resolution`) mismatches that of the anatomy (`Anatomy↦resolution`), we either perform bilinear interpolation or increase the resolution of the kernels to match that of the anatomy.

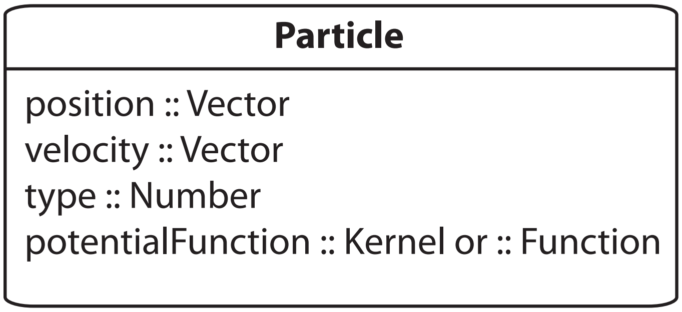


Figure 3.6: Particle class definition: a particle object has the following attributes: position, velocity, type and potentialFunction. In the figure, the entries follow this structure: <Attribute name>:: <type>:: <subtype>.

3.2.4 Optimization

Depending on the application, kinetic particles are initialized with an initial speed. For the initial PSO phase, we loop over all kinetic particles one at a time. The current particle is excluded from the current total potential field (U). We compute the negative gradient of the potential field ($-\nabla U$) and update the position of the current particle using the motion equations 3.2, 3.4, and 3.5. After updating all particles, we compute the total distance variation of the whole particle distribution; if it is less than a given threshold, the process stops, otherwise the process repeats.

```
cvx_begin
     variable x(n) nonnegative;
     variable s(m) nonnegative;
     variable t(m) nonnegative;
     minimize norm(x, 2) + ...
              norm(lambda_M*s,2) +...
              norm(lambda_m*t,2);
     subject to
         D*x - s <= ideal_dose_max;
         D*x + t >= ideal_dose_min;
 cvx_end
```

Table 3.2: CVX syntax for deterministic optimization.

These previous calculations are performed in parallel taking advantage of Matlab®'s parallel for-loop, `parfor`. In the event that particles' trajectories are important, they are stored in a trajectory matrix, and smoothed using a third-degree polynomial in x and y coordinates, using the z coordinate as the characteristic parameter.

Finally, we use CVX syntax (see table 3.2) and formulate an optimization problem to solve for dwell times (`x`) such that the final dose distribution (`D*x`) is within the prescribed ideal dose (`ideal_dose_min` and `ideal_dose_max`). We tried different solvers from those available in CVX, and picked MOSEK for its impressive running times. For more details on the interior point method implemented by MOSEK please refer to [76].

3.2.5 Plan Generation

In this phase, we filter out the dwell times that are negligible and compute a final dose distribution. We generate a dose-volume histogram from the final dose distribution and produce a plot. In addition, we generate plots for the final 3D distribution using our 3D distribution plotter script and save the final dose distribution, the dwell times vector, and the kinetic particle vector. We stop after saving this file, since we are not currently producing a final plan for a specific machine.

3.2.6 Other tools

- **3D Distribution plotter:** this script receives an anatomy object, final dose distribution three-dimensional matrix as parameters and draws the 3D surfaces of each structure, and produces an iso-surface of the targeted dose. If particles locations are provided, it shows the particles and their trajectories.

- **Anatomy loader and Phantom generator:** this script loads anatomy files compatible to our anatomy file syntax. In addition, it can save an Anatomy object into the same syntax.

- **Dose-volume histogram plotter and verifier:** this script calculates and plots a DVH. In addition, provided the path of a previous DVH, it overlays both DVHs plots for comparison.

- **Kernel interpolator:** this script has been used to perform dose kernel interpolation in the event that the kernel object and the anatomy object have a different resolution.

- **Kernel loader:** this script has been used load a precomputed kernel object that has been stored as a text file under our own syntax.

Chapter 4

Gamma Knife® radiosurgery

> “Well, I must endure the presence of a few caterpillars if I wish to become acquainted with the butterflies.”
>
> —Antoine de Saint-Exupéry, The Little Prince.

4.1 Overview

In section 2.1.2, we mentioned that current gamma knife® radiosurgery is performed as a two-step process: tumor filling and optimization. For dose calculation, we use precomputed dose kernels. We map Gamma Knife® dose kernels to a potential function associated to each particle in the swarm and simulate the interaction among them until our swarm reaches a stable configuration, i.e., the energy-stable state of the system. Particles’ locations are then translated to dwell locations and then the non-negative least squares algorithm is used to calculate the dwell time of each kernel aiming to meet the dose prescription.

4.2 Mapping

Patient Representation:

Let T and C be the set of tumor voxels and critical structure voxels, and ∂T and ∂C be their respective surface voxels.

Preparation for Dose Calculation:

For the initialization, we randomly place static particles on the surface ∂T and ∂C.

Each Gamma Knife® high dose kernel has a dose falloff which is a complementary error function, $erfc(r)$, as shown in 4.1:

$$erfc(r) = 1 - \frac{2}{\sqrt{\pi}} \int_0^r e^{-t^2} dt. \tag{4.1}$$

Since the calculation of the $erfc(r)$ function is a time consuming task, instead we use $\frac{1}{r}$ as its surrogate in order to speed-up computations. Please refer to figure 4.1 where we show a comparison between the real dose falloff of the 4 $[mm]$ Gamma Knife® dose kernel versus a scaled $\frac{1}{r}$ function.

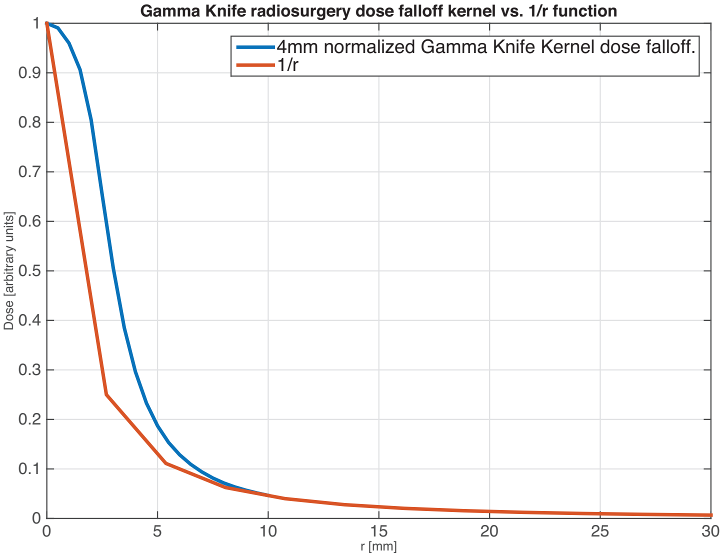


Figure 4.1: Gamma Knife® 4$[mm]$ dose kernel lateral falloff versus the $\frac{1}{r}$ function. $\frac{1}{r}$ serves as a surrogate to the real dose falloff of Gamma Knife® dose kernels.

We assign to each particle a $\frac{1}{\alpha \cdot r}$ potential function, where r is the distance from the particle to a voxel, and α is a constant scaling factor directly proportional to the dose kernel radio to be used in the simulation.

The intuition behind such an arrangement is that the kinetic particles, i.e. the spherical high dose volumes, should not be too close to each other, and should not be too close to the boundary or the critical structures.

Optimization:
The kinetic particles are initialized at random positions with zero initial velocities. The quantity and potential functions associated to each particle are calculated from an approximate solution obtained from the following constrained integer-linear problem

$$\begin{aligned} \underset{y_i \in Z^+}{\text{minimize}} \quad & |a \cdot y - b| \\ \text{subject to} \quad & \|y\|_1 \leq n, \end{aligned} \tag{4.2}$$

where a is a row vector containing the volume each kernel can cover with a high density dose, y is a column vector representing the distinct kernel spot sizes to be used, b is the cardinality of T, and n is the cardinality of K. The potentials U_S and U_K created by static and kinetic particles respectively are computed. In each iteration, the locations of the kinetic particles are updated based on the forces from the potential field until they converge. The PSO phase outputs the locations of each spherical high dose volume. To complete a radiosurgery plan, we also need the dwell or beam-on times for each location. Let $\dot{D}_j$ and t_j be the dose rate and beam-on times for each particle/spherical high dose volume, and D^* be the optimal desired dose distribution. We use non-negative least squares

$$\min \left\| \sum_j \dot{D}_j t_j - D^* \right\|_2^2 \quad \text{subject to } t_j \geq 0 \tag{4.3}$$

to calculate the optimal beam-on times.

4.3 Results

We have applied our optimization algorithm to a challenging phantom study. The phantom consists of a C-shaped 3D phantom surrounding a spherical critical structure (see figure 4.2). The prescription for this case targets 100 arbitrary dose units to the tumor while delivering 0 arbitrary dose units to the critical organ. We found that convergence is typically achieved within 20 to 40 iterations.

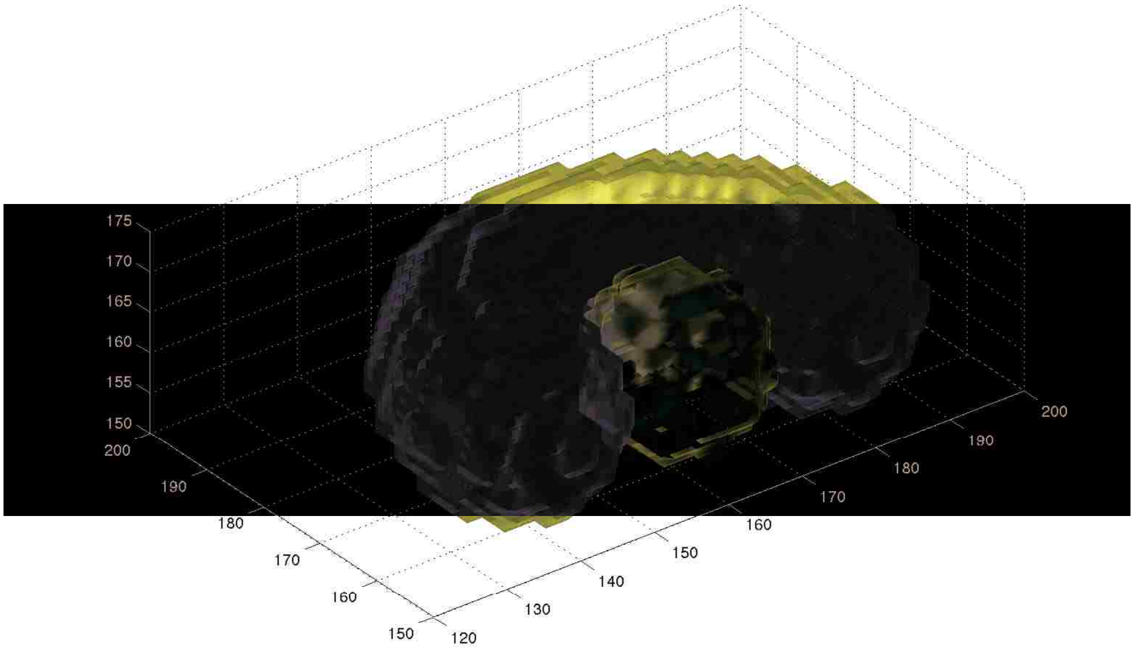


Figure 4.2: Phantom used for Gamma Knife® case study: C-shaped tumor is shown in yellow. The critical organ is shown in blue.

We started our tests by using homogeneous kernel spot sizes; in particular, we used the 4 $[mm]$ dose kernel. Figure 4.3 shows the evolution of particles in time.

Kinetic particles start at random locations inside the tumor and at each iteration they evolve, spreading evenly throughout the volume through the dynamic interactions among them and the potential field from the structure's surfaces. Particles get pushed in the opposite direction from the critical organ, which reflects the geometrical awareness imposed onto them by the potential field associated with the critical organ.

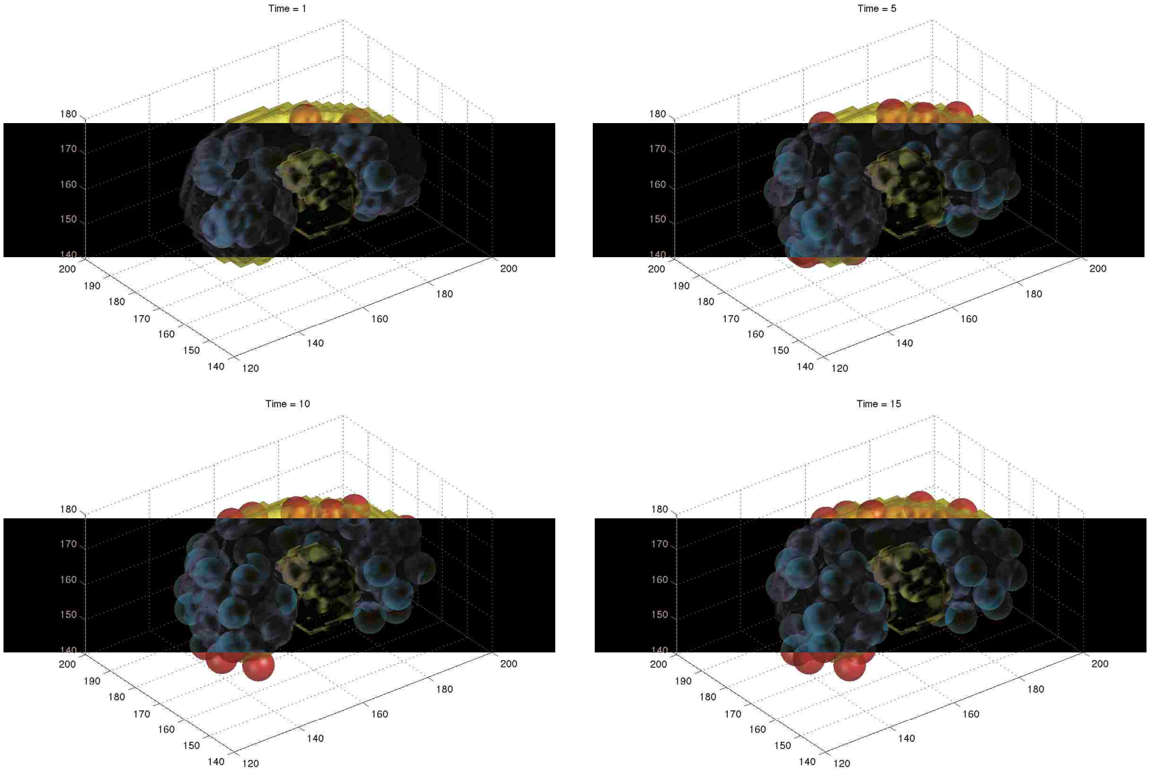


Figure 4.3: Evolution: Particles (shown as red spheres) distribute evenly along the tumor's volume while avoiding those regions close to the critical organ. The 1^{st}, 5^{th}, 10^{th}, and 15^{th} iterations are shown.

After 25 iterations (see figure 4.4), the spherical high dose volumes have converged to a steady location configuration. Now, the final swarm configuration is used as the 4 $[mm]$ kernel centers and non-negative least squares is used to filter out those locations that are either redundant or that do not contribute to meeting the prescription.

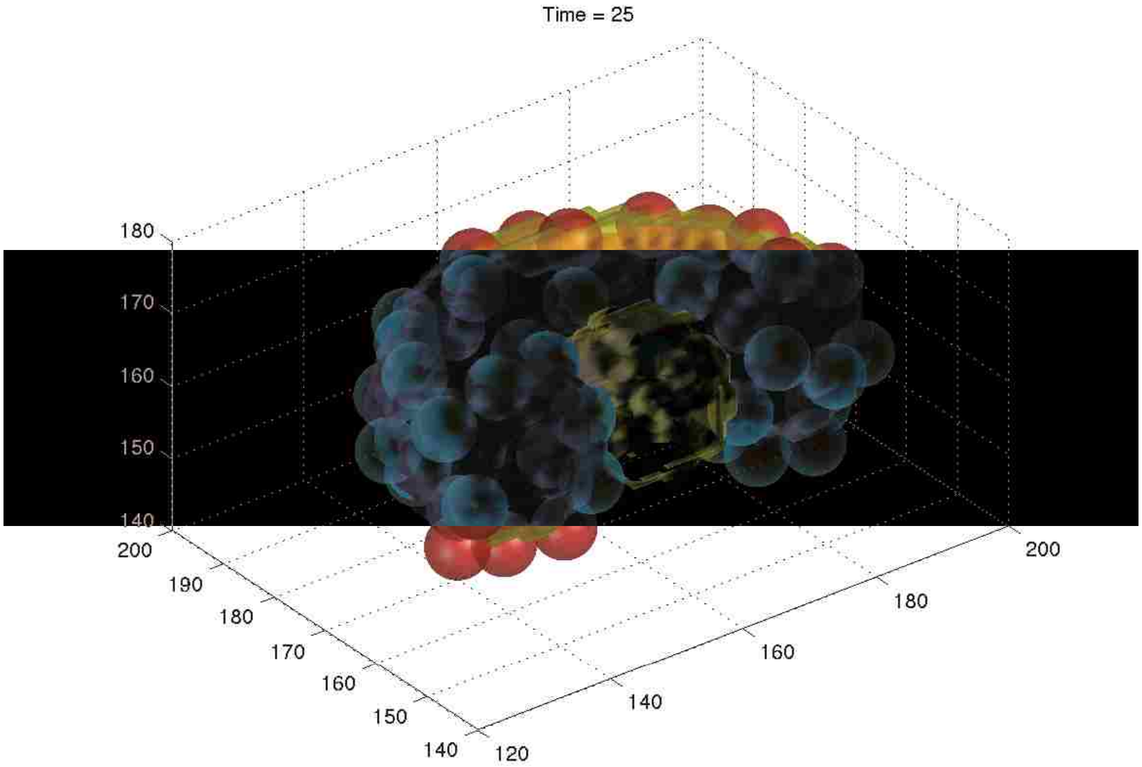


Figure 4.4: Particles (shown as red spheres) distribute evenly along the tumor's volume while avoiding those regions close to the critical organ.

Figure 4.5 shows a 2D cross-section of the phantom's final dose distribution after NNLS (on top), and also shows the dose-volume-histogram of the final plan after numerical optimization (bottom). In the DVH plots, the horizontal axis is the dose, while the vertical axis is the percent volume. There is a DVH line for each structure of interest, which describes the amount of dose delivered to a given percent volume. The quality of our plan is comparable to the manually obtained clinical plans, which aims to cover the target tumor volume with 50% of the maximum dose.

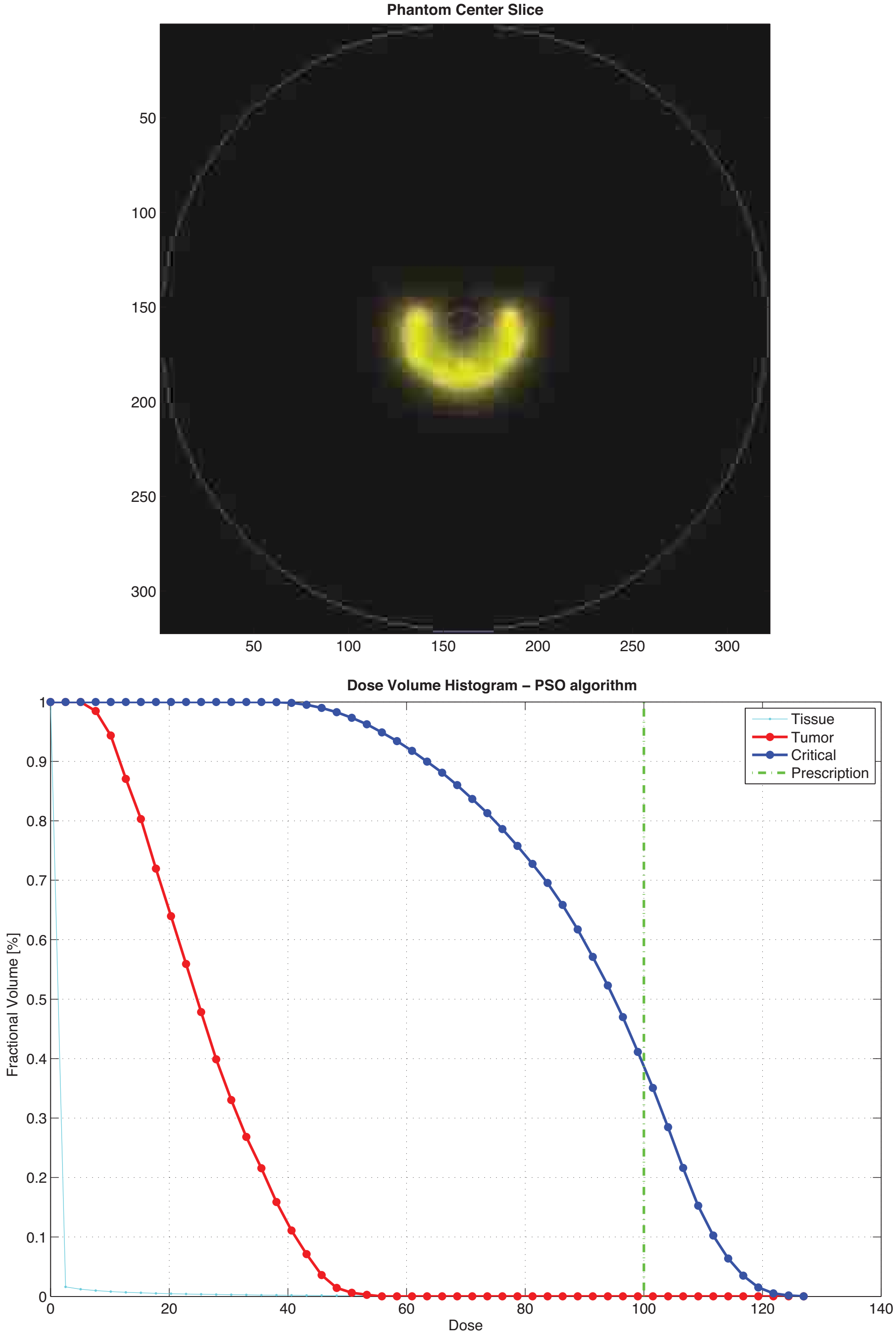


Figure 4.5: Phantom center cross-section: final dose distribution is shown in yellow (on top). Shown in blue are the tumor and critical organ contours. Gamma Knife® final plan DVH using 4 [mm] dose kernels (bottom).

We also tested our algorithm using heterogeneous kernel spot sizes. We obtained from equation 4.2 the kernel spot sizes that approximately cover our C-shaped tumor and use four 4 $[mm]$ and fourteen 8 $[mm]$ dose kernels. We run our simulation and find that particles present a similar behavior as for the homogeneous case; nonetheless, bigger dose kernels do not necessarily occupy those locations with bigger volumes, and smaller dose kernels do not necessarily fill the gaps between bigger dose kernels nor get pushed towards the boundary to balance those regions missed by bigger dose kernels. Figure 4.6 shows the final dose distribution as well as the final plan DVH. It is worth mentioning, that since larger dose kernels are used, the plan is less uniform, but still meets clinical goals.

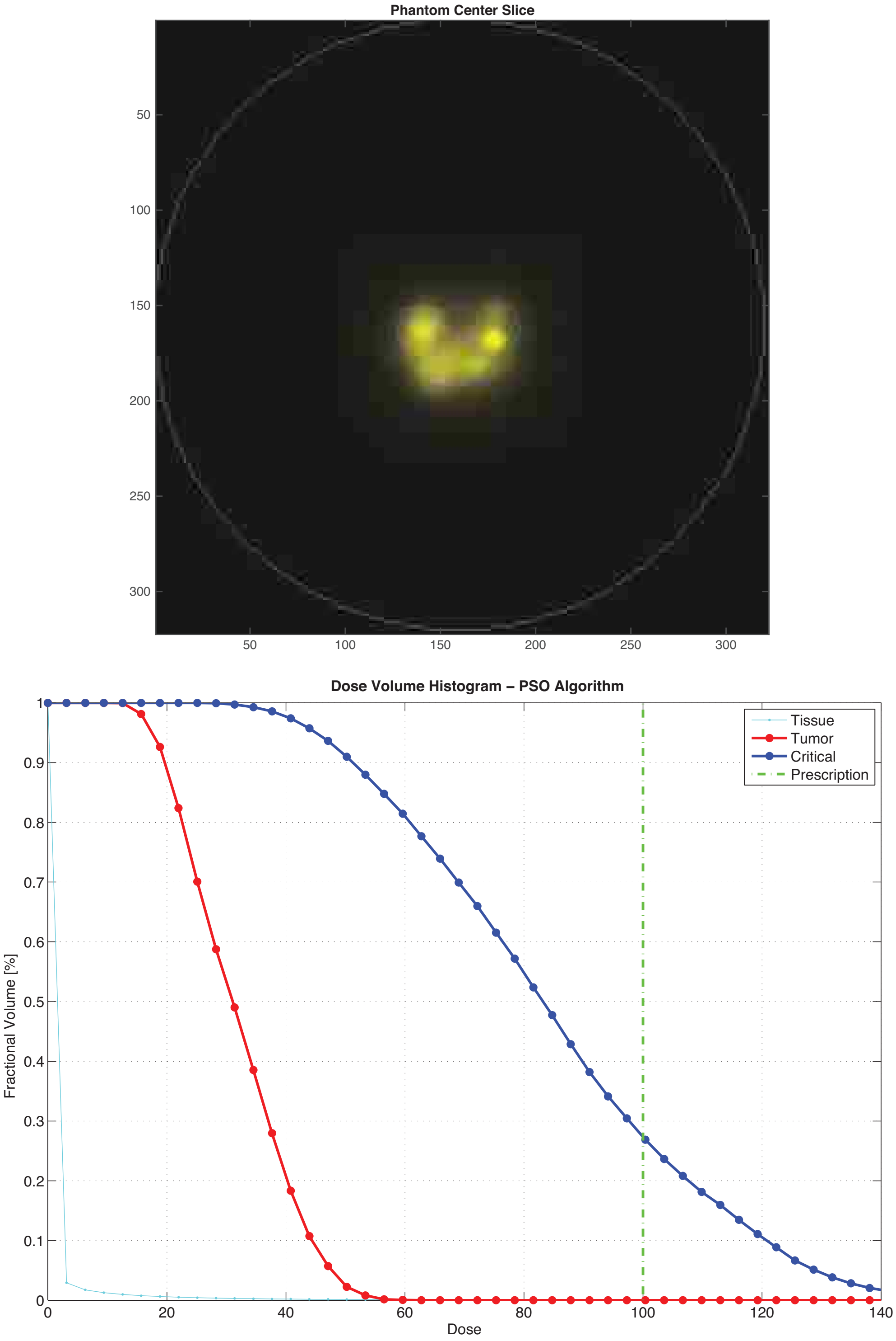


Figure 4.6: Phantom center cross-section: final dose distribution is shown in yellow (on top). It corresponds to an heterogeneous use of dose kernels. Shown in blue are the tumor and critical organ contours. Gamma Knife® final plan DVH using 4 [mm] and 8 [mm] dose kernels (bottom).

Chapter 5

High-dose rate (HDR) brachytherapy

> "It is the time you have wasted for your rose that makes your rose so important."
>
> —Antoine de Saint-Exupéry, The Little Prince.

5.1 Overview

In section 2.1.3, we mentioned that HDR brachytherapy is performed as a four-step process: catheter or implant placement, image contouring, dwell time optimization, and dose delivery. We map each kinetic particle to a radiation source and impose a set of initial conditions, which include defining their starting positions and their initial velocities in the direction of previous clinical implants. The key idea of this simulation is to let the particles in the swarm to traverse the tumor and record its particles' trajectories. We map the trajectory of each particle as the implant trajectory, and the final dose delivered is calculated using each particle position as the center of a radiation source, further optimized using least-distance programming.

5.2 Mapping

Patient Representation:

Let T and C be the set of tumor voxels (i.e., prostate) and critical structure voxels (i.e., rectum, bladder, and urethra), and ∂T and ∂C be their respective surface voxels.

Preparation for Dose Calculation:

We modeled each needle as the trajectory of a kinetic particle with the potential function $\frac{1}{r^2}$, where r is the distance to the particle. We have chosen this potential function, $\frac{1}{r^2}$, because it is a good surrogate of the dose fall-off rate of the ^{192}Ir source (please refer to The AAPM Task Group 43 point source dose calculation approximation [47, 48]). Figure 5.1 shows the lateral dose falloff of a HDR brachytherapy ^{192}Ir radiation source versus the $\frac{1}{r^2}$ function.

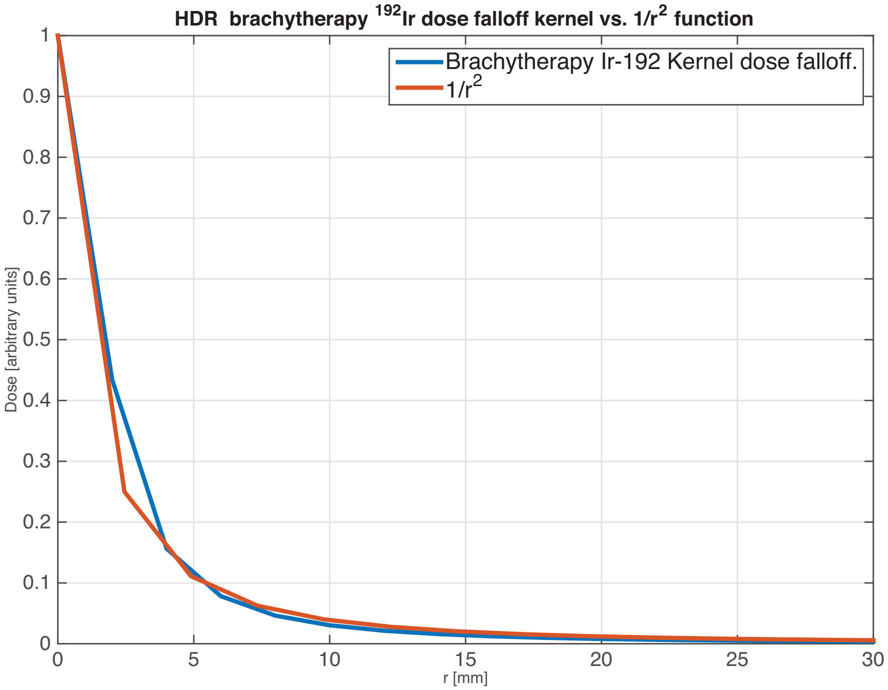


Figure 5.1: HDR brachytherapy ^{192}Ir lateral dose falloff versus the $\frac{1}{r^2}$ function. Notice how $\frac{1}{r^2}$ serves as a close approximation to the real dose falloff.

Similarly to the Gamma Knife® case, we randomly place static particles on the surface of the prostate and various critical structures such as the rectum, urethra, and bladder. The potential function of the static particles is $\frac{1}{r^2}$. The key to this PSO simulation is the initial velocity of the kinetic particles. Clinically, the bevel needles used have a curvature constraint and can bend up to a certain degree value.The initial velocity of the kinetic particles ensures that the curvature constraint will not be violated during the simulation. The total number of kinetic particles, i.e., the number of needles, is specified by the user.

Optimization:

To determine the initial position of the kinetic particles, we run another PSO optimization by placing the kinetic particles randomly in a cross-section of the tumor with zero velocity and constrained to move within that cross-section. Once the particles stabilize, they are given a velocity vector parallel to the principal needle direction typically used in a clinical setting. After the trajectories converge, a third-degree polynomial regression is applied to smooth the final particles' trajectories, which are the final needle positions. To calculate the dwell time, we use the following formulation of a least-distance programming (LDP) problem

$$\begin{array}{ll} \text{minimize} & ||x||_2^2 + \|\Lambda_{\max} s\|_2^2 + \|\Lambda_{\min} t\|_2^2 \\ \text{subject to} & Dx - s \leq b_{\max}, \\ & Dx + t \geq b_{\min}. \end{array} \tag{5.1}$$

The vectors b_{min} and b_{max} specify the minimum and maximum dose for each voxel. The matrix D is the dose matrix of each discretized position along the needles. Ideally, one would like to meet all the prescription constraints and achieve $Dx \geq b_{min}$ and $Dx \leq b_{max}$, while minimizing the total treatment time $||x||$. However, this is never feasible. To overcome this, slack variables s and t are added to the constraints to ensure feasibility. The weighting variables $\Lambda_{\max}$ and $\Lambda_{\min}$ (both are diagonal square matrices) reflect the “importance” of the various constraints. The goal of the objective function is to minimize the total treatment time while making sure that as few as possible voxels are violating the prescription constraints.

5.3 Results

We focused on the following aspects of the algorithm:

1. Given a specified number of implants, can the algorithm find the optimal trajectories to meet the clinical goals?
2. Can the algorithm reduce the number of implants?
3. We expect future implant procedures will be performed under the guidance of such pre-calculated trajectories. If the pre-calculated implant trajectories are not followed perfectly, is the impact negligible?
4. Can the algorithm obtain homogeneous dose plans? Would there be any trade-offs involved in such plans?

Our experimental results reveal that the answers to all of the above questions are affirmative.

We experimented with our PSO based optimization framework on previously treated prostate cases. Figure 5.2 shows the anatomy of one such case. The target prostate is shown in red. There are three main critical structures. The urethra is shown in blue, the rectum is shown in yellow, and the bladder is shown in green. The prescribed dose for the prostate requires that at least 98% of its volume be covered by 1050 $[cGy]$, where 1 $[cGy] = 0.01$ $[J/Kg]$ is the unit for radiation energy depositions.

The maximum dose tolerance for the urethra is 1150 [cGy], for the rectum is 760 [cGy] and for the bladder is 1050 [cGy].

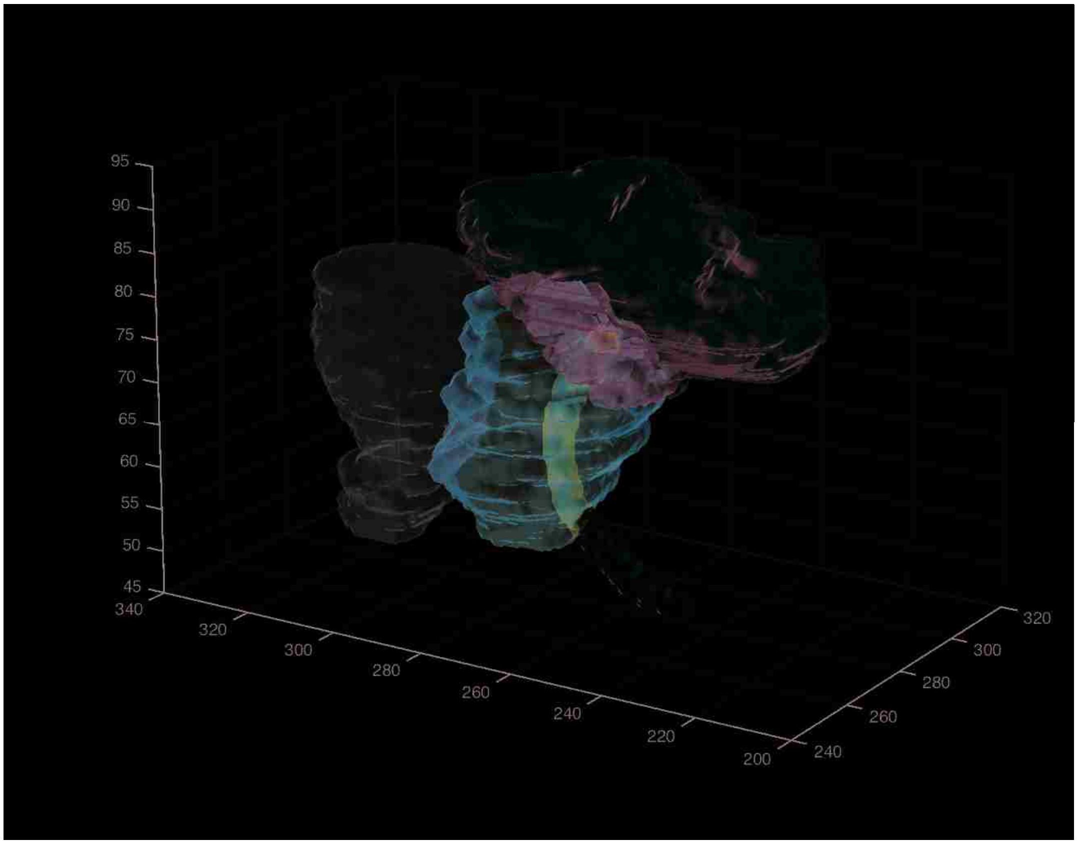


Figure 5.2: Prostate case. Red: prostate. Blue: urethra. Yellow: rectum. Green: bladder. The goal is to cover the prostate with at least 98% of the prescribed dose, 1050 [cGy].

Figure 5.3 (top) shows the final position of the initial swarm, i.e. the white circles, that move constrained to a 2D-plane, i.e., the area marked in red on the cross-section of this case CT-scan. In this figure, the rectum contours are also shown in yellow, the prostate contours in red, and the urethra in blue. Figure 5.3 (bottom) shows the final needle positions computed by our optimization. Needles spread evenly and tend to bend along the prostate's boundary, avoiding the urethra and other critical organs.

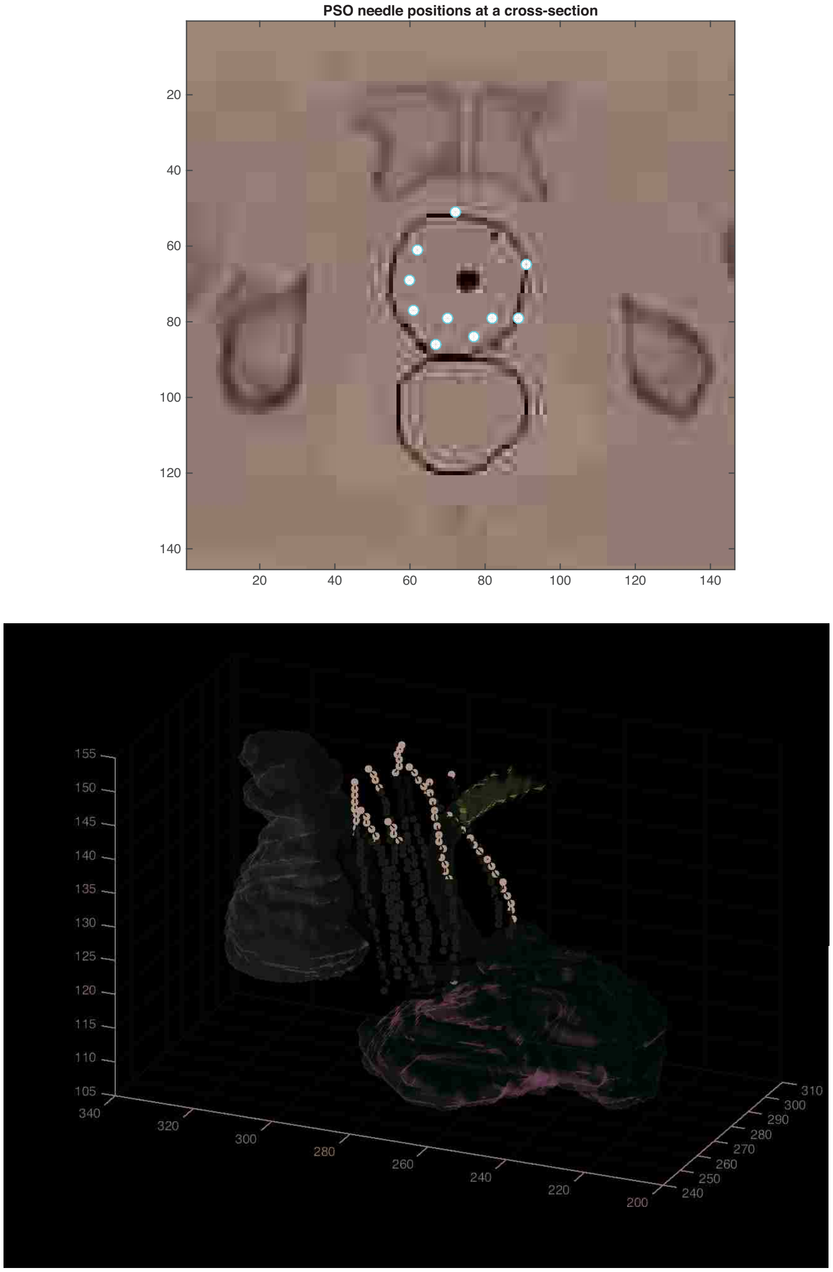


Figure 5.3: Top: Kinetic particles' positions (in white) at a cross-section of the prostate (red contour). The rectum is shown as a yellow contour and the urethra as a blue area. Bottom: Prostate case final needle trajectories in black. Urethra, rectum and bladder are shown in blue, yellow and green. The prostate is shown as a grey shadow.

Figure 5.4 shows the iso-dose surface of the targeted dose calculated by our least-distance programming model. Figure 5.5 shows the dose-volume histogram (DVH) comparisons between our plan and the previously used clinical plan. Recall that the DVH describes the amount of dose delivered to a given percent volume. Take the DVH line for the prostate in figure 5.5 as an example. It denotes that 98% of the prostate receives at least the prescribed 1050 [cGy] of dose. As shown in figure 5.5, our plan achieves all the prescription goals and is comparable to the previously used clinical plan. It is also worth pointing out that our plan is less uniform than the clinical plan. Nonetheless, this could indicate that our plan is in fact superior because these dose inhomogeneities are mainly localized inside the target, which could allow higher cell damage and tumor control [22, 77], while meeting the radiation tolerance constraints for all critical structures.

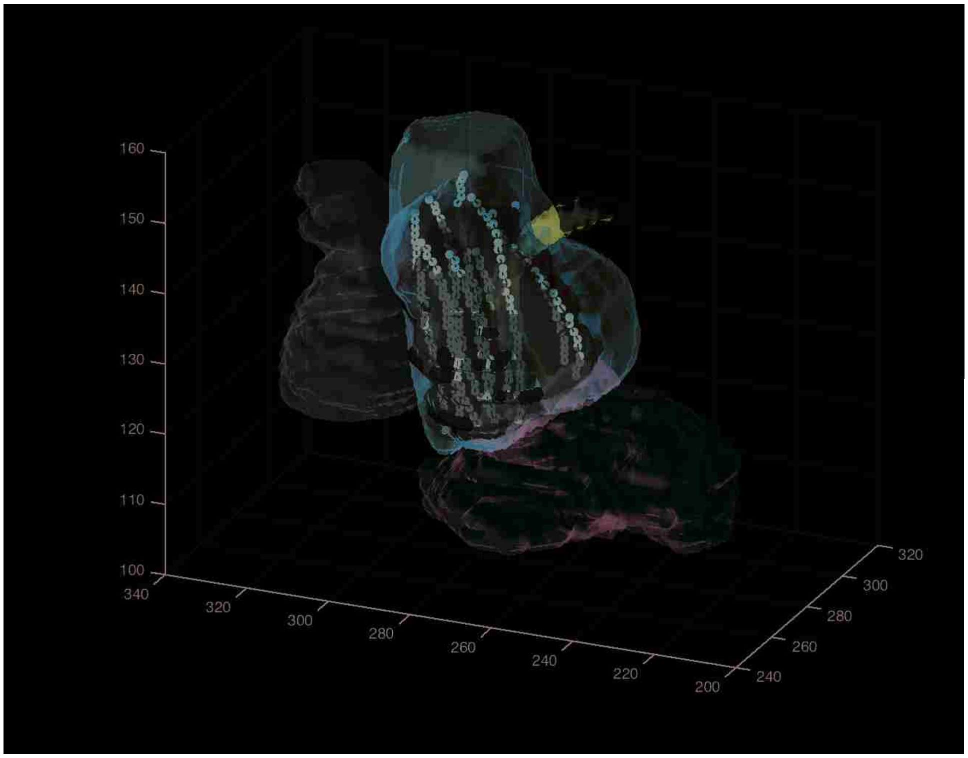


Figure 5.4: Final dose delivered (1050 [cGy] iso-dose surface) to this prostate case by our algorithm (in red). The prostate is shown as a white shadow.

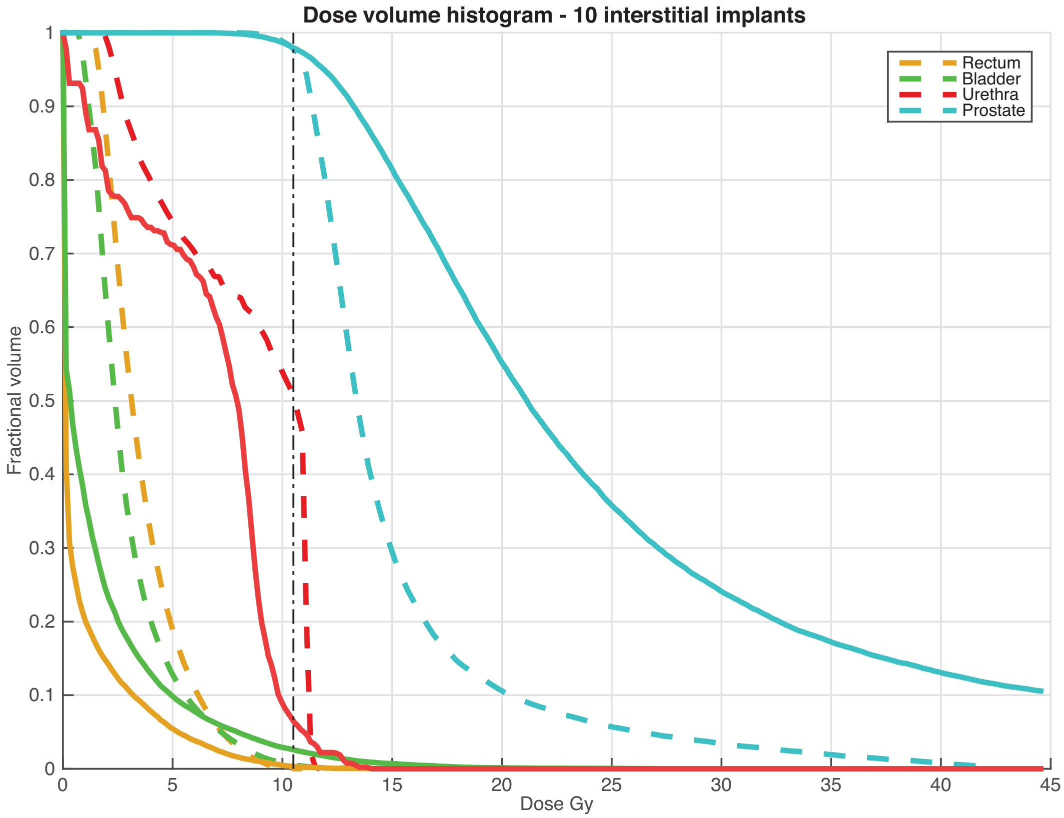


Figure 5.5: (Prostate case DVH comparison. Dashed lines correspond to the clinical plan. Solid lines correspond to our PSO algorithm.

In terms of the number of needles used, our plan is able to calculate different needle setups (see DVH comparisons in figures 5.5, 5.6 (a), 5.6 (b), 5.6 (c), and 5.6 (d) for plans using 10, 12, 14, 16, and 17 needles respectively). Our algorithm has been able to produce high quality plans for all these setups. All prescription goals are achieved and are comparable to the previously used clinical plan. As a result of using less implants, our plans suffer from dose nonuniformity at the target, which is an expected side-effect. In addition, a higher implant density is able to lower dose impact to the urethra, rectum and bladder. Finally, while the clinical plan for this case used 17 needles, our plan has been able to produce high quality plans using 10 needles, which corresponds to a 41% reduction.

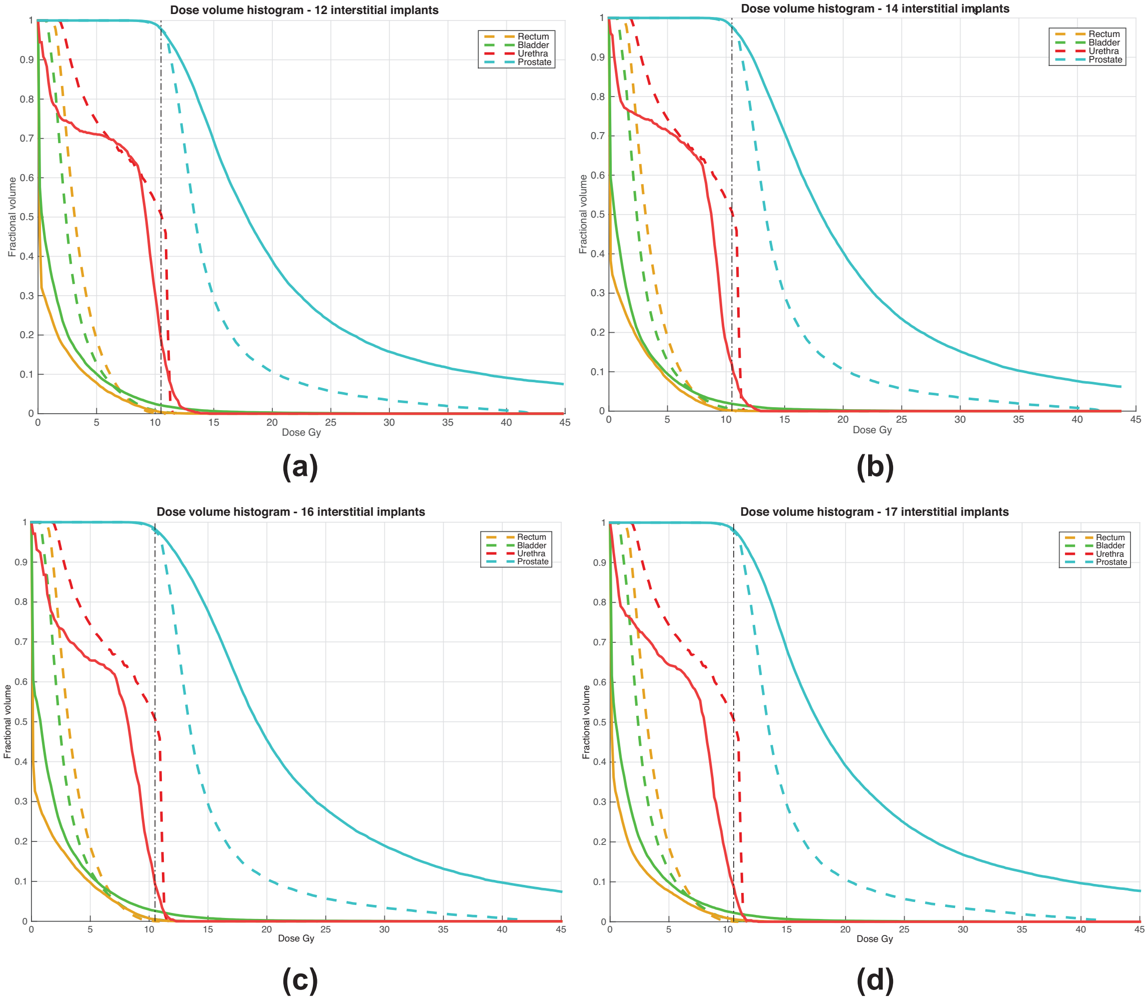


Figure 5.6: (a) Dose volume histogram for 12 interstitial implants. (b) Dose volume histogram for 14 interstitial implants. (c) Dose volume histogram for 16 interstitial implants. (d) Dose volume histogram for 17 interstitial implants.

We expect future implant placement procedures will be performed under the guidance of pre-calculated trajectories. In order to assess the applicability of our generated plans, we randomly perturb the trajectories to mimic the manual implant errors. We generate 30 random modified needle trajectories for 10 implants and compute the final dose distribution using least distance programming. We calculate the average and standard deviations for each bin at each dose volume histogram and show in figure 5.7 the error tolerance after these perturbations. As it can be seen, these perturbations are negligible. The reason is that least-distance programming is able to compensate for these errors.

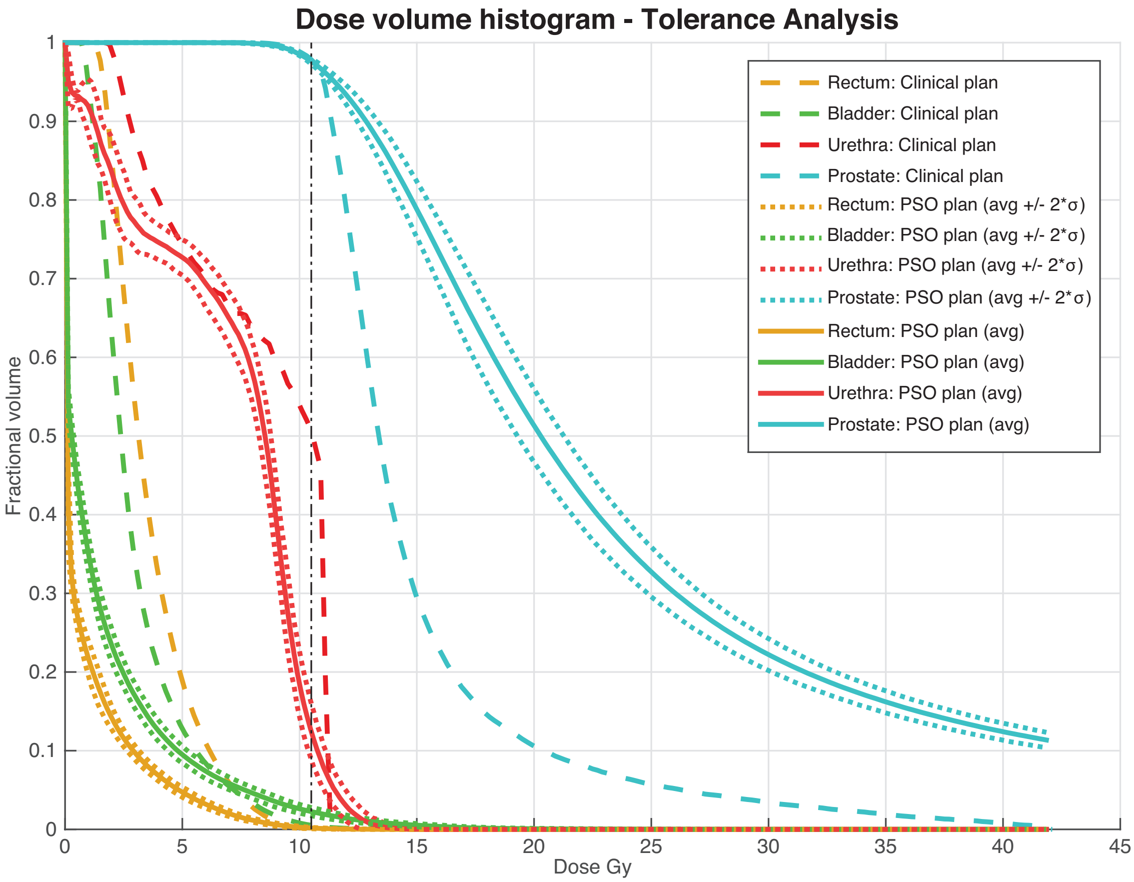


Figure 5.7: Error tolerance. Solid lines show the original plan, while dotted lines show the 95% confidence interval (up to 2σ) of the plans after perturbation with $\sigma = 2.2$mm.

Even though we understand that HDR brachytherapy plans inherently have inhomogeneous dose distributions, due to the nature of the radiation sources that deliver a very high dose to those tissues close to the implants, we wanted to test how versatile our optimization technique was by trying to obtain uniform dose distributions.

Initially, we vary the weighting variables $\Lambda_{\max}$ and $\Lambda_{\min}$ to reflect the importance of lower doses deposited inside the target. We observed that higher dose homogeneity can be obtained, but with a slight change in the objective function of our least-distance programming optimization, i.e.,

$$\begin{array}{ll} \text{minimize} & \lambda\left(||x||_2^2\right) + (1-\lambda)\left(\|\Lambda_{\max} s\|_2^2 + \|\Lambda_{\min} t\|_2^2\right) \\ \text{subject to} & Dx - s \leq b_{\max}, \\ & Dx + t \geq b_{\min}, \\ \text{where} & 0 \leq \lambda \leq 1, \end{array} \tag{5.2}$$

we noticed that better results are achieved. These results are displayed in figures 5.8 and 5.9.

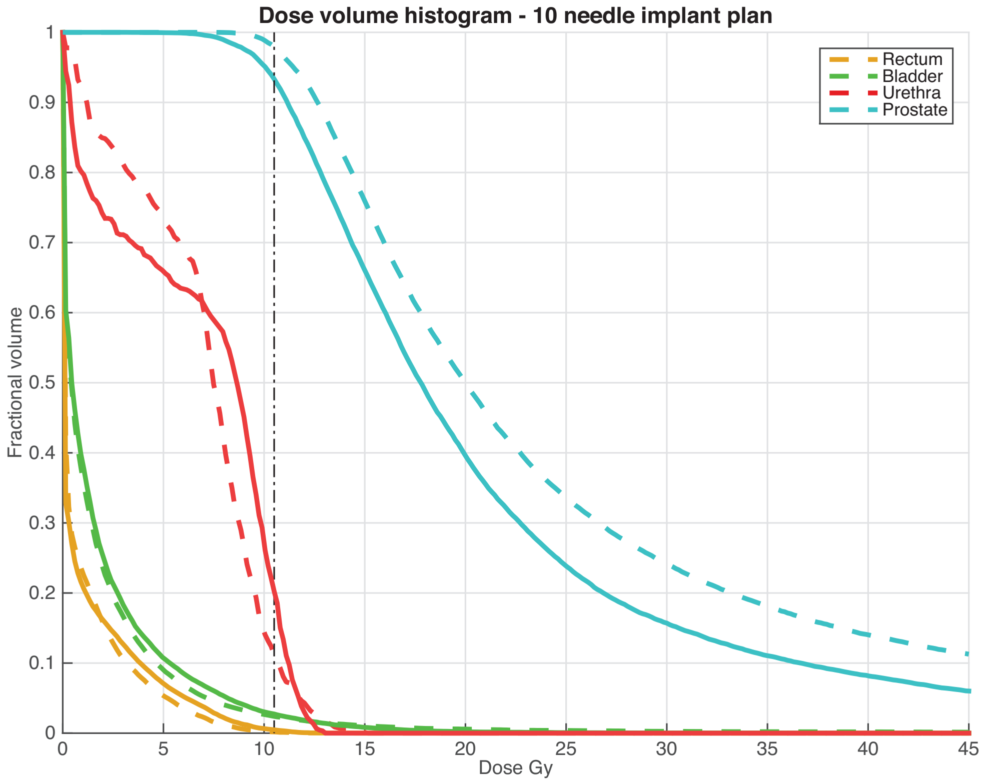


Figure 5.8: Dose Volume Histogram for a 10 interstitial implant plan using the modified objective function compared to our initial calculated 10 interstitial implant plan. Dotted lines correspond to our original plan. Solid lines correspond to our new plan.

Equation 5.2 is the result of realizing that the values in the terms, $\|\Lambda_{\max}s\|_2^2 + \|\Lambda_{\min}t\|_2^2$, were up to three-orders of magnitude bigger than those from $||x||_2^2$. The convex combination $\lambda\,(\cdots) + (1-\lambda)\,(\cdots)$ allows control over these two unrelated parts of our new objective function.

We notice in figures 5.8 and 5.9 that the use of higher implant density improves homogeneity, but since our algorithm is able to generate implant trajectories that bend intelligently along the prostate boundary, the critical organs rectum and bladder always remain underdosed (with respect to the clinical plan), while the urethra suffers a slight overdosage due to a better implant utilization inside the prostate.

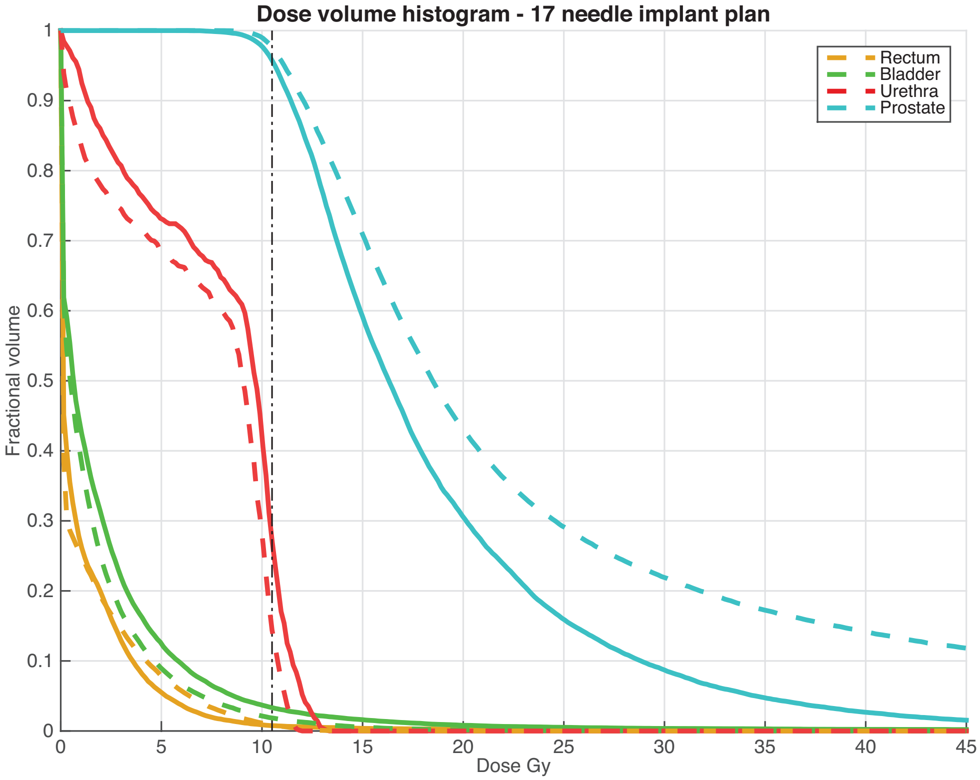


Figure 5.9: Dose Volume Histogram for a 17 interstitial implant plan using the modified objective function compared to our initial calculated 17 interstitial implant plan. Dotted lines correspond to our original plan. Solid lines correspond to our new plan.

Chapter 6

Outlook

> “Here is my secret.
> It is very simple: It is only with the heart that one can see rightly;
> what is essential is invisible to the eye.”
> —Antoine de Saint-Exupéry, The Little Prince.

6.1 Future Work

We have identified six projects: Gamma Knife® core integration, Gamma Knife® non-spherical dose kernels integrations, fully automatic HDR brachytherapy planning and treatment system, the application of our method to particle therapy treatment planning, radiofrequency ablation, and cryotherapy. The first three are a direct consequence of our results and aim to enhance the modalities discussed in this work; the rest are research proposals for a new radiation therapy modality and two localized cancer treatments.

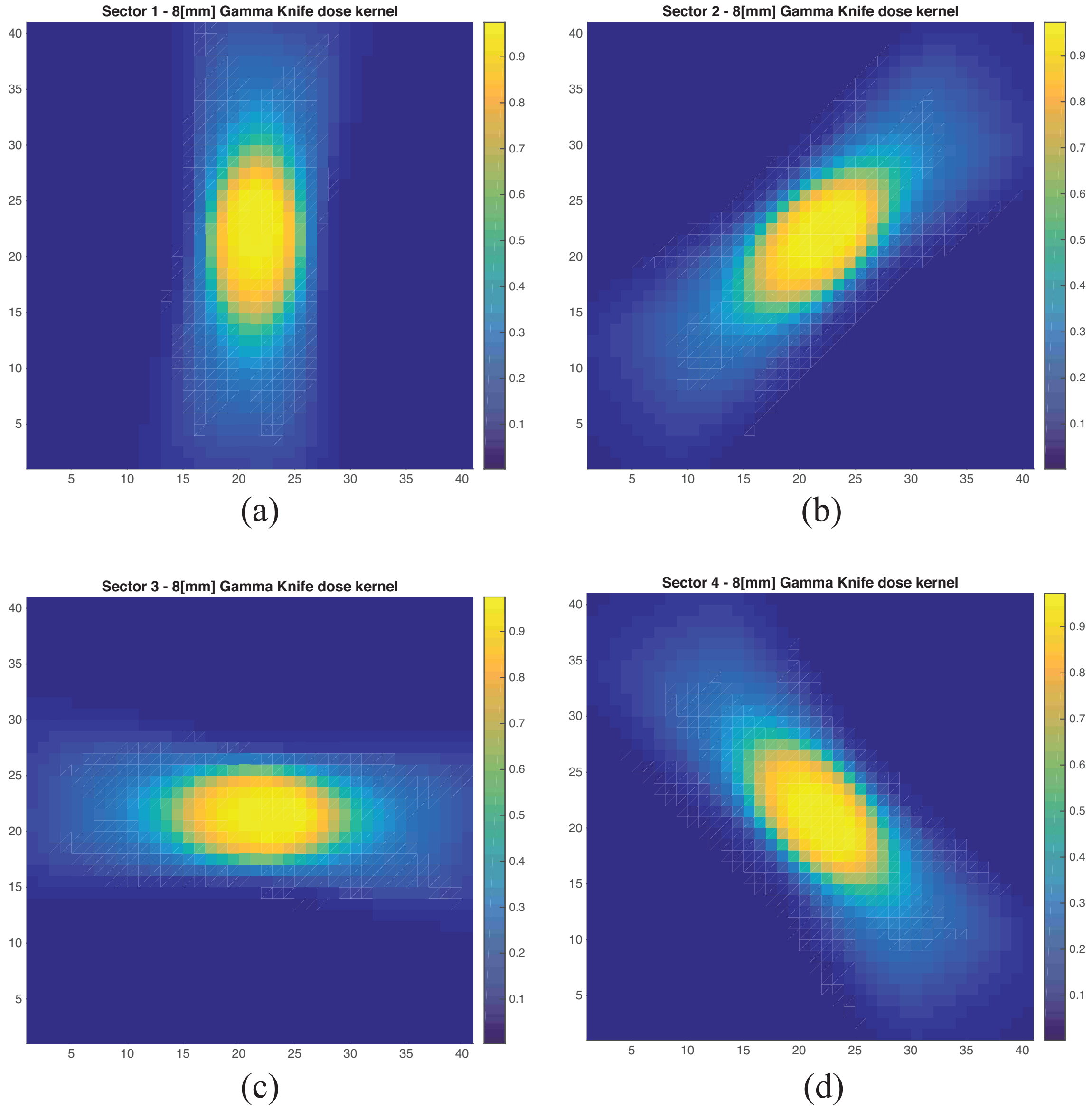
6.1.1 Gamma Knife® core integration

We expect to be able to use previous research agreements with Elekta in order to combine our PSO based Gamma Knife® optimization to the Elekta Gamma Knife® unit. This would grant us access to the kernel of their current treatment planning system. This project will demand the following enhancements to our current prototype:

- Full C++ implementation of our method.
- Calculation of static and kinetic potentials in parallel. The use of graphics processing units (GPUs) is possible due to the great amount of matrix summations.
- Translation from euclidean coordinates to machine coordinates.

6.1.2 Gamma Knife® non-spherical dose kernels integration

We would like to extend the scope of the application of our method to the use of non-spherical dose kernels for Gamma Knife® . We mentioned in section 2.1.2, that when all the collimators of the several ^{60}Co sources are open in the Elekta Gamma Knife® machine, spherical high dose volumes can be generated. Nonetheless, the Elekta Gamma Knife® machine can also produce different non-spherical dose kernels using only some open collimators. In fact, the newer machine versions, e.g., the Leksell Gamma Knife® Perfexion™, is divided into 8-sector locations and can produce 8-sector beam-lets for 4, 8, and 16 [*mm*] kernels. Figure 6.1 shows the center cross-section of these beam-lets for sectors 1, 2, 3, and 4 for the 8 [*mm*] kernel. Notice that these kernels have an ellipsoidal shape.


Sector 1 - 8[mm] Gamma Knife dose kernel

Sector 2 - 8[mm] Gamma Knife dose kernel

Sector 3 - 8[mm] Gamma Knife dose kernel

Sector 4 - 8[mm] Gamma Knife dose kernel

(a) (b)

(c) (d)

Figure 6.1: Center cross-sections of beam-lets for sectors 1, 2, 3, and 4 for the 8 $[mm]$ kernel. Values are normalized and units are arbitrary.

Our intention is to modify our current algorithm to use these beam-lets by restricting the movement a particle with this type of potential function can make. Sector information kernels are particular useful to ensure avoiding critical structures while delivering treatment to a patient. We suspect that restricting both particle movement within a certain iso-potential margin and particle movement within a certain tumor partition/region (according to the closeness to the open collimators) will generate clinically acceptable plans.

6.1.3 Fully automatic HDR brachytherapy planning and treatment system

We mentioned that for HDR brachytherapy treatment, three phases take place: catheter placing, treatment planning, and treatment delivery. From these phases, catheter placing and treatment planning are heavily manual processes where the surgeon is responsible to cover the tumor with "enough" catheters to achieve a good conformal dose to it. Our intention is to combine the current research that tries to move away from these manual processes, e.g., N. Hungr et al. in their paper "Design of an ultrasound-guided robotic brachytherapy needle-insertion system" [44], and our collaborators, J. Zhou et al. in their paper "Real-time catheter tracking for high-dose-rate prostate brachytherapy using an electromagnetic 3D-guidance device: A preliminary performance study" [43] with our algorithm to create the very first fully automatic HDR brachytherapy planning system. Ideally, catheters would be implanted by a robot following the trajectories calculated by our optimization algorithm controlled by a real-time tracking system.

6.1.4 Particle therapy treatment planning

We showcased our optimization algorithm as a general optimization framework for radiation therapy planning. In this project, our intention is to use our optimization technique for particle therapy treatment planning. In this section, we discuss in a broader detail our initial insights of how to map our method to this new modality.

Particle therapy

The therapeutic use of heavier charged particles was first proposed by Robert Wilson, in his 1946 paper: "Radiological Use of Fast Protons", where he described the potential benefits of accelerated protons for human radiation therapy [50, 78]. His principal argument was based on the depth-dose profile of a proton beam (see figure 6.3), which compared to that of photons exhibits a low entrance dose, a high dose concentration at some depth (Bragg peak), and a steep dose fall-out after the peak. X-rays or photon depth-dose profiles are characterized by a high entrance dose and a long logarithmic decrease dose tail (see figure 6.2). Since Wilson's early studies, further research has provided better insights into the clinical benefits of charged particles and their higher radiobiological effectiveness (RBE) compared to that of photons [79, 80, 81, 82, 83] which has led to the construction of big infrastructures to treat cancer patients.

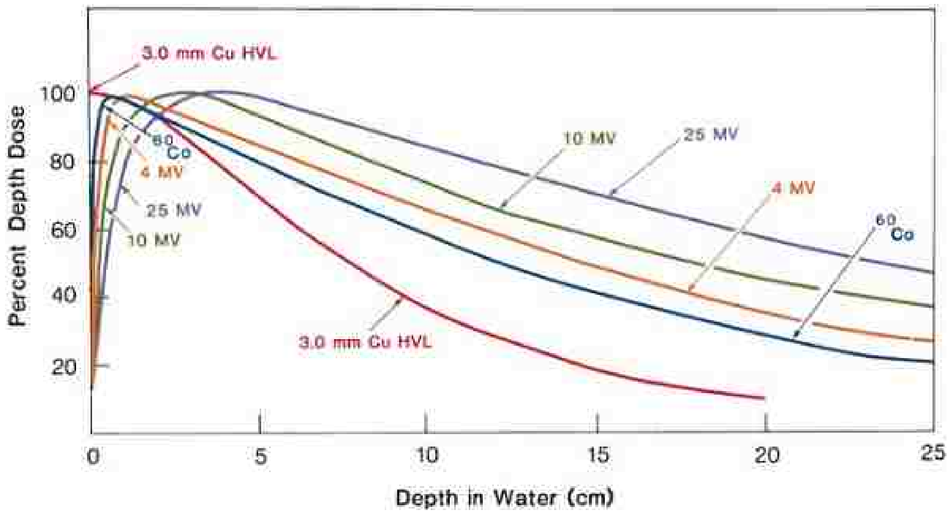


Figure 6.2: Photon beams - depth-dose diagram [2]. Dose deposition along the photon path in water are shown for photon energies of 25 MeV, 10 MeV, 4 MeV, 1.33 MeV (i.e., ^{60}Co), and 511 keV (i.e., 3.0 mm Cu Half Value Layer)

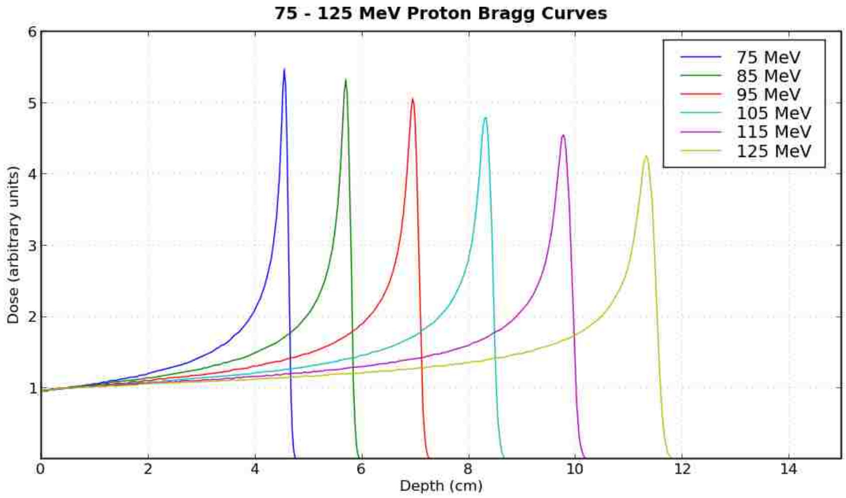


Figure 6.3: Proton beams - depth-dose diagram. Dose deposition along the path of proton beams in water are shown for proton beam energies of 75, 85, 95, 105, 115, and 125 MeV.

The current most advanced delivery techniques for particle therapy is active scanning. In active scanning particle therapy, the target is first partitioned into layers so that each layer can be reached using the same energy. For each layer, the beam spot (ranging from a few millimeters in diameter to a few centimeters) scans through the layer as shown in Figure 6.4) to deliver the prescribed dose.

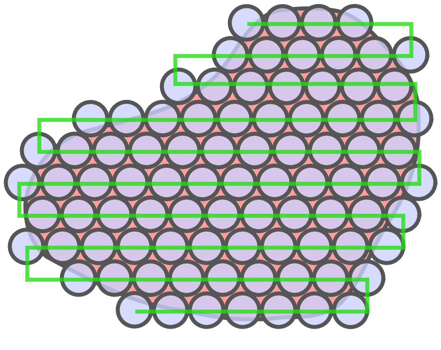


Figure 6.4: Scanning beam therapy: *circles* represent a targeted tumor voxel, and *line connecting circles* represents the scanning path used to “paint” the tumor.

For readers interested in the treatment planning and dose calculations of particle therapy, we refer them to the paper by Krämer et. al.: "Treatment planning for heavy-ion radiotherapy: physical beam model and dose optimization" [84].

Mapping

Notice that the current delivery technique for particle therapy, active scanning (see figure 6.4), can be seen as a bin-packing problem. Basically, a proton or heavy-ion particle beam can be seen as a high-dose volume shot packed around the Bragg Peak. Our initial idea is to keep the *patient representation* invariant from the previous radiation therapy modalities. For *preparation for fose calculation*, we think that a similar approach to the one used for Gamma Knife® radiosurgery would work to pack shots up-to a iso-dose contour. Then for *optimization*, we can follow a similar approach to the one used for Gamma Knife® but in addition, we plan to connect all particles that lie on an iso-potential surface constrained by the possible treatment angles in order to produce a continuous delivery plan that minimizes the energy changes.

6.1.5 Radiofrequency ablation

Radiofrequency ablation is a medical procedure that can be used to treat localized tumors. It consists on the use of the heat produced by a high frequency alternating current that is delivered through a catheter-like mechanism to the tumor with the intention to destroy (i.e., burn) it. Similarly to the Gamma Knife® application, the heat produced by this mechanism is highly concentrated and can be applied at different radii. The mapping of our algorithm would occur similarly as in the Gamma Knife® provided that suitable access to all regions of the tumor is available.

6.1.6 Cryotherapy

Cryotherapy ablation is a medical procedure that can be used to treat localized tumors. It consists on the use of extreme cold produced by liquid nitrogen or argon gas delivered to a tumor through a cryoprobe (i.e., a hollow catheter). Similarly to the Gamma Knife® application, the extreme cold forms a ball of ice that freezes nearby tumor cells killing them. The mapping of our algorithm would occur similarly as in the Gamma Knife® provided that suitable access to all regions of the tumor is available; otherwise, the mapping for HDR brachytherapy could produce suitable trajectories for cryoprobes.

6.2 Summary and Conclusions

Modern radiation therapy heavily depends on both mathematical optimization and computer models in order to find and calculate radiation source configurations that deliver a uniform high dose to a tumor while protecting all healthy tissues around.

We worked on two world widely used modalities: Gamma Knife® radiosurgery and High-Dose Rate brachytherapy for prostate cancer. We were originally introduced to the current limitations in HDR brachytherapy for prostate cancer: manual implant placement, inefficient implant trajectories and high number of implants used. As a consequence, we developed optimization algorithms and techniques, which we have shown to be general for those radiation therapy modalities that use high-dose volumes to cover a tumor. As an example of the general characteristics of our algorithms and techniques, we applied them to Gamma Knife® radiosurgery.

We implemented a general optimization method that can be used to find the optimal locations of radiation sources in radiation therapy inverse planning using a combination of Particle Swarm Optimization framework and deterministic optimization (non-negative least squares, least distance programming, etc.) in a constrained geometrical environment. Our algorithm consists of placing a uniform potential function on the surface of tumors and critical organs so that kinetic charged particles inside the tumor could be aware of their geometrical environment.

We observed that our inverse planning approach allows us to optimally find radiation sources' locations for Gamma Knife® radiosurgery. We use the Elekta Gamma Knife® spherical dose kernels mapped to kinetic particles with a $\frac{1}{r}$ potential function. The final configuration of the swarm of particles is further filtered using non-negative squares. We noticed that for heterogeneous and homogeneous kernel spot sizes, our algorithm returns high quality plans that are clinically acceptable. Nonetheless, our method requires higher control statements to enhance the quality of the heterogeneous kernel spot sizes results. We strongly believe that the techniques suggested in section 6.1.2 will not only handle the use of non-spherical kernels but also will handle heterogeneous kernel spot sizes increasing the current plans quality.

HDR Brachytherapy for prostate cancer inverse planning is undoubtedly a very challenging optimization problem. We use our general method to generate optimal implant trajectories, reduce the number of implants used, and provide optimal conformal dose delivered to the prostate while minimizing the impact on the urethra, rectum and bladder. To the best of our knowledge, we have implemented the first inverse planning system for HDR brachytherapy for prostate cancer that provides optimal implant trajectories for minimal number of implants (in fact, our algorithm provides a natural path to select the best possible plan using any number of implants within the prescribed dose constraints). In addition, we have proposed the use of least-distance programming to find optimal conformal dose distribution to the prostate while meeting all prescribed goals.

We compare the results of our simulations against the best clinical practices for both Gamma Knife® radiosurgery and HDR brachythreapy and we claim that our optimization approach is able to achieve high-quality treatment plans which are clinically superior to those from commercial clinical treatment planning systems.

We recognize the potential of our general optimization algorithm and expect in the near future to incorporate it inside current clinical treatment planning systems for Gamma Knife® radiosurgery and HDR brachytherapy.

References

[1] D. Schardt, T. Elsässer, and D. Schulz-Ertner, *Heavy-ion tumor therapy: Physical and radiobiological benefits*, Rev. Mod. Phys. **82** (2010), 383–425.

[2] F. M. Khan, *The Physics of Radiation Therapy*, Lippincott Williams & Wilkins, 2010.

[3] ELEKTA *Elekta gamma knife radio-surgery dvd*, 2008.

[4] L. Gunderson and J. Tepper, *Clinical Radiation Oncology*, Elsevier Health Sciences, 2011.

[5] R. Nath, L. L. Anderson, G. Luxton, K. A. Weaver, J. F. Williamson, and A. S. Meigooni, *Dosimetry of interstitial brachytherapy sources: recommendations of the aapm radiation therapy committee task group no. 43. american association of physicists in medicine*, Med Phys **22** (1995), 209–34.

[6] D. Granero, J. Vijande, F. Ballester, and M. J. Rivard, *Dosimetry revisited for the hdr 192ir brachytherapy source model mhdr-v2*, Med Phys **38** (2011), 487–94.

[7] K. Parsopoulos and M. Vrahatis, *Particle Swarm Optimization and Intelligence: Advances and Applications*, Premier reference source, Information Science Reference, 2010.

[8] D. J. Griffiths, *Introduction to Elementary Particles*, Wiley, John & Sons, Inc., 1987.

[9] P. Mayles, A. Nahum, and J. Rosenwald, *Handbook of Radiotherapy Physics: Theory and Practice*, CRC Press, 2007.

[10] J. Cassidy, D. Bissett, and R. Spence, *Oxford Handbook of Oncology*, Oxford Medical Handbooks, OUP Oxford, 2002.

[11] K. Krane, *Modern Physics*, Modern Physics, Wiley, 2012.

[12] W. Wieszczycka and W. H. Scharf, *Proton Radiotherapy Accelerators*, World Scientific, 2001.

[13] A. C. Society, *Cancer facts & figures 2015*, Atlanta: American Cancer Society.

[14] ———, *Cancer facts & figures 2014*, Atlanta: American Cancer Society.

[15] M. Goitein, *Radiation Oncology: A Physicist's-Eye View*, Biological and Medical Physics Series, Springer, 2010.

[16] G. Delaney, S. Jacob, C. Featherstone, and M. Barton, *The role of radiotherapy in cancer treatment*, Cancer **104** (2005), 1129–1137.

[17] R. Souhami and J. Tobias, *Cancer and its Management*, Wiley, 2008.

[18] J. Tobias and D. Hochhauser, *Cancer and Its Management*, Wiley, 2009.

[19] M. Tubiana, *The role of local treatment in the cure of cancer.*, European journal of cancer (Oxford, England : 1990) **28A** (1992), 2061–9.

[20] M. Joiner and A. van der Kogel, *Basic Clinical Radiobiology: Fourth Edition*, Hodder Arnold, 2009.

[21] J. Cox and K. Ang, *Radiation Oncology: Rationale, Technique, Results*, Elsevier Health Sciences, 2009.

[22] A. Dicker, G. Merrick, L. Gomella, R. Valicenti, and F. Waterman, *Basic and Advanced Techniques in Prostate Brachytherapy*, Taylor & Francis, 2005.

[23] F. H. Attix, *Introduction to Radiological Physics and Radiation Dosimetry*, WILEY-VCH, 2004.

[24] B. J. Slotmann, T. D. Solberg, and D. Verellen, *Extracranial Stereotactic Radiotherapy and Radiosurgery*, Taylor & Francis, 2006.

[25] J. C. Flickinger, L. Dadeˆ Lunsford, A. Wu, A. H. Maitz, and A. M. Kalend, *Treatment planning for gamma knife radiosurgery with multiple isocenters*, International journal of radiation oncology, biology, physics **18** (1990), 1495–1501.

[26] C. for Disease Control, D. o. C. P. Prevention, and Control *Prostate cancer statistics*, 2010.

[27] E. Lessard and J. Pouliot, *Inverse planning anatomy-based dose optimization for HDR-brachytherapy of the prostate using fast simulated annealing algorithm and dedicated objective function*, Medical Physics **28** (2001), 773–779.

[28] R. A. Novelline, *Squire's Fundamentals of Radiology: Sixth Edition*, Harvard University Press, 2004.

[29] H. Becquerel, *Sur les radiations émises par phosphorescence (on the invisible rays emitted by phosphorescent bodies)*, Comptes rendus de l'Académie des Sciences **122** (1896), 420–421.

[30] M. P. C. P. Curie and G. Bémont, *Sur une nouvelle substance fortement radioactive, contenue dans la pechblende)*, Comptes rendus de l'Académie des Sciences **127** (1898), 1215–1217.

[31] J. M. H. M. L. V. A. Gerbaulet, R. Potter, *The GEC ESTRO handbook of brachytherapy. European Society of Therapeutic Radiology and Oncology*, Brussels: 2002.

[32] D. Baltas, L. Sakelliou, and N. Zamboglou, *The Physics of Modern Brachytherapy for Oncology*, Series in Medical Physics and Biomedical Engineering, Taylor & Francis, 2010.

[33] E. J. Hall and A. J. Giaccia, *Radiobiology for the Radiologist*, Lippincott Williams & Wilkins, 2006.

[34] P. M. Pardalos and H. Romeijn, *Handbook of Optimization in Medicine*, Springer, 2008.

[35] Z. Chen *Dynamic photon painting* Master's thesis, University of New Mexico, 2010.

[36] S. Luan, N. Swanson, Z. Chen, and L. Ma, *Dynamic gamma knife radiosurgery*, Physics in Medicine and Biology **54** (2009), 1579.

[37] D. M. Shepard, M. C. Ferris, R. Ove, and L. Ma, *Inverse treatment planning for Gamma Knife radiosurgery*, Medical Physics **27** (2000), 2748–2756.

[38] J. C. Flickinger, D. Kondziolka, L. D. Lunsford, R. J. Coffey, M. L. Goodman, E. G. Shaw, W. R. Hudgins, R. Weiner, G. R. H. IV, P. K. Sneed, and D. A. Larson, *A multi-institutional experience with stereotactic radiosurgery for solitary brain metastasis*, International Journal of Radiation Oncology*Biology*Physics **28** (1994), 797 – 802.

[39] D. M. Shepard, M. C. Ferris, R. Ove, and L. Ma, *Inverse treatment planning for Gamma Knife radiosurgery*, Medical Physics **27** (2000), 2748–2756.

[40] ELEKTA, *Inverse planning in leksell gammaplan® 10 - white paper.*

[41] J. Dutreix, M. Tubiana, and B. Pierquin, *The hazy dawn of brachytherapy*, Radiother Oncol **49** (1998), 223–32.

[42] M. Tubiana, *The role of local treatment in the cure of cancer*, Eur J Cancer **28A** (1992), 2061–9.

[43] J. Zhou, E. Sebastian, V. Mangona, and D. Yan, *Real-time catheter tracking for high-dose-rate prostate brachytherapy using an electromagnetic 3D-guidance device: A preliminary performance study*, Medical Physics **40** (2013), 021716+.

[44] N. Hungr, J. Troccaz, N. Zemiti, and N. Tripodi, *Design of an ultrasound-guided robotic brachytherapy needle-insertion system*, in Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, Sept 2009, pp. 250–253.

[45] G. M. Daskalov, E. Löffler, and J. F. Williamson, *Monte carlo-aided dosimetry of a new high dose-rate brachytherapy source*, Med Phys **25** (1998), 2200–8.

[46] M. J. Rivard, *Refinements to the geometry factor used in the aapm task group report no. 43 necessary for brachytherapy dosimetry calculations. american association of physicists in medicine*, Med Phys **26** (1999), 2445–50.

[47] M. J. Rivard, B. M. Coursey, L. A. DeWerd, W. F. Hanson, M. S. Huq, G. S. Ibbott, M. G. Mitch, R. Nath, and J. F. Williamson, *Update of aapm task group no. 43 report: A revised aapm protocol for brachytherapy dose calculations*, Med Phys **31** (2004), 633–74.

[48] M. J. Rivard, W. M. Butler, L. A. DeWerd, M. S. Huq, G. S. Ibbott, A. S. Meigooni, C. S. Melhus, M. G. Mitch, R. Nath, J. F. Williamson, and American Association of Physicists in Medicine, *Supplement to the 2004 update of the aapm task group no. 43 report*, Med Phys **34** (2007), 2187–205.

[49] S.-T. Chiu-Tsao, D. R. Schaart, C. G. Soares, R. Nath, and AAPM Therapy Physics Committee Task Group No 149, *Dose calculation formalisms and consensus dosimetry parameters for intravascular brachytherapy dosimetry: recommendations of the aapm therapy physics committee task group no. 149*, Med Phys **34** (2007), 4126–57.

[50] T. F. DeLaney and H. M. Kooy, *Proton and Charged Particle Radiotherapy*, Lippincott Williams & Wilkins, 2008.

[51] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, SIAM, 1995.

[52] J. Nocedal and S. Wright, *Numerical Optimization*, Springer series in operations research and financial engineering, Springer, 1999.

[53] S. Mikki and A. Kishk, *Particle Swarm Optimization: A Physics-based Approach*, Synthesis lectures on computational electromagnetics, Morgan & Claypool, 2008.

[54] S. Rao and S. Rao, *Engineering Optimization: Theory and Practice*, Wiley, 2009.

[55] R. Sarker, M. Mohammadian, and X. Yao, *Evolutionary optimization*, Vol. 48, Springer Science & Business Media, 2002.

[56] T. Weise, *Global Optimization Algorithms – Theory and Application*, http://www.it-weise.de, 2009.

[57] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[58] J. Kennedy and R. Eberhart, *Particle swarm optimization*, in Neural Networks, 1995. Proceedings., IEEE International Conference on, Vol. 4, Nov 1995, pp. 1942–1948 vol.4.

[59] (A. Lazinica, ed.), *Particle Swarm Optimization*, InTech, January 2009.

[60] M. Clerc, *Particle Swarm Optimization*, ISTE, Wiley, 2010.

[61] S. Kiranyaz, T. Ince, and M. Gabbouj, *Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition*, Adaptation, Learning, and Optimization, Springer-Verlag New York Incorporated, 2013.

[62] R. Eberhart and Y. Shi, *Particle swarm optimization: developments, applications and resources*, in Evolutionary Computation, 2001. Proceedings of the 2001 Congress on, Vol. 1, 2001, pp. 81–86 vol. 1.

[63] S. Ebbesen, P. Kiwitz, and L. Guzzella, *A generic particle swarm optimization matlab function*, in American Control Conference (ACC), 2012, June 2012, pp. 1519–1524.

[64] A. Banks, J. Vincent, and C. Anyakoha, *A review of particle swarm optimization. part i: background and development*, Natural Computing **6** (2007), 467–484.

[65] T. Blackwell and J. Branke, *Multi-swarm optimization in dynamic environments*, in Applications of Evolutionary Computing, (G. Raidl, S. Cagnoni, J. Branke, D. Corne, R. Drechsler, Y. Jin, C. Johnson, P. Machado, E. Marchiori, F. Rothlauf, G. Smith, and G. Squillero, eds.), Vol. 3005 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2004, pp. 489–500.

[66] T. M. Blackwell and P. J. Bentley, *Dynamic search with charged swarms*, in Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '02, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002, pp. 19–26.

[67] A. Banks, J. Vincent, and C. Anyakoha, *A review of particle swarm optimization. part ii: hybridisation, combinatorial, multicriteria and constrained optimization, and indicative applications*, Natural Computing **7** (2008), 109–124.

[68] I. Aljarah and S. Ludwig, *Parallel particle swarm optimization clustering algorithm based on mapreduce methodology*, in Nature and Biologically Inspired Computing (NaBIC), 2012 Fourth World Congress on, Nov 2012, pp. 104–111.

[69] M. Jin and H. Lu, *Parallel particle swarm optimization with genetic communication strategy and its implementation on gpu*, in Cloud Computing and Intelligent Systems (CCIS), 2012 IEEE 2nd International Conference on, Vol. 01, Oct 2012, pp. 99–104.

[70] M. Clerc and J. Kennedy, *The particle swarm - explosion, stability, and convergence in a multidimensional complex space*, Evolutionary Computation, IEEE Transactions on **6** (2002), 58–73.

[71] Y. Li, D. Yao, and W. Chen, *Adaptive particle swarm optimizer for beam angle selection in radiotherapy planning*, in Mechatronics and Automation, 2005 IEEE International Conference, Vol. 1, July 2005, pp. 421–425 Vol. 1.

[72] J. O. Deasy *Cerr: A computational environment for radiotherapy research*, 2007.

[73] M. Grant and S. Boyd *CVX: Matlab software for disciplined convex programming, version 2.1* `http://cvxr.com/cvx`, 2014.

[74] J. O. Deasy *Cerr: A computational environment for radiotherapy research - google forum*, 2015.

[75] DICOM *Dicom - website*, 2015.

[76] E. Andersen, C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming **95** (2003), 249–277.

[77] C. Ling, J. Roy, N. Sahoo, K. Wallner, and L. Anderson, *Quantifying the effect of dose inhomogeneity in brachytherapy: Application to permanent prostatic implant with 125i seeds*, International Journal of Radiation Oncology*Biology*Physics **28** (1994), 971 – 977.

[78] R. R. Wilson, *Radiological Use of Fast Protons*, Radiology **47** (1946), 487+.

[79] N. Bassler, O. Jäkel, C. S. Sondergaard, and J. B. Petersen, *Dose- and let-painting with particle therapy.*, Acta Oncol **49** (2010), 1170–6.

[80] J. J. Wilkens and U. Oelfke, *Three dimensional let calculations for treatment planning of proton therapy.*, Z Med Phys **14** (2004), 41–6.

[81] J. Wilkens and U. Oelfke, *Analytical linear energy transfer calculations for proton therapy.*, Med Phys **30** (2003), 806–15.

[82] J. J. Wilkens and U. Oelfke, *A phenomenological model for the relative biological effectiveness in therapeutic proton beams*, Physics in Medicine and Biology **49** (2004), 2811.

[83] C. Grassberger, A. Trofimov, A. Lomax, and H. Paganetti, *Variations in linear energy transfer within clinical proton therapy fields and the potential for biological treatment planning*, International Journal of Radiation Oncology*Biology*Physics **80** (2011), 1559 – 1566.

[84] M. Krämer, O. Jäkel, T. Haberer, G. Kraft, D. Schardt, and U. Weber, *Treatment planning for heavy-ion radiotherapy: physical beam model and dose optimization*, Physics in Medicine and Biology **45** (2000), 3299–3317.